Single Cell Chemical Proteomics (SCCP) Interrogates the Timing and Heterogeneity of Cancer Cell Commitment to Death

Ákos Végvári*, Jimmy E Rodriguez, Roman A Zubarev*

Division of Physiological Chemistry I, Department of Medical Biochemistry & Biophysics, Karolinska Institutet, Biomedicum A9, Solnavägen 9, SE-171 77 Stockholm, Sweden

Contact to corresponding authors: akos.vegvari@ki.se (phone: +46-8-524 877 07) and roman.zubarev@ki.se (phone: +46-8-524 875 94)

Abstract

Chemical proteomics studies the effects of drugs upon cellular proteome. Due to the complexity and diversity of tumors, the response of cancer cells to drugs is also heterogeneous, and thus proteome analysis at single cell level is needed. So far, single cell proteomics techniques have not been quantitative enough to tackle the drug effects on the target proteins. Here, we developed the first single cell chemical proteomics (SCCP) technique and studied with it the time-resolved response to anticancer drugs methotrexate, camptothecin and tomudex of individual adenocarcinoma A549 cells. For each drug SCCP identified and quantified 1,000-2,000 proteins across >150 single cells in a time course of treatment up to 48 h, revealing the early emergence of cellular subpopulations committed and uncommitted to death. SCCP represents a novel approach to exploring the heterogeneous response to drugs of cancer cells.

Keywords: TMT labeling; Drug sensitivity; Resistance to drugs; Cell death; Methotrexate; Camptothecin; Tomudex

31 **1. Introduction**

32 Chemical proteomics studies the effects of drugs on cellular proteomes with the purpose of
33 deciphering the targets and mechanisms of action (MOAs) of these molecules. [1–3] When
34 sensitive cells are treated with toxic compounds for extended period of time, mechanistic target
35 proteins become significantly regulated, and their profiling provides the first hint on the
36 compound's targets and MOA.[4,5] This approach has been employed in functional identification
37 of target by expression proteomics (FITExP)[6], which laid ground for the online chemical
38 proteomics ProTargetMiner tool.[1] In FITExP, cells are treated at a LC_{50} concentration for 48 h,
39 by which time half of the cells die. The dying cells detach from the substrate (for adherent cell
40 types) and are found floating on the flask surface. In the remaining (surviving) cells, the drug
41 target's expression level is significantly and specifically regulated up or down, which serves a
42 basis for drug target identification in FITExP. As an example, when cancer cells undergo
43 treatment with methotrexate (MTX), the target protein dihydrofolate reductase (DHFR)
44 becomes highly upregulated before the cells undergo programmed cell death.[7–9] Interestingly,
45 while the proteomes of the dying and surviving cells are very different (lending support to the
46 notion that cell death is the ultimate case of cell differentiation), the drug target behaves in a
47 similar manner in both types of cells. [2]

48 The adherent cells usually start losing their attachment to the surface after 24 h of
49 treatment at LC_{50} concentration, but the decision to survive or become dying must be made by
50 the cell well before that.[1,10,11] The intricate details of this decision making process are of great
51 scientific interest, as they possibly hold keys to drug resistance mechanisms.[1,6] These decision
52 making processes can only be studied at the single cell level, while all so far reported chemical
53 proteomics studies relied on bulk cell analysis.[12] Cellular heterogeneity is currently analyzed
54 routinely by single cell transcriptomics (SCT),[13] with mRNA levels assumed to be proportional
55 to the protein expression levels. However, at any given moment the concentration of both

56 mRNA and proteins reflects the balance between their corresponding expression and
57 degradation, and while mRNA transcription and protein expression are linked together rather
58 well, the degradation processes for mRNA and proteins are completely decoupled. As a result,
59 in the biological processes driven mostly by protein expression, mRNA levels provide excellent
60 proxy for protein concentrations, but this correlation seems to break down already at steady
61 states of the cell.[14] In cell death processes mediated by protein degradation (*e.g.,* via caspase
62 proteases), a correlation between mRNA and protein levels cannot be presumed. Therefore,
63 cell heterogeneity in death-related processes can best be studied with single cell proteomics
64 (SCP).

65 Compared to the rather well developed SCT approaches, SCP methods are still
66 emerging. While some targeted antibody-based immunoassays have been applied to
67 characterize proteins in single cells,[15,16] these approaches are limited to a few dozen proteins
68 per experiment and exhibit strong bias in quantification. Mass spectrometry (MS)-based
69 proteomics can in principle overcome these limitations, but lacking the benefit of PCR, MS
70 proteomic analysis at a single cell level is very challenging due to the extremely low amounts
71 of proteins (ca. 0.2 ng in a mammalian cell), the high dynamic range of protein expression (7
72 orders of magnitude versus 3-4 orders for mRNAs),[17,18] and the inevitable sample loss during
73 protein extraction, digestion and chromatographic separation of the peptide digest.[19]
74 Consequently, despite the introduction of such ground-breaking SCP methods as SCoPE-MS[19],
75 SCoPE2[20,21] and nanoPOTS[22] that have been able to analyze between 500 and 2000 proteins in
76 diverse cell lines, SCP has not been sufficiently quantitative to apply the techniques of chemical
77 proteomics, such as FITExP. Thus, we have embarked on a painstaking enterprise of solving
78 this issue to develop Single Cell Chemical Proteomics (SCCP).

79 Most SCP studies so far have considered two different types of cells (*e.g.,* monocytes
80 *vs* macrophage cells or Jurkat *vs* U-937 cells)[19,21] with vastly different proteomes. Separation

81 of these cells by SCP was relatively straightforward as it could be done using a few most
82 abundant proteins. In contrast, in cells influenced by a drug the most significantly regulated
83 proteins (drug targets) are seldom highly abundant, being frequently found in the abundance-
84 sorted list below the 1000th position. Therefore, our SCCP-development efforts were directed
85 to achieving the following two intermediate objectives. First, average protein abundances in a
86 homogeneous cell population measured by SCP must correlate with the abundances in bulk
87 proteome analysis. This goal was achieved by starting from analyzing as bulk a relatively high
88 number of cells and gradually reducing this number down to single cells, monitoring the
89 correlation with the bulk analysis and systematically troubleshooting when this correlation
90 broke down. A number of issues have been found and resolved related to protein extraction,
91 digestion, labeling with isobaric regents, LC separation, MS acquisition and statistical analysis.
92 At the end, satisfactory correlations between SCP and bulk proteomics results were consistently
93 obtained. The second intermediate goal objective was to detect with SCP the known strong
94 regulation of the drug targets, as in FITExP, with high statistical significance. This again
95 required systematic studies and optimizations.

96 Here we present the developed SCCP workflow applied to studying in a time-course
97 manner the proteome effects of anticancer drugs MTX, camptothecin (CPT) and tomudex
98 (TDX), also known as raltitrexed. These drugs were applied at LC_{50} concentration to A549
99 human lung adenocarcinoma cells, leading half of the cells to death in 48 h. Our workflow
100 comprises the isolation of cells using fluorescence-activated cell sorting (FACS), minimal
101 sample preparation including tryptic digestion, tandem mass tag (TMT) isobaric labeling for
102 protein quantification, incorporation of a carrier proteome (CP) to boost the MS signal,
103 chromatographic separation at a low flow rate, MS/MS data acquisitions and SCCP-optimized
104 data processing (see **Figure 1**). The goal of the study was to identify the time scale of the
105 decision-making dying/surviving process, *i.e.*, to reveal at what time the homogeneous cell

106 population started to differentiate under the influence of a drug into cells committed to
107 surviving or dying.
108
109 **2. Results**

110 **2.1. Improving critical steps of the SCCP workflow**

111 To establish a robust analytical method, a human lung cancer cell line A549 was selected as a
112 cellular model because of its easy cultivation and handling. The cell size (ca. 10 µm in diameter)
113 permitted isolation of single cells by FACS without using surface markers for gating. The cells
114 treated with MTX (LC_{50} = 1.15 µM), CPT (LC_{50} = 3 µM), TDX (LC_{50} = 50 µM) and vehicle
115 (DMSO) exhibited in FACS homogeneous populations that only changed shape when drug was
116 applied (**Figure S1**). We consider FACS homogeneity (single population) an important element
117 of any SCP study that justifies single cell analysis, as otherwise cell subpopulations can be
118 counted, isolated by FACS and conventionally analyzed as bulk.

119 All drugs were known to strongly regulate their targets after 48 h incubation in A549
120 cells.[1] Since cell lines are known to "drift" with time, we corroborated the regulation of DHFR,
121 topoisomerase 1 (TOP1) and thymidylate synthase (TYMS) as the targets of MTX, CPT and
122 TDX, respectively, in our cellular model using bulk proteome analysis.

123 The sample preparation optimization started from the SCoPE-MS strategy with
124 TMT10plex labeling and 200 cells in the CP channel,[19] and was adapted to maximize the
125 number of detected proteins and the precision of their quantification. Among the optimized
126 parameters was the volume of buffer to capture cells on plate. Smaller volumes give higher
127 protein concentration but are more difficult to handle, which potentially leads to larger losses.
128 The results were comparable for volumes in the 1-5 µL range (16,000-20,000 peptides and
129 2200-2400 proteins quantified), and for practical reasons 5 µL was chosen for all subsequent
130 experiments.

131 We found that the property of the ABIRD device to significantly (2-3 times) reduce the
132 chemical noise was especially beneficial in SCP, and utilized this device in all experiments.
133 The flow rate across the EASY-Spray™ C18 analytical columns was reduced from the nominal
134 ~300 nL/min to 100 nL/min to increase the signal abundance. Further reduction in the flow rate
135 reflected negatively on spray stability. We also increased the accumulation time in the Orbitrap
136 for MS/MS above the recommended maximum 86 ms to 150 ms. These measures allowed us
137 to identify and quantify over 2000 proteins in a stable way across 64 cells.

138 Significant improvements were required in data analysis. PCA of raw SCP data gave no
139 separation between the treated and untreated cells; the data clustered instead according to the
140 TMT sets. To address this inherent for SCP batch effect, we employed the batch-compensation
141 function in the *SVA* package using empirical Bayesian approach. We also found that it was
142 important to normalize the peptide and protein abundances by the median abundance in a given
143 cell. Using these approaches, we started to obtain treated-untreated group separation in PCA.

144

145 **2.2. Introducing split semi-bulk CP**

146 In order to verify that the separation between the cell groups in PCA was related to biology
147 and was not an uncompensated artefact, we introduced a CP in two separate channels. In
148 general, CP channel is commonly used in TMT-SCP to increase the identification rate and
149 reduce the risk of a catastrophic sample loss during sample preparation.[19,23,24] Most often, 200
150 cells or their equivalent are used for CP, and since its effect on ion signal is additive[25], we split
151 the CP between two channels, with one channel carrying 100 treated cells and another one –
152 100 untreated cells. The expectation was that the average (or median) protein abundances in
153 single treated cells will correlate with the protein abundances in treated CP, and the untreated
154 cells will correlate with untreated CP.

155 Failure to observe such correlation between single cell and bulk proteome data in our first
156 such experiments made us reconsider the way how bulk proteome is obtained. In previous
157 experiments, the bulk proteomic analysis was performed in a conventional proteomics way with
158 the use of detergents, S-S bond reducing and alkylating agents, *etc*., which significantly differed
159 from SCP sample preparation. Then we prepared CPs the same way as single cell proteomes,
160 calling such CPs semi-bulk. We used semi-bulk CPs for analysis of 84 MTX-treated and 84
161 vehicle-treated single cells in 24 TMT11plex sets. Each of these sets these sets carried 7 single
162 cells (3 or 4 for each condition), an untreated semi-bulk CP in the channel 131C and MTX-
163 treated semi-bulk CP in the channel 131N, while two channels (130N and 130C) remained
164 empty. In total, 13,142 peptides were quantified belonging to 1825 proteins. A satisfactory
165 correlation (R=0.64) was obtained between the log2 values of the protein treated-control
166 regulations in the semi-bulk CPs and the median SC data (**Figure 2A**), which was similar to
167 the correlation between the semi-bulk CPs and the bulk-analyzed CPs prepared in the SCP
168 manner from large number of cells (R=0.81, data not shown). In an OPLS-DA model where
169 treated CP, control CP, treated single cell and untreated single cell data were pooled in four
170 respective groups, single cell data clustered together with their corresponding CPs (data not
171 shown). When only the top 25, 200, 900 most abundant proteins as well as all 1825 proteins
172 were used for OPLS-DA, the Q^2 quality coefficient increased with the number of proteins,
173 proving that RIAs from most proteins in single cells were responsible for the observed
174 separation and not just a few abundant molecules. This was an encouraging message for SCCP,
175 as in Chemical proteomics most drug targets have medium abundance.
176 Importantly, the SCCP dataset contained the MTX target DHFR that was found on the 6th
177 position from the top among the most significantly upregulated proteins (**Figure 2B**). These
178 results gave us confidence that the SCP data can be used in Chemical proteomics for drug target
179 determination and monitoring. In subsequent experiments we continued to use for carrier and

180 bulk proteomes the SCP protocol – that is, all subsequent CPs and bulk proteomes are "semi-
181 bulk". Also, in subsequent experiments we obtained CPs by sorting 100 or 200 cells into one
182 well, thus ensuring that CP cells undergo the same FACS treatment as single cells.
183
184 **2.3. TMTpro™ in SCCP workflow**

185 The recently introduced 16-plex TMTpro™ offers the obvious advantage of multiplexing more
186 single cells into one TMT set (12 vs. 7 in TMT11plex™). However, this advantage might come
187 with a price paid in form of a lower single cell reporter ion signal in MS/MS, as it is distributed
188 over larger number of channels. To test whether the trade-off is acceptable for SCCP, we
189 analyzed 96 single A549 cells treated with MTX and 96 vehicle-treated cells with split CPs (in
190 channels 126 and 127N) and obtained somewhat lower number of protein IDs but still
191 comparable with the TMT11plex™ analysis. In total, 1515 proteins were TMTpro-quantified
192 with 8870 peptides in single cells with 133 significantly regulated proteins, and 1874 proteins
193 with 12,317 peptides in the CPs with 326 regulated proteins. The correlation between SC data
194 and CPs was satisfactory (R=0.71), and significant up-regulation of DHFR was also detected
195 for both SCs and CPs. This result confirmed that SCP can be performed using the TMTpro
196 labelling, but as 11-plex TMT provided somewhat deeper proteome coverage, we continued to
197 use it for SCCP purposes. To accommodate more single cells in one TMT11plex set, for single
198 cell analysis subsequently we used one CP (unless specified otherwise), composed of a 1:1
199 mixture of treated and untreated cells, while separately analyzing these proteomes for
200 calculating the correlation with single cell data.
201
202 **2.4. Time-course SCCP analysis**

203 The goal of the experiment was to determine the time point at which the attached cells make
204 the decision to die, so that their proteome becomes altered to resemble that of the end-point
205 detached (dying) cells rather than the end-point attached (surviving) cells. For that purpose, the

206 cells were treated with MTX for 3, 6, 12, 24 and 48 h at LC_{50} concentration. The attached cells
207 at each time point and the detached cells at 48 h were collected, and FACS-isolated 96 cells of
208 each type were analyzed with SCCP using a single CP representing a mixture of the 48 h
209 attached and 48 h detached cells. The bulk proteomes of 48 h detached and attached treated
210 cells were analyzed separately. On average, over 1500 proteins and 10,000 peptides were
211 identified and quantified in single cells at each incubation time. **Figure 3** shows how the
212 attached treated and untreated cell populations, being almost indistinguishable on a PCA plot
213 at 3 h treatment, become gradually separated with time, achieving nearly full separation at 12
214 h.
215 In order to identify at each timepoint the cell subgroup that was committed to death and the one
216 committed to survival in the attached treated population, a hierarchical cluster analysis of
217 protein abundances was performed (**Figure S2**). It was assumed that the two most abundant
218 cell clusters represent the subgroups of the future surviving and dying cell. The hypothesis was
219 that, being put on a PCA plot together with the 48 h attached and detached cells representing
220 the two ultimate cell destinies, the two subgroups will reveal their identities by being closer to
221 the respective destiny type. For time points earlier than the commitment event, cell clustering
222 into the two subgroups will be random, and thus both subgroups would end up in the middle of
223 the OPLS-DA plot close to each other.
224 Both these predictions were confirmed when the median abundances of all 1170 quantified
225 proteins and 100 most abundant proteins in group 1 (G1) and group 2 (G2) separated by
226 clustering analysis of attached cells were used for building an OPLS-DA model. The model
227 also included the data on 48 h attached cells and 48 h detached cells, which represented the
228 final destinations of the survival and dying subpopulations (**Figure 4A**). For 3 h and 6 h
229 treatments, there was an overlap of the dots representing the G1 and G2 clusters, with a
230 separation between them in the direction of the destiny points at 12 h and longer treatment

231 times. G1 was thus acquiring a proteome profile corresponding to the dying fate, while G2
232 represented the surviving subpopulation. As expected, the OPLS-DA separation between these
233 two subpopulations grew with time. Similarly, the number of proteins with significantly
234 changed abundances between the vehicle- and MTX-treated populations increased with time
235 from 32 and 15 proteins at 3 and 6 h to 38, 121 and 134 proteins at 12, 24 and 48 h, respectively.
236 Therefore, the A549 cell commitment to death occurs between 6 and 12 h past MTX treatment.
237 This time scale is consistent with the earlier reports on dynamic proteomics measurements in
238 cells treated with a drug at LC_{50} – in the first hours past treatment, the cells try to overcome the
239 encountered difficulty, activating survival pathways, and only commit to death after such an
240 attempt fails.[1]
241 Interestingly, at 12 h more separation was seen for the whole proteome, while at 24 h the 100
242 most abundant proteins showed bigger separation. This observation agreed well with the notion
243 that the cell path to death starts with the inner mechanism altering lower-abundant mechanistic
244 proteins first, followed by the altering household proteins that change cell morphology.
245 Consistent with this scenario, when the main OPLS-DA coordinates of the individual cells were
246 plotted on a scale normalized such that the attached cells treated for 48 h had x = 1 and the
247 corresponding detached cells had x = -1, the obtained distributions of G1 and G2 cells were
248 separated in 12 h for the full proteome, but less so for 400 most abundant proteins and not at
249 all for top 100 proteins (**Figure 4B**). At the same time, for 24 h treatment the G1 and G2
250 proteomes gave broad distributions separated more for highly abundant proteins, suggesting
251 that cell morphology alteration is well underway.
252 Pathway analysis of 179 proteins with significantly different abundances in G1 versus G2 at 12
253 h past MTX treatment revealed that they preferentially belong to metabolic, carbon metabolism,
254 ribosome- and proteasome-related pathways (**Figure S3)**.
255

256 **2.5. SCCP with camptothecin (CPT) and tomudex/raltitrexed (TDX)**

257 Similar results as with MTX were obtained with CPT and TDX, with the targets TOP1
258 (downregulated) and TYMS (upregulated) emerging among the top proteins in the respective
259 areas of the volcano plot (**Figure 5**). While these drugs have different MOAs and targets, the
260 A549 cells have clearly formed two well separated clusters in PCA.

261

262 **2.6. Target percolation by OPLS-DA of MTX, CPT and TDX data**

263 The ultimate goal of a Chemical proteomics experiment is drug target identification, which can
264 be obtained by contrasting a specific treatment against all other treatments and controls. While
265 designing ProTargetMiner[2], we found that on average it takes 30-50 contrasting treatments to
266 identify ("percolate") the target uniquely among thousands of proteins in the proteome as the
267 most specifically up- or down-regulated protein. Here, we merged the MTX, CPT and TDX
268 SCCP data (treatment *vs.* untreated control) at 48 h of treatment and contrasted one drug against
269 the other two (**Figure S4**). For MTX, the target DHFR was 4th most specifically upregulated
270 protein; for CPT, TOP1 was 15th most specifically downregulated protein; and for TDX, TYMS
271 was 10th most specifically upregulated protein. These results demonstrate that SCCP has the
272 potential for unique drug target identification, provided enough contrasting treatments are
273 obtained.

274

275 **3. Discussion**

276

277 Considering that cell to cell heterogeneity is a fundamental property of highly complex cellular
278 systems[16], the analysis of proteomes at a single cell level is essential for understanding the
279 complex diseases, such as cancer, where diverse phenotypes contribute to the survival and
280 progression,[26] as well as for studying the mechanisms of cell resistance to anticancer treatment.

281 Here we demonstrated that SCP can be sufficiently quantitative for enabling Chemical
282 proteomics approaches for drug target identification and monitoring. One important milestone
283 was to achieve significant correlation between the average protein regulation factors in SCCP
284 between the treated and untreated cells with those obtained with conventional bulk proteomics
285 of these cells. Another important milestone was to demonstrate similar behavior of the drug
286 target proteins in SCCP as in bulk Chemical proteomics. Finally, the possibility to "percolate"
287 by a contrasting OPLS-DA analysis the drug targets has a paramount importance for the use of
288 such a powerful drug target deconvolution method as ProTargetMiner.
289 While these achievements were the SCCP proofs of principle, time course analysis provided
290 new biologically relevant information, confirming that cell commitment to death can now be
291 studied at a proteome level for individual cells. Between 6 h and 12 h past treatment, a large
292 group of attached drug-treated cells already committed to detach and form a floating dying
293 population. Importantly, these changes were detected among the lower-abundant proteins,
294 while highly abundant proteins remained at that point unaffected. It was even possible to
295 determine the pathways and parts of cell machinery participating in the decision-making
296 process.
297

298 **4. Conclusions**

299 After the quantitative aspect of single cell proteomics has been improved, chemical proteomics
300 at the level of single cells became reality. The detailed profiling with SSCP of the heterogeneity
301 of cancer cell response to drugs or treatments and the mechanistic analysis with cellular
302 resolution of resistance to therapy is now possible. Moreover, the FITExP method of chemical
303 proteomics is now applicable to single cells. The remaining challenges are however vast. For
304 example, SCCP needs to provide deeper proteome analysis, targeting the benchmark of 5,000
305 proteins quantified with ≥2 peptides. A great achievement would be if complementary tools of

306 chemical proteomics, such as proteome-wide integral solubility alteration (PISA) assay[27] could
307 be implemented for single cells. With PISA, one could monitor the protein target engagement
308 of the drug molecule. This, however, requires significant efforts in improving the methods of
309 handling and analyzing ultra-small protein amounts.
310
311
312 **5. Methods**
313 **5.1. Cell culturing treatment**
314 Human A549 lung adenocarcinoma cells obtained from ATCC (Manassas VA) were grown in
315 Dulbecco's Modified Eagle's Medium (DMEM – Lonza, Wakersville MD) supplemented with
316 10% FBS superior (Biochrom, Berlin, Germany), 2 mM L-glutamine (Lonza) and 100 U/mL
317 penicillin/streptomycin (Thermo, Waltham MA) at 37°C in 5% CO_2.
318 The LC_{50} values for the drugs (MTX, CPT and TDX) were determined by CellTiter-Blue® cell
319 viability assay (Promega). Cells were seeded into 96 plates at a density of 3000 cells per well,
320 and after 24 h of culture they were treated with serial concentrations of the respective drug:
321 MTX (0 – 100 μM), CPT (0 – 100 μM) and TDX (0 – 100 μM). After 48 h, the media were
322 discarded and replaced with 100 μL of fresh culture media. In each well, 20 μL of resazurin
323 (CellTiter-Blue® Cell Viability Assay kit – Promega) were added to perform the viability assay.
324 After 4 h of incubation at 37°C, fluorescence of wells was measured in Infinite F200 Pro
325 fluorometer (Tecan) by detecting the ratio between the excitation at 560 nm and emission at
326 590 nm. The LC_{50} values were determined from the dose-response curves by calculating the
327 concentration causing the 50% fluorescence reduction compared with the untreated control.
328 Cells were then cultured and treated with MTX, CPT and TDX at LC_{50} concentrations in 75
329 cm^2 flasks for 3, 6, 12, 24 and 48 h (for CPT and TDX only 12, 24 and 48 h treatment were
330 performed). Control cells were treated with the vehicle (10 mM dimethyl sulfoxide - DMSO).

331 After each incubation time point, the supernatant was collected and the attached cells were
332 disconnected from the surface with TrypLE (Gibco) for 5 min, after which they were harvested
333 by centrifugation at 1000 rpm for 3 min. Both types of cells (detached and adhered) were
334 washed twice with cold 1X phosphate buffered saline (PBS).
335

336 **5.2. Isolation of single cells by FACS**

337 Collected attached cells were subjected to FACS analysis in FACSAria™ Fusion (BD
338 Biosciences), in which cells were sorted based on the forward and side scatter
339 (FSC/SSC) parameters only. Individual singlet cells were collected in a 96-well Lo-Bind plate
340 (Eppendorf, Hamburg, Germany) containing 5 µL of 100 mM triethylammonium bicarbonate
341 (TEAB) per well. A total of 96 single cells were sorted for each condition/drug (untreated and
342 treated cells) using separate plates. In addition, a third plate was prepared, being dedicated only
343 to CP. On that plate, 200 cells (100 treated plus 100 control) were collected per well in the first
344 two rows. Altogether, 24 wells of CP cells were collected for each treatment and time point.
345

346 **5.3. Protein extraction and digestion**

347 Proteins from single cells and CPs were extracted in four freeze – thaw cycles. Plates were
348 frozen for 2 min in liquid nitrogen and immediately heated at 37°C for 2 min. Proteins were
349 denatured by heating the plates at 90°C for 5 min. The resulted protein solutions were
350 centrifuged at 1000 rpm for 2 min to spin down all the volume present in the wells. Finally, 1
351 µL of 25 ng/µL sequencing grade trypsin (Promega, Madison WA) in 100 mM TEAB was
352 dispensed using a MANTIS® automatic dispenser (Formulatrix, Bedford MA). In the case of
353 CP, digestion was achieved with 2 µL of trypsin solution added. Plates were incubated at 37°C
354 overnight (16 h).
355

356 **5.4. TMT labeling**

357 Both TMT10plex™ (or TMT11plex™ including channel 131C) and 16-channel TMTpro™
358 (ThermoFisher Scientific, Rockford IL) were used in this study. Unless specified, each
359 TMT10plex™ set contained 4 control cells and 4 treated cells with tags interspaced, as well as
360 a single channel with CP (200 cells in channel 131). TMT 130N was not used because of the
361 cross contamination with the CP channel. Peptides were TMT-labeled by dispensing 1 µL of
362 the respective TMT reagent dissolved in dry acetonitrile (ACN) at a concentration of 10 µg/µL
363 using the MANTIS robot. Plates were incubated at room temperature (RT) for 2 h and then the
364 reaction was quenched by adding 1 µL of 5% hydroxylamine (also with automatic liquid
365 handler), following incubation at RT for 15 min. In the case of TMTpro™ labeling, each set
366 contained 6 control cells and 6 treated cells with the tags interspaced. The CP was split in two
367 channels, 126 and 127N, one composed of 100 control cells and the other one 100 treated cells.
368 Channels 127C and 128N were left empty to prevent cross-contamination from CP channels.
369 The labeled samples were pooled together using a 10-µL glass syringe (VWR, Japan), starting
370 always with the CP samples in each TMT set in order to minimize sample loss during the
371 pooling.[19] Samples were pooled into MS-sample vials with glass insert (TPX snap ring vial
372 from Genetec, Sweden) and dried in a speed vacuum concentrator (Concentrator Plus,
373 Eppendorf). Dry peptides were resuspended in 7 µL of 2% ACN, 0.1% formic acid (FA) prior
374 to LC-MS/MS analysis.

375

376 **5.5. Bulk proteome analysis**

377 Unless specified, $1x10^6$ were harvested and washed twice with PBS. To each cell pellet, 150
378 µL of 8M urea in 50 mM ammonium bicarbonate (AmBic) pH 8.5 were added, tubes were
379 sonicated in bath at 4°C for 5 min and probe-sonicated for 40 s with a Vibra-cell™ ultrasonic
380 liquid processor (VCX 130, Sonics & Materials Inc., Newtown CT) using pulse 2/2 with amp

381 20. Each cell lysate was transferred to a new tube, while the original tube was washed with 150
382 µL of 50 mM AmBic pH 8.5 and the wash was added to the new tube. Samples were centrifuged
383 for 10 min at 13,000 *g* and 4°C. Protein concentration was measured with BCA protein assay.
384 Proteins were reduced by adding 20 mM dithiothreitol in 50 mM AmBic and incubated at 37°C
385 for 1 h, followed by alkylation with 40 mM iodoacetamide at RT for 30 min in the dark. Finally,
386 proteins were digested by adding trypsin in a 1:50 ratio and incubating overnight at 37°C. The
387 resulting peptides were cleaned-up on HyperSep™ filter plate (ThermoFisher Scientific,
388 Rockford IL) then dried in a speed vacuum concentrator. Dried samples were dissolved in 100
389 mM TEAB and labeled with TMT10plex™ reagents that were dissolved in dry ACN and then
390 mixed with the peptide solutions and incubated at RT for 1 h. The reaction was quenched by
391 adding 11 µL of 5% hydroxylamine and incubating at RT for 15 min. Finally, samples with
392 different TMT labels were pooled into a single tube, cleaned-up on HyperSep™ filter plate and
393 dried in a speed vacuum concentrator prior to LC-MS/MS analysis.
394
395 **5.6. RPLC-MS/MS analysis**
396 Peptide samples were separated on a Thermo Scientific™ Ultimate™ 3000 UHPLC
397 (ThermoFisher Scientific) using a 10 min loading at 3 µL/min flow rate to a trap column
398 (Acclaim™ PepMap™ 100, 2 cm x 75 µm, 3 µm, 100 Å - ThermoFisher Scientific). The
399 separation was performed on an EASY-Spray™ C18 analytical column (25 cm x 75 µm, 1.9
400 µm, 300 Å – ES802A, ThermoFisher Scientific). A constant flow rate of 100 nL/min was
401 applied during sample separation achieved in a linear gradient ramped from 5%B to 27%B over
402 120 min, with solvents A and B being 2% ACN in 0.1% FA and 98% ACN in 0.1% FA,
403 respectively. LC-MS/MS data were acquired on an Orbitrap Fusion™ Lumos™ Tribrid™ mass
404 spectrometer (ThermoFisher Scientific, San José CA), using nano-electrospray ionization in
405 positive ion mode at a spray voltage of 1.9 kV. To reduce atmospheric background and enhance

406 the peptide signal-to-noise ratio, an Active Background Ion Reduction Device (ABIRD - ESI
407 Source Solutions, LLC, Woburn MA) was used on the nanospray ion source. Data dependent
408 acquisition (DDA) mode parameters were set as follows: isolation of top 20 precursors in full
409 mass spectra at 120,000 mass resolution in the *m/z* range of 375 – 1500, maximum allowed
410 injection time (IT) of 100 ms, dynamic exclusion of 10 ppm for 45 s, MS2 isolation width of
411 0.7 Th with higher-energy collision dissociation (HCD) of 35% at resolution of 50,000 and
412 maximum IT of 150 ms in a single microscan. The mass spectrometry proteomics data are
413 deposited to the ProteomeXchange Consortium via the PRIDE partner repository[28] with the
414 data set identifier PXD025481.
415
416 **5.7. Data analysis**
417 Raw data from LC-MS/MS were analyzed on Proteome Discoverer v2.4 (ThermoFisher
418 Scientific), searching proteins against SwissProt human database (release July 30, 2019 with
419 20,373 entries) and known contaminants with Mascot Server v2.5.1 (MatrixScience Ltd, UK)
420 allowing for up to two missed cleavages. Mass tolerance for precursor and fragment ions was
421 10 ppm and 0.05 Da, respectively. Oxidation of methionine, deamidation of asparagine and
422 glutamine, as well as TMTpro or TMT6plex lysine and N-termini were set as variable
423 modifications. Percolator node[29] in Proteome Discoverer was set to target false discovery rate
424 (FDR) at 1% with validation based on *q*-value.
425 The TMT reporter ion abundances (RIA) at a peptide level were extracted from the search
426 results. The subsequent analyses were performed in the RStudio (version 1.3.1073)
427 programming language environment, the software for multivariate data analytics SIMCA (v.
428 15.0.2.5959, Sartorius) and Perseus software platform.[30] Peptides from single cells with RIAs
429 exceeding 10% of the abundance values for the respective carrier channel were filtered out,
430 being considered a result of co-isolation or other interferences, resulting in about 30% of the

431 peptides discarded for further analysis. After filtering, the remaining RIAs were arranged into
432 a matrix of peptide IDs (rows) *vs.* single cells (columns). All RIAs were log2-transformed and
433 the data were normalized in columns by subtracting their median values computed, ignoring
434 the missing values. Peptides quantified in less than ten cells were discarded (usually <0.05%
435 peptides per dataset). Protein-level quantification was achieved by attributing each unique
436 peptide to its respective top ranked protein within a protein group. As protein relative
437 abundance, the median RIA value among the peptides belonging to that protein was taken. The
438 new relative abundance matrix (protein IDs *vs.* single cells) was again normalized by
439 calculating the median value for each column (or single cell), and then, subtracting the median
440 value calculated to each abundance on the respective column. Missing values in the resulting
441 matrix were imputed based on the normal distribution of valid values (method available in
442 Perseus software platform[31]), using a width of 0.3 standard deviations of the Gaussian
443 distribution of the valid values and a downshift of 1.8 standard deviations. Finally, the batch
444 effects across the TMT sets were corrected by applying an empirical Bayesian framework in
445 the SVA package[32].
446 The obtained matrix of relative protein abundances was used for statistical analysis. Principal
447 component analysis (PCA) was performed to determine the separation degree between the
448 control and drug-treated cells and to identify the outliers (single cells outside the limits of the
449 PCA diagram with $p < 0.05$), which were removed from subsequent analysis. Resulting data
450 were analyzed by orthogonal partial least squares discriminant analysis (OPLS-DA) and
451 clustering analysis, and the fold changes were presented as volcano plots.
452
453
454
455

6. Acknowledgements

We would like to express our gratitude to Bogdan Budnik (Harvard Center for Mass Spectrometry, Harvard University, Cambridge MA) for discussion and advice on experimental design and suggesting the Mantis liquid handling robot, John Neveu (ESI Source Solutions, Woborn MA) and Erik Verschuuren (MS Wil B.V., Aarle-Rixtel, The Neatherlands) for ABIRD connection; Ujjwal Neogi and Flora Mikeloff (Karolinska Institutet) for valuable help with R algorithms, Indira Pla (Lund University) for useful advices in statistics and R algorithms, Willy Björklund, Erik Damgård and many others at Thermo Scientific for valuable technical support. The study was funded by the Swedish Research Council (grant 2018-06156, VR).

7. Authors contributions

RAZ conceptualized the study. AV and JER contributed to the conception and design of the study. AV and JER performed experiments and data analysis. RAZ wrote the manuscript and all authors contributed to the article and approved the submitted version.

8. Competing interest statement

The authors declare no competing interest.

474 **9. References**

475

476 1. Saei, A. A. *et al.* Comparative Proteomics of Dying and Surviving Cancer Cells
477 Improves the Identification of Drug Targets and Sheds Light on Cell Life/Death
478 Decisions. *Mol. Cell. Proteomics* **17**, 1144–1155 (2018).

479 2. Saei, A. A. *et al.* ProTargetMiner as a proteome signature library of anticancer
480 molecules for functional discovery. *Nat. Commun.* **10**, 5715 (2019).

481 3. Gaetani, M. & Zubarev, R. New Promises of Chemical Proteomics for Drug
482 Development. *Nov. Approaches Drug Des. Dev.* **2**, 12–14 (2017).

483 4. Moffat, J. G., Vincent, F., Lee, J. A., Eder, J. & Prunotto, M. Opportunities and
484 challenges in phenotypic drug discovery: an industry perspective. *Nat. Rev. Drug*
485 *Discov.* **16**, 531–543 (2017).

486 5. Ye, C. *et al.* DRUG-seq for miniaturized high-throughput transcriptome profiling in
487 drug discovery. *Nat. Commun.* **9**, 4307 (2018).

488 6. Chernobrovkin, A., Marin-Vicente, C., Visa, N. & Zubarev, R. A. Functional
489 Identification of Target by Expression Proteomics (FITExP) reveals protein targets and
490 highlights mechanisms of action of small molecule drugs. *Sci. Rep.* **5**, 11176 (2015).

491 7. Hsieh, Y.-C. *et al.* Enhanced Degradation of Dihydrofolate Reductase through
492 Inhibition of NAD Kinase by Nicotinamide Analogs. *Mol. Pharmacol.* **83**, 339–353
493 (2013).

494 8. Banerjee, D. *et al.* Novel aspects of resistance to drugs targeted to dihydrofolate
495 reductase and thymidylate synthase. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1587**,
496 164–173 (2002).

497 9. Marin-Vicente, C., Lyutvinskiy, Y., Romans Fuertes, P., Zubarev, R. A. & Visa, N.
498 The Effects of 5-Fluorouracil on the Proteome of Colon Cancer Cells. *J. Proteome Res.*

499 **12**, 1969–1979 (2013).

500 10. Ameisen, J. C. On the origin, evolution, and nature of programmed cell death: a
501 timeline of four billion years. *Cell Death Differ.* **9**, 367–393 (2002).

502 11. Mansilla, S., Llovera, L. & Portugal, J. Chemotherapeutic Targeting of Cell Death
503 Pathways. *Anticancer. Agents Med. Chem.* **12**, 226–238 (2012).

504 12. Richards, A. L., Merrill, A. E. & Coon, J. J. Proteome sequencing goes deep. *Curr.*
505 *Opin. Chem. Biol.* **24**, 11–17 (2015).

506 13. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual
507 Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).

508 14. Liu, Y., Beyer, A. & Aebersold, R. On the Dependency of Cellular Protein Levels on
509 mRNA Abundance. *Cell* **165**, 535–550 (2016).

510 15. Bendall, S. C. *et al.* Single-Cell Mass Cytometry of Differential Immune and Drug
511 Responses Across a Human Hematopoietic Continuum. *Science (80-.).* **332**, 687–696
512 (2011).

513 16. Heath, J. R., Ribas, A. & Mischel, P. S. Single-cell analysis tools for drug discovery
514 and development. *Nat. Rev. Drug Discov.* **15**, 204–216 (2016).

515 17. Zubarev, R. A. The challenge of the proteome dynamic range and its implications for
516 in-depth proteomics. *Proteomics* **13**, 723–726 (2013).

517 18. Schwanhäusser, B. *et al.* Correction: Corrigendum: Global quantification of
518 mammalian gene expression control. *Nature* **495**, 126–127 (2013).

519 19. Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCoPE-MS: mass spectrometry of
520 single mammalian cells quantifies proteome heterogeneity during cell differentiation.
521 *Genome Biol.* **19**, 161 (2018).

522 20. Emmott, E. *et al.* How to perform quantitative single cell proteomics with SCoPE2. *J.*
523 *Biomol. Tech.* **31**, S36–S37 (2020).

524 21. Specht, H. *et al.* Single-cell proteomic and transcriptomic analysis of macrophage
525 heterogeneity using SCoPE2. *Genome Biol.* **22**, 50 (2021).
526 22. Zhu, Y. *et al.* Nanodroplet processing platform for deep and quantitative proteome
527 profiling of 10-100 mammalian cells. *Nat. Commun.* **9**, 882 (2018).
528 23. Zhang, P. *et al.* Carrier-Assisted Single-Tube Processing Approach for Targeted
529 Proteomics Analysis of Low Numbers of Mammalian Cells. *Anal. Chem.* **91**, 1441–
530 1451 (2019).
531 24. Tsai, C.-F. *et al.* An Improved Boosting to Amplify Signal with Isobaric Labeling
532 (iBASIL) Strategy for Precise Quantitative Single-cell Proteomics. *Mol. Cell.*
533 *Proteomics* **19**, 828–838 (2020).
534 25. Cheung, T. K. *et al.* Defining the carrier proteome limit for single-cell proteomics. *Nat.*
535 *Methods* **18**, 76–83 (2021).
536 26. Oh, B. Y. *et al.* Intratumor heterogeneity inferred from targeted deep sequencing as a
537 prognostic indicator. *Sci. Rep.* **9**, 4542 (2019).
538 27. Gaetani, M. *et al.* Proteome Integral Solubility Alteration: A High-Throughput
539 Proteomics Assay for Target Deconvolution. *J. Proteome Res.* **18**, 4027–4037 (2019).
540 28. Deutsch, E. W. *et al.* The ProteomeXchange consortium in 2020: Enabling 'big data'
541 approaches in proteomics. *Nucleic Acids Res.* **48**, D1145–D1152 (2020).
542 29. Spivak, M., Weston, J., Bottou, L., Käll, L. & Noble, W. S. Improvements to the
543 Percolator Algorithm for Peptide Identification from Shotgun Proteomics Data Sets. *J.*
544 *Proteome Res.* **8**, 3737–3745 (2009).
545 30. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of
546 (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
547 31. Yu, S.-H., Kyriakidou, P. & Cox, J. Isobaric Matching between Runs and Novel PSM-
548 Level Normalization in MaxQuant Strongly Improve Reporter Ion-Based

549 Quantification. *J. Proteome Res.* **19**, 3945–3954 (2020).

550 32. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The sva package
551 for removing batch effects and other unwanted variation in high-throughput
552 experiments. *Bioinformatics* **28**, 882–883 (2012).

553

554

555 **10. Figure legends**

556

557 **Figure 1.** SCCP workflow. The workflow developed for SCCP included cell culturing and
558 treatment with drugs, isolation of individual cells by FACS, protein extraction and digestion,
559 TMT labeling of thus obtained tryptic peptides followed by multiplexing, LC-MS/MS and
560 statistical data analysis. All steps are optimized for achieving the desired proteome depth and
561 quantitative correlation with bulk analysis. In the figure, split carrier proteome occupies two
562 channels (131N and 131C) in a TMT11plex set, with two other channels (130N and 130C)
563 remaining empty (doted lines). Identification of peptides is achieved via matching masses of
564 sequence-specific fragments, and quantification is performed by the abundances of the low-
565 mass TMT reporter ions.

566

567 **Figure 2.** Benchmarking SCCP data. A) Correlation between the protein regulations in the
568 semi-bulk CPs (y-axis) and the median SC data (x-axis). B) Volcano plot of protein regulations
569 obtained from SCCP data showing MTX target DHFR to be among the most significantly
570 upregulated proteins. The curved lines separate the areas with significantly regulated proteins.

571

572 **Figure 3.** Time-course results upon treatment with MTX. PCA plots of SCCP data as a time
573 course demonstrating the emergence of separation between the MTX-treated and untreated
574 attached cells with incubation time, and the corresponding volcano plots of regulated proteins
575 showing the emergence of DHFR among the top regulated proteins.

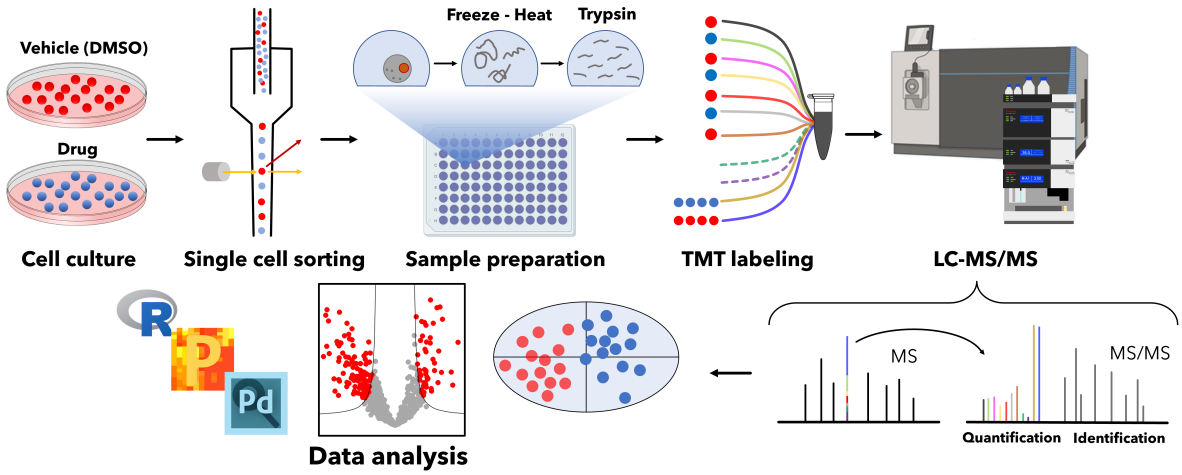
576

577 **Figure 4.** Statistical analysis of single cells treated with MTX. A) OPLS-DA analysis of SCCP
578 data on median protein abundances in G1 and G2 cell groups from MTX treated cells at
579 different time points together with bulk CP abundances for the total proteome (1170 proteins)

580 and top 100 most abundant proteins. B) Distribution of the main OPLS-DA coordinates of G1
581 and G2 groups of MTX-treated attached cells at 12 h and 24 h past treatment for total proteome,
582 top 400 and top 100 proteins. The x-coordinates were normalized such that the coordinates of
583 the attached and detached cell bulk-analyzed proteomes after 48 h treatment are +1 and -1,
584 respectively.
585
586 **Figure 5.** Time-course results upon treatment with CPT and TDX. PCA plots of SCCP data as
587 a time course demonstrating the emergence separation between the untreated cells and the
588 attached cells treated with A) CPT and B) TDX with incubation time, and the corresponding
589 volcano plots of regulated proteins showing the emergence of the known drug target among the
590 top regulated proteins.
591

592 **Figures**

593
594



595
596
597
598 **Figure 1.**
599

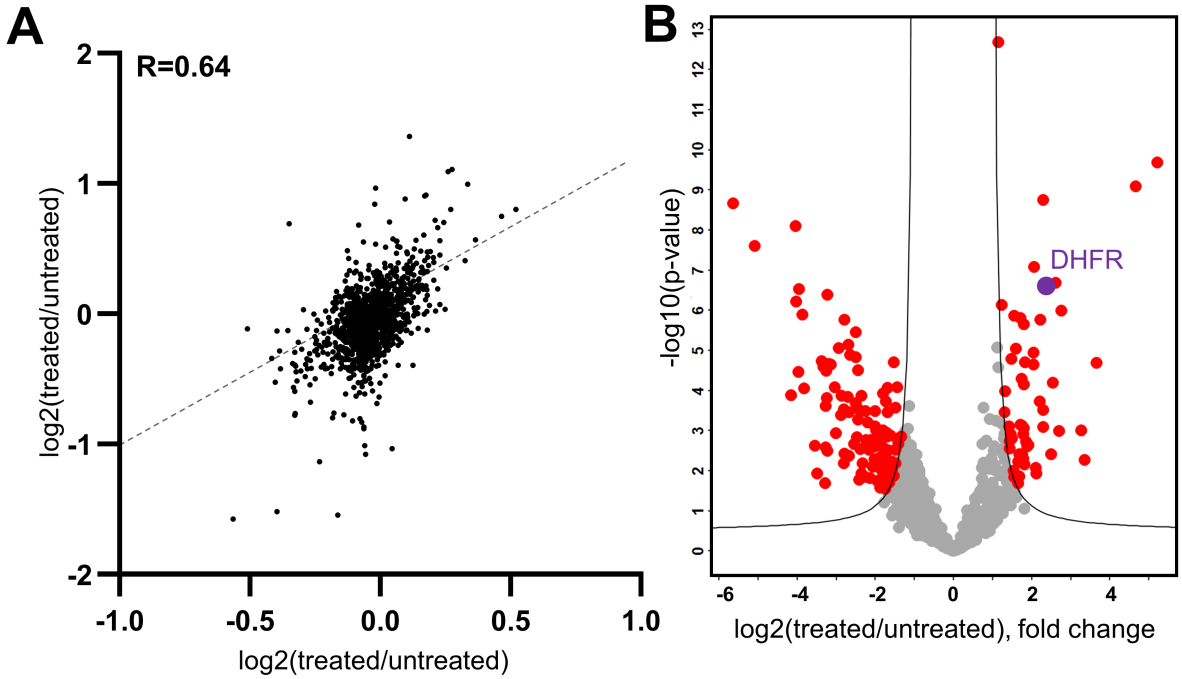

A
R=0.64
log2(treated/untreated)
log2(treated/untreated)
B
DHFR
-log10(p-value)
log2(treated/untreated), fold change

Figure 2.

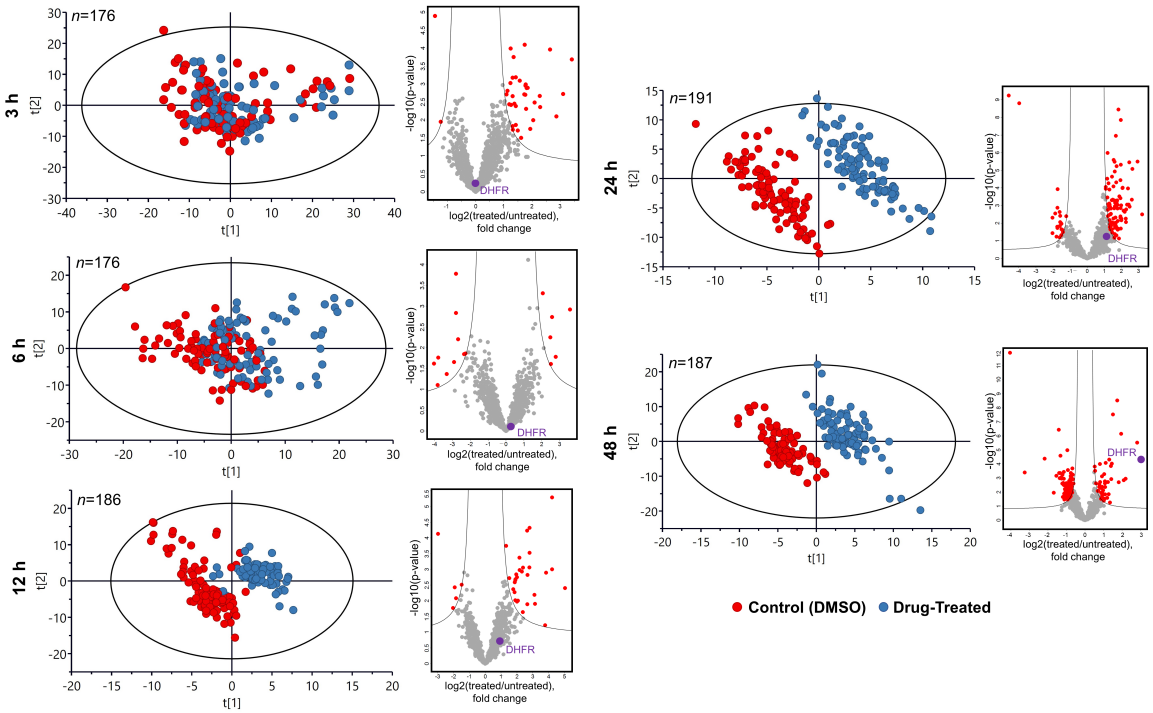

3 h
n=176
t[1]
t[2]
-log10(p-value)
log2(treated/untreated), fold change
DHFR
6 h
n=176
12 h
n=186
24 h
n=191
48 h
n=187
Control (DMSO)
Drug-Treated

Figure 3.

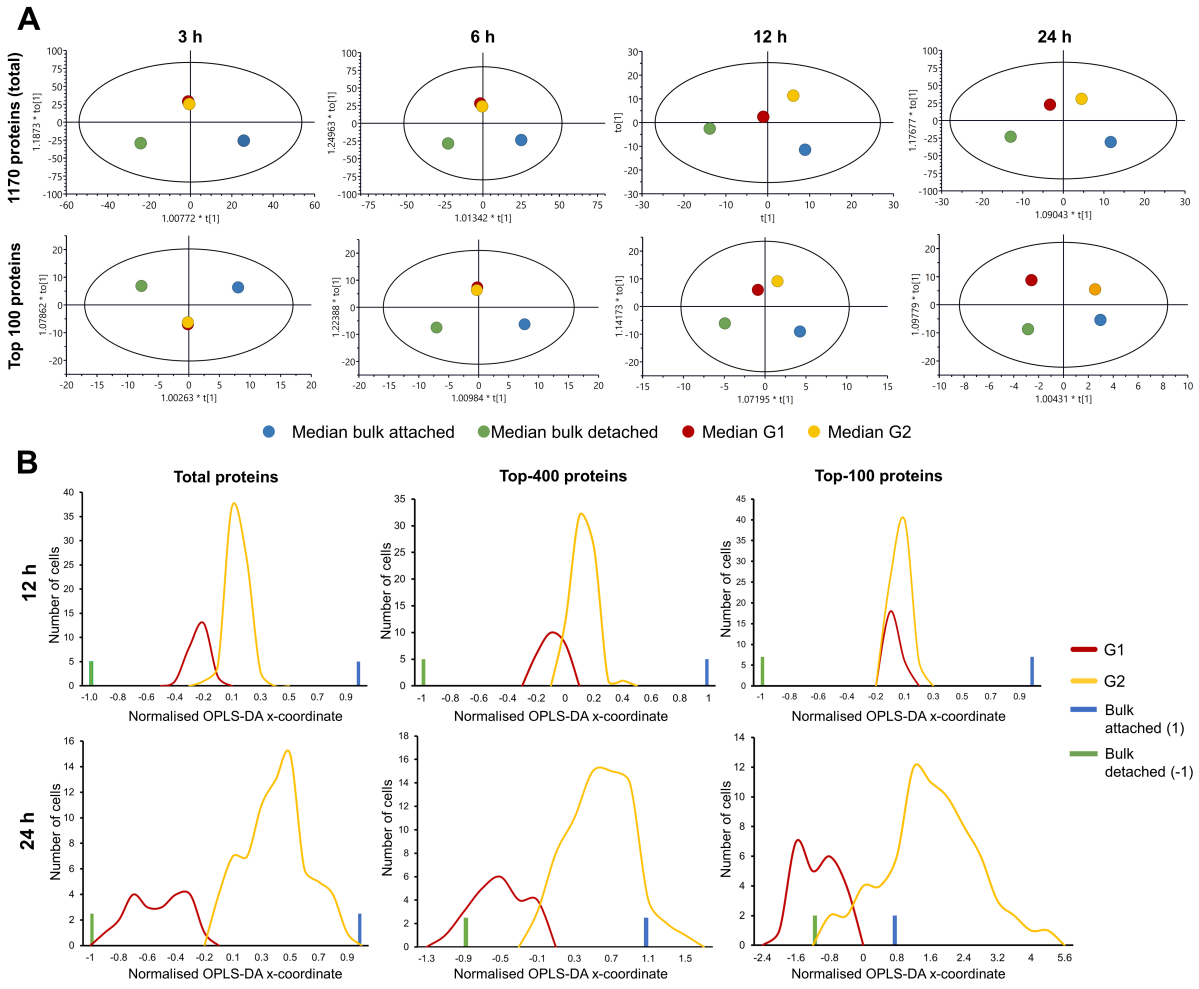
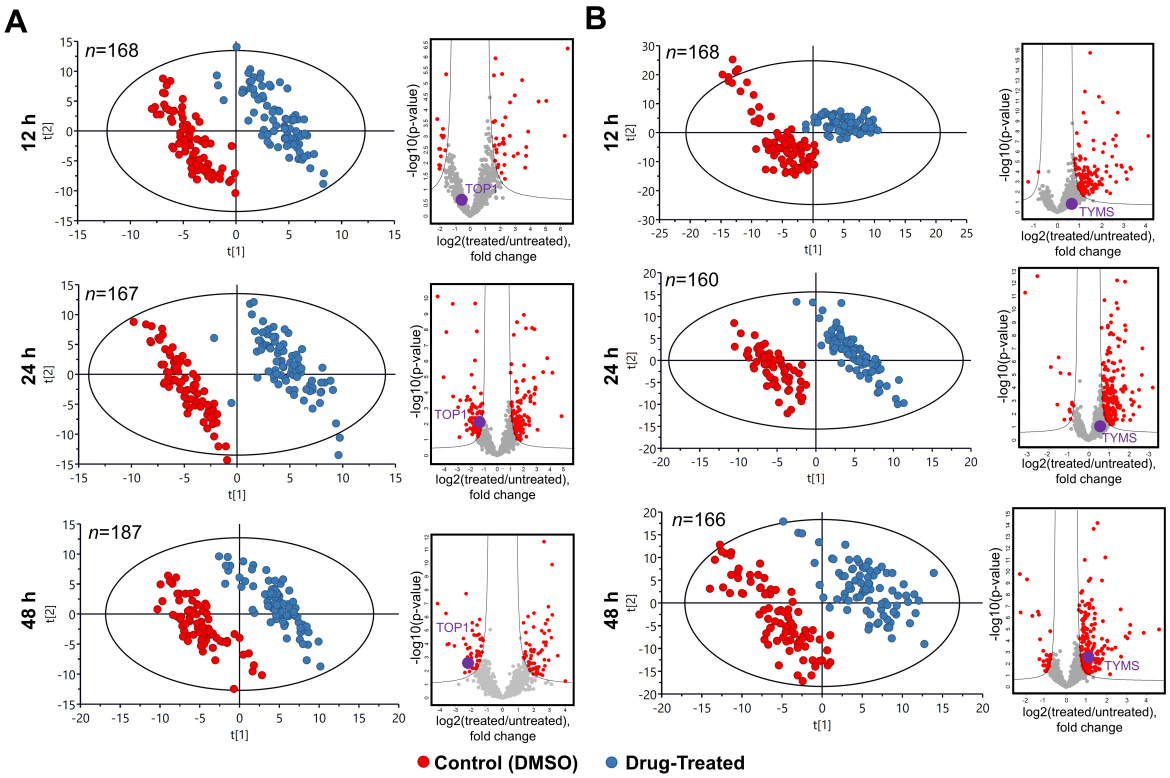

A
3 h
6 h
12 h
24 h
1170 proteins (total)
Top 100 proteins
Median bulk attached
Median bulk detached
Median G1
Median G2
B
Total proteins
Top-400 proteins
Top-100 proteins
12 h
24 h
Number of cells
Normalised OPLS-DA x-coordinate
G1
G2
Bulk attached (1)
Bulk detached (-1)

Figure 4.

613



614
615

616 **Figure 5.**

617