Published in final edited form as:

ACS Chem Biol. 2017 February 17; 12(2): 389–397. doi:10.1021/acscchembio.6b01018.

Proteoform-Specific Protein Binding of Small Molecules in Complex Matrices

Geuncheol Gil¹, Pan Mao¹, Bharathi Avula², Zulfikar Ali², Amar G. Chittiboyina², Ikhlas A. Khan^{2,3}, Larry A. Walker^{2,4}, and Daojing Wang^{1,*}

¹Newomics Inc., Emeryville, California 94608, United States

²National Center for Natural Products Research, The University of Mississippi, University, Mississippi 38677, United States

³Division of Pharmacognosy, Department of BioMolecular Sciences, School of Pharmacy, The University of Mississippi, University, Mississippi 38677, United States

⁴Division of Pharmacology, Department of BioMolecular Sciences, School of Pharmacy, The University of Mississippi, University, Mississippi 38677, United States

Abstract

Characterizing the specific binding between protein targets and small molecules is critically important for drug discovery. Conventional assays require isolation and purification of small molecules from complex matrices through multi-step chromatographic fractionation which may alter their original bioactivity. Most proteins undergo posttranslational modification and only certain proteoforms have the right conformation with accessible domains and available residues for small molecule binding. We developed a top-down mass spectrometry (MS) centric workflow for rapid evaluation of the bioactivity of crude botanical extracts after a one-step reaction. Our assay distinguished covalent from non-covalent binding, and mapped the residue for covalent binding between bioactive constituents and specific proteoforms of the target protein. We augmented our approach with a nanoflow liquid chromatography-selected reaction monitoring (SRM)-MS assay for simultaneous identification and label-free multiplex quantitation of small molecules in the crude botanical extracts. Our assay was validated for various proteoforms of human serum albumin, which plays a key role in pharmacokinetics of small molecules *in vivo*. We demonstrated the utility of our proteoform-specific assay for evaluating thymoquinone in crude botanical extracts, studying its pharmacokinetics in human blood, and interpreting its toxicity to human breast cancer cells in tissue culture.

*Correspondence should be addressed to: Daojing Wang, Ph.D., Newomics Inc., 5980 Horton Street, Suite 525, Emeryville, CA 94608, wang@newomics.com.

Author Contributions

G.G. and P.M. designed and performed experiments and wrote the manuscript. B.A., Z.A., A.G.C., I. A. K., and L.A.W. provided crude botanical extracts and contributed to the writing of the manuscript. D.W. directed the experiments and wrote the manuscript.

Conflict of Interest Disclosure

Geuncheol Gil, Pan Mao, and Daojing Wang are employees of Newomics Inc., which is commercializing some of the technologies described in this work with pending patent applications.

Supporting Information

This article contains Supplemental Figs. S1, S2, S3, and Table S1.

Graphical Abstract

Introduction

Drug discovery from natural products (NPs) is undergoing a renaissance.^{1, 2} Identification of the specific binding between small molecules and proteins is a critical step in drug and its target discovery. However, high-throughput analysis of small molecules, such as bioactive constituents of botanical species and their molecular targets, remains a key challenge in the field.^{3, 4} Currently, there is no rapid method to directly identify and quantify the protein binding of bioactive constituents. Traditionally, NPs are purified through multi-step preparative liquid chromatographic (LC) fractionations, subsequently determined for their concentrations in each fraction, and characterized for their activities using protein binding assays such as thermal shift assay (TSA),^{5, 6} and to a lesser extent, surface plasmon resonance (SPR).⁷ This workflow may introduce significant experimental errors due to the following: (i) the sample loss as a result of solubility changes as solvent compositions change during the fractionation steps; (ii) the potential degradation and physicochemical conversion of labile compounds during the lengthy process; (iii) the varying distribution of the same compound in multiple LC fractions; and (iv) the potential competition and interference from closely-related compounds. More importantly, because TSA and SPR lack detailed molecular characterizations, it is challenging to differentiate specific from non-specific binding, and covalent from non-covalent binding. On the other hand, majority of proteins undergo posttranslational modifications (PTMs) such as phosphorylation and acetylation. PTMs may lead to changes in the protein conformation and accessibility of certain domains and residues, depending on the identity and location of the PTMs. As a result, the protein activity and binding to small molecules can be altered. Therefore it is important to study small molecule binding at the proteoform level.

Herein, we report the development, validation, and application of a new workflow for directly studying the bioactivity of crude botanical extracts using specific proteoforms (Figure 1a). Recent technological developments in high-resolution mass spectrometry have enabled the top-down mass spectrometry analyses of full-length proteins.^{8–14} We here demonstrate a nanoelectrospray ionization mass spectrometry (nanoESI-MS) assay of full-length protein targets for characterizing their proteoforms and binding to bioactive constituents in crude botanical extracts without chromatographic pre-fractionations. The assay is supplemented with top-down and bottom-up ESI-MS/MS to identify the covalent binding site, and a multiplex nanoflow liquid chromatography-selected reaction monitoring-

MS (nanoLC-SRM-MS) method for simultaneous identification and label-free quantitation of multiple NPs. There are at least three novel aspects of our overall workflow. First, it directly identifies and measures the stoichiometry of the specific binding between small molecules and their protein targets at the proteoform level in a proteoform-specific manner; furthermore, it differentiates covalent from non-covalent binding. Second, it does not require lengthy preparative fractionation procedures, and hence enables rapid analysis of NPs using small amounts (ng- μ g) of botanical crude extracts as compared to the typically larger (μ -mg) amount needed for conventional methods. Third, it correlates the binding stoichiometry between NPs and their protein targets with the amount of the bioactive constituents in whole crude extracts rather than their various fractions, thus minimizing the experimental errors in quantitation.

Results and Discussion

We first developed a top-down ESI-MS assay for identification and quantitation of protein binding of NPs (Figure 1b). Human serum albumin (HSA) was chosen as a representative protein target in this study. HSA is the most abundant protein in human serum and plays a key role in pharmacokinetics of NPs. It is highly post-translationally modified and serves as a good model system for validating our proteoform-specific binding assay. In fact, HSA is a critically important protein for drug discovery and development, because of its propensity for binding many exogenous compounds (covalent and non-covalent, reversible and irreversible), e.g. drugs and dietary constituents, in a highly specific and stereo-selective fashion, at physiological or therapeutic concentrations. The *in vivo* binding can alter pharmacokinetics of these small molecules and affect their efficacies and toxicities.

We optimized the mobile phase (20 mM ammonium acetate, pH 5.75) by titrating the pH and salt concentrations to maximize the nanoelectrospray ionization (nanoESI) efficiency of full-length proteins at their native states. Under our direct infusion native nanoESI-MS conditions, we observed both the non-native (denatured and partially-denatured) and native forms of HSA, with the former having the higher charge states (with the highest peak intensity at 42+), while the latter shows lower charge states (15+ to 18+) (Figure 1b(i)). At each charge state (zoom-in showing the 42+ and 16+ charge states, respectively), we observed the unmodified (Un) HSA, and different adducts including cysteinylated (+Cys), glycosylated (+Glyc), and cysteinylated plus glycosylated (+CysGlyc) proteoforms of HSA. Although we performed mass spectrometry under the solvent conditions preferable to native isoforms, we detected both non-native (the most intense charge state: 42+) and native isoforms depending on the protein charge states (charge states: 15+, 16+); this finding is consistent with those native MS studies published previously that focused mostly on low charge state regions for native isoforms^{15–17}. We also performed mass spectrometry under denaturing conditions using a solvent containing a high organic content (50% acetonitrile and 0.2% formic acid in water) (Supplemental Figure S1) and observed the denatured isoforms at even higher charge states (the most intense charge state: 50+). We identified and quantified the relative abundance of each proteoform of HSA on both native and non-native regions, based on the deconvoluted top-down ESI-MS spectra (Figure S1), confirming what we had previously described in more detail.¹⁸ By looking at both native and non-native regions, our assay directly differentiated non-binding, covalent-binding, and non-covalent

binding between HSA and small molecules including NPs. In addition, the binding stoichiometry (the number of molecules per HSA) can be determined by quantifying the relative abundance of their proteoforms. If there is no adduct observed in either region, it suggests that there is no binding between the small molecule and HSA under our conditions. However, if the adduct is observed in the native region but not in the non-native region, it suggests that non-covalent binding between HSA and the small molecule is formed. For example, non-covalent binding between Rhein and HSA did not change the non-native region of HSA, while the native region showed the binding of 1 to 6 of Rhein molecules per HSA based on the mass increases of each HSA adduct (Figure 1b(ii)). Indeed, an X-ray crystallographic and fluorometric analysis showed that Rhein strongly binds to the IIA subdomain of HSA through hydrogen bonds and salt bridges.¹⁹ It is possible that the conformational changes of HSA after Rhein binding may facilitate binding of additional Rhein molecules in the solution phase under our ESI-MS conditions. Follow-up studies using alternative approaches such as isothermal titration calorimetry (ITC) and microscale thermophoresis (MST) will help confirm the binding stoichiometry and understand the mechanism.

On the other hand, if the adduct is observed in both native and non-native regions, it suggests that stable binding between the small molecule and HSA takes place. Because a small molecule could bind non-covalently to a denatured or partially-denatured protein, to determine whether it is covalent binding, another experiment was performed under the denaturing ESI-MS conditions using a solvent containing 50% acetonitrile and 0.2% formic acid in water. If the same mass shift still exists, it suggests covalent binding, or that it is otherwise stable enough under the denaturing ESI solvent conditions. To further differentiate covalent binding from stable non-covalent binding, we used MS/MS fragmentation (top-down and bottom-up MS/MS). Thymoquinone (TQ) is a natural product that has generated significant interests due to its potential for anti-tumor, anti-diabetes, anti-inflammation, and many other multifaceted activities.^{20–24} However, the role of TQ in each disease and the underlying mechanism remain elusive. As shown in Figure 1b(iii), the covalent binding between TQ to HSA increased the protein masses in both the non-native and native regions, with the populations of unmodified HSA dramatically decreased and TQ-bound HSA significantly increased, respectively. We also observed the same mass increase under the denaturing ESI-MS conditions for the fully-denatured TQ-bound HSA, confirming the covalent binding of TQ to HSA (Figure S1). Furthermore, different buffer conditions such as pH and ionic strength could be selected in our assays to study the binding modes and potentially assess the binding kinetics and thermodynamics of the non-covalent binding between NPs and their protein targets. For example, depending on the isoelectric point of the protein, changing the pH may alter the charge states of the full protein and/or the specific residues for binding. Given the resolution of the QE-Plus at m/z 1,000–2,000 range where the HSA charge envelope was observed, the lower limit on small molecules that can be detected as binding to HSA (detected MW at 66 kDa) will be around 10 Da. In fact, we were able to resolve a minor oxidized HSA (16 Da difference) from the un-modified HSA (data not shown). The lower limit of the size of small molecules would increase with larger proteins, since MS resolution is reduced at higher m/z due to the QE-Plus resolution constraint. To further validate the binding specificity, docking may be performed to predict

the binding domain for non-covalent binders, followed by experimental validation. For example, H/D exchange experiments can be performed to pinpoint the binding site and residues involved, using either the full-length protein or a protein fragment containing the potential binding domain.

We next validated the specificity of our assays by probing the binding between HSA and TQ (Figure 2). TQ is a major active component of *Nigella sativa* (black seed). Carvacrol (CR) and thymol (THY) are bioactive isomers that are putative precursors of TQ in biosynthetic pathways, and in *N. sativa*, CR is thought to be the main precursor of TQ.²⁵ The production of CR, THY, and TQ in essential oils can vary substantially according to climate, geographical location, and vegetative stage for different plants.²⁶ Therefore, a sensitive and robust assay for studying the bioactivity of TQ and its related compounds such as CR and THY will speed up the mechanistic studies by rapid evaluation of their bioavailability and stability. We first confirmed that our assay could quickly differentiate TQ from closely-related compounds in HSA binding (Figure 2a). We confirmed that THY (and CR, an isomer of THY, see Supplemental Figure S2) did not bind HSA. On the other hand, closely-related methyl-benzoquinone and dimethyl-benzoquinone bound HSA; however they could be quickly differentiated from TQ because of the difference in mass increases for HSA adducts due to the covalent binding. We next studied whether other small molecules in the biological matrices could compete or interfere with TQ for covalent binding to HSA (Figure 2b). In the presence of an equimolar amount of free cysteine (Cys) or glutathione (GSH), the binding of TQ to HSA was dramatically reduced, since TQ reacted with both of them, resulting in TQ-Cys or TQ-GSH adducts (Figure 2b(iii) and Figure 2b(v)), respectively. This suggested that the bioactive TQ in the biological matrices such as cell culture media and blood samples could be overestimated from the initial TQ added into the system, because the reactive sulfur compounds are rich in these matrices. To check if our assay could be applied in physiological matrices such as human plasma and determine the free TQ in *N. sativa* extracts, we analyzed the binding between HSA in pooled human plasma and two commercial dietary supplements of black seed oils from *N. sativa* (Figure 2c). We detected a significantly higher concentration of free TQ in seed oil A than that in seed oil B by our nanoLC-SRM-MS assay, which is also consistent with the specification provided by the two manufacturers, respectively. (More detailed quantitation of free TQ in crude extracts using nanoLC-SRM-MS assay is described below in Figure 4b).

We further confirmed the covalent binding between HSA and TQ and conclusively identified Cys34 as the only binding residue for TQ by combining top-down and bottom-up MS/MS (Figure 3). Although Cys34 is the only free cysteine residue of HSA, to the best of our knowledge, no literature has demonstrated the covalent binding of TQ to Cys34 of HSA. We performed top-down MS/MS for various HSA proteoforms, including unmodified (HSA), glycosylated (HSA-Glyc), cysteinylated (HSA-Cys), and TQ-bound (HSA-TQ), respectively. The charge state of +60 for each proteoform was chosen as the precursor ion for fragmentation because it gave the best series of N-terminal fragment ions including b34 which contains Cys34 (Figure 3a(i)). The observed charge states for b ions were mostly +4 to +7 (Figure 3a(ii)–(v)). The m/z and isotopic peak patterns (inserts) of each b34 (+7) ion matched with the expected values for each proteoform. The monoisotopic peaks for b34 ions from unmodified HSA were observed at m/z 651.00 (+6) and m/z 558.15 (+7) (Figure

3a(ii)). The same b34 ion (m/z 558.15, +7) was observed for HSA-Glyc even though its precursor ion (m/z 1110.97, +60) had a different m/z from that of HSA (m/z 1108.31, +60) (Figure 3a(iii)). A relatively lower signal was detected for HSA-Glyc than for HSA because of its lower abundance at the proteoform level (Figure S1). Furthermore, the ionization and fragmentation efficiencies could be slightly different for HSA-Glyc and HSA. Importantly, glycated b34 ions were not detected at either m/z 678.01 (+6) or m/z 581.30 (+7), indicating glycation was unlikely on Cys34 or any other residue within the b34 ion. In fact, the glycation of HSA is known to be predominantly at its Lys525 residue¹⁸. In contrast, the fragmentation of HSA-Cys produced a b34 ion at m/z 575.15 (+7), suggesting cysteinylolation was most likely on Cys34 (Figure 3a(iv)). We also observed a minor peak of the unmodified b34 ion at m/z 558.15 (+7) from HSA-Cys (Figure 3a(iv)). The peak was not from the unmodified b34 ion of HSA because the precursor ion of HSA-Cys was chosen for MS/MS. It was consistent with the cysteinylated b34 ion of HSA-Cys after the Cys was released from Cys34, because the disulfide bond was not stable for the short peptide in the gas phase. Much less likely, it could come from the residual HSA-Glyc, because the 0.68 Da mass difference between the precursor ions of HSA-Cys (m/z 1110.29, +60) and HSA-Glyc (m/z 1110.97, +60) might not completely exclude the other ion, given the narrowest isolation window (m/z , ± 0.2 Da) on the Q Exactive Plus. On the other hand, the fragmentation of HSA-TQ (m/z 1110.97, +60) produced the TQ-bound b34 ion at m/z 581.30 (+7) (Figure 3a(v)). The monoisotopic mass of TQ is 164, but its covalent binding to Cys34 resulted in a mass shift of 162 due to the loss of two protons for the formation of the S-C bond. The measured mass difference of 162 Da between the unmodified and TQ-modified b34 ions confirmed the TQ modification on Cys34. Furthermore, we did not observe the unmodified b34 ion at m/z 575.15 (+7) for HSA-TQ, indicating: (a) The Cys34 was completely modified by TQ for HSA; and (b) HSA-Glyc which had the same precursor mass (m/z 1110.97, +60) as that of HSA-TQ was also modified by TQ and converted to HSA-Glyc-TQ. These data are consistent with our top-down MS analysis showing that HSA-Glyc was completely converted to HSA-Glyc-TQ (Figure 2b(ii)). Our data also suggest the strong covalent bond between TQ and Cys34, and furthermore, that Cys34 is the only Cys residue in HSA bound to TQ. Otherwise, we would have observed the unmodified b34 (+7) ion as the major peak for HSA-TQ, as was the case for HSA-Glyc (Figure 3a(iii)). We also performed bottom-up LC-MS/MS analysis of the tryptic digests of HSA with and without the bound TQ. Based on the MS/MS fragmentation patterns of the tryptic peptides ALVLIAFAQYLQQC₃₄PFEDHVK, we conclusively identified the residue Cys34 of HSA as the covalent binding residue for TQ, and furthermore, the binding stoichiometry was one TQ per HSA (Figure 3b).

We then demonstrated the utility of our assays for studying the bioavailability, pharmacokinetics, and toxicity of NPs using TQ as an example (Figure 4). The HSA-TQ level (HSA-TQ%) was calculated by the peak intensities on the charge-deconvoluted spectra (Figure S1), and the calibration curve was obtained for HSA-TQ% as a function of the molar ratio of TQ/HSA (Figure 4a). We obtained a linear correlation between HSA-TQ% and TQ/HSA molar ratio in the range of 0.1–1.0. We also developed a multiplex nanoLC-SRM-MS method for simultaneous identification and quantitation of multiple compounds including TQ, CR, and THY in the crude botanical extracts (Figure S2). The accuracy,

precision, reproducibility, recovery, limit of detection (LOD), and limit of quantification (LOQ) of the nanoLC-SRM-MS method were determined (Supplemental Figure S3 and Table S1). The recovery calculated in Table S1 was used to validate our SRM assay for accurate quantitation of NPs in the crude botanical extracts. We then used this method to determine the concentrations of the free TQ in the crude extracts of five plants and two dietary supplements, and correlated them to the binding between TQ and HSA (HSA-TQ%) as quantified by top-down ESI-MS (Figure 4b). A strong correlation between the free TQ and HSA-TQ% ($r=0.97$, $p<0.05$) was observed, suggesting our top-down assay of proteoform sensors could directly link bioactivity with bioavailability of NPs such as TQ in the crude botanical extracts. We further investigated the feasibility of using our assay for studying the pharmacokinetics and toxicity of TQ in the biological matrices such as human blood samples and cell culture media (Figure 4c–4d). There have been strong interests in TQ's anti-tumor activities²³. TQ has been shown to affect 9 out of 10 well-accepted cancer-relevant pathways²². This implies that TQ may impact a wide range of molecular targets *in vivo*. However, there have been conflicting results in the literature for both *in vitro* and *in vivo* studies of TQ. Because we unexpectedly detected the covalent binding between Cys34 of HSA and TQ, we hypothesized that: (1) the HSA binding affects TQ pharmacokinetics in the human blood; and (2) the effective TQ concentration could be masked by bovine serum albumin (BSA) contained in the 10% FBS media typically used for *in vitro* cell culture studies, which may explain the discrepancy in IC₅₀ values of TQ obtained from various studies. We confirmed that *ex vivo* incubation of either male or female human blood samples with TQ standards or crude botanical extracts resulted in the covalent binding of HSA to TQ (Figure 4c). Interestingly, much higher populations of unmodified HSA were observed in the two whole blood samples than that in the HSA standard (Figure 4c(i)). As expected, we did not detect TQ-bound HSA after TQ was removed from the crude extract of *N. sativa* prior to its reaction with human plasma (Figure 4c(ii)). Unmodified HSA in the plasma was completely modified by the extract of *M. fistulosa*, but only partially by that of *N. sativa*, consistent with the data in Figure 4b where HSA standard was used. We then performed toxicity studies of TQ on a human breast cancer cell line MDA-MB-231 using cell culture media containing 10%, 5%, and 0% FBS for 48 hours, respectively (Figure 4d). We determined that IC₅₀ was 0.7 μ M and 10 μ M for 0% and 5% FBS, respectively, with high statistical significance. Strikingly, the apparent IC₅₀ in the 10% FBS medium was above 20 μ M, much higher than that in the 0% FBS medium. However, we could not obtain a full dose-response for 10% FBS because TQ covalently bound to BSA. Using our top-down ESI-MS assay, we confirmed that around 1/3 of BSA in the 10% FBS medium was TQ bound (data not shown), confirming the significant masking of effective TQ concentrations by BSA in the 10% FBS medium. Our results suggested that caution should be taken to evaluate the toxicity and efficacy of bioactive NPs using cell-based assays with different medium compositions.

By combining top-down and bottom-up ESI-MS studies of the proteoforms of protein targets, we have developed a rapid assay of the binding between specific proteoforms and NPs for directly evaluating the bioactivity of crude botanical extracts. The method is augmented by a multiplex nanoLC-SRM-MS assay for direct analysis of bioactive NP constituents and their closely-related compounds in crude botanical extracts. Our workflow

diminishes the lengthy preparative fractionation steps of NPs, thus enabling rapid analysis and minimizing experimental errors. Our study may have practical significance. For example, our method can be used to quickly screen bindings between a protein target (not limited to HSA in the current study) and NPs in a wide range of plants using small amounts (ng- μ g) of their crude extract or minimally-prepared fractions. If a hit is identified, then efforts can be focused on those crude extracts and/or fractions for further purification and isolation of target NPs. In contrast, conventional methods typically start with larger amounts (μ g-mg) of crude extracts in the multi-step chromatographic fractionation in order to generate many NP fractions for downstream activity evaluation. Our approach will dramatically increase the throughput of drug discovery for NPs to target particular proteins. For botanical drug candidates, identification of active constituents is highly recommended by FDA (though not always essential) for clinical trials and for QC purposes. Our method can be used to quickly confirm and quantify the bioactive constituents in botanical drugs. One of the key goals for pharmacokinetics studies is to determine the bioavailability of drugs by measuring the concentrations of free and protein-bound drugs. Given that HSA accounts for ~60% of total proteins in the human plasma and is the dominant drug binder, our method coupled with the aforementioned nanoLC-SRM-MS assay could greatly facilitate the pharmacokinetics studies of NPs and their bindings to HSA. Some NPs including TQ and Rhein are regarded by some to be promiscuous (or pan-assay interference compounds, PAINS) in their behavior during screening campaigns.²⁷ In spite of the label “promiscuous”, the behavior of such compounds is highly target- and assay-dependent. Given the sensitivity and specificity of our assays, we will be able to clarify the modes and stoichiometry of binding *in vitro* against different targets, thereby providing more mechanistic understanding of the selectivity and specificity of their behaviors *in vitro* and *in vivo*. By comparing them to different classes of compounds, we will gain new mechanistic insights into PAINS and other NPs. Finally, our method may find applications in studying the binding between proteins and small molecules other than NPs, for example, antibody-drug conjugates (ADCs) and their pharmacokinetics *in vitro* and *in vivo*, or reciprocal screening for a specific small molecule targeting a protein library, or for a specific protein targeted by a drug library. Thus it may contribute to mechanistic studies and rational drug design for drug screening and discovery.

Methods

Materials

Human serum albumin (HSA), thymoquinone (TQ), carvacrol (CR), thymol (THY), Rhein, and caffeine-¹³C₃ were purchased from Sigma-Aldrich. Black seed dietary supplement A (Lot#80814, expiration date 08/2017) and B (Lot#AUSTL168980) were obtained from Amazing Herbs Nutraceuticals and Hab Shifa, respectively, via Amazon.com. Sequencing grade trypsin was purchased from Promega. LC-MS grade acetonitrile, ammonium acetate, dimethyl sulfoxide (DMSO), formic acid, methanol, trifluoroethanol, and water were purchased from Sigma-Aldrich. Human pooled plasma (Lot#14168) and human whole blood samples collected from unidentified healthy male and female individuals were obtained from Innovative Research Inc. Breast cancer cell line MDA-MB-231 was obtained from ATCC and cell culture media were obtained from Thermo Fisher Scientific.

Sample Preparation for top-down ESI-MS

Compound standards, crude botanical extracts, and dietary supplements were prepared in DMSO or methanol. For direct infusion ESI-MS analysis of native proteins, HSA standard, human plasma, or blood was diluted in the mobile phase-1 (MP-1, 20 mM ammonium acetate, pH 5.75) and mixed with compound standards, crude botanical extracts, or dietary supplements. The mixture was incubated at 37°C for 2 hours in the dark and then loaded into a Microcon 30 kDa MWCO filter (Millipore) for centrifugation at $14,000 \times g$ for 15 min. The sample retained in the filter was washed three times with MP-1 and recovered with 50 μ L of MP-1. The filtered sample was centrifuged at $14,000 \times g$ for 15 min to remove any precipitate prior to ESI-MS analysis. For direct infusion ESI-MS analysis of denatured proteins, after following the same incubation procedures, the incubated sample was loaded into a GHP NanoSep MF centrifugal device (0.45 μ m filter, Pall Bioscience) and centrifuged at $14,000 \times g$ for 1 min. The flow-through was loaded into a MacroSpin column (C4, 5 μ m, Harvard Apparatus), centrifuged at $110 \times g$ for 2 min, and washed three times with 400 μ L of 20% (v/v) acetonitrile, 0.2% (v/v) formic acid in water at $110 \times g$ for 2 min. The protein was eluted with 100 μ L of mobile phase-2 (MP-2, 50% (v/v) acetonitrile, 0.2% (v/v) formic acid in water) by centrifugation at $110 \times g$ for 2 min. The eluate was centrifuged at $14,000 \times g$ for 5 min to remove any precipitate prior to ESI-MS analysis.

Sample Preparation for bottom-up LC-MS/MS

HSA standard and human plasma with and without reaction with TQ standard and crude botanical extracts, respectively, were reconstituted with 50 mM ammonium bicarbonate, mixed (1:1) with trifluoroethanol (TFE), and incubated at 90°C for 15 min for heat denaturation. The denatured protein was diluted 10 $\times$ with 50 mM ammonium bicarbonate and digested with trypsin (enzyme: substrate mass ratio at 1:40) with overnight incubation at 37°C. After the first digestion, the same amount of trypsin was added for the second digestion at 37°C for 4 hours and then formic acid was added (final concentration at 1%) to stop the digestion. The tryptic digests were dried under SpeedVac and reconstituted with 3% acetonitrile, 0.2% formic acid in water for LC-MS/MS analysis.

Sample Preparation for natural product analysis

The plant material of *Monarda bradburiana*, *Monarda fistulosa*, *Monarda punctata*, and *Thymus caucasicus* was obtained from Missouri botanical garden, St. Louis, Missouri, with taxonomic verification. The seeds of *N. sativa* were purchased from Frontier Natural Products Co-op. Crude extracts from different plants were prepared at the National Center for Natural Products Research (University of Mississippi) as follows: the seeds of *N. sativa* (~30 g) were extracted with hexanes (300 mL $\times$ 10 min $\times$ 3 times) at 21°C and the solvent was removed in vacuum to afford an oily residue. The whole plant samples of *M. bradburiana*, *M. fistulosa*, *M. punctata*, and *T. caucasicus* (10–30 g) were extracted with 95% ethanol (200–300 mL $\times$ 10 min $\times$ 3 times) at 37°C. After solvent removal by vacuum, the gummy materials were freeze-dried, yielding 3–10% of dry weight on average.

Standard TQ, CR, and THY stock solutions were prepared in methanol with a concentration of 200 μ M. It was further diluted four-fold (1:4) with 0.2% formic acid in water to make a solution of 50 μ M in 25% methanol/0.05% formic acid in water, which was the solvent

optimized for NPs in our LC-MS analysis for best recovery and sensitivity. Further dilutions were made as needed with 25% methanol/0.05% formic acid in water. Lyophilized crude extracts from different plants were dissolved in methanol to make 2 $\mu\text{g}/\mu\text{L}$ stock solutions. They were diluted four-fold with 0.2% formic acid in water to make a solution of 0.5 $\mu\text{g}/\mu\text{L}$ in 25% methanol/0.05% formic acid in water. Again, further dilutions were made with 25% methanol/0.05% formic acid in water if necessary. Dietary supplement products of black seed oil were dissolved using a ratio of 0.04% oil in methanol and vigorously vortexed for 30 min. After centrifugation at $2000 \times g$ for 10 min, the supernatant was further diluted (1:4) with 0.2% formic acid in water to make 25% methanol/0.05% formic acid in water. The diluted samples of crude extracts and dietary supplements were filtered using 0.45 μm GHP NanoSep MF Centrifugal devices (Pall Life Sciences) to remove insoluble small particles and precipitates from plants. The resultant samples were further diluted (1:1) with or without different concentrations of standards prior to the LC-MS analysis.

To measure the recovery of TQ, CR, THY during our sample preparation, three separate sample sets were prepared. In the first set (S1), TQ, CR, and THY standards were spiked into the crude extracts or in the dietary supplement and then samples underwent serial dilutions and filtration prior to LC-MS. The second set (S2) was prepared in the same way as the first set, except that the standard compounds were spiked after the serial dilutions and filtration and before the LC-MS runs. The third set (S3) was prepared in the same way as the first two, but without the standard compounds spiked in. Caffeine- $^{13}\text{C}_3$ was spiked into all three sets of samples prior to LC-MS. The run order of the three sets of samples were randomized and LC-MS analysis was performed in triplicate for each sample.

Top-down ESI-MS analysis of proteins

For characterizing different proteoforms of HSA, proteins samples were directly infused to a Q Exactive Plus mass spectrometer (QE-plus, Thermo Fisher Scientific) using a syringe pump at 300 nL/min using a Thermo metal emitter (30 μm ID) on a Nanospray Flex source. Fragmentation was performed with AGC target of 1×10^6 , 140,000 resolving power at m/z 200, precursor isolation window of ± 0.2 Da, and higher energy collisional dissociation (HCD) at CE 20. The spray voltage was set to 2.0 kV, the temperature of the heated capillary was held at 250°C, and S-Lens RF level was 55.

For quantifying the NP binding of HSA, protein samples were directly infused to a hybrid quadrupole/orthogonal TOF (Q-TOF) API US mass spectrometer integrated with a CapLC (Waters) at 600 nL/min using a New Objective Picotip emitter (30 μm ID). For the native protein ESI-MS, capillary voltage was set at 2.1 kV, sample cone voltage 40 V, extraction cone voltage 1.5 V, source temperature 60°C. For denatured protein ESI-MS, MS parameters were the same as above except that the source temperature was kept at 120°C.

NanoLC-MS/MS analysis of peptides

Tryptic digests were analyzed on a hybrid quadrupole/orthogonal TOF (Q-TOF) API US mass spectrometer integrated with a CapLC (Waters Corp.). Capillary voltage was 2.1 kV, sample cone voltage 40 V, extraction cone voltage 1.5 V, source temperature 120°C. Peptides were separated on a Thermo PepMap C18 column (75 μm ID $\times$ 150 μm L, 5 mm)

with a Thermo PepMap C18 trap column (0.3 mm ID $\times$ 5 mm L, 5 μ m). The mobile phase consisted of 3% acetonitrile/0.2% formic acid in water (A) and 97% acetonitrile/0.2% formic acid in water (B). Flow rate was 600 nL/min in the LC gradient as follows: 7.5% B for 3 min, 7.5–40%B for 40 min, 40–70%B for 2 min. New Objective Picotip emitter (10 μ m ID) was used.

NanoLC-SRM-MS analysis of natural products

SRM assay was performed on a TSQ Quantiva Triple Quadrupole Mass Spectrometer interfaced with an UltiMate 3000 nanoUPLC system (Thermo Fisher Scientific). Samples were separated on a Halo Peptide ES-C18 column (0.1 mm ID $\times$ 150 mm L, 5 μ m, Advanced Materials Technology Inc.) with a PepMap C18 trap column (0.3 mm ID $\times$ 5 mm L, 5 μ m, Thermo Fisher Scientific). A Thermo stainless steel emitter (30 μ m ID) was used and the injection volume was 1 μ L. The mobile phase consisted of 3% acetonitrile/0.2% formic acid in water (A) and 97% acetonitrile/0.2% formic acid in water (B). Flow rate was 1 μ L/min unless otherwise mentioned in the LC gradient as follows: 5%B for 3 min, 5–40%B for 5 min, 40%B isocratic for 6 min at 600 nL/min, 40–60%B for 2 min, 99%B for 3 min, and 5%B for 6 min. A hyperbolic gradient (curve 1) was applied to the 5–40%B to achieve a complete separation of TQ and the isomeric CR and THY. Tandem MS spectra of CR, THY, and TQ were acquired using the product ion scanning mode. Transitions of m/z 151/109 and m/z 151/91 for CR and THY, and transitions of m/z 165/137 and m/z 165/109 for TQ were selected based on the tandem MS spectra. The multiplex SRM assay was developed by optimizing the parameters as follows: collision energies were 10 V for transition m/z 151/109 and 20 V for m/z 151/91, and 15 V for m/z 165/137 and 25 V for m/z 165/109, ion transfer tube temperature 350 $^{\circ}$ C, and RF lens 50 for CR and THY, and 52 for TQ.

Treatment of MDA-MB-231 cells with TQ standards

MDA-MB-231 breast cancer cells were cultured in DMEM (1X) + GlutaMAX-I (Gibco) with heat-inactivated FBS (Gibco) and 100 U/mL penicillin-streptomycin (Gibco) at 37 $^{\circ}$ C and 5% CO₂. After seeding 4000 cells in 100 μ L media per well into a 96-well microplate, MDA-MB-231 cells were treated with 100 μ L media containing 0, 5, or 10% FBS with TQ standard at different concentrations. After 48 hour treatment, 10 μ L Cell Counting Kit-8 (CCK-8, Sigma) reagent was added to the media, followed by incubation at 37 $^{\circ}$ C for 2 hours. The cell viability was determined by the absorbance difference at 450 nm and 650 nm (A₄₅₀ – A₆₅₀) following vendor's protocol using a VersaMax Microplate Reader (Molecular Devices). Dose response curves were plotted and fitted with a 3-parameter nonlinear regression by GraphPad Prism 6.0. IC₅₀ value for each medium containing different FBS% was determined and correlated to the BSA-TQ% measured by top-down ESI-MS, as described for HSA-TQ%.

Data analysis

Error bars in all figures represent standard deviations of at least three technical replicates. Raw data of intact protein ESI-MS were processed using MassLynx 4.0 software. MS spectra were charge-deconvoluted using MaxEnt1 with the following parameters: the output mass range of 65,000 to 68,000, resolution of 0.1 Da/channel, uniform Gaussian damage

model with 0.4 Da width at the half-height, minimum intensity ratio of 80% for both left and right, and iteration to convergence. The deconvoluted data were further processed for background subtraction with a polynomial order of 25 with 40% below curve, spectrum smoothing with ± 6 smooth window using Savitzky-Golay method, and then centering spectrum. Given that TQ bound HSA has the same mass increase (+162) as glycation, we could not differentiate them simply by the mass increase. Instead, we calculated the percentage of TQ bound HSA (HSA-TQ%) using the approach shown in Supplemental Figure S1. Briefly, deconvoluted spectra were used to obtain the intensities of five prominent proteoforms (higher than 5% of Y-axis scale) from the HSA control: unmodified (Un, I_1), cysteinylated (+Cys, I_2), glycated (+Glyc, I_3), cysteinylated plus glycated (+CysGlyc, I_4), and doubled-glycated (+GlycGlyc, I_5) HSA. We next obtained the intensities of unmodified (Un, J_1), cysteinylated (+Cys, J_2), TQ modified or glycated (+TQ/Glyc, J_3), cysteinylated plus glycated (+CysGlyc, J_4), and TQ modified plus glycated or double-glycated (+GlycTQ/GlycGlyc, J_5) for HSA reacted with TQ standard or crude extracts. We calculated the percentage of unmodified and glycated proteoforms from the HSA control (background peaks) and that of unmodified one from HSA reacted with samples. The TQ bound HSA was calculated as the percentage of TQ modified or glycated (+TQ/Glyc, J_3) and TQ modified plus glycated or two-glycated (+GlycTQ/GlycGlyc, J_5), subtracted by the percentage of remaining glycated proteoforms in HSA reacted with samples ($I_{\text{Glyc}} \times (J_{\text{un}}/I_{\text{un}})$). Three basic assumptions were made for our calculation: (i) different HSA proteoforms with and without TQ bound had similar ionization efficiencies for top-down ESI-MS; (ii) the binding kinetics of TQ on the Cys34 residue of the unmodified HSA was the same as that on its glycated proteoforms; and (iii) the five prominent proteoforms that we have identified accounted for the majority of the HSA populations and the other modification levels were negligible.

For top-down MS on the Q Exactive Plus, MS/MS raw data underwent charge deconvolution using Xtract on the Qual browser (Xcaliber 4.0) and were exported as Xtract raw files. Then the post Xtract raw files were loaded onto ProSightPC3.0 to perform the search in the single protein mode against HSA sequence with a 10 ppm fragment tolerance.

For nanoLC-SRM-MS analysis, SRM raw data were processed using Skyline 3.1 (U of Washington) and Pinpoint 1.4 (Thermo Fisher Scientific) to integrate each peak area. The external calibration curve for TQ was obtained by calculating the peak areas (responses) for different amounts of the standard (0.05–50 pmole). The limit of detection (LOD) of TQ was determined at $S/N=3$ and the limit of quantification (LOQ) was determined at $S/N=10$. The interday and intraday variation was analyzed using multiple LC-MS runs of 1.5 pmole TQ standards injected between samples. The internal calibration curves of CR, THY, and TQ were also obtained by measuring spiked standards of three different amounts in each sample (See an example in Figure S3). The CR, THY, and TQ amounts in each sample were calculated using measured intensities of the endogenous compounds in each sample divided by the slopes obtained from the internal calibration curves. To calculate the recovery of the spiked-in TQ, CR, and THY standards, their intensities (peak areas) were first normalized with those of the spiked-in $^{13}\text{C}_3$ -caffeine using the sum of its SRM transitions at m/z 198/140 and m/z 198/112. The CVs for the intensity of $^{13}\text{C}_3$ -caffeine peak for the same plant samples were 4–5% during the LC-MS runs, indicating that our LC-MS was highly reproducible from run to run. The normalized peak area (response: R_{S_i}) in each sample set

(R_{S1} for S1, R_{S2} for S2, and R_{S3} for S3) for each compound spiked into a particular plant was used to calculate the overall recovery (%) as below:

$$\frac{R_{S1} - R_{S3}}{R_{S2} - R_{S3}} \times 100$$

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank M. Jacob for facilitating the access to the Natural Products Repository at National Center for Natural Products Research (NCNPR), and the colleagues at Newomics Inc. for technical assistance and helpful discussions. The NCNPR plant repository is partially supported by a cooperative agreement with the Agriculture Research Service, US Dept. of Agriculture (58-6060-6-015).

Funding Sources

The work was supported by the National Institutes of Health under the awards AT008297, ES022360, ES023529, GM109682, AG046025, AI106100, and HHSN261201300033C (to Newomics Inc.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Newman DJ, Cragg GM. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod*. 2012; 75:311–335. [PubMed: 22316239]
2. Luo Y, Cobb RE, Zhao H. Recent advances in natural product discovery. *Curr Opin Biotechnol*. 2014; 30:230–237. [PubMed: 25260043]
3. Mishra KP, Ganju L, Sairam M, Banerjee PK, Sawhney RC. A review of high throughput technology for the screening of natural products. *Biomed Pharmacother*. 2008; 62:94–98. [PubMed: 17692498]
4. Betz JM, Brown PN, Roman MC. Accuracy, precision, and reliability of chemical measurements in natural products research. *Fitoterapia*. 2011; 82:44–52. [PubMed: 20884340]
5. Vedadi M, Niesen FH, Allali-Hassani A, Fedorov OY, Finerty PJ Jr, Wasney GA, Yeung R, Arrowsmith C, Ball LJ, Berglund H, Hui R, Marsden BD, Nordlund P, Sundstrom M, Weigelt J, Edwards AM. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci U S A*. 2006; 103:15835–15840. [PubMed: 17035505]
6. Martinez Molina D, Jafari R, Ignatushchenko M, Seki T, Larsson EA, Dan C, Sreekumar L, Cao Y, Nordlund P. Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science*. 2013; 341:84–87. [PubMed: 23828940]
7. Rich RL, Myszka DG. Advances in surface plasmon resonance biosensor analysis. *Curr Opin Biotechnol*. 2000; 11:54–61. [PubMed: 10679342]
8. Tran JC, Zamborg L, Ahlf DR, Lee JE, Catherman AD, Durbin KR, Tipton JD, Vellaichamy A, Kellie JF, Li M, Wu C, Sweet SM, Early BP, Siuti N, LeDuc RD, Compton PD, Thomas PM, Kelleher NL. Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature*. 2011; 480:254–258. [PubMed: 22037311]
9. Han X, Wang Y, Aslanian A, Bern M, Lavalley-Adam M, Yates JR 3rd. Sheathless capillary electrophoresis-tandem mass spectrometry for top-down characterization of *Pyrococcus furiosus* proteins on a proteome scale. *Anal Chem*. 2014; 86:11006–11012. [PubMed: 25346219]
10. Sarsby J, Griffiths RL, Race AM, Bunch J, Randall EC, Creese AJ, Cooper HJ. Liquid Extraction Surface Analysis Mass Spectrometry Coupled with Field Asymmetric Waveform Ion Mobility

- Spectrometry for Analysis of Intact Proteins from Biological Substrates. *Anal Chem.* 2015; 87:6794–6800. [PubMed: 26066713]
11. Gault J, Donlan JA, Liko I, Hopper JT, Gupta K, Housden NG, Struwe WB, Marty MT, Mize T, Bechara C, Zhu Y, Wu B, Kleanthous C, Belov M, Damoc E, Makarov A, Robinson CV. High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nat Methods.* 2016; 13:333–336. [PubMed: 26901650]
 12. Cai W, Guner H, Gregorich ZR, Chen AJ, Ayaz-Guner S, Peng Y, Valeja SG, Liu X, Ge Y. MASH Suite Pro: A Comprehensive Software Tool for Top-Down Proteomics. *Mol Cell Proteomics.* 2016; 15:703–714. [PubMed: 26598644]
 13. Ntai I, Kim K, Fellers RT, Skinner OS, Smith ADt, Early BP, Savaryn JP, LeDuc RD, Thomas PM, Kelleher NL. Applying label-free quantitation to top down proteomics. *Anal Chem.* 2014; 86:4961–4968. [PubMed: 24807621]
 14. Toby TK, Fornelli L, Kelleher NL. Progress in Top-Down Proteomics and the Analysis of Proteoforms. *Annu Rev Anal Chem (Palo Alto Calif).* 2016; 9:499–519. [PubMed: 27306313]
 15. Benkestock K, Edlund PO, Roeraade J. Electrospray ionization mass spectrometry as a tool for determination of drug binding sites to human serum albumin by noncovalent interaction. *Rapid Commun Mass Spectrom.* 2005; 19:1637–1643. [PubMed: 15915474]
 16. Kleinova M, Belgacem O, Pock K, Rizzi A, Buchacher A, Allmaier G. Characterization of cysteinylolation of pharmaceutical-grade human serum albumin by electrospray ionization mass spectrometry and low-energy collision-induced dissociation tandem mass spectrometry. *Rapid Commun Mass Spectrom.* 2005; 19:2965–2973. [PubMed: 16178042]
 17. Li H, Wongkongkathep P, Van Orden SL, Ogorzalek Loo RR, Loo JA. Revealing ligand binding sites and quantifying subunit variants of noncovalent protein complexes in a single native top-down FTICR MS experiment. *J Am Soc Mass Spectrom.* 2014; 25:2060–2068. [PubMed: 24912433]
 18. Mao P, Wang D. Top-Down Proteomics of a Drop of Blood for Diabetes Monitoring. *J Proteome Res.* 2014; 13:1560–1569. [PubMed: 24533899]
 19. Li M, Lee P, Zhang Y, Ma Z, Yang F, Zhou Z, Wu X, Liang H. X-ray crystallographic and fluorometric analysis of the interactions of rhein to human serum albumin. *Chem Biol Drug Des.* 2014; 83:167–173. [PubMed: 24033934]
 20. Banerjee S, Padhye S, Azmi A, Wang Z, Philip PA, Kucuk O, Sarkar FH, Mohammad RM. Review on molecular and therapeutic potential of thymoquinone in cancer. *Nutr Cancer.* 2010; 62:938–946. [PubMed: 20924969]
 21. Bamosa AO, Kaatabi H, Lebdaa FM, Elq AM, Al-Sultanb A. Effect of *Nigella sativa* seeds on the glycemic control of patients with type 2 diabetes mellitus. *Indian J Physiol Pharmacol.* 2010; 54:344–354. [PubMed: 21675032]
 22. Schneider-Stock R, Fakhoury IH, Zaki AM, El-Baba CO, Gali-Muhtasib HU. Thymoquinone: fifty years of success in the battle against cancer models. *Drug Discov Today.* 2014; 19:18–30. [PubMed: 24001594]
 23. Darakhshan S, Bidmeshki Pour A, Hosseinzadeh Colagar A, Sisakhtnezhad S. Thymoquinone and its therapeutic potentials. *Pharmacol Res.* 2015; 95–96C:138–158.
 24. Amin B, Hosseinzadeh H. Black Cumin (*Nigella sativa*) and Its Active Constituent, Thymoquinone: An Overview on the Analgesic and Anti-inflammatory Effects. *Planta Med.* 2016; 82:8–16. [PubMed: 26366755]
 25. Jukic M, Politeo O, Maksimovic M, Milos M, Milos M. In vitro acetylcholinesterase inhibitory properties of thymol, carvacrol and their derivatives thymoquinone and thymohydroquinone. *Phytother Res.* 2007; 21:259–261. [PubMed: 17186491]
 26. Lukas B, Schmiderer C, Franz C, Novak J. Composition of essential oil compounds from different Syrian populations of *Origanum syriacum* L. (Lamiaceae). *J Agric Food Chem.* 2009; 57:1362–1365. [PubMed: 19178273]
 27. Baell J, Walters MA. Chemistry: Chemical con artists foil drug discovery. *Nature.* 2014; 513:481–483. [PubMed: 25254460]

Figure 1.

A proteoform-specific assay for direct analysis of bioactive constituents in crude botanical extracts. **(a)** A scheme showing our workflow for analysis of the binding, including non-binding, covalent binding, and non-covalent binding, between the natural products (NPs) in crude botanical extracts and specific proteoform targets. (i) Top-down MS identifies NP binding to specific proteoforms after a one-step reaction; (2) Top-down and bottom-up MS/MS confirm the covalent binding and identify the binding site; and (3) SRM-MS quantifies the NPs and correlate their concentrations to proteoform binding. Symbols with different colors and shapes represent individual compounds. PTM: posttranslational modification. **(b)** Top-down ESI-MS assay for the binding between natural products and HSA. Both the non-native and native isoforms of HSA were detected at different charge states. Zoom-in regions show the different proteoforms of HSA at charge state 42+ and 16+, respectively. Representative ESI-MS spectra show HSA reacted with (i) DMSO control for non-binding; (ii) Rhein for non-covalent binding; and (iii) TQ for covalent binding, respectively. HSA: human serum albumin; Un: unmodified; +Cys: cysteinylolation; +Glyc: glycation; +CysGlyc: cysteinylolation plus glycation; TQ: thymoquinone.

Figure 2.

Validation of the specificity of our top-down mass spectrometry centric assay. **(a)** Deconvoluted ESI-MS spectra of HSA reacted with (i) thymol; (ii) methyl-benzoquinone; and (iii) dimethyl-benzoquinone. **(b)** ESI-MS spectra of (i) TQ standard; (iii) TQ-Cys adduct; and (v) TQ-GSH adduct (left panels) along with the deconvoluted ESI-MS spectra of HSA reacted with (ii) TQ standard; (iv) TQ standard in the presence of free Cys; and (vi) TQ standard in the presence of GSH (right panels). Cys: cysteine; GSH: glutathione. **(c)** Deconvoluted ESI-MS spectra of HSA in human plasma reacted with (i) Control (DMSO); (ii) Seed oil A; and (iii) Seed oil B of *N. sativa*, respectively.

Figure 3.

Confirmation of the covalent binding of TQ to HSA through Cys34 by top-down and bottom-up mass spectrometry. **(a)** Top-down MS. (i) N-terminal fragment ions matched with HSA top-down data; and MS/MS spectra of the +60 charged precursor ions for: (ii) unmodified HSA (HSA, m/z 1108.31); (iii) glycosylated HSA (HSA-Glyc, m/z 1110.97); (iv) cysteinylated HSA (HSA-Cys, m/z 1110.29); and (v) TQ modified HSA (HSA-TQ, m/z 1110.97). The inserts show the zoom-in view of the isotopic peaks for b_{34} (+7) ion corresponding to each proteoform. b ions are from the N-terminal. **(b)** Bottom-up MS. ESI-MS/MS spectra of Cys34 containing tryptic peptides of: (i) unmodified HSA control; and (ii) TQ modified HSA, respectively. y ions are from the C-terminal and b ions are from the N-terminal.

Figure 4.

Applications of our top-down mass spectrometry centric assay in evaluating the pharmacokinetics of NPs and interpreting their toxicity data. **(a)** Titration curve in molar ratios for TQ standard reacted with HSA standard. Zoom-in region shows the linear response of the percentage of TQ bound HSA (HSA-TQ%) to the TQ/HSA ratio between 0.1–1.0. The maximum HSA-TQ% is ~25%, consistent with the unmodified HSA at ~30% calculated from Figure 1b for the same HSA standard. **(b)** HSA-TQ% measured for the crude extracts of (i) 5 different plants, and (ii) two dietary supplements of *N. Sativa*, compared to their corresponding endogenous TQ concentrations shown in (iii) and (iv), respectively. The mass ratios between HSA and total crude extracts were 1:1. The correlation between HSA-TQ% and free TQ concentration was shown with a Pearson's correlation curve ($p < 0.05$). **(c)** Deconvoluted ESI-MS spectra of HSA in (i) both male and female whole blood samples reacted with DMSO control or TQ standard, and (ii) human plasma samples, after reaction with DMSO control (the 1st panel), crude extract of *N. sativa* with the removal of TQ (the 2nd panel), crude extract of *N. sativa* without removal of TQ (the 3rd panel), and crude extract of *M. fistulosa* (the 4th panel). **(d)** (i) Dose response curves for MDA-MB-231 breast cancer cells treated with TQ standard in cell culture media containing 0%, 5%, and 10% FBS, respectively; and (ii) The percentage of TQ bound BSA (BSA-TQ%, black bars) and the measured IC₅₀ values for TQ (orange dot line) in culture media containing 0%, 5%, and 10% FBS, respectively.