Published in final edited form as:
Nat Methods. 2015 November ; 12(11): 1055–1057. doi:10.1038/nmeth.3590.

Proteome-wide small molecule and metabolite interaction mapping

Kilian V. M. Huber[1,3,4], **Karin M. Olek**[1], **André C. Müller**[1], **Chris Soon Heng Tan**[1,5,6], **Keiryn L. Bennett**[1], **Jacques Colinge**[1,7], and **Giulio Superti-Furga**[1,2]

[1]CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria

[2]Center for Physiology and Pharmacology, Medical University of Vienna, Vienna, Austria

Abstract

Thermal stabilization of proteins upon ligand binding provides an efficient means to assess binding of small molecules to proteins. We show here that in combination with quantitative mass spectrometry the approach allows for the systematic survey of protein engagement by cellular metabolites and drugs. The profiling of methotrexate, (*S*)-crizotinib and 2′3′-cGAMP in intact cells identified the respective cognate targets including the transmembrane receptor STING involved in innate immune signalling.

Keywords

target deconvolution; proteomics; drug discovery; metabolite; network

The systematic matching of small chemical compounds such as metabolites, xenobiotics, natural products or synthetic drugs to the repertoire of cellular proteins with which they entertain functional relationships has been a major technical hurdle and paramount objective of chemical biology. Many successful target deconvolution approaches use chemical modifications[1] to confer properties to the compound that allow for efficient affinity purification or target labelling[2, 3]. An ideal method to monitor the engagement of targets by

Correspondence should be addressed to G.S.-F. (gsuperti@cemm.oeaw.ac.at), K.V.M.H. (kilian.huber@sgc.ox.ac.uk), J.C. (jacques.colinge@inserm.fr).
[3]Present address: Structural Genomics Consortium, University of Oxford, Oxford, UK
[4]Present address: Target Discovery Institute, University of Oxford, Oxford, UK
[5]Present address: Institute of Molecular & Cell Biology, Agency for Science, Technology and Research (A*STAR), Singapore, Singapore
[6]Present address: Institute of Medical Biology, Agency for Science, Technology and Research (A*STAR), Singapore, Singapore
[7]Present address: Institut de Recherche en Cancérologie de Montpellier Inserm U1194, Université Montpellier, Montpellier, France

Author contributions. K.V.M.H. designed and jointly with K.M.O. performed experiments; A.C.M. and K.L.B. performed mass spectrometry; C.S.H.T. and J.C. performed bioinformatics analysis; K.V.M.H. and G.S.-F. conceived the study and wrote the manuscript. All authors contributed to the discussion of results and participated in manuscript preparation.

Accession codes. The raw mass spectrometric data used in this study are available via PeptideAtlas (www.peptideatlas.org) using the identifier PASS00693.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

Competing financial interests. The authors declare no competing financial interests.

Europe PMC Funders Author Manuscripts

small molecules would not require modification of the chemical agent, be unbiased and thus not rely on previous assumptions, and operate with intact cells and proteins in their native environment. Ligand binding can induce a conformational change in the target protein which in turn can alter for instance the accessibility of protease cleavage sites[4, 5]. Protein thermal stability *in vitro* has been used for many years in crystallography and discovery programmes as a robust method to identify ligands for various protein classes[6, 7]. The method is based on the frequent observation that binding of a ligand to a protein yields a complex with increased thermodynamic stability as compared to the free protein. This difference in stability is reflected in a change of the apparent melting temperature (T_m) of the protein (Fig. 1a). A recent breakthrough study has demonstrated that this concept can be expanded to monitor drug target engagement in live cells (cellular thermal shift assay (CETSA))[8]. However, so far the approach has only been used to investigate binding of compounds to a given cognate target. As the concept of the assay should be in general applicable to any protein target that is stabilized upon binding of a ligand, we hypothesized that the methodology could be further developed to identify any cellular targets, including unknown and unsuspected, that are stabilized upon engagement of a small molecule. Combined with modern mass spectrometry the approach should provide a comprehensive view of ligand-protein interactions through the highly multiplexed determination of individual T_m shifts (Fig. 1b,c).

To put our hypothesis to the test we chose to apply the procedure to the MTH1 inhibitor (*S*)-crizotinib. We and others have recently shown that disruption of nucleotide pool homeostasis by MTH1 inhibition may provide an effective new means to treat fast-replicating tumours and that (*S*)-crizotinib is a highly specific inhibitor for the Nudix hydrolase MTH1[3, 9]. In the aforementioned study, we profiled (*S*)-crizotinib both by classical chemoproteomics and extensive biochemical screening. These results should therefore provide an excellent dataset to evaluate the efficacy and validity of the approach. Thus, we heated aliquots of SW480 colon carcinoma cells treated either with (*S*)-crizotinib or DMSO to different temperatures and digested equal volumes of supernatants after cell lysis and centrifugation with trypsin (Fig. 1b). To enable reliable and comparative quantification of protein abundance in the samples with varying temperatures we employed a labelling strategy based on isobaric tandem mass tags (TMT)[10]. The pooled 10-plex samples were fractionated offline before analysis on a Q-Exactive mass spectrometer which provided data for more than 3,400 proteins after stringent filtering. Data were normalized (Supplementary Fig. 1) and DMSO- versus (*S*)-crizotinib-treated samples compared by fitting a sigmoid model to each and scoring the thermal shift (ΔT_m, Fig. 1c). We identified MTH1 (*NUDT1*) as the top cellular target of (*S*)-crizotinib as indicated by a pronounced stabilizing effect which notably was absent for the related family member *NUDT5* (Fig. 2a). We also observed shifts for other proteins linked to DNA repair such as POLA[11] (Supplementary Fig. 2, Supplementary Table 1). To validate the thermal stabilization of MTH1 protein by (*S*)-crizotinib as detected by mass spectrometry under these experimental conditions, we performed an analogous experiment using SW480 colon carcinoma cells and analyzed the supernatants by immunoblot. Indeed, (*S*)-crizotinib protected the MTH1 protein from unfolding whereas DMSO-treated control cells showed a sharp decline of MTH1 protein content (Fig. 2b, Supplementary Fig. 3).

To investigate potential differences between intact cells and lysates we also performed an analogous experiment using SW480 cell lysates (Supplementary Table 2). Similar to the Western blot method[8], we applied (*S*)-crizotinib at a higher final concentration (100 μM) to avoid potential dilution effects in the lysate. Interestingly, in contrast to the previous intact cell experiment the analysis revealed not only MTH1 but also several kinases as targets, many of which had been highlighted by the previous *in vitro* kinase screen (Supplementary Fig. 4,5, Supplementary Table 3). This supports the notion that indeed compartmentalization may affect a compound's target profile and that our method may be more likely to reflect the actual physiological context.

Next, we chose to evaluate the cellular target profile of the antimetabolite methotrexate (MTX) in K562 cells. Our analysis identified the cognate and well-studied target dihydrofolate reductase (DHFR) as the top candidate (Supplementary Table 4, Fig. 2c). Similar to (*S*)-crizotinib, we could confirm the stabilization by immunoblot (Fig. 2d, Supplementary Fig. 6). We also detected a highly reproducible shift for thymidylate synthase (*TYMS*), an important mediator of MTX action. MTX itself only exhibits inhibitory activity toward *TYMS* at concentrations above 10 μM whereas it has been shown that MTX polyglutamate metabolites are active in the nanomolar range[12]. As our treatment conditions are compatible with the formation of MTX polyglutamate species the data indicate that the method - in contrast to classical target identification approaches - could enable tracking of both drug and drug metabolite effects.

Our results prompted us to further investigate the possibility of using the methodology to elucidate the molecular interaction partners of cellular metabolites and thus for applications beyond drug target deconvolution. To challenge our hypothesis we decided to examine if the approach would be able to recover the cellular targets of the recently reported cGAMP synthase (cGAS) metabolite 2′3′-cGAMP[13]. Binding of 2′3′-cGAMP to its receptor STING (stimulator of interferon genes) elicits the transcription of pro-inflammatory cytokines. To ensure sufficient STING target engagement by its endogenous ligand we transfected RAW macrophages with 2′3′-cGAMP and analysed the samples by mass spectrometry. Bioinformatic filtering revealed a reproducible shift for STING (*Tmem173*) alongside well-known downstream mediators of innate immune signaling such as NFκB and STATs (Fig. 2e, Supplementary Table 5)[14]. We also confirmed the stabilizing effect on STING by immunoblot (Fig. 2f, Supplementary Fig. 7) not only demonstrating that the approach can recapitulate the activation of the inflammatory response by the cGAS metabolite but also that the procedure is compatible with transmembrane proteins.

During consideration of this manuscript a study reported thermal profiling coupled to mass spectrometry[15]. The method was efficiently used to identify the targets of kinase inhibitors highlighting the feasibility of the methodology for drug discovery. However, the approach of the study also revealed some important limitations as examination of the BCR-ABL inhibitor dasatinib paradoxically failed to alter the cognate target BCR-ABL in K562 cells. It is possible that the technology has problems when target proteins may not be stabilized sufficiently by a ligand, or when the targeted domain is only but one out of many such as in BCR-ABL.

Our study clearly extends the applicability of the method to enzymes other than kinases such as hydrolases and oxidoreductases, as well as transmembrane proteins. Moreover, as our procedure differs both in sample processing, mass spectrometric analysis as well as computational and statistical interpretation of the data (Supplementary Note), the robustness highlights the general validity of the approach and heralds several future flavours of the methodology, tailored to specific target classes, subcellular compartments or tissues. The choice of temperature steps may be critical in order to be able to capture differences in individual stability within particular groups of proteins. Therefore, several experiments with different temperature ranges may be required to reveal the unknown targets of less well-studied compounds. Future systematic work will undoubtedly reveal correlations between ligand pocket architecture and many other parameters. Moreover, thorough charting of the general thermal stability of the proteins in a cell should uncover many interesting features regarding thermal stability of cellular proteins and their subcellular localization, abundance, multi-protein complex formation, and posttranslational modifications. Finally, there is an increasing appreciation of the fundamental role of cellular metabolism in human disease and an exact mapping of the functional relationships between metabolites and their protein counterparts may not only increase our mechanistic understanding but also reveal new potential drug targets. Importantly, the approach should allow for target engagement in human biopsies. Combined with targeted quantitative proteomic approaches the method may focus on specific biomarkers and thus synergize with other efforts in this area[16].

In summary, here we present a methodology which takes advantage of alterations in protein thermal stability upon ligand binding providing a straightforward and low-effort approach to identify the cellular proteins that are engaged by a metabolite or a drug in intact, living cells.

Online methods

Cell culture

K562 and RAW 264.7 cells were obtained from DMSZ and ATCC, respectively. SW480 were kindly provided by W. Berger. All cells were cultured in the recommended media containing 10% fetal bovine serum and 10 U ml^{-1} penicillin/streptomycin (Gibco) and checked for mycoplasma by PCR or ELISA prior to experimental use.

Immunoblotting

The following antibodies were used according to manufacturer's instructions: mouse anti-DHFR (A-9, sc-377091, Santa Cruz Biotechnology)[8], rabbit anti-MTH1 (NB100-109, Novus Biologicals)[3], rabbit anti-STING (3337, Cell Signaling)[17].

Sample preparation

SW480 cells were cultured in 12-well cell culture plates to 80% confluency and treated with media containing either DMSO or 5 μM (*S*)-crizotinib for 3 h. For K562, cells were cultured in T75 cell culture flasks and upon reaching about 80% confluency cells were aliquoted into 12-well cell culture plates and treated with media containing either DMSO or 10 μM methotrexate for 3 h. After treatment, cells were detached with trypsin (only SW480), collected by centrifugation and subsequently resuspended in PBS. Cell suspensions were

transferred into 0.2 ml PCR-tubes and heated for 3 min. After a subsequent incubation of 3 min on room temperature, cells were lysed by adding 30 µl lysis buffer and three repeated cycles of freeze-thawing using liquid nitrogen. Precipitated proteins were separated from the soluble fraction by centrifugation at 17,000g for 20 min at 4 °C. For metabolite experiments RAW cells were cultured in 100 mm culture plates to 80% confluency and transfected with either water or 2′3′-cGAMP (10 µM, InvivoGen) using Lipofectamine 2000. After four hours cells were harvested and treated as described above. For cell lysate experiments, cells were lysed with lysis buffer (50 mM Tris-HCl, 100 mM NaCl, 0.2% NP-40, 5% glycerol, 1.5 mM $MgCl_2$, 25 mM NaF, 1 mM Na_3VO_4, 1 mM phenylmethylsulfonyl fluoride, 1 mM dithiothreitol (DTT), 10 µg/mL TLCK, 1 µg/mL leupeptin, 1 µ g/mL aprotinin, and 10 µg/mL soybean trypsin inhibitor (Sigma), pH 7.5) and after centrifugation incubated with DMSO or (*S*)-crizotinib (100 µM) for 30 min (6 mg total protein per condition). Six aliquots were prepared for each sample and heated for 3 min in a PCR machine. Precipitated proteins were separated from the soluble fraction by centrifugation as described above.

In solution tryptic digestion, TMT derivatisation and 2D-RPRP-liquid chromatography

Protein concentration of cell lysates was determined using the Bradford method (Biorad). Thirty (SW480), fifty (K562), and 80 (RAW) micrograms total protein were denatured with 8 M urea and reduced with dithiothreitol (DTT, 10 mM), cysteine residues alkylated with iodoacetamide (55 mM) and digested with modified porcine trypsin (1:100, Worthington Biochemical) overnight at 37 °C following dilution of urea below 2M. After quenching with TFA, peptides were purified by solid phase extraction using C18 SPE columns (The Nest Group). Eluted peptides were lyophilised in a vacuum centrifuge and re-suspended in TEAB buffer to a final concentration of 100 mM prior to labelling with the TMT reagents (ThermoFisher Scientific). Labelling was performed according to the instructions provided by the manufacturer. After quenching, all samples were pooled, purified by solid phase extraction and separated at pH 10 by reversed-phase liquid chromatography as described[18] using an Agilent 1200 series HPLC system. Separation was performed at a flow rate of 100 µl min^{-1} on a Phenomenex column (150 × 2.0 mm Gemini-NX 3 µm C18 110 Å, Phenomenex) with a 50 min linear gradient from 5–70% (vol/vol) acetonitrile containing 20 mM ammonium formate. Twenty time-based fractions were automatically collected, peptides lyophilised in a vacuum centrifuge and reconstituted in 5% (vol/vol) formic acid for on-line LCMS. Peptide abundance per fraction was estimated based on the UV trace and samples diluted accordingly to avoid overloading the nano-LCMS system. Fractions were injected onto a Dionex Ultimate 3000 system (ThermoFisher Scientific) coupled to a Q-Exactive mass spectrometer (ThermoFisher Scientific). Peptides were loaded onto a trap column (Zorbax 300SB-C18 5 µm, 5 × 0.3 mm, Agilent Biotechnologies) at 45 µl min^{-1} and separated on a customised 16 cm × 50 µm i.d. reversed-phase column packed with 3 µm C18 particles (ReproSil-Pur 120 C18-AQ, Dr. Maisch GmbH). Peptides were separated at a constant flow rate of 100 nl min^{-1} with a 85 min gradient from 3–35% (vol/vol) solvent B, increased to 70% (vol/vol) solvent B within 12 min and, finally, to 100% (vol/vol) solvent B for 10 min before re-equilibrating to 3% solvent B. Solvent A and solvent B consisted of 0.4% (vol/vol) formic acid and 70% (vol/vol) methanol, 20% (vol/vol) isopropanol and 0.4% (vol/vol) formic acid, respectively. Analysis on the Q-Exactive was performed in data dependent mode with the top ten most intense ions selected for fragmentation. MS and

higher-energy collision-induced dissociation (HCD) MS^2 spectra were acquired at 70,000 and 35,000 resolution (at *m*/z 200), respectively. Automatic gain control was used to prevent overfilling of the ion trap and was set to 3E6 and 2E5 ions for MS and HCD, respectively. Maximum ion accumulation times were set to 250 and 120 ms for MS and HCD, respectively. Dynamic exclusion for selected ions was 60 s. The singly-charged siloxane mass at *m/z* 445.120024 was used for lock mass correction[19]. The threshold for switching from MS to MS/MS was 1% of the target ion value. Normalized collision energy was 29%, with fixed first mass of 100 *m*/z. Software versions for the operation of the Q-Exactive were Tune 2.3 SP1 and Xcalibur 3.0.63.

Bioinformatic analysis

All MS/MS spectra were searched against UniprotKB/Swissprot human with Mascot[20] and Phenyx[21]. Results were merged and false discovery rate imposed a maximum of 1% according to a procedure already described[22]. Spectra providing a signal above estimated background noise[23] in all the 10 TMT channels were considered only and we further limited the analysis to proteins identified by two such spectra at least.

Data normalization

Classical normalization approaches could not be applied since the total amount of proteins in samples at increasing temperatures was reduced on purpose. Nonetheless, it can be appreciated that unnormalized data, especially upon DMSO treatment, suffer from a departure from the expected sigmoid curve at the first three temperatures following the lowest temperature (Supplementary Fig. 1). This was already observed in the original Western blot method[8] and we hypothesise that this phenomenon depicts an artefact due to protein crowding effects at the lower temperatures. We thus reasoned that data could be normalized through their shape instead of their values. To this aim we summed 50,000 randomly selected spectra in each sample separately, modelling the average temperature-dependent abundance decrease, which we fitted by a sigmoid of the form

$$1/(1+e\hat{}(b(a-t)))$$

with *t* the temperature and *a, b* the curve parameters (determined by nonlinear least square optimization). At each temperature, the sum spectrum was compared with the sigmoid and a relative shift determined that was subsequently applied to all the proteins of the corresponding sample as a mean of normalization (Supplementary Fig. 1).

Target scoring

Two scores were computed to capture different features of the thermal shift that should be observed in targets: a temperature shift at 50% reduction of protein abundance in the sample; a signed area between the DMSO- or vehicle-treated sample curve and the compound-treated curve (Fig. 1c and Supplementary Fig. 8). The sigmoid curves required to estimate the thermal shift at 50% concentration reduction were of the form above. The two scores were necessary since more than 3,000 proteins had to be scored in each sample comparison and odd curves were able to yield good values according to one of the two scores by chance.

In addition, it was required that a potential drug target featured at least 25% protein abundance reduction at the highest temperature under compound treatment and 75% under vehicle for MTX; 25% and 50% for (*S*)-crizotinib; 10% and 10% for 2′3′-cGAMP. These non-stringent conditions reduced the analysis to proteins where the assay generated meaningful, interpretable data. This procedure also eliminated some oddly-shaped melting profiles that generated tentatively decent scores by coincidence and would be discarded by expert inspection immediately.

Target selection

Exploiting the two replicates available in each cell line for each compound and vehicle control, we created two pairs of vehicle versus compound-treated samples. Proteins identified in a pair were scored as described above. In each pair we retained proteins in the top p% of both scores and intersected such selections across the two biological replicates. For MTX we used p=10, for (*S*)-crizotinib p=15, and for 2′3′-cGAMP p=20. The CV values reported in the Supplementary Tables were calculated by applying the classical definition (standard deviation divided by the mean) to the two replicates of a given score (temperature shift or signed area). Namely, if s1 and s2 are the replicates of one of the two scores, CV=sqrt(2)|s1-s2|/(s1+s2).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We are grateful to W. Berger, Institute of Cancer Research, for providing SW480 cells and P. Majek, CeMM, for assistance with data processing. This work was supported by the Austrian Academy of Sciences, the European Union (FP7 259348, ASSET), and the Austrian Science Fund (FWF F4711, MPN).

References

1. Lee J, Bogyo M. Curr. Opin. Chem. Biol. 2013; 17:118–126. [PubMed: 23337810]
2. Kambe T, Correia BE, Niphakis MJ, Cravatt BF. J. Am. Chem. Soc. 2014; 136:10777–10782. [PubMed: 25045785]
3. Huber KVM, et al. Nature. 2014; 508:222–227. [PubMed: 24695225]
4. Feng Y, et al. Nat. Biotechnol. 2014; 32:1036–1044. [PubMed: 25218519]
5. Lomenick B, et al. Proc. Natl. Acad. Sci. USA. 2009; 106:21984–21989. [PubMed: 19995983]
6. Niesen FH, Berglund H, Vedadi M. Nat. Protocols. 2007; 2:2212–2221. [PubMed: 17853878]
7. Fedorov O, et al. Proc. Natl. Acad. Sci. USA. 2007; 104:20523–20528. [PubMed: 18077363]
8. Molina DM, et al. Science. 2013; 341:84–87. [PubMed: 23828940]
9. Gad H, et al. Nature. 2014; 508:215–221. [PubMed: 24695224]
10. Dayon L, et al. Anal. Chem. 2008; 80:2921–2931. [PubMed: 18312001]
11. Parlanti E, Locatelli G, Maga G, Dogliotti E. Nucleic Acids Res. 2007; 35:1569–1577. [PubMed: 17289756]
12. Allegra CJ, et al. J. Biol. Chem. 1985; 260:9720–9726. [PubMed: 2410416]
13. Ablasser A, et al. Nature. 2013; 498:380–384. [PubMed: 23722158]
14. Cai X, Chiu Y-H, Chen Zhijian J. Mol. Cell. 2014; 54:289–296. [PubMed: 24766893]
15. Savitski MM, et al. Science. 2014; 346
16. Lambert J-P, et al. Nat. Methods. 2013; 10:1239–1245. [PubMed: 24162924]

Methods-only references

17. Yoh, Sunnie M, et al. Cell. 2015; 161:1293–1305. [PubMed: 26046437]
18. Gilar M, Olivova P, Daly AE, Gebler JC. J. Sep. Sci. 2005; 28:1694–1703. [PubMed: 16224963]
19. Olsen JV, et al. Mol. Cell. Proteomics. 2005; 4:2010–2021. [PubMed: 16249172]
20. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. Electrophoresis. 1999; 20:3551–3567. [PubMed: 10612281]
21. Colinge J, Masselot A, Giron M, Dessingy T, Magnin J. Proteomics. 2003; 3:1454–1463. [PubMed: 12923771]
22. Burkard TR, et al. BMC Syst. Biol. 2011; 5:17. [PubMed: 21269460]
23. Breitwieser FP, et al. J. Prot. Res. 2011; 10:2758–2766.

Europe PMC Funders Author Manuscripts

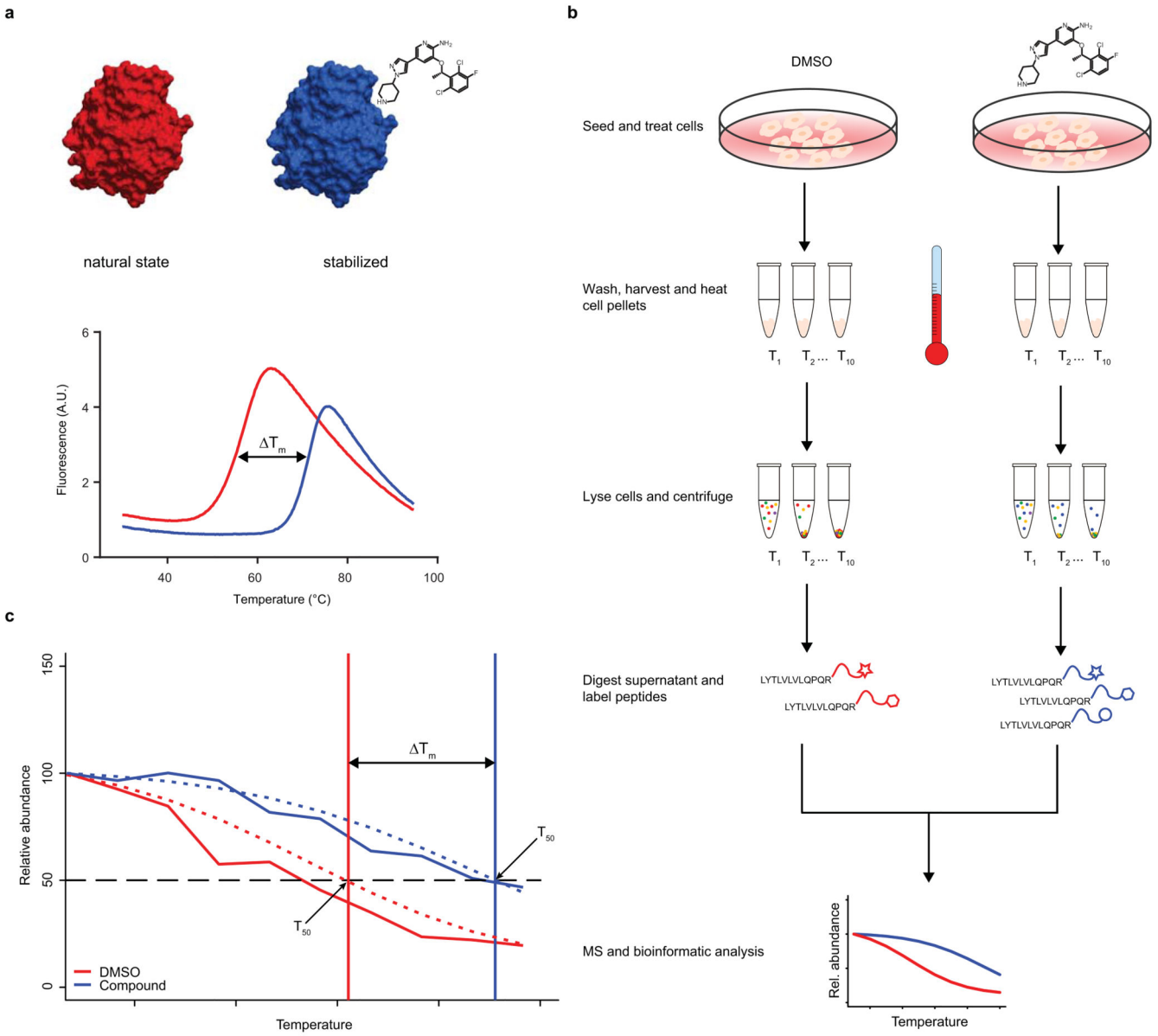

Figure 1.

Schematic representation of the methodology. (**a**) Binding of a ligand to its target protein (*top*) increases the enthalpy required for unfolding. As a result the melting temperature (T_m) is shifted which can be exploited by e.g. differential scanning fluorimetry (DSF, "Tm shift assay") (*bottom*). (**b**) Representative workflow. (**c**) Schematic illustrating the bioinformatic scoring parameters used for data analysis. Continuous lines depict raw data, dashed lines are curve fitting results.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

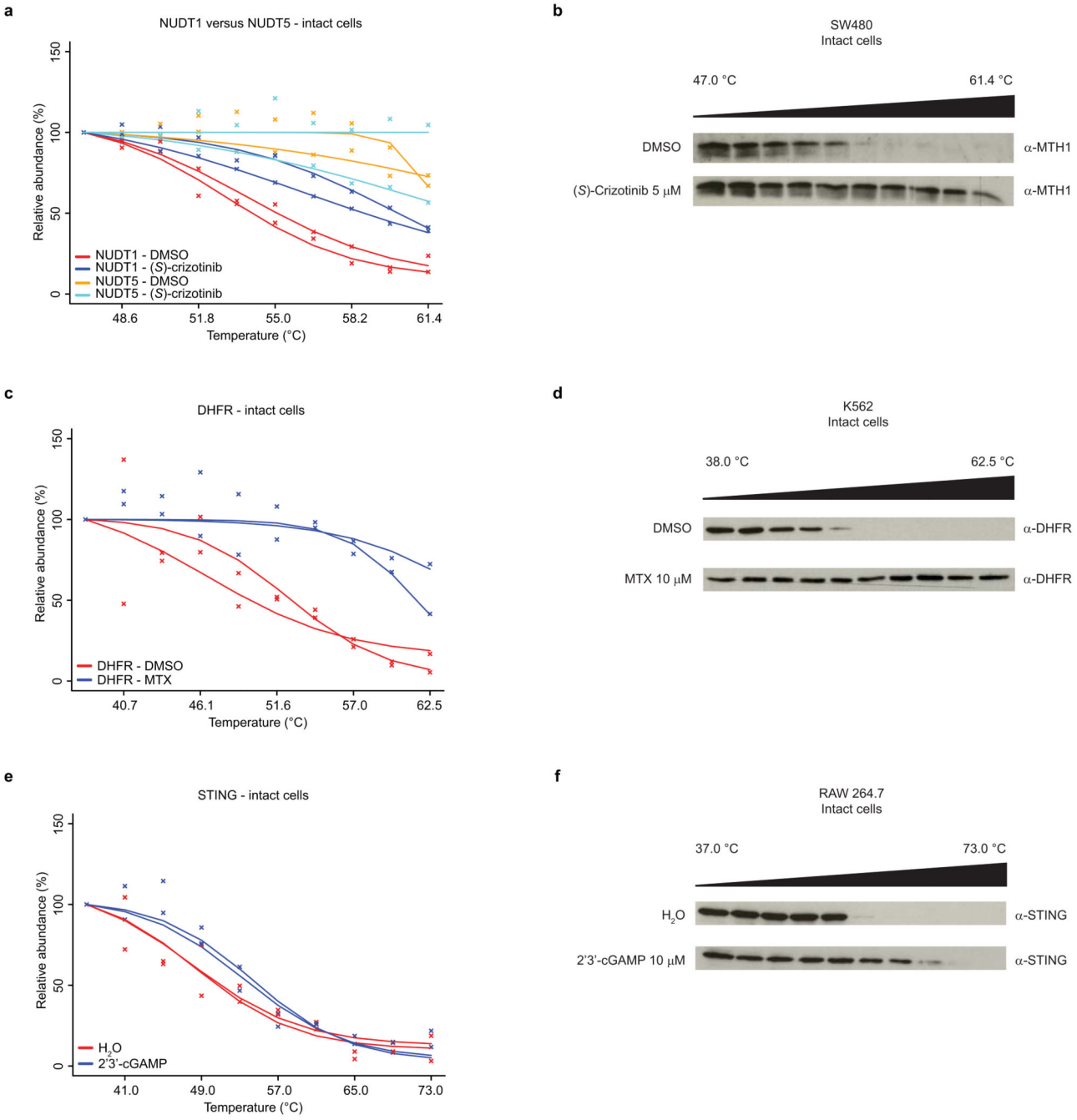

Figure 2.

Thermal profiling results for the MTH1 inhibitor (*S*)-crizotinib, methotrexate (MTX), and 2′3′-cGAMP. (**a**) Protein abundance graphs obtained from MS-TMT-10-plex experiments. In contrast to the cognate target MTH1 (*NUDT1*), the close family member *NUDT5* is not stabilized upon treatment of intact cells with the specific inhibitor (*S*)-crizotinib. Data indicate two independent experiments (n = 2). (**b**) Western blot confirming the stabilization of MTH1 protein by (*S*)-crizotinib as compared to DMSO using intact SW480 cells. Images are cropped for clarity; full-length blots are presented in Supplementary Figure 3. (**c**) Protein

abundance graphs obtained from MS-TMT-10-plex experiments demonstrating a strong shift in thermal stability for the well-established MTX target DHFR. Data indicate two independent experiments (n = 2). (**d**) Immunoblot confirming the stabilization of DHFR by methotrexate (MTX) compared to DMSO using intact K562 cells. Images are cropped for clarity; full-length blots are presented in Supplementary Figure 6. (**e**) MS analysis results of RAW macrophages treated with 2′3′-cGAMP reveal stabilization of the transmembrane receptor STING (*Tmem173*). Data indicate two independent experiments (n = 2). (**f**) Western blot validation of the increased thermal stability observed for STING after treatment with 2′3′-cGAMP. Images are cropped for clarity; full-length blots are presented in Supplementary Figure 7.

Europe PMC Funders Author Manuscripts