Supporting Information

A straightforward method for measuring binding affinities of ligands to proteins of unknown concentration in biological tissues

Bin Yan*[1] and Josephine Bunch*[1, 2]

1. National Centre of Excellence in Mass Spectrometry Imaging, National Physical Laboratory, Hampton Road, Teddington, TW11 0LW, U.K.

2. Department of Metabolism, Digestion and Reproduction, Imperial College London, Hammersmith Hospital, Du Cane Road, London, W12 0NN, U.K.

***Corresponding author**: bin.yan@npl.co.uk; josephine.bunch@npl.co.uk

METHODS

Materials

Except for recombinant mouse liver fatty acid binding protein (L-FABP) and cytidine-5'-diphosphate (CDP) disodium salt (95%) (purchased from UK Bio-Techne and Alfa Aesar, respectively), all other protein samples including lysozyme from chicken egg white, ribonuclease A (RNase A) from bovine pancreas, carbonic anhydrase I (CA1) from human erythrocytes, myoglobin from equine heart, ubiquitin from bovine erythrocytes, and small molecule ligands N,N',N''-Triacetylchitotriose (NAG_3, ≥93%), indapamide (analytical standard), acetazolamide (≥99%), fenofibric acid (95%), prednisolone (≥99%), gemfibrozil (pharmaceutical secondary standard), as well as MS grade salt ammonium acetate were obtained from Sigma-Aldrich (Gillingham, UK). Deionized water (ρ = 18 MΩ cm^{-1}) was generated by an Elga Purelab system (High Wycombe, UK).

Unless otherwise stated, protein samples were prepared in 10 mM ammonium acetate aqueous solution. The aqueous stock solution of L-FABP was desalted twice using a centrifugal filter with a 7 kDa cutoff (Zeba™ Spin Desalting Columns 2 mL, Fisher Scientifi c, UK) (~90% protein recovery rate). The ligand solutions of NAG_3, acetazolamide, prednisolone and CDP disodium salt were also prepared as aqueous solution while indapamide, fenofibric acid and gemfibrozil were dissolved in methanol. The protein and ligand are mixed and incubated in either 10 mM ammonium acetate aqueous solution or with 5% methanol added for a sufficient time (30 mins in this study) to allow protein-ligand binding to reach equilibrium before MS measurement. For binding affinity measurements from surface, 0.2 µL aliquots of each protein solution (lysozyme: 1mM, RNAse A: 0.5mM, CA1: 1mM and 0.5mM for indapamide and acetazolamide binding study, repectively) were spotted onto untreated clean glass microscope slides and left to air dry to form sample spots with diameters of ~1 mm. For titration MS measurments, protein concentrations were 1 µM, 0.5 µM, 5 µM, and 0.36 µM in the binding studies of Lysozyme-NAG_3, RNAseA-CDP, hCA I-indapamide, and hCA I-acetazolamide, respectively.

For animal tissue studies, wild type mouse liver tissue provided by collaborators at the Francis Crick Institute (London, UK) was sectioned using the CryoStar™ NX70 cryostat microtome (Thermo Scientific, Waltham, USA) to thickness of 10 µm. The tissue sections were then thaw mounted onto standard glass slides (Thermo Menzel-Glaser Superfrost) and analysed without further processing.

MS studies

Except for the hCA I-acetazolamide complex studied by an Orbitrap Exploris™ 480 Mass Spectrometer (Thermo Scientific, Massachusetts, USA), MS measurements were mainly performed on a Synapt G2-Si Q-ToF mass spectrometer (Waters, Manchester, UK) coupled with a Triversa Nanomate chip-based nano-electrospray ionization source (Advion Biosciences, Inc., Ithaca, USA). Set at resolution mode (FMHM of 28,000 measured at m/z 922, $Na_7I_6^+$) for the mass analyser, the mass spectrometer was calibrated in postive ion mode with RMS error less than 1ppm at *m/z* 100-4000 using the mixture solution containing sodium iodide and cesium iodide. For optimum transmission of noncovalent protein complexes from ESI plume to the mass spectrometer, the MS cone voltage was at 80v and the ion source tempereture was 80 °C. Other MS parameters were left as default.

For infusion ESI-MS measurements of protein-ligand mixtures, 3 µL of sample solution was aspirated into a conductive, carbon-impregnated pipette tip and delivered to an ESI chip with an array of 5 µm inner diameter nozzles. Through applying the high voltage of 1.4-1.7 kV and the backing pressure of 0.3-0.5 psi, nano-electrospray is formed in which highly-charged ions of proteins and their complexes are included.

For surface analysis, the Triversa Nanomate was programmed using Advion ChipSoftX with the new Developers Kit to automate a series of operations including surface sampling, serial dilution of the extracted protein solution, mixing solutions of protein and ligand, and starting ESI-MS measurements. During surface sampling, the pipette held by the Nanomate robotic arm first aspirates 5 µL of extraction solvent (10 mM ammonium acetate aqueous solution doped with the ligand of interest) from a well of half a 384-well plate placed on the right of a LESA universal adapter plate. Then the pipette tip is moved to a height of about 0.5 mm above the surface of sample (glass slide placed on the left of the LESA adapter plate) and 2 µL solvent is dispensed to form a liquid micro-junction between the tip and the surface. After a delay of 15-30 s, 2.2 µL of the extraction solvent was re-aspirated into the pipette and the total volume of 5.2 µL was stored in 384 well plate which was further diluted to 10 µL with the same solvent. Subsequently, serial dilutions of the prepared protein-ligand mixture were performed to create another minimum two additional solutions in which the protein concentration was reduced by 10-fold and 100-fold, respectively while the ligand concentration was fixed. After 30 mins time of incubation, the mixture solutions were analyzed by infusion ESI-MS. The scan time of each MS measurement was 1 s and each spectrum consists of 60-100 scans. At least triplicate measurements were conducted for each sample mixture. Mass spectra were recorded by Waters Masslynx software V4.1. Data analysis and figure preparation were performed using Masslynx and OriginPro 2020.

Determination of binding affinity

1. Single-point approach

With native mass spectrometry, rapid determination of protein-ligand binding affinity can be achieved by the single-point approach in which relative intensities of free, unbound and ligand bound proteins are measured of a protein-ligand mixture sample. Knowing the initial concentrations of protein and ligand, dissociation constant K_d can be obtained from the integrated peak areas for the ligand-bound and unbound protein ions at all charge states using Equation 1

$$K_d = \frac{\sum_n P^{n+}}{\sum_n PL^{n+}} \left([L]_0 - \frac{\sum_n PL^{n+}}{\sum_n P^{n+} + \sum_n PL^{n+}} [P]_0 \right) \quad (1)^{[1]}$$

where $[L]_0$ and $[P]_0$ correspond to the concentrations of ligand and protein; P^{n+} and PL^{n+} represent the abundances of unbound protein and ligand-bound protein ions with *n+* charges, respectively.

2. Titration method

In some circumstances due to dissociation of the non-covalent interaction assembled ligand-bound complexes or large discrepancy between the response factors (ionization, transmition and detection) for the bound and unbound proteins, the relative ratio of the bound and unbound proteins in the solution phase may not be effectively maintained during sample transfer to the gas phase. As a result, the single-point approach often results in large deviations in the obtained K_d values and titration-based experiments should be performed. In a titration study using native MS, the ion abundance ratio *R* of the bound and unboud protein is directly monitored as a function of the ligand concentration while the initial protein concentration is fixed. The binding affinity (K_d) can be acquired by fitting Equation 2 or Equation 3 to a plot of the bound fraction *F* ($F = R/(R+1)$) versus the initial ligand concentration.

$$F = \frac{[P]_0 + [L]_0 + K_d - \sqrt{(K_d - [L]_0 + [P]_0)^2 + 4K_d[L]_0}}{2[P]_0} \quad (2)^{[2]}$$

3. Measurement from complex samples with unknown protein concentrations

As reported in literatures,[3] to achieve accurate measurement of binding affinity the protein concentration commonly needs to be kept well below the dissociation constant to be determined. Otherwise, essentially all added ligand is depleted due to binding to free protein until all binding sites are occupied, which could cause measured value deviate from the "true" K_d by orders of magnitude. Under such circumstance, the parameter of initial protein concentration P_0 in Equation 1 has little impact on the calculation of K_d hence could be eliminated to simplify the single-point method (Equations 3), as well as the titration method (Equation 4).

$$K_d = \frac{\sum_n P^{n+}}{\sum_n PL^{n+}}[L]_0 \quad (3)$$

$$F = \frac{R}{R+1} = \frac{\sum_n PL^{n+}}{\sum_n P^{n+}} / \left(\frac{\sum_n PL^{n+}}{\sum_n P^{n+}} + 1\right) = \frac{[L]_0}{K_d} / \left(\frac{[L]_0}{K_d} + 1\right) = \frac{[L]_0}{[L]_0 + K_d} \quad (4)$$

For complex sample systems with unknown protein concentrations, a serial dilution method can be used to ensure that accurate binding affinities are obtained at sufficiently low protein concentrations. Assuming no significant change is observed in measured K_d after a dilution (such as K_{d2} and K_{d3} in Fig. 1a), the protein concentration is believed to be well below the true K_d value prior to that dilution hence accurate determination of binding affinity can be achieved in both cases (i.e. $K_{d2} \approx K_{d3} \approx$ True K_d).

4. Measuring K_d values for multiple proteins simultaneously

Refering to the process of deriving the equation to obtain K_d values for a protein in the presence of multiple, competing ligands, the following formula was created to calculate K_d values when multiple proteins bind to the same ligand.

$$K_{d,i} = \frac{\sum_n P_i^{n+}}{\sum_n P_i L^{n+}} \left([L]_0 - \sum_i \frac{\sum_n P_i L^{n+}}{\sum_n P_i^{n+} + \sum_n P_i L^{n+}} [P_i]_0 \right) \quad (5)^{[1]}$$

Similarly, the calculation could be simpllified to the following equation if the concentrations of all proteins are very small.

$$K_{d,i} = \frac{\sum_n P_i^{n+}}{\sum_n P_i L^{n+}}[L]_0 \quad (6)$$

Supporting Figures & Tables

Table S1. K_d values for the binding of FABP to the drug ligand fenofibric acid measured from mouse liver tissue sections and recombinant protein solution samples: a) determined by intergrating all charge states; b) calcualted at different charge state

a)

	$[P]_0$	K_{d1} (µM)	K_{d2} (µM)
Tissue sample	Surface extracted	66.2 ± 2.5	69.0 ± 6.4
	2x diluted	44.9 ± 5.2	45.7 ± 7.3
	4x diluted	42.7 ± 4.5	48.5 ± 5.8
	Determined K_d	**44.0 ± 5.0**	**46.9 ± 6.8**
Solution sample	**5 µM (5% methanol)**	**42.4 ± 4.0**	**45.0 ± 5.3**
	5 µM (2% methanol)	**39.1 ± 6.5**	**42.5 ± 7.8**

b)

	$[P]_0$	K_{d1} (µM)		K_{d2} (µM)	
		7+	**8+**	**7+**	**8+**
Tissue sample	Surface extracted	65.3 ± 3.7	69.4 ± 4.5	64.4 ± 6.6	70.5 ± 3.2
	2x diluted	41.5 ± 6.1	47.6 ± 3.9	41.8 ± 6.2	46.5 ± 5.6
	4x diluted	39.1 ± 2.8	47.9 ± 3.9	45.8 ± 5.4	56.4 ± 2.5
	Determined K_d	**40.4 ± 5.1**	**47.7 ± 3.9**	**43.5 ± 6.2**	**50.7 ± 6.7**
Solution sample	**5 µM (5% methanol)**	**41.1 ± 3.9**	**47.0 ± 2.5**	**41.5 ± 5.1**	**49.5 ± 4.3**
	5 µM (2% methanol)	**36.4 ± 5.7**	**39.9 ± 8.0**	**40.3 ± 6.4**	**44.6 ± 9.5**

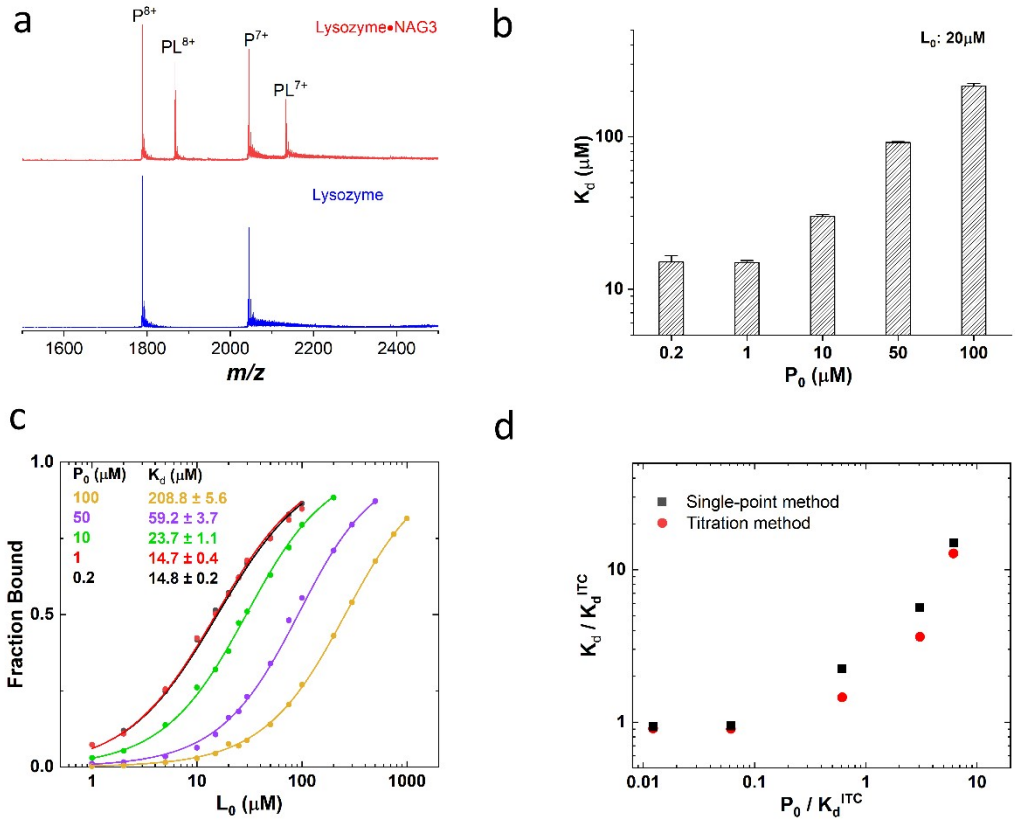

Figure S1. Effect of protein concentration on the measurement accuracy of binding affinity. (a) Infusion nanoESI native mass spectra of 5 μM hen egg white lysozyme without and with 15 μM ligand of NAG_3. (b) Bar plot of lysozyme-NAG_3 binding dissociation constant K_d measured using single-point method and Equation 1 at different initial protein concentrations. (c) Bound lysozyme fraction as a function of initial concentration of NAG_3 ligand. K_d values were obtained through nonlinear fitting of titration data using Equation 2. Colors represent different protein concentrations used in the titration study. (d) scatter plot of normalized K_d measured as a function of normalized protein concentration. Both K_d and protein concentration were normalized to the K_d value measured by isothermal titration calorimetry (ITC).[4]

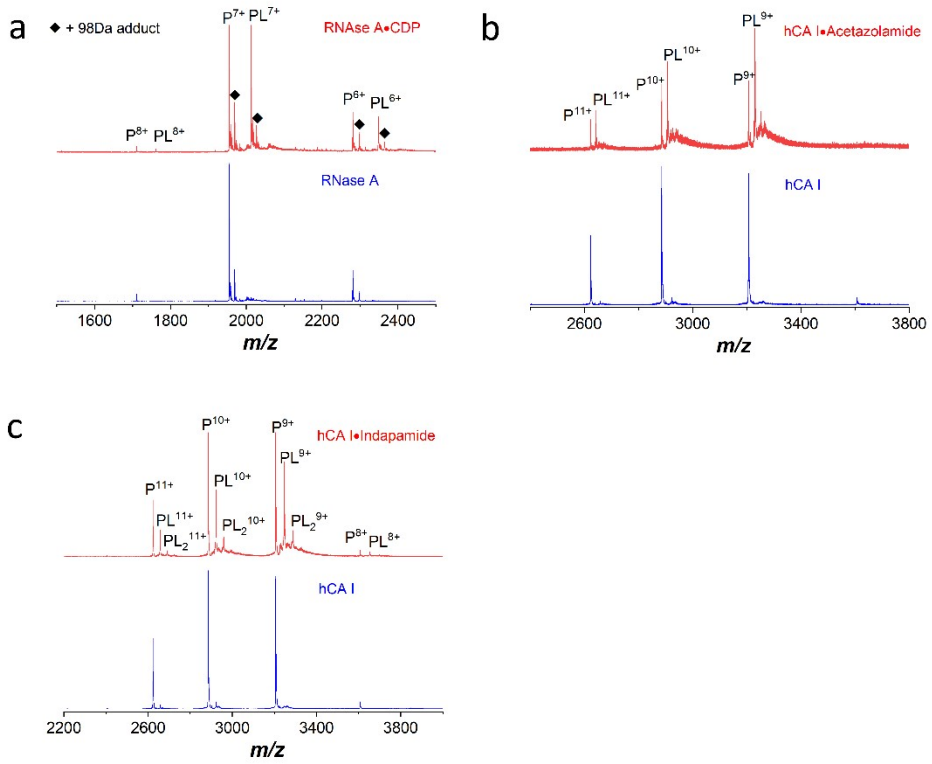

Figure S2. Infusion nanoESI native mass spectra of (a) 0.5 µM RNase A without and with 5 µM ligand of CDP; (b) 0.36 µM human carbonic anhydrase 1 without and with 4 µM ligand of acetazolamide; (c) 2 µM human carbonic anhydrase 1 without and with 12 µM ligand of indapamid. Samples were prepared in 10 mM ammonium acetate aqueous solution except that CA1-indapamide mixture was in the same solution with 5% methanol added.

Table S2. Comparison of K_d values determined between with and without protein concentration using the single-point method and titration method at low protein concentrations. Protein RNase A: Ribonuclease A; CA1: Carbonic Anhydrase 1; For single-point method, K_d: calculated using the Equation 1; K_d': calculated using Equation 3 without P_0. For titration method, K_d: fitted using the Equation 2; K_d': fitted using Equation 4 without P_0. *Ion intensities of both 1:1 and 1:2 protein-ligand complexes were taken into account to calculate the summed bound protein fraction *F*

	Single-point method			***Titration method***		
	$[P]_0$ (µM)	**K_d (µM)**	**K_d' (µM)**	**$[P]_0$ (µM)**	**K_d (µM)**	**K_d' (µM)**
Lysozyme-NAG_3	0.2	14.8 ± 0.2	14.9 ± 0.2	1	14.8 ± 0.2	15.1 ± 0.4
	1	14.7 ± 0.4	15.1 ± 0.4			
RNase A-CDP	0.2	4.0 ± 0.2	4.1 ± 0.2	0.5	4.1 ± 0.2	4.3 ± 0.2
	0.5	3.8 ± 0.1	4.1 ± 0.1			
CA1-Acetazolamide	0.01	0.361 ± 0.049	0.362 ± 0.049	0.02	0.348 ± 0.015	0.357 ± 0.015
	0.02	0.373 ± 0.041	0.375 ± 0.042			
CA1-Indapamide 1:1	0.5	15.6 ± 2.0	16.0 ± 2.0	5	13.1 ± 0.7*	15.2 ± 0.5*
	2	16.6 ± 1.7	18.4 ± 1.6			
CA1-Indapamide 1:2	0.5	43.8 ± 4.2	44.9 ± 4.2			
	2	37.4 ± 4.1	41.5 ± 4.1			

Table S3. Comparison of protein concentrations in surface extracted and diluted samples with the sought-after K_d values. Assuming the surface sampling efficiency is 100%, the maximum protein concentrations in surface extracted samples are higher than the "true " K_d values, resulting in underestimation of binding affinities (larger dissociation constants measured). After 10x dilution, except for the study of hCA I-Acetazolamide, protein concentrations were reduced to well below K_d values hence accurate determination of binding affinities could be achieved.

	Deposited spot (nM)	Surface extraction (μM/L)	10x dilution (μM/L)	True K_d (μM/L)
Lysozyme	0.2	20 (max.)	2 (max.)	~15[4-5]
RNAse A	0.1	10 (max.)	1 (max.)	1-3[5b, 6]
hCA I (Indapamide)	0.2	20 (max.)	2 (max.)	~ 10[1]
hCA I (Acetazolamide)	0.1	10 (max.)	1 (max.)	0.2-0.5[1, 7]

Table S4. Dissociation constants of studied protein-ligand binding systems determined from surface-deposited protein samples using the method illustrated in Fig. 1a and Equation 3. The corresponding measured native MS spectra are shown as Fig. 2(a-d). * K_{d1} of 15.3 ± 1.8 μM and K_{d2} of 45.2 ± 4.3 μM were also acquired for human carbonic anhydrase 1 binding to the first and second ligand of indapamide. ** K_d was determined from 500x and 1000x diluted samples by Orbitrap MS.

	Lysozyme-NAG_3	RNAseA-CDP	hCA I-Indapamide*	hCA I-Acetazolamide
Surface extracted	94.8 ± 2.6	35.4 ± 2.1	49.5 ± 4.2	24.3 ± 0.6
10x diluted	17.2 ± 2.3	3.7 ± 0.2	13.1 ± 1.1	0.375 ± 0.042**
100x diluted	15.5 ± 1.2	3.9 ± 0.2	13.0 ± 1.1	0.362 ± 0.049**
Determined K_d	**16.3 ± 2.0** **16.4 ± 2.1 (7+)** **15.4 ± 1.8 (8+)**	**3.8 ± 0.2** **4.1 ± 0.5 (6+)** **3.2 ± 0.2 (7+)**	**13.1 ± 1.1** **11.2 ± 1.1 (9+)** **14.4 ± 1.0 (10+)** **16.2 ± 1.4 (11+)**	**0.368 ± 0.046** **0.331 ± 0.030 (10+)** **0.405 ± 0.088 (11+)**

The dissociation constant of lysozyme-NAG_3 binding meaured from surface samples is 16.3 ± 2.0 μM, which is in good agreement with that measured in previous works using mass spectrometry and other biophysical methods.[4-5]

Ribonuclease A is a type of nuclease that catalyzes the degradation of RNA into smaller molecules. Together with its homologs, they play a critical role in many biological processes including organ growth, vascular and rheumatoid diseases and cancer development. Therefore, Ribonuclease have become important targets for many drug design, in which one of the fundamental questions it to determine the binding affinity between the target protein and its potential drug ligand. The isoform RNase A is known to bind its nucleotide ligands in 1:1 stoichiometry with the dissociation constant of 1-3 μM.[5b, 6] Using infusion ESI MS, we studied the RNase A-CDP binding from solution samples. The dissociation constant K_d determined by the single-point method (Supplementary Table 1) and the titration MS method (Figure 2e) were 3.9 ± 0.2 μM and 4.1 ± 0.2 μM, respectively. Using the developed binding affinity measurement method without the information of protein information, the K_d of 3.8 ± 0.2 μM for RNase A-CDP binding was also obtained from the surface deposited protein sample.

As ubiquitous enzymes, carbonic anhydrases play important physiological and physio-pathological functions by catalyzing one of the most important reactions in life: reversible hydration of carbon dioxide to bicarbonate. Clinically, CA inhibitors including acetazolamide and indapamide have been widely used to treat various diseases such as glaucoma, epilepsy, high blood pressure, cancers etc. In this work using native MS (spectra shown in Supplementary Fig. 2), the binding stoichometry of CA1-acetazolamide and CA1-indapamide was determined to be 1:1 and 1:2, respectively. Through curve fitting of the titration data (Fig. 2e), correponding dissociation constants were also obtained, i.e. 0.348 ± 0.015 μM for CA1-acetazolamide binding and 13.1 ± 0.7 μM for CA1-indapamide binding. Using single-point approach, K_{d1} of 16.1 ± 1.9 μM and K_{d2} of 40.6 ± 5.2 μM could be acquired for CA1 binding to the first and second ligand of indapamide. Compared with that reported in literatures, the determined binding affinities of CA1-indapamide and CA1-acetazolamide are very similar.[1, 7] Besides, using the new method developed for surface measurment, the K_d values of CA1 binding to acetazolamide (0.368 ± 0.046 μM) and indapamide (15.3 ± 1.8 μM and 45.2 ± 4.3 μM for 1st and 2nd ligand binding respectively) were determined, which were highly consistent with those measured from solution phase.

Table S5. Comparison of K_d values determined between with and without protein concentration using other biophysical techniques based on titration measurements. K_d: calculated using the Equation 7, where *Int* is the observed experimental signal, Int_f is the experimental signal for unbound protein and Int_b is the experimental signal for completely bound protein (these parameters should be fitted); K_d': calculated using Equation 8 without P_0

		$[P]_0$ (µM)	K_d (µM)	K_d' (µM)
Fluorescence spectroscopy	Lysozyme-NAG_3[8]	2.85	7.7 ± 1.0	9.6 ± 1.2
	HusA-Haem[9]	1	4.3 ± 0.8	4.9 ± 0.9
Microscale thermophoresis	SAM-II riboswitch-SMA[10]	0.05	0.14 ± 0.03	0.16 ± 0.03
	AhR-ARNT complex – 1-Hydroxyphenazine[11]	0.25	0.88 ± 0.23	1.03 ± 0.24

$$Int = Int_f + \left(Int_b - Int_f\right) * \frac{[P]_0 + [L]_0 + K_d - \sqrt{\left(K_d - [L]_0 + [P]_0\right)^2 + 4K_d[L]_0}}{2[P]_0} \quad (7)^{[8]}$$

$$Int = Int_f + \left(Int_b - Int_f\right) * \frac{[L]_0}{[L]_0 + K_d} \quad (8)$$

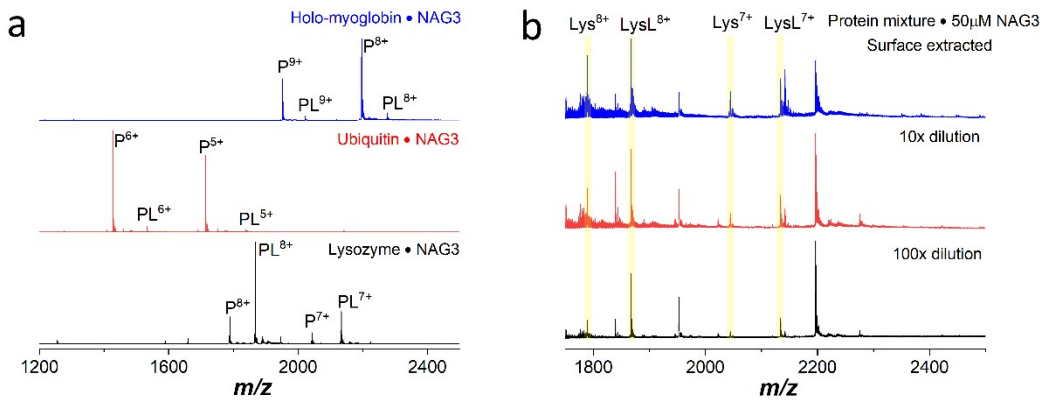

Figure S3. Evaluation of off-target binding on K_d measurement from surface (a) Native mass spectra of 50 μM NAG_3 ligand mixed with 10 μM protein of myoglobin (top), ubiquitin (middle) and lysozyme (bottom). A large amount of lysozyme was bound to NAG_3 while most of myoglobin and ubiquitin were unbound, indicating that the binding affinity of NAG_3 to lysozyme is much stronger. (b) Comparison of bound/unbound lysozyme ratios with dilution of protein mixtures extracted from surface. The bound fraction of lysozyme increased at 10x dilution and did not change significantly from 10-fold to 100-fold dilution. The maxium concentration of each protein in the mixture sample prepared by liquid extraction is 30 μM assuming 100% surface sampling efficiency.

References

[1] G. T. H. Nguyen, T. N. Tran, M. N. Podgorski, S. G. Bell, C. T. Supuran, W. A. Donald, *ACS Cent. Sci.* **2019**, *5*, 308.

[2] E. G. Báez Bolivar, D. T. Bui, E. N. Kitova, L. Han, R. B. Zheng, E. J. Luber, S. Y. Sayed, L. K. Mahal, J. S. Klassen, *Anal. Chem.* **2021**, *93*, 4231.

[3] a) J. L. Bennett, G. T. H. Nguyen, W. A. Donald, *Chem. Rev.* **2022**, *122*, 7327; b) I. Jarmoskaite, I. AlSadhan, P. P. Vaidyanathan, D. Herschlag, *Elife* **2020**, *9*, e57264.

[4] A. F. M. Gavriilidou, B. Gülbakan, R. Zenobi, *Anal. Chem.* **2015**, *87*, 10378.

[5] a) T. Liu, T. M. Marcinko, R. W. Vachet, *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 1544; b) L. Jaquillard, F. Saab, F. Schoentgen, M. Cadene, *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 908; c) B. López-Méndez, S. Uebel, L. P. Lundgren, A. Sedivy, *Eur. Biophys. J.* **2021**, *50*, 653.

[6] a) S. Zhang, C. K. Van Pelt, D. B. Wilson, *Anal. Chem.* **2003**, *75*, 3010; b) C. L. Jones, F. Fish, D. D. Muccio, *Anal. Biochem.* **2002**, *302*, 184.

[7] M. C. Jecklin, S. Schauer, C. E. Dumelin, R. Zenobi, *J. Mol. Recognit.* **2009**, *22*, 319.

[8] X. Li-Blatter, L. Zweifel, T. Sharpe, in *Protein-Ligand Interactions: Methods and Applications* (Eds.: T. Daviter, C. M. Johnson, S. H. McLaughlin, M. A. Williams), Springer US, New York, NY, **2021**, pp. 47.

[9] A. Yammine, J. Gao, A. H. Kwan, *Bio-protocol* **2019**, *9*, e3253.

[10] M. H. Moon, T. A. Hilimire, A. M. Sanders, J. S. Schneekloth, Jr., *Biochem.* **2018**, *57*, 4638.

[11] A. Stinn, J. Furkert, S. H. Kaufmann, P. Moura-Alves, M. Kolbe, *Biosensors* **2021**, *11*, 60.