A Method for Removing Effects of Nonspecific Binding on the Distribution of Binding Stoichiometries: Application to Mass Spectroscopy Data

Liat Shimon,[†] Michal Sharon,[†*] and Amnon Horovitz[‡*]
[†]Department of Biological Chemistry and [‡]Department of Structural Biology, Weizmann Institute of Science, Rehovot, Israel

ABSTRACT There is often an interest in knowing, for a given ligand concentration, how many protein molecules have one, two, three, etc. ligands bound in a specific manner. This is a question that cannot be addressed using conventional ensemble techniques. Here, a mathematical method is presented for separating specific from nonspecific binding in nonensemble studies. The method provides a way to determine the distribution of specific binding stoichiometries at any ligand concentration when using nonensemble (e.g., single-molecule) methods. The applicability of the method is demonstrated for ADP binding to creatine kinase using mass spectroscopy data. A major advantage of our method, which can be applied to any protein-ligand system, is that no previous information regarding the mechanism of ligand interaction is required.

INTRODUCTION

Ligand-protein interactions are the major driving force for many cellular processes. Determining the equilibrium constants and energetics of such interactions is therefore of fundamental importance. Traditional biochemical methods can provide reliable estimates of affinities, but they cannot reveal distributions of binding stoichiometries, which can be valuable for deciphering reaction mechanisms. However, such distributions can be extracted using the more modern, single-molecule (1) and mass spectroscopy (MS) techniques (2–4). In recent years, MS has proven to be a very valuable tool for characterizing noncovalent interactions (5–13). The power of this approach originates from its high sensitivity, high mass accuracy, low sample requirements and speed of analysis.

Analysis of binding reactions must often take into account the issue of nonspecific binding. Classical biochemical methods can deal with this issue in a fairly straightforward manner (14), but for single-molecule and MS techniques it is a major problem. In the case of MS, nonspecific binding can occur throughout the electrospray process during the desolvation of droplets, leading to an increase in ligand concentration. Consequently, the protein/ligand ratio is altered and the gas-phase measurements no longer reflect the true stoichiometries (15). Several approaches have been suggested for overcoming the problem of nonspecific binding in MS studies (16–21). For example, controlled dissociation of nonspecific gas-phase interactions by blackbody infrared radiation was utilized by Wang et al. (21), but complex dissociation may be induced when the nonspecific binding is strong. Another suggested strategy involves the addition of a reference protein to monitor the appearance of nonspecific complexes (18,19). In this method, there is uncertainty regarding the generality of the approach (20), and the ionization efficiency of the target protein may be compromised, especially in the case of large protein complexes. Finally, van der Rest and co-workers have suggested a mathematical model for distinguishing between specific and nonspecific binding (17), but in their approach, noncooperative binding of the ligand was assumed. Given that nonspecific binding is a problem not only in MS studies, we were motivated to develop a new method for dealing with this issue that is simple, straightforward, and of general applicability.

THEORY

Our analysis of specific versus nonspecific binding is general in the sense that it does not assume any particular reaction mechanism and is applicable for studying weak complexes. Two assumptions made are that 1), the number of specific binding sites, N_s, is known, and 2), the nonspecific binding is noncooperative and can be described by a single binding constant. The first step in the analysis involves extracting the value of the nonspecific binding constant from the ratio of the areas of the peaks (referred to as intensities in what follows) corresponding to binding numbers that are larger than N_s. For example, in the case of a dimer with two specific binding sites, we can extract the value of the nonspecific association binding constant, K_n, from the intensities I_3 and I_4, corresponding to populations with three and four bound ligand molecules, respectively. The ratio I_4/I_3 for this example is given by

$$\frac{I_4}{I_3} = \frac{[E]K_1K_2K_n^2[S]^4 + [E]K_1K_n^3[S]^4 + [E]K_n^4[S]^4}{[E]K_1K_2K_n^1[S]^3 + [E]K_1K_n^2[S]^3 + [E]K_n^3[S]^3} = K_n[S], \quad (1)$$

where K_1 and K_2 are the specific association binding constants, [S] is the free ligand concentration, and [E] is

Submitted April 9, 2010, and accepted for publication June 24, 2010.

*Correspondence: michal.sharon@weizmann.ac.il or amnon.horovitz@weizmann.ac.il

Editor: Kathleen B. Hall.

0006-3495/10/09/1645/5 $2.00

doi: 10.1016/j.bpj.2010.06.062

the free protein concentration. The nonspecific binding constant can be evaluated at different [S] values and for different charge states.

Given the value of the nonspecific binding constant from the intensities corresponding to binding numbers that are larger than N_s, we can now determine the values of the specific binding constants from the other peaks in the spectra. The ratio of the intensities I_1 and I_0, corresponding to populations with one and zero bound ligand molecules, respectively, is given by

$$\frac{I_1}{I_0} = \frac{[\mathrm{E}]K_1[\mathrm{S}] + [\mathrm{E}]K_n[\mathrm{S}]}{[\mathrm{E}]} = K_n[\mathrm{S}] + K_1[\mathrm{S}]. \quad (2)$$

Hence, $K_1[\mathrm{S}] = (I_1 - I_0K_n[\mathrm{S}])/I_0$. Likewise, the ratio of the intensities I_2 and I_0, corresponding to populations with two and zero bound ligand molecules, respectively, is given by

$$\frac{I_2}{I_0} = K_1K_2[\mathrm{S}]^2 + K_n[\mathrm{S}](K_n[\mathrm{S}] + K_1[\mathrm{S}]). \quad (3)$$

Hence, by combining Eqs. 2 and 3, we get $K_1K_2[\mathrm{S}]^2 = (I_2 - I_1K_n[\mathrm{S}])/I_0$. It can be seen, therefore, that the population of protein species that have ligand bound only to the specific sites can be determined from the appropriate measured intensities and the value of the nonspecific binding constant.

The protein species with ligand bound also to nonspecific sites are considered next. For example, in the case where intensities corresponding to binding numbers from 0 to 3 are observed, the total concentration of species with one ligand bound at a specific site, C_1, is equal to:

$$C_1 = [\mathrm{E}]K_1[\mathrm{S}] + [\mathrm{E}]K_1K_n[\mathrm{S}]^2 + [\mathrm{E}]K_1K_n^2[\mathrm{S}]^3. \quad (4)$$

It can be readily shown that $C_1/[\mathrm{E}]$ can be expressed as

$$C_1/[\mathrm{E}] = \left(\frac{I_1 - K_n[\mathrm{S}]I_0}{I_0}\right) + \left(\frac{K_n[\mathrm{S}]I_1 - K_n^2[\mathrm{S}]^2I_0}{I_0}\right) + \left(\frac{K_n^2[\mathrm{S}]^2I_1 - K_n^3[\mathrm{S}]^3I_0}{I_0}\right), \quad (5)$$

where the first, second, and third terms in parentheses correspond to $K_1[\mathrm{S}]$, $K_1K_n[\mathrm{S}]^2$, and $K_1K_n^2[\mathrm{S}]^3$, respectively. In general, in the case where intensities corresponding to binding numbers from zero to α are observed, one can express $C_1/[\mathrm{E}]$ as

$$C_1/[\mathrm{E}] = \sum_{i=1}^{\alpha} \frac{(K_n[\mathrm{S}])^{i-1}I_1 - (K_n[\mathrm{S}])^iI_0}{I_0}. \quad (6)$$

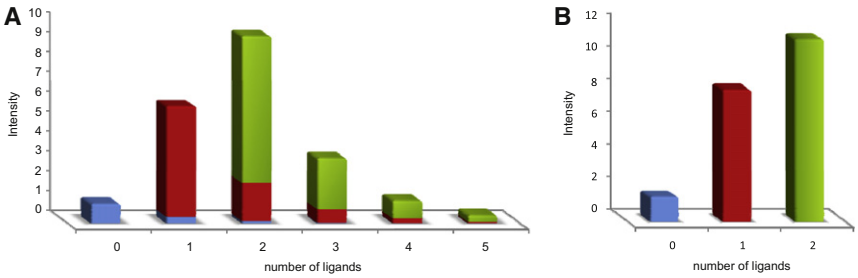
Likewise, it can be readily shown that one can express $C_2/[\mathrm{E}]$ as

$$C_2/[\mathrm{E}] = \sum_{i=1}^{\alpha} \frac{(K_n[\mathrm{S}])^{i-1}I_2 - (K_n[\mathrm{S}])^iI_1}{I_0}. \quad (7)$$

In general, one can show that $C_N/[\mathrm{E}]$ is given by:

$$C_N/[\mathrm{E}] = \sum_{i=1}^{\alpha} \frac{(K_n[\mathrm{S}])^{i-1}I_N - (K_n[\mathrm{S}])^iI_{N-1}}{I_0}, \quad (8)$$

where N is an integer between zero and the total number of specific binding sites, N_s, ($0 \leq N \leq N_s$) and α is, as before, the total number of intensity peaks observed (each of which corresponds to a different number of bound ligand molecules, including zero). By definition, $I_{N-1} = 0$ when $N = 0$. It is important to note that, in the above treatment, mixed species with ligand molecules bound to both specific and nonspecific sites are not eliminated but are counted together with other species that have the same number of bound ligands at specific sites. For example, a protein population with, say, one ligand molecule bound to a specific site and two ligand molecules bound to nonspecific sites contributes to I_3 but is counted in our above treatment together with the populations of other species that have only one ligand molecule bound to a specific site. In other words, our mathematical treatment makes it possible to correct the detected intensity of each ligand-bound state by 1), subtracting the artificial increase in intensity due to nonspecific binding, and 2), adding the different intensities after subtraction to those that correspond to species with the same number of ligand molecules bound to specific sites (Fig. 1).



I_0	I_1	I_2	I_3	I_4	I_5
1	$K_n[S]$ $K_1[S]$	$K_1K_2[S]^2$ $K_1K_n[S]^2$ $K_n^2[S]^2$	$K_1K_2K_n[S]^3$ $K_1K_n^2[S]^3$ $K_n^3[S]^3$	$K_1K_2K_n^2[S]^4$ $K_1K_n^3[S]^4$ $K_n^4[S]^4$	$K_1K_2K_n^3[S]^5$ $K_1K_n^4[S]^5$ $K_n^5[S]^5$

C_0	C_1	C_2
1	$K_1[S]$	$K_1K_2[S]^2$
$K_n[S]$	$K_1K_n[S]^2$	$K_1K_2K_n[S]^3$
$K_n^2[S]^2$	$K_1K_n^2[S]^3$	$K_1K_2K_n^2[S]^4$
$K_n^3[S]^3$	$K_1K_n^3[S]^4$	$K_1K_2K_n^3[S]^5$
$K_n^4[S]^4$	$K_1K_n^4[S]^5$	
$K_n^5[S]^5$		

FIGURE 1 An illustration demonstrating the calculated specific and nonspecific components that contribute to the peak intensities of the 19^+ charge state in the case of a solution containing 4 μM CK and 50 μM ADP. (*A*) Histogram showing the experimentally observed intensities corresponding to the different ligand-bound states. Both specific and nonspecific binding are reflected in each intensity. (*B*) A histogram generated after correcting for the contribution of nonspecific binding, so that the intensities reflect only specific binding. Listed below each intensity are the factors that contribute to it. The intensities are normalized with respect to I_0. The relative contributions of the free protein and protein bound with one or two ligand molecules at the specific sites before (*A*) and after (*B*) the correction are indicated in blue, red, and green, respectively.

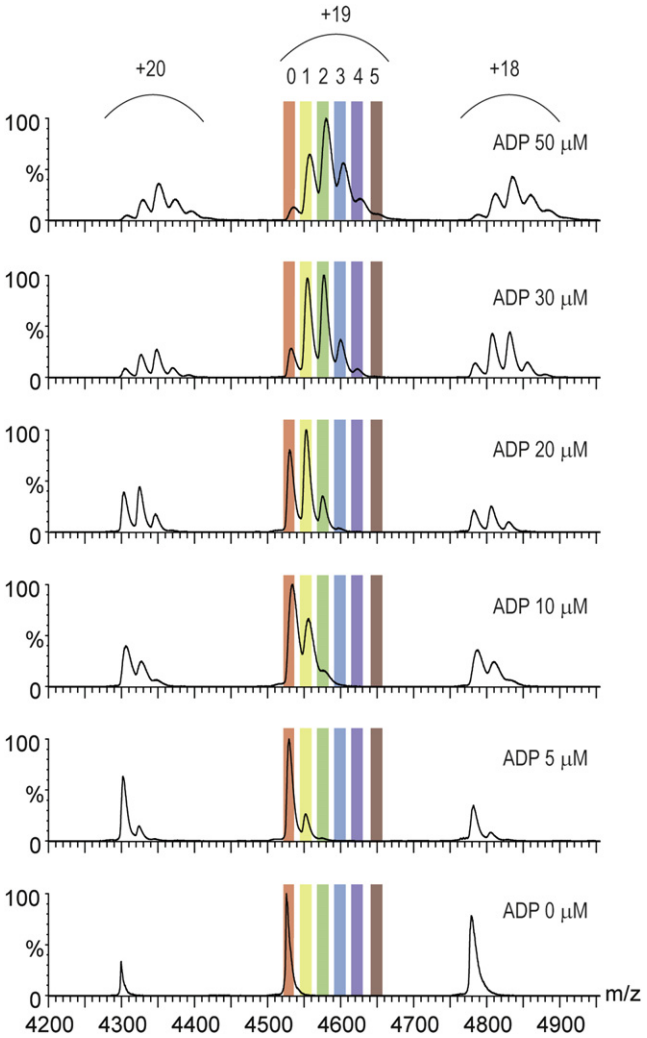


FIGURE 2 Nanoelectrospray mass spectra of CK obtained at increasing concentrations of ADP. The number of bound ADP molecules is highlighted for the 19^{+} charge state. The spectra were acquired in the presence of 4 μM CK in 250 ammonium acetate (pH 7).

The final step needed to determine the C_N values is to substitute [E] in Eq. 8 with a term that can be calculated. Given that the total protein concentration, C_T, is $C_T = \sum_{N=0}^{N_s} C_N$, we can express [E] as $[E] = C_T / \sum_{N=0}^{N_S} C_N/[E]$, where the terms in the denominator are calculated using Eq. 8. Hence, in general, C_N can be calculated from

$$C_N = \frac{(C_N/[E])C_T}{\sum_{N=0}^{N_S} C_N/[E]}. \quad (9)$$

It can be seen by inspection of Eq. 8 that $C_N/[E]$ is related to $C_{N-1}/[E]$, as follows:

$$C_N/[E] = K_N[S]C_{N-1}/[E], \quad (10)$$

where K_N is the binding constant of the ligand to the Nth specific site. Hence, the value of K_N can be obtained from the slope of a plot of $(C_N/[E])/(C_{N-1}/[E])$ versus [S] for different substrate concentrations by assuming that $[S] \approx [S]_{total}$. This assumption is not valid in the case of tight binding. It is important to note, however, that it is not required for determining the distribution of binding stoichiometries.

MATERIALS AND METHODS

Creatine kinase (CK) from rabbit muscle and adenosine 5′-diphosphate sodium salt (ADP) were purchased from Sigma (St. Louis, MO). CK was further purified on a gel filtration column in 20 mM Tris-HCl (pH 7.5). Before MS analysis, the protein was buffer-exchanged into 250 mM ammonium acetate (pH 7.5) using microbiospin 6 columns (BioRad, Hercules, CA). Titration experiments were performed in the presence of 4 μM CK, 1 mM magnesium acetate and ADP concentrations ranging between 0 and 100 μM. CK/ADP complexes were incubated for 5 min at room temperature before MS analysis.

MS measurements were performed in positive ion mode using a nanoelectrospray ionization quadrupole time-of-flight (ESI-Q-TOF) instrument (Applied Biosystems, Foster City, CA) modified for high mass detection (22,23). Aliquots of 2 μl were electrosprayed from gold-coated borosilicate capillaries prepared in-house as described previously (24). The mass spectrometer was operated at a capillary voltage of 1100 V, declustering potential of 140 V, and second-declustering potential of 15 V. External calibration of the mass spectra was achieved using solutions of 100 mg/ml cesium iodide in water. For each mass spectrum, the peak areas for the free CK and its ligand-bound states were calculated using deconvolution software (peakfit v4, Jandel Scientific, San Rafael, CA).

RESULTS AND DISCUSSION

To test our methodology, we applied it to CK homodimer (25) and its ligand ADP. Nanoelectrospray (nanoES) mass spectra were acquired for 4 μM of CK and increasing concentrations of ADP ranging from 0 to 100 μM. Fig. 2 shows the mass spectra of CK acquired at different concentrations of ADP. The CK charge-state distribution is centered at 4550 m/z and the measured mass of the free CK, 85,990 $\pm$ 34 Da, is found to be highly consistent with the calculated mass of the homodimer complex (85,960 Da). When titrated with ADP, no change in charge state was obtained, but additional peaks corresponding to the various ligand-bound states appeared (Fig. 2). Due to the relatively small difference in mass between apo and ADP-bound CK (<1%), the binding of the nucleotide

TABLE 1 Corrected concentrations of apo CK (C_0), and CK with one (C_1) and two (C_2) specifically bound ADP ligands

ADP concentration	C_0	C_1	C_2
0	4.00	0.00	0.00
5	3.28	0.72	0.00
10	2.31	1.38	0.31
15	2.28	1.51	0.21
20	1.71	1.81	0.48
30	1.09	1.89	1.02
40	0.87	1.71	1.42
50	0.29	1.50	2.22
100	0.13	1.41	2.46

Corrected concentrations were obtained from an average of three charge states in a typical experiment. All values are in μM.

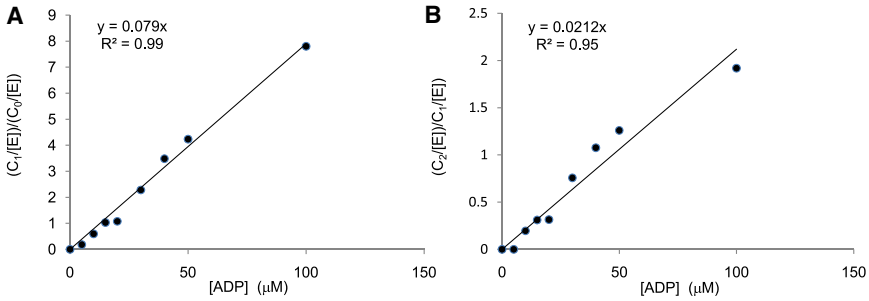


FIGURE 3 Determination of the values of K_1 (*A*) and K_2 (*B*) by plotting $(C_N/[E])/(C_{N-1}/(E]))$ versus [S] and fitting the data using Eq. 10, where [S] is the ADP concentration. Each data point is calculated from the average of the three observed charge states (18^+, 19^+, and 20^+) in four independent experiments.

does not appear to alter the ionization efficiency, as observed by others (10,17,26).

For each mass spectrum, the peak areas for free CK and its ligand-bound states were calculated using a deconvolution algorithm (Jandel Scientific peakfit v4) (Table S1 in the Supporting Material). In all calculations, it was assumed that the measured peak areas (referred to as I_n) correlate with the solution concentrations (6,9–13). Initially, we determined the value of K_n by calculating the average value of the I_4/I_3 and I_5/I_4 ratios (Table S2) as in Eq. 1. It is important to note that the values of K_n determined from the I_4/I_3 and I_5/I_4 ratios are very similar, thus supporting our assumption that the nonspecific binding is not cooperative. Using an average value of $0.007 \pm 0.003\ \mu M^{-1}$ for K_n and the measured intensities, we next corrected the distribution of binding stoichiometries using Eq. 8 (Table S3). The absolute values of C_0, C_1, and C_2 were then calculated using Eq. 9 (Table 1 and Table S4). Finally, utilizing Eq. 10, we determined the values of the binding (association) constants K_1 and K_2 to be 0.08 ± 0.002 and $0.02 \pm 0.001\ \mu M^{-1}$, respectively (Fig. 3). These values correspond to dissociation constants of 12.6 ± 0.3 and $47 \pm 2\ \mu M$, respectively, which are in very good agreement with previously reported values (17,27–31). In particular, they are in excellent agreement with the respective values of 11.8 ± 1.5 and $48 \pm 6\ \mu M$ for K_1 and K_2 determined more recently (17). The finding that $K_1 < K_2$ is in agreement with the previously reported negative cooperativity of ADP binding to CK (31).

In summary, the method presented in this article makes it possible to separate specific from nonspecific binding when using nonensemble methods, so that the distribution of specific binding stoichiometries at any ligand concentration can be determined. The method can be applied to data for any ligand-protein system obtained using an experimental approach that can distinguish between populations with different numbers of bound ligand molecules. Major advantages of our method are that no previous information regarding the mechanism of ligand interaction is required and determination of the values of the binding constants is straightforward.

SUPPORTING MATERIAL

Details of the distribution binding calculations are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(10)00806-4.

This work was supported by grant 153/08 of the Israel Science Foundation (to A.H.) and by funding from the European Research Council (ERC) under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement No. 239679 (to M.S.).

REFERENCES

1. Jiang, Y., Q. Wang, …, W. E. Moerner. 2008. Hardware-based anti-Brownian electrokinetic trap (ABEL trap) for single molecules: control loop simulations and application to ATP binding stoichiometry in multi-subunit enzymes. *Proc. Soc. Photo Opt. Instrum. Eng.* 7038: 1–12.
2. Benesch, J. L., B. T. Ruotolo, …, C. V. Robinson. 2007. Protein complexes in the gas phase: technology for structural genomics and proteomics. *Chem. Rev.* 107:3544–3567.
3. Sharon, M., and C. V. Robinson. 2007. The role of mass spectrometry in structure elucidation of dynamic protein complexes. *Annu. Rev. Biochem.* 76:167–193.
4. van den Heuvel, R. H. H., and A. J. R. Heck. 2004. Native protein mass spectrometry: from intact oligomers to functional machineries. *Curr. Opin. Chem. Biol.* 8:519–526.
5. Daneshfar, R., E. N. Kitova, and J. S. Klassen. 2004. Determination of protein-ligand association thermochemistry using variable-temperature nanoelectrospray mass spectrometry. *J. Am. Chem. Soc.* 126:4786–4787.
6. Jorgensen, T. J. D., P. Roepstorff, and A. J. R. Heck. 1998. Direct determination of solution binding constants for noncovalent complexes between bacterial cell wall peptide analogues and vancomycin group antibiotics by electrospray ionization mass spectrometry. *Anal. Chem.* 70:4427–4432.
7. Li, Z., F. Song, …, K. S. Anderson. 2009. Monitoring enzyme catalysis in the multimeric state: direct observation of Arthrobacter 4-hydroxybenzoyl-coenzyme A thioesterase catalytic complexes using time-resolved electrospray ionization mass spectrometry. *Anal. Biochem.* 394:209–216.
8. McCammon, M. G., and C. V. Robinson. 2004. Structural change in response to ligand binding. *Curr. Opin. Chem. Biol.* 8:60–65.
9. Peschke, M., U. H. Verkerk, and P. Kebarle. 2004. Features of the ESI mechanism that affect the observation of multiply charged noncovalent protein complexes and the determination of the association constant by the titration method. *J. Am. Soc. Mass Spectrom.* 15:1424–1434.
10. Sannes-Lowery, K. A., R. H. Griffey, and S. A. Hofstadler. 2000. Measuring dissociation constants of RNA and aminoglycoside antibiotics by electrospray ionization mass spectrometry. *Anal. Biochem.* 280:264–271.
11. Wortmann, A., M. C. Jecklin, …, R. Zenobi. 2008. Binding constant determination of high-affinity protein-ligand complexes by electrospray ionization mass spectrometry and ligand competition. *J. Mass Spectrom.* 43:600–608.
12. Wortmann, A., F. Rossi, …, R. Zenobi. 2005. Determination of zinc to β-peptide binding constants with electrospray ionization mass spectrometry. *J. Mass Spectrom.* 40:777–784.

13. Yu, Y., C. E. Kirkup, …, J. A. Leary. 2004. Characterization of noncovalent protein-ligand complexes and associated enzyme intermediates of GlcNAc-6-O-sulfotransferase by electrospray ionization FT-ICR mass spectrometry. *J. Am. Soc. Mass Spectrom.* 15:1400–1407.

14. van Zoelen, E. J. J., R. H. Kramer, …, J. H. Veerkamp. 1998. The use of nonhomologous Scatchard analysis in the evaluation of ligand-protein interactions. *Trends Pharmacol. Sci.* 19:487–490.

15. Wang, W., E. N. Kitova, and J. S. Klassen. 2003. Influence of solution and gas phase processes on protein-carbohydrate binding affinities determined by nanoelectrospray Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* 75:4945–4955.

16. Ayed, A., A. N. Krutchinsky, …, H. W. Duckworth. 1998. Quantitative evaluation of protein-protein and ligand-protein equilibria of a large allosteric enzyme by electrospray ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 12:339–344.

17. Daubenfeld, T., A. P. Bouin, and G. van der Rest. 2006. A deconvolution method for the separation of specific versus nonspecific interactions in noncovalent protein-ligand complexes analyzed by ESI-FT-ICR mass spectrometry. *J. Am. Soc. Mass Spectrom.* 17:1239–1248.

18. Sun, J., E. N. Kitova, and J. S. Klassen. 2007. Method for stabilizing protein-ligand complexes in nanoelectrospray ionization mass spectrometry. *Anal. Chem.* 79:416–425.

19. Sun, N., N. Soya, …, J. S. Klassen. 2010. Nonspecific interactions between proteins and charged biomolecules in electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* 21:472–481.

20. Touboul, D., L. Maillard, …, R. Zenobi. 2009. How to deal with weak interactions in noncovalent complexes analyzed by electrospray mass spectrometry: cyclopeptidic inhibitors of the nuclear receptor coactivator 1-STAT6. *J. Am. Soc. Mass Spectrom.* 20:303–311.

21. Wang, W., E. N. Kitova, …, J. S. Klassen. 2005. Blackbody infrared radiative dissociation of nonspecific protein-carbohydrate complexes produced by nanoelectrospray ionization: the nature of the noncovalent interactions. *J. Am. Soc. Mass Spectrom.* 16:1583–1594.

22. Chernushevich, I. V., and B. A. Thomson. 2004. Collisional cooling of large ions in electrospray mass spectrometry. *Anal. Chem.* 76: 1754–1760.

23. Sobott, F., H. Hernández, …, C. V. Robinson. 2002. A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal. Chem.* 74:1402–1407.

24. Hernández, H., and C. V. Robinson. 2007. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat. Protoc.* 2:715–726.

25. Rao, J. K. M., G. Bujacz, and A. Wlodawer. 1998. Crystal structure of rabbit muscle creatine kinase. *FEBS Lett.* 439:133–137.

26. Loo, J. A. 1997. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.* 16:1–23.

27. Borders, Jr., C. L., M. J. Snider, …, P. L. Edmiston. 2002. Determination of the affinity of each component of a composite quaternary transition-state analogue complex of creatine kinase. *Biochemistry.* 41:6995–7000.

28. Burbaum, J. J., and J. R. Knowles. 1989. Internal thermodynamics of enzymes determined by equilibrium quench: values of Kint for enolase and creatine kinase. *Biochemistry.* 28:9306–9317.

29. Hornemann, T., D. Rutishauser, and T. Wallimann. 2000. Why is creatine kinase a dimer? Evidence for cooperativity between the two subunits. *Biochim. Biophys. Acta.* 1480:365–373.

30. McLaughlin, A. C. 1974. The interaction of 8-anilino-1-naphthalenesulfonate with creatine kinase. Evidence for cooperativitiy of nucleotide binding. *J. Biol. Chem.* 249:1445–1452.

31. Price, N. C., and M. G. Hunter. 1976. Non-identical behaviour of the subunits of rabbit muscule creatine kinase. *Biochim. Biophys. Acta.* 445:364–376.