JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

NMR Spectroscopy-Based Metabolic Profiling of Drug-Induced Changes In Vitro Can Discriminate between Pharmacological Classes

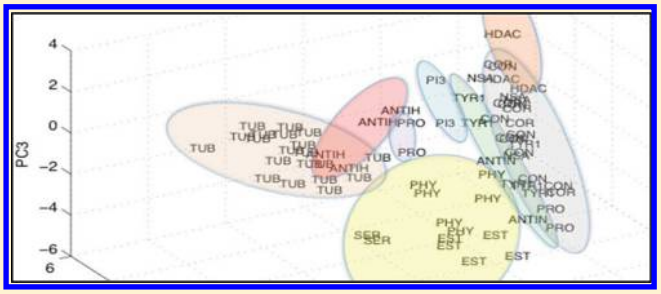
Obaid Aftab,[†] Mikael K. R. Engskog,[‡] Jakob Haglöf,[‡] Albert Elmsjö,[‡] Torbjörn Arvidsson,[‡] Curt Pettersson,[‡] Ulf Hammerling,*[,†] and Mats G. Gustafsson*[,†,§]

[†]Department of Medical Sciences, Cancer Pharmacology and Computational Medicine, Uppsala University, Uppsala University Hospital, SE-751 85, Uppsala, Sweden

[‡]Division of Analytical Pharmaceutical Chemistry, Uppsala University, Box 574, SE-751 23 Uppsala, Sweden

Supporting Information

ABSTRACT: Drug-induced changes in mammalian cell line models have already been extensively profiled at the systemic mRNA level and subsequently used to suggest mechanisms of action for new substances, as well as to support drug repurposing, i.e., identifying new potential indications for drugs already licensed for other pharmacotherapy settings. The seminal work in this field, which includes a large database and computational algorithms for pattern matching, is known as the "Connectivity Map" (CMap). However, the potential of similar exercises at the metabolite level is still largely unexplored. Only recently, the first high-throughput metabolomic assay pilot study was published, which involved screening the metabolic response to a set of 56 kinase inhibitors in a 96-well format. Here, we report results from a separately developed metabolic profiling assay, which leverages ^{1}H NMR spectroscopy to the quantification of metabolic changes in the HCT116 colorectal cancer cell line, in response to each of 26 compounds. These agents are distributed across 12 different pharmacological classes covering a broad spectrum of bioactivity. Differential metabolic profiles, inferred from multivariate spectral analysis of 18 spectral bins, allowed clustering of the most-tested drugs, according to their respective pharmacological class. A more-advanced supervised analysis, involving one multivariate scattering matrix per pharmacological class and using only 3 spectral bins (3 metabolites), showed even more distinct pharmacology-related cluster formations. In conclusion, this type of relatively fast and inexpensive profiling seems to provide a promising alternative to that afforded by mRNA expression analysis, which is relatively slow and costly. As also indicated by the present pilot study, the resulting metabolic profiles do not seem to provide as information-rich signatures as those obtained using systemic mRNA profiling, but the methodology holds strong promise for significant refinement.



INTRODUCTION

Over the past decade, systemic molecular profiling technologies have emerged as highly useful tools to acquire a deeper understanding of drug actions as well as disease mechanisms. This particularly applies to mRNA gene expression profiling where the "Connectivity Map" (CMap), which is a database and core concept based on a highly standardized *in vitro* setting, represents a major landmark.[1] CMap has enabled the identification of novel clinical areas for several already licensed drugs (drug repurposing), thereby significantly expanding their applicability.[2−4] It has also been very useful for the elucidation of drug mechanisms for several uncharacterized drug candidates.[5−14] Despite an impressive body of interesting reports, this platform is still hampered by certain inherent insufficiencies, especially high costs per assay and a relatively slow and elaborative experimental procedure.[15] Very recently, a low-cost and more-high-throughput alternative to the original CMap database was launched under the label LINCS (Library of Integrated Network-based Cellular Signatures, http://lincs.hms.harvard.edu/). In contrast to mRNA microarrays covering ∼22 000 probe sets, as in the current layout (second generation) of the CMap database, LINCS mRNA profiling is compacted to 1000 carefully selected mRNA probe sets that have been determined to be the most informative, in terms of explaining (predicting) values of the larger set.

Other alternative routes to drug profiling, which are based on platforms that are faster and less costly than comprehensive mRNA profiling, are attracting mounting interest among basic research and pharmaceutical development communities alike. Metabolic profiling (metabolomics) is one such experimental format.[16,17] Most early reports on this general platform involve toxicology testing and characterization of inborn metabolic errors, but the application range has expanded to encompass a range of additional areas, as typified by subtle pathological alterations in complex diseases and pharmacological response to drug cues.[18] Metabolomics has similarly demonstrated applicability in diagnostic and prognostic biomarker discovery,

Received: August 17, 2014
Published: October 16, 2014

dx.doi.org/10.1021/ci500502f | J. Chem. Inf. Model. 2014, 54, 3251−3258

as well as drug discovery and development.[19−22] Besides low profiling cost as well as fast and relatively straightforward experimental procedures, metabolic profiling has added value, compared to mRNA profiling, because metabolites are inherently of high biological relevance.[23] Both (proton) nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) settings are mainstay experimental platforms for metabolomics-oriented studies; each has its own merits and drawbacks. NMR spectroscopy is being acknowledged for its robustness and adaptation to automation, whereas MS affords higher sensitivity.[24] However, the potential of a relatively fast and inexpensive comparative *in vitro* metabolic profiling of sets of substances, which can help suggesting mechanism(s) of actions for unknown substances and to propose new therapies largely by analogy with the CMap concept, is still largely unexplored. Only recently, the first pilot study reporting a high content metabolomic assay for screening of the metabolomic response of drug-treated mammalian cells in a 96-well format was published and the single application introduced was to monitor the cellular metabolomic response in the cell line CCRF-CEM to a library of 56 preselected kinase inhibitors.[25]

Here, we report results from using an independently developed metabolic semihigh throughput profiling assay, which is based on NMR spectroscopy. A major discrepancy, relative to the already published pilot study, is the metabolite extraction protocol: our procedure involves the separation of metabolites into one polar and one nonpolar fraction prior to NMR analysis (see the Materials and Methods section). This entails a slightly more elaborate experimental protocol, but can help alleviating problems with overlapping spectral peaks originating from different compounds. Such overlaps are commonly seen in ^{1}H NMR spectra of complex samples, such as cell lysates, and may entail difficulties with respect to identification and quantification of individual metabolites present in the sample. In the proof-of-concept study reported here, ^{1}H NMR spectra were collected only for the polar fraction of the cell lysates and used to study changes in the polar metabolome in the colorectal cancer cell line HCT116, following exposure to 26 different market-licensed or investigational drugs distributed across 12 different pharmacological classes. The subsequent multivariate data processing reveals that many among the 26 accordingly tested agents clustered in compliance with their respective major pharmacology classification. Thus, the results presented support the notion that this type of metabolite profiling offers a straightforward and inexpensive method to predict the drug mode of action. Similarly analogous with the original CMap approach, the presented metabolite profiling opens a new experimental avenue to suggesting alternative (individualized) pharmacotherapies, based on inverse pharmacology response versus disease signature patterns.

MATERIALS AND METHODS

Experimental Design. All drug experiments were performed twice, resulting in two different spectra. The experiments were organized into five different batches, each containing multiple dimethyl sulfoxide (DMSO) controls, multiple sterile water controls, and multiple wells containing one of the two drugs: Vincristine or Vinblastine.

Drug Classes. The study was based on profiling of 26 investigational or market-approved drugs that may be annotated into the following 12 pharmacological classes: TUB, EST, PHY, SER, ANTIN, ANTIH, PI3, HDAC, PRO, COR, NSA, and TYR1. [For details, see the Supporting Information (Table S1).] All drugs were purchased from Sigma−Aldrich (Stockholm, Sweden) and dissolved in DMSO with a DMSO final concentration equivalent to 0.01% or less.

Chemicals for the Metabolite Extraction Protocol and NMR Analysis. Chloroform and 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt (DSS) were obtained from Sigma−Aldrich (Steinheim, Germany) and deuterium oxide (D_2O, 99% D) from Armar Chemicals (Döttingen, Switzerland). Na_2HPO_4 and NaH_2PO_4 were from Fluka (Buchs, Switzerland), while methanol (HPLC grade) was purchased from Fisher Scientific (Zurich, Switzerland). All chemicals were of analytical grade. The water was purified using a Milli-Q water system (MilliPore, Bedford, MA, USA).

Cell Culturing and Sample Harvesting. The human colon carcinoma cell-line HTC 116 (from Horizon Discovery, Ltd., Cambridge, U.K.) was used for this study. The cells were cultured in an incubator at 37 °C under 5% CO_2, grown in 21.5 cm^2 culture dishes from Thermo Fisher Scientific (Denmark) and supplemented with McCoy's 5A medium (Catalog No. M8403) with 10% heat inactivated fetal calf serum (Catalog No. F6131), 2 mM glutamine (Catalog No. G7513), 100 μg/mL streptomycin, and 100 U/mL penicillin (Catalog No. P0781). Cell culture reagents were purchased from Sigma−Aldrich, unless stated otherwise. Cells were harvested at ∼75% confluence, equivalent to ∼5 million cells, using cold methanol. All cell sample harvesting was performed on ice. The growth medium was removed and cells were rapidly washed three times with cold Dulbecco's phosphate buffered saline (PBS, Nordic Biolabs), followed by detachment of cells using a rubber-tipped cell scraper (Sarstedt). The detached cells were collected using cold methanol (2 mL) and transferred to polypropylene tubes for storage at −80 °C until metabolite extraction.

Metabolite Extraction. Harvested cells were thawed at room temperature and subjected to centrifugation (10 min, 4 °C, 2200 RCF) to remove precipitated cellular debris. The supernatant was transferred to a fresh extraction tube, followed by the addition of chloroform and Milli-Q water in the proportions 4:4:2.85 (MeOH:$CHCl_3$:Milli-Q water).[26] The resulting two-phase system was gently mixed and then left at 4 °C for 30 min before centrifugation (2200 RCF, 20 min, 4 °C). The aqueous phase was transferred to a new tube and evaporated under N_2 at 40 °C until dryness. The aqueous phase was subsequently reconstituted in 0.7 mL phosphate buffered D_2O (150 mM, pD 7) containing 34 μM DSS as a chemical shift reference. The $CHCl_3$ phases were stored for future studies. For the analysis, pooled quality control (QC) samples were also created by pooling an equal volume of cell extracts prior to extraction, i.e., generating one QC sample for each batch.

NMR Spectroscopy Measurements. NMR measurements were carried out at 298 K on a Bruker Model Avance 600 MHz system (Bruker BioSpin GmbH, Rheinstetten, Germany) that was equipped with a cryoprobe. For each sample, the 1D NOESYPR1D standard pulse sequence (−RD−90°−t_1−90°−t_m−90°−ACQ) was used. Each pulse had a pulse length of 90°; the total amount of FIDs recorded were 256, collected into 64K data points that was zero-filled to 128K data points. The spectra width was set to 7183.91 Hz, giving a spectral acquisition time of 4.56 s. The longitudinal relaxation time (t_1) and mixing time (t_m) was set to 6 μs and 180 ms, respectively, and the relaxation delay (D_1) was 3 s, resulting in a total acquisition time of 33

min.[27] The one-dimensional (1D) spectra were manually phased and baseline-corrected and the ^{1}H chemical shifts were referenced to added DSS using the ACDLABS software (version 12.01, Advanced Chemistry Development, Inc., Canada).

Basic Spectral Preprocessing (Phase Correction, Baseline Correction, Binning). Each ^{1}H NMR spectrum, within a range of 0−10 ppm, was reduced to 873 bins of fixed width (0.01 ppm) excluding resonance regions for water (δ 5.15−4.67 ppm), DMSO (δ 2.70−2.72 ppm), and the internal standard (DSS, δ 0.65−0.00, 1.77−1.72, and 2.92−2.88 ppm). The signal intensity in each bin was integrated using the ACDLABS software. Because of the absence of peaks in one large noise region of the spectral range, the spectra were reduced to 350 bins before any further processing.

Spectral Normalization and Removal of Noisy Bins. Each spectrum was normalized by a trimmed mean procedure in which the spectrum was scaled by the area under the 10% of bins having amplitudes most similar to the median level. This procedure was used in order to achieve a robust normalization step, which may be regarded as a compromise between classical area normalization and median normalization. Area normalization may be misleading if there are single outstanding peaks (outliers). Median normalization may be noisy, since only a single amplitude value (the median), which may be contaminated with experimental noise, is used to scale the spectrum.

Metabolite Identification. Assignments of NMR peaks were performed according to the Metabolomics Standards Initiative[28] by comparing with literature[26] and the Human Metabolome Database (V 2.5).[26,28−30] The metabolites identified should be seen as putatively annotated compounds, according to the Metabolomics Standards Initiative nomenclature.

Determination of Induced Differences by Subtracting a Batch Control Average. For each drug experiment, the differences induced by the drug were determined by subtracting a reference spectrum, which was defined as the arithmetic average across all DMSO control spectra for the particular batch in which the drug experiment was performed.

Elimination of Unreliable Bins. After determination of the induced differences for all drug experiments, the coefficient of variation (CV), which was determined for each of the 350 bins, was performed based on the repeated experiments performed for the drug Vinblastine performed across all five batches. A simple and reasonable criterion was employed: Only bins showing a CV value of <25% were kept; the remaining bins were considered to have too much experimental variability to be regarded as reliable. Only 18 bins passed this quality control criterion.

Principal Component Analysis. Only for visualization purposes, the 18-dimensional differential profiles were subject to conventional principal component analysis (PCA) and subsequent dimensionality reduction into a three-dimensional (3-D) point (profile) for each drug.

Maximal Class Separability via Within-Class and Between-Class Scatter Matrices. A subset of 3 features was selected by a comprehensive screen of all possible subsets of 3 bins among the 18 bins showing sufficiently low CV values. The criterion used for selection of the most discriminatory subset of bins was

$$J = \frac{|S_b|}{|S_w|}$$

where $|S_b|$ is the determinant of the 3 × 3-dimensional between-class scatter matrix S_b and $|S_w|$ is the determinant of the 3 × 3-dimensional within-class scatter matrix S_w. These two scatter matrices were defined and determined as follows. S_b is simply the covariance matrix of the mean spectra for each of the 14 different experimental classes consisting of the 12 pharmacological classes as well as CON (sterile water controls) and QC (pooled samples used only for quality control purposes) (not shown). S_w is simply the average covariance matrix of all the 14 different covariance matrices obtained for each experimental class. Some slightly different forms of scatter matrices and performance criteria have been proposed in the literature; however, for current purposes, we found the present one sufficient.

Nearest-Neighbor Search. The differential metabolic profile corresponding to each drug measurement was used as a query to determine the 7 closest instances in the profile library created. This was simply achieved by first calculating the Euclidean distance between the 18-dimensional query profile and each of the corresponding library profiles and then sorting the resulting distances in ascending order.

RESULTS

In total, metabolic changes, as induced by each of 26 approved or investigational drugs and covering the following 12 major pharmacological classes, were measured and analyzed:

- tubulin modulators,
- steroidal estrogen agonists,
- phytoestrogens,
- a selective estrogen response modulator,
- anthelminths,
- antinematodicums,
- proteasome inhibitors,
- a suppressor of the phosphoinositide 3-kinase (PI3)/mammalian target of rapamycin (mTOR),
- histone diacetylase inhibitors,
- nonsteroidal anti-inflammatroy drugs (NSAIDs),
- corticosteroids, and
- tyrosine kinase inhibitors

(See Table S1 in the Supporting Information.) Thus, this set of compounds covers a broad spectrum of bioactivity, such as estrogen interference, suppression of inflammation, receptor tyrosine kinase inhibition, and chromosome modification.

Each of the 26 substances was applied to microcultures of HCT116 cells, grown in a culture medium. Subsequent to 24 h of exposure to any one of the set of drugs, cells were harvested and processed to enrich polar metabolites, which are the most abundant class and generally perceived as a priority metabolite fraction for bioanalytical purposes. The selection of the time point of 24 h was based on some previous exploratory analyses and motivated by the fact that the cells should be perturbed long enough for maximum cellular processes to be triggered. Subsequent to a series of regular preprocessing steps at the instrument level (phase correction, baseline correction, binning), the NMR spectra obtained accordingly were normalized (for more details, see theMaterials and Methods section). Thereafter, spectral changes associated with each drug measurement were obtained by subtracting a corresponding average control spectrum obtained for each of five different

batches. Finally, only those 18 bins found to have a CV value of <25%, according to replicates of one drug (Vinblastine) being used across all five batches, were retained for downstream analysis. For almost all 26 substances, two independent wells containing HCT116 cells were treated with the same drug. Key steps of the experimental outline are summarized in Figure 1 (for more details, see theMaterials and Methods section).

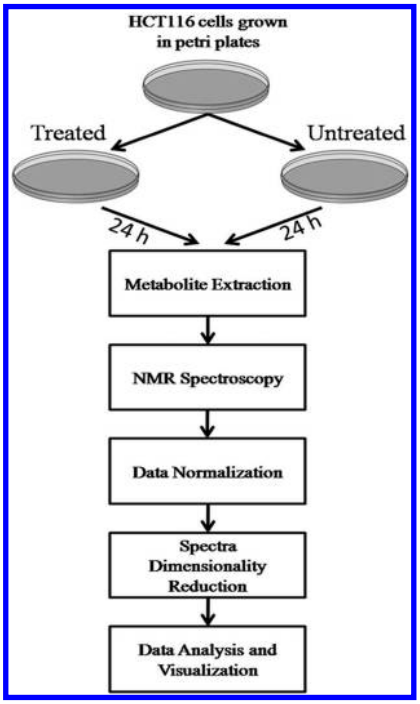


Figure 1. Illustrated outline of the proof-of-concept study. (a) The cell line HCT116 was exposed to one drug at a time for 24 h. (b) Treated cells and control cell samples were subject to a fast and straightforward metabolite extraction protocol. (c) NMR spectra were obtained, followed by basic preprocessing steps (phase correction, baseline correction, binning, normalization). (d) For each drug metabolite profile, a differential profile was obtained by subtracting with the average of all the profiles measured for DMSO controls in the same batch. (e) Using spectra for the drug Vinblastine, which was used in all five experimental batches analyzed, the spectral profiles were reduced into 18-dimensional profiles incorporating only a subset of bins with sufficiently low experimental variability across all 5 batches used. (f) Visualization and nearest-neighbor analysis of the resulting profiles, to determine how well the profiles group, according to the known pharmacological classes involved.

As an introductory multidimensional computational analysis of the preprocessed 18-dimensional NMR spectra, dimensionality reduction into a 3-D space was conducted by conventional PCA. The result, visualized in the top portion of Figure 2, reveals several distinct groups of drugs, as highlighted by semitransparent colored ellipses. Each ellipse harbors a coherent pharmacological/mechanistic class, except for one large and heterogeneous group of agents sitting close to a set of sterile water controls (gray). Tubulin-modifying (light brown) and estrogen-interfering (yellow) compounds are perhaps the most obvious subclusters of the chart, although each class occupies a relatively large space. Tubulin modulators (TUB) are coherently located, but also include two instances of one anthelmintic drug (ANTH). As can be identified from the Supporting Information (Figure S2), this antihelmintic drug is mebendazole, which is known to intercept tubulin polymerization. The estrogen-interfering compounds (yellow) encompass three functional subclasses: an archetype steroidal estrogen agonist (EST), phytoestrogens (PHY), and a selective estrogen response modulator (SER). Furthermore, four smaller aggregations are seen: anthelminths (one of them partly overlapping the tubulin modulator space as already mentioned) (colored red); each of a single antinematodicum (ANTN), together with a subset of tyrosine kinase inhibitors (TYR1), forming a narrow elongated cluster (colored green); each of a histone diacetylase (HDAC) inhibitor (colored light orange); and each of a proteasome inhibitor (PRO) (colored light violet); and, finally each of a suppressor of phosphoinositide 3-kinase (PI3)/mammalian target of rapamycin (mTOR) (colored light blue). In addition, there is also a large cluster (colored gray) containing all sterile water controls (CON). This cluster also houses several instances of various diverse pharmacological classes: some from the TYR1 group, each of another histone diacetylase (HDAC) inhibitor, all NSAIDs (NSA), and all corticosteroids (COR). Thus, this set of agents in the gray cluster seems to have induced insufficiently strong changes to appear as discriminative NMR spectra.

We also employed a semisupervised computational data processing technique, designed to select the optimal subset of three bins and maximizing a class separability score, using within-class and between-class scatter matrices as defined in the Materials and Methods section. The search was comprehensive and thus covered all possible subsets of three bins. The best triplet found was (221, 240 and 276), as depicted in a 3-D scatter plot shown at the bottom of Figure 2. Metabolite interpretations of the triplet found was performed as explained in the Materials and Methods section. Bin 276 did not lend itself to our construal, whereas bin 221 might correspond to L-leucine or iso-leucine, and bin 240 seems to correspond to 3-hydroxybutyric acid.

The resulting subclusters are strikingly similar to those obtained by compression with PCA shown in the top panel of Figure 2. The main difference is the emergence of a new subcluster, corresponding to the NSAID (NSA) subgroup of drugs, as a relatively distinct aggregation (dark violet) at the outskirts of the control cluster (gray). Another obvious change is the appearance of a coherent (green) cluster of the upper panel as two formations, one containing only TYP1 observations and the other only ANTIN observations. In summary, the bottom panel of Figure 2 shows that slightly enhanced separability between the pharmacological classes can be achieved by incorporating only three bins, rather than all the 18 bins with a CV value of <25%. Notably, since this feature (bin) selection was derived from all spectra available, and their known class labels, the observation that only a few bins (metabolites) provide roughly the same, actually even better, class separation relative to that based on all 18 bins should not be overinterpreted. Therefore, more independent data must be collected to validate this preliminary conclusion.

Finally, we also used a nearest-neighbor search procedure in which the NMR profile for each measurement was used as a query to find the 7 closest (Euclidean distance) neighbors in the 18-dimensional bin-space that corresponds to the upper panel of Figure 2. The resulting table of hits (see the Supporting Information) was analyzed to determine, in greater detail, how well the different pharmacological classes were recovered. In conclusion, and as can be predicted from Figure 2, for most queries, the closest member was either the corresponding replicate recording or that of the same overall

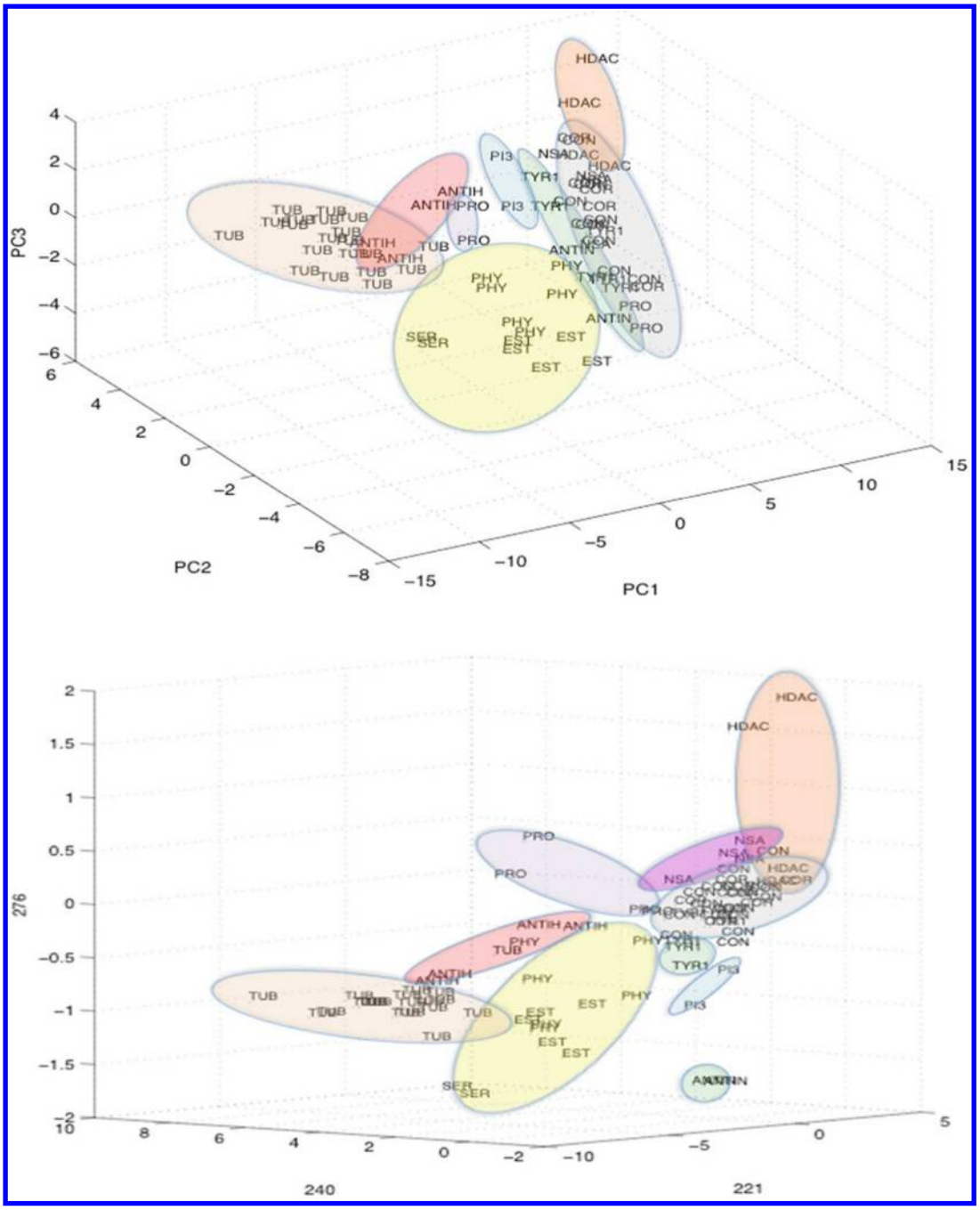


Figure 2. (Top) Visualization of the geometrical relationships between the 18-dimensional differential metabolite profiles after PCA-based compression into the 3-dimensional (3-D) space of maximal variance. Among the totally 12 pharmacological classes studied, a few distinct groups of drugs are revealed. [Legend: light brown, tubulin modulators (TUB); yellow, estrogen modulators (EST, PHY, SER); red, antihelminthics (ANTH); green, antinematodicums (ANTN) and typosine kinase receptor modulators (TYR1); light violet, proteasome inhibitor (PRO); light blue, PI3K modulator (PI3); and light orange, HDAC inhibitor (HDAC).] There is also one large group of substances (colored gray) sitting close to the sterile water controls (CON) consisting of all samples from the two subgroups NSA and COR with some of the HDAC and TYR1 observations also included. (Bottom) Visualization of the geometrical relationships between the 3-D differential metabolite patterns found by means of a comprehensive feature selection procedure where the bin triplet (221, 240, 276) was found to be best at discriminating between the 12 pharmacological classes extended with two control classes, sterile water controls (CON) and pooled quality control samples (not shown). [The same color code and notation is used here as in the panel above.] Here, the original cluster of TYR1 and ANTN observations (colored green) has separated into two isolated clusters and the NSA group (colored dark violet) has also formed a distinct cluster located very close to the control cluster (colored gray).

pharmacological class. Moreover, among the remaining 6 closest neighbors, 2 or 3 of them typically belong to the same category, strongly supporting the notion that it would be possible to infer the correct pharmacological class in most cases, even in the absence of replicate instances. As also inferred from Figure 2, in a few cases (the metabolite profiles sitting inside or close to the gray region), some of the closest neighbors were sterile water control samples (CON). In these

dx.doi.org/10.1021/ci500502f | J. Chem. Inf. Model. 2014, 54, 3251−3258

cases, of course, one cannot rule out the possibility that the query profiles reflect very limited effects, at least in the spectral band studied.

Taken together, the results presented here suggest that inexpensive and fast metabolic profiling, although not offering complete separation into all the pharmacological classes involved in the study presented here, is indeed applicable to pharmacological (mechanistic) classification of drugs already when based on such a small number of spectral bins. Our findings, as well as those reported in a recent pilot study referred to in the Introduction,[25] also suggest that there is strong promise for significant refinement via a few realizable experimental improvements, notably reduced experimental variability (only 18 bins showed a CV value of <25%), higher sensitivity and complementary metabolite profiling of the nonpolar fraction of the cell lysates studied.

DISCUSSION

Metabolomics, generally perceived as the science devoted to the global analysis of all (or as many as possible) metabolites in a sample, undergoes fast progress alongside several directions and recently has gained recognition as a key discipline of systems biology. Moreover, the proximity of metabolites to organism or cellular phenotype is clearly an outstanding merit of metabolomics, relative to other major profiling technologies. Similar to genomics and proteomics, an array of database resources have emerged in support of metabolite identification and annotation, as well as broad data interpretation, including metabolite network patterns, and a high-quality reconstruction of human metabolism is underway.[30−32] Thus, the realm of omics-science has gradually expanded beyond transcriptomics and proteomics to embrace metabolomics as a viable profiling format. Yet, a major focus is laid on biofluids, but *in vitro* cultured cells have gradually attracted mounted interest among researchers.[33,34] Such settings have revealed several pertinent drug-response patterns, including subtle dissimilarities between those of similar structure.[25,35−38] However, to the best of our knowledge, application of metabolomics to broad drug bioactivity mapping, as introduced here, following the CMap concept, has not yet been reported.

Unsupervised (PCA) data analysis of drug-stimulated HCT116 cells, as determined by ^{1}H NMR spectral analysis, disclosed several well-segregated drug clusters, each housing members with shared pharmacology activity. Because of weak or otherwise nondiscriminatory proton NMR signals, some agents appear in a rather wide cluster (gray) encompassing the sterile water control samples. The most prominent drug segregation patterns include those of estrogen mimetics and other compounds interacting with this signaling pathway (EST, PHY, and SER), tubulin modifiers (TUB), histone deacetylate inhibitors (HDAC), a PI3K/mTOR inhibitor (PI3), and one among the pair of the proteasome inhibitors (PRO), see Figure 2 (top). (Individual drug identities are not shown in the figure, but are available in the Supporting Information.) Notably, the TUB aggregation incorporates, apart from the classical tubulin targeting drugs, one among the two anthelminths (mebendazole) of the dataset, whereas the remaining class member (albendazole) is located outside but is in close proximity to the main TUB cluster. Interestingly, both anthelminths are supposed to interfere with tubulin polymerization and data from the NCI 60 cell line panel (http://dtp.nci.nih.gov/) suggest differences in their respective mechanism of action. This was recently also experimentally confirmed in a screening campaign.[5] Furthermore, the cluster of drugs modulating estrogen signaling occupies a relatively large space. Of some interest in this context, one should note that the bioactivity of tamoxifen (an archetypal SERM and the only representative of this class in the dataset), located highly isolated in the chart, varies with tissue. For example, it exerts estrogen antagonist activity in the mammary gland, but is an agonist in the endometrium and in bone tissue.[39] Furthermore, the histone deacylators (HDAC) vorinostat and trichostatin A fall into a cluster at an appreciable distance from the tubulin modulators, in compliance with mutually distinct activities by these two drug classes on chromosomal proteins.

Also according to the PCA results in Figure 2 (top), drugs of some pharmacological classes were seemingly only partially resolved by ^{1}H NMR spectral analysis. This is manifested by the appearance of one member of each category either on the border of or within the sterile water control samples (gray). First, bortezomib, which is an archetype proteasome targeting drug, occupies a small distinctive cluster (indigo area), whereas the other member of the PRO class, thiostrepton, sits in a peripheral location of the sterile water control cluster (CON). Notably, apart from proteasome inhibition, this compound also features antibactericidal and general protein synthesis inhibition activities. Second, in the set of two HDAC inhibitors vorinostat and trichostatin A (orange area), the latter drug sits at a pheripheral location in the control cluster, while vorinostat clearly stands out. Third, a group consisting of the two ANTN instances and some of the TYR1 instances (green) formed a joint elongated cluster also located in the outskirts of the control cluster. Fourth, the control cluster contains all instances of the NSAIDSs (NSA) and the corticosteroids (COR) as well as some instances of the TYP1 category. As shown in the bottom part of Figure 2, after careful selection of the three most discriminative bins the resolution is slightly improved with the NSAIDs becoming an isolated new cluster and the ANTH +TYP1 cluster being separated into two natural aggregations.

Although the proof-of-concept study reported here shows promising results, the experimental setup used can be improved in various ways. First of all, the incorporation of the nonpolar fraction, as extracted from each sample, would theoretically extend significantly on the current information content. Actually, a two-part metabolite extraction protocol, followed by sample merge and subsequent ^{1}H NMR data acquisition, has been reported in a successful profiling of acute myeloid leukemia cells.[35] Moreover, a systematic analysis aiming at identifying and eliminating the main factors causing the high experimental variability presently observed (only 18 bins having a CV value of <25%) is also expected to enhance resolution significantly. NMR spectra of cellular metabolites typically include large regions with low signal-to-noise ratio, because of relatively low sensitivity inherent to the signal capture technique (signals from metabolites in the high micromolar (μM) to millimolar (mM) range). Therefore, these regions were excluded, because of the inherent randomness of such signals. The removed bins may nonetheless contain useful information, but it is our notion that an important strength lies in the fact that we were able to discriminate between drug classes even after removing those bins. Additional directions toward higher resolution of cell-based drug profiling may involve NMR spectrometry by means of an even more sensitive instrument design, e.g., 800 MHz rather than 600 MHz, as in the current study. Furthermore, replacement of the signal acquisition protocol applied to samples investigated here with

any one among those more-sophisticated schemes reported recently holds similar potential for refinement with impact on drug profiling patterns.[40−42] Lastly, transition to a liquid chromatography mass spectrometry (LC-MS) platform would principally take measurements of ∼2 orders of magnitude higher sensitivity, which is arguably beneficial to the identification of subtle signal levels beyond detection by ^{1}H NMR.[43] Nonetheless, the NMR format holds certain obvious strengths, notably high reproducibility, adaptation to high-throughput automation, and still with potential for sensitivity enhancement.

Finally, one should note that the NMR format explored here is at least 1 order of magnitude less expensive and faster than currently available standard microarray-based mRNA gene expression analysis. The cost for consumables (no costs for instrument, bench fee, laboratory technician, etc.) per sample, including everything from cell harvesting to NMR data acquisition, is approximately $5 and the procedure takes only 44 min (see Table S2 in the Supporting Information), whereas the corresponding cost for running a mRNA gene expression analysis is approximately $350 and this procedure will require several hours of work, extending over 1−3 days.

CONCLUSION

The experimental results show that NMR-spectroscopy-based metabolic profiling of drug-induced changes in vitro can discriminate between multiple pharmacological classes. Since there is still many possibilities for improvements of our current assay for NMR-spectroscopy-based *in vitro* metabolomics, the reasonable drug agent classification achieved in this proof-of-concept study is very encouraging. Most of the 26 agents profiled were nicely segregating into their respective relevant pharmacological classes, tubulin inhibitors, estrogen-interfering substances being the most salient examples. However, some compounds did not produce NMR spectra that were much different from those derived from sterile water control samples. Considering the fact that only 18 bins showed sufficiently high signal-to-noise ratio to be reliable (CV < 25%), crisp discrepancies between any set of pharmacological classes, as well as between them and the control wells, cannot be expected. However, several refinements of our basic assay, including, for example, a more robust experimental protocol and more sensitive instruments/protocols, this type of metabolic profiling of drug effects analogous with the CMap concept holds strong promise as an alternative to as-yet relatively costly and slow mRNA profiling.

ASSOCIATED CONTENT

Supporting Information

This file includes the following: (1) a list of 26 compounds and their pharmacological classes (Table S1); (2) detailed results of the exploratory principal component analysis (PCA) shown in the top portion of Figure 2; (3) detailed results of the feature extract shown in the bottom portion of Figure 2; (4) the results of a nearest-neighbor search procedure in which the NMR profile for each measurement was used as a query to find the seven closest (Euclidean distance) neighbors in the 18-dimensional bin-space; and (5) estimated cost per sample from cell harvesting through data acquisition and data preprocessing. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: Ulf.Hammerling@medsci.uu.se.
*E-mail: mats.gustafsson@medsci.uu.se.

Present Address

[§]Cancer Pharmacology and Computational Medicine, Department of Medical Sciences Academic Hospital, Uppsala University, Entrance 61, Third Floor, 751 85 Uppsala, Sweden.

Author Contributions

U.H. and M.G. conceived the investigational outline and supervised the project. O.A. planned and performed the drug exposure experiments and the cell harvesting steps. M.E., J.H., and A.E. performed the metabolite extraction steps as well as the NMR experiments and the preprocessing of the raw data files. O.A. and M.G. performed the data analysis steps followed after the initial raw data preprocessing. M.E., J.H., A.E., C.P., and T.A. provided expertise input regarding the NMR experiments and subsequent interpretations. O.A., U.H., and M.G. drafted the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Higher Education Commission of Pakistan (O.A.).

ABBREVIATIONS

TEM = time-evolving morphology
CMap = connectivity map
NMR = nuclear magnetic resonance
MS = mass spectrometry
PCA = principal component analysis
CV = coefficient of variation
mTOR = mammalian target of rapamycin
DMSO = dimethyl sulfoxide
DSS = 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt
QC = quality control
D_2O = deuterium oxide
NSAIDs = nonsteroidal anti-inflammatroy drugs

REFERENCES

(1) Lamb, J. The Connectivity Map: A new tool for biomedical research. *Nat. Rev. Cancer* **2007**, *7*, 54−60.
(2) Lamb, J.; Crawford, E. D.; Peck, D.; Modell, J. W.; Blat, I. C.; Wrobel, M. J.; et al. The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* **2006**, *313*, 1929−1935.
(3) Iorio, F.; Bosotti, R.; Scacheri, E.; Belcastro, V.; Mithbaokar, P.; Ferriero, R.; et al. Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 14621−14626.
(4) Dudley, J. T.; Deshpande, T.; Butte, A. J. Exploiting drug−disease relationships for computational drug repositioning. *Brief Bioinform.* **2011**, *12*, 303−311.
(5) Nygren, P.; Fryknas, M.; Agerup, B.; Larsson, R. Repositioning of the anthelmintic drug mebendazole for the treatment for colon cancer. *J. Cancer. Res. Clin. Oncol.* **2013**, *139*, 2133−2140.
(6) Fryknas, M.; Gullbo, J.; Wang, X.; Rickardson, L.; Jarvius, M.; Wickstrom, M.; et al. Screening for phenotype selective activity in multidrug resistant cells identifies a novel tubulin active agent insensitive to common forms of cancer drug resistance. *BMC Cancer* **2013**, *13*, 374.
(7) Jarvius, M.; Fryknas, M.; D'Arcy, P.; Sun, C.; Rickardson, L.; Gullbo, J.; et al. Piperlongumine induces inhibition of the ubiquitin−

proteasome system in cancer cells. *Biochem. Biophys. Res. Commun.* **2013**, *431*, 117−123.
(8) Hassan, S.; Laryea, D.; Mahteme, H.; Felth, J.; Fryknas, M.; Fayad, W.; et al. Novel activity of acriflavine against colorectal cancer tumor cells. *Cancer Sci.* **2011**, *102*, 2206−2213.
(9) D'Arcy, P.; Brnjic, S.; Olofsson, M. H.; Fryknas, M.; Lindsten, K.; De Cesare, M.; et al. Inhibition of proteasome deubiquitinating activity as a new cancer therapy. *Nat. Med.* **2011**, *17*, 1636−1640.
(10) Gullbo, J.; Fryknas, M.; Rickardson, L.; Darcy, P.; Hagg, M.; Wickstrom, M.; et al. Phenotype-based drug screening in primary ovarian carcinoma cultures identifies intracellular iron depletion as a promising strategy for cancer treatment. *Biochem. Pharmacol.* **2011**, *82*, 139−147.
(11) Karlsson, H.; Fryknas, M.; Larsson, R.; Nygren, P. Loss of cancer drug activity in colon cancer HCT-116 cells during spheroid formation in a new 3-D spheroid cell culture system. *Exp. Cell Res.* **2012**, *318*, 1577−1585.
(12) Zhang, X.; Fryknas, M.; Hernlund, E.; Fayad, W.; De Milito, A.; Olofsson, M. H.; et al. Induction of mitochondrial dysfunction as a strategy for targeting tumour cells in metabolically compromised microenvironments. *Nat. Commun.* **2014**, *5*, 3295.
(13) Andersson, C. R.; Fryknas, M.; Rickardson, L.; Larsson, R.; Isaksson, A.; Gustafsson, M. G. In vitro drug sensitivity-gene expression correlations involve a tissue of origin dependency. *J. Chem. Inf. Model.* **2007**, *47*, 239−248.
(14) Fryknas, M.; Rickardson, L.; Wickstrom, M.; Dhar, S.; Lovborg, H.; Gullbo, J.; et al. Phenotype-based screening of mechanistically annotated compounds in combination with gene expression and pathway analysis identifies candidate drug targets in a human squamous carcinoma cell model. *J. Biomol. Screen* **2006**, *11*, 457−468.
(15) Koltai, H.; Weingarten-Baror, C. Specificity of DNA microarray hybridization: characterization, effectors and approaches for data correction. *Nucleic Acids Res.* **2008**, *36*, 2395−2405.
(16) Zhang, A. H.; Sun, H.; Qiu, S.; Wang, X. J. NMR-based metabolomics coupled with pattern recognition methods in biomarker discovery and disease diagnosis. *Magn. Reson. Chem.* **2013**, *51*, 549−556.
(17) Kell, D. B.; Goodacre, R. Metabolomics and systems pharmacology: why and how to model the human metabolic network for drug discovery. *Drug Discov Today* **2014**, *19*, 171−182.
(18) Ma, Y.; Zhang, P.; Yang, Y.; Wang, F.; Qin, H. Metabolomics in the fields of oncology: A review of recent research. *Mol. Biol. Rep.* **2012**, *39*, 7505−7511.
(19) Fernie, A. R.; Trethewey, R. N.; Krotzky, A. J.; Willmitzer, L. Metabolite profiling: from diagnostics to systems biology. *Nat. Rev. Mol. Cell. Biol.* **2004**, *5*, 763−769.
(20) Clayton, T. A.; Lindon, J. C.; Cloarec, O.; Antti, H.; Charuel, C.; Hanton, G.; et al. Pharmaco-metabonomic phenotyping and personalized drug treatment. *Nature* **2006**, *440*, 1073−1077.
(21) Gupta, S.; Chawla, K. Oncometabolomics in cancer research. *Expert Rev. Proteomics* **2013**, *10*, 325−336.
(22) Robertson, D. G.; Frevert, U. Metabolomics in drug discovery and development. *Clin. Pharmacol. Ther.* **2013**, *94*, 559−561.
(23) Keun, H. C.; Athersuch, T. J. Application of metabonomics in drug development. *Pharmacogenomics* **2007**, *8*, 731−741.
(24) Zhang, A.; Sun, H.; Wang, P.; Han, Y.; Wang, X. Modern analytical techniques in metabolomics analysis. *Analyst* **2012**, *137*, 293−300.
(25) Tiziani, S.; Kang, Y.; Choi, J. S.; Roberts, W.; Paternostro, G. Metabolomic high-content nuclear magnetic resonance-based drug screening of a kinase inhibitor library. *Nat. Commun.* **2011**, *2*, 545.
(26) Teng, Q.; Huang, W.; Collette, T. W.; Ekman, D. R.; Tan, A. A direct cell quenching method for cell-culture based metabolomics. *Metabolomics* **2009**, *5*, 199−208.
(27) Beckonert, O.; Keun, H. C.; Ebbels, T. M.; Bundy, J.; Holmes, E.; Lindon, J. C.; et al. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat. Protoc.* **2007**, *2*, 2692−2703.
(28) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; et al. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **2007**, *3*, 211−221.
(29) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; et al. HMDB: The Human Metabolome Database. *Nucleic Acids Res.* **2007**, *35*, D521−D526.
(30) Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; et al. HMDB 3.0—The Human Metabolome Database in 2013. *Nucleic Acids Res.* **2013**, *41*, D801−D807.
(31) Ellinger, J. J.; Chylla, R. A.; Ulrich, E. L.; Markley, J. L. Databases and Software for NMR-Based Metabolomics. *Curr. Metabolomics* **2013**, 1.
(32) Thiele, I.; Swainston, N.; Fleming, R. M.; Hoppe, A.; Sahoo, S.; Aurich, M. K.; et al. A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* **2013**, *31*, 419−425.
(33) Cuperlovic-Culf, M.; Barnett, D. A.; Culf, A. S.; Chute, I. Cell culture metabolomics: Applications and future directions. *Drug Discovery Today* **2010**, *15*, 610−621.
(34) Zhang, A.; Sun, H.; Xu, H.; Qiu, S.; Wang, X. Cell metabolomics. *Omics* **2013**, *17*, 495−501.
(35) Tiziani, S.; Lodi, A.; Khanim, F. L.; Viant, M. R.; Bunce, C. M.; Gunther, U. L. Metabolomic profiling of drug responses in acute myeloid leukaemia cell lines. *PLoS One* **2009**, *4*, e4251.
(36) Triba, M. N.; Starzec, A.; Bouchemal, N.; Guenin, E.; Perret, G. Y.; Le Moyec, L. Metabolomic profiling with NMR discriminates between biphosphonate and doxorubicin effects on B16 melanoma cells. *NMR Biomed.* **2010**, *23*, 1009−1016.
(37) Bai, J.; Wang, M. X.; Chowbay, B.; Ching, C. B.; Chen, W. N. Metabolic profiling of HepG2 cells incubated with *S*(−) and *R*(+) enantiomers of anti-coagulating drug warfarin. *Metabolomics* **2011**, *7*, 353−362.
(38) Massimi, M.; Tomassini, A.; Sciubba, F.; Sobolev, A. P.; Devirgiliis, L. C.; Miccheli, A. Effects of resveratrol on HepG2 cells as revealed by (1)H-NMR based metabolic profiling. *Biochim. Biophys. Acta* **2012**, *1820*, 1−8.
(39) Pinkerton, J. V.; Thomas, S. Use of SERMs for treatment in postmenopausal women. *J. Steroid. Biochem. Mol. Biol.* **2014**, *142*, 142−154.
(40) Furihata, K.; Tashiro, M. BASHD-J-resolved-COSY: A new method for measuring proton−proton spin coupling constants of multiplet signals. *Magn. Reson. Chem.* **2012**, *50*, 713−716.
(41) Furihata, K.; Tashiro, M. Selective COSY-J-resolved-HMBC, a new method for improving sensitivity of cross peaks of methine proton signals attached to a methyl group. *Magn. Reson. Chem.* **2012**, *50*, 409−414.
(42) Lupulescu, A.; Aharon, H.; Frydman, L. Two-dimensional RF pulses: A new approach to selectively exciting J-coupled spins in nuclear magnetic resonance. *J. Chem. Phys.* **2012**, *139*, 144204.
(43) Wang, X.; Sun, H.; Zhang, A.; Wang, P.; Han, Y. Ultra-performance liquid chromatography coupled to mass spectrometry as a sensitive and powerful technology for metabolomic studies. *J. Sep. Sci.* **2011**, *34*, 3451−3459.