This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. This Accepted Manuscript will be replaced by the edited, formatted and paginated article as soon as this is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
Ultracentrifugation-Based Multi-Target Affinity Selection Mass Spectrometry

Y. Jin, X. H. Cheng, F. Z. Yang, and L. Fu*

We have developed a new approach of affinity selection based on the strategy of rate-zonal density gradient centrifugation combined with the application of ultra-performance liquid chromatography coupled quadrupole time-of-flight mass spectrometer. In this method, a discontinuous gradient of sucrose solution is used as the centrifugal medium, and the mixed proteins and compounds are laid on the top of it; an applied centrifugal driving force is then used to separate the mixed proteins and their respective ligands in the gradient. Ligand binding ability is defined by comparing the concentration distribution of compounds with the respective concentration distribution of targets after centrifugation once the solution is fractionated. Ideally, a specific ligand would essentially distribute identically with its targeted proteins. This method could be used to screen multiple targets simultaneously, and it would be especially helpful to screen multi-target directed ligands that can interact with multiple targets for the specific pathogenesis.

Introduction

In the past decades, a number of fairly efficient fast affinity selection-mass spectrometry (AS-MS) approaches, such as pulsed-ultrafiltration-MS, frontal affinity chromatography-MS, and size exclusion chromatography, have been developed for the discovery of biologically active compounds from combinatorial libraries and natural product extracts. Unlike the screening approaches combined with UV, or fluorescence, or radioactivity requiring labelled competitors, AS-MS approaches are able to perform screening assays in a label-free mode because the equipped mass spectrometric detection, which has the advantages of high sensitivity and selectivity, relies entirely on the mass-to-charge ratio (m/z) of analytes. Additionally, AS-MS approaches can facilitate the natural product screening programs because it is much more difficult to discover active compounds from natural product extracts by using the traditional isolation-to-bioactivity evaluation or bioassay-guided isolation, which are time-consuming and laborious. Certainly, each AS-MS approach has its own advantages and drawbacks with respect to screening throughput, specialized protein requirements, and specialized library design requirements, for examples. Currently, no one-size-fits-all AS-MS approach is expected to satisfy all the screening demands. It is known that most AS-MS approaches at present are basically designed to satisfy the requirement of screening single targeted ligands. However, ligands that affect a single target might not always affect our complex physiological systems in the desired way, even if they completely change the behaviour of their immediate target. For example, drug discovery in the multifaceted neurodegenerative disorder named Alzheimer’s disease is gradually moving from the development of molecules able to modulate the biological function of a single target to the multi-target-directed-ligands (MTDLs). In term of the MTDLs, recent researches raise the perception that MTDLs cannot be rationally designed but rather discovered accidentally by screening approach. In this case, it is our aim to develop an AS-MS approach that can be a valuable addition to traditional drug discovery methods and especially can satisfy the requirement of screening MTDLs.

Harlan and his colleagues developed an AS-MS approach to screen ligands of macromolecules depending on the differential centrifugation. This method they described, which takes advantage of the different sedimentation rates of molecules of differing molecular weight (S20, w) in a centrifugal force field, that can be performed using virtually any standard laboratory ultracentrifuge. It can be applied to almost any system of interacting molecules showing a sufficient differential in sedimentation behaviour. It is reported that tight binding ligands would essentially be distributed identically to the protein target after sufficient centrifugation while weaker or non-binder would not. Although this
differential centrifugation-based AS-MS method can only apply to screen ligands of a single target, it provides a potential way for screening MTDLs based on centrifugal force. As it is known, rate-zonal density gradient centrifugation is another well-established strategy of centrifugation depending on different sedimentation rates of subcellular particles in liquid density gradients for the high-resolution separation \(^9\). In the studies present here, we describe a new AS-MS approach, especially for the use of rapidly screening of MTDLs based on the strategy of rate-zonal density gradient centrifugation. A schematic diagram of this approach is shown in Fig. 1. The incubated mixture of proteins and compounds is well mixed and laid on the top of the discontinuous gradient of sucrose solution, which is a widely used centrifugal medium. Subsequently, the specific ligands are redistributed separately in the sucrose solution with their target proteins by centrifugation, respectively. After that, the samples are divided into 5 fractions, and the compounds in each fraction are released from the sucrose solution and analysed by ultra-performance liquid chromatography coupled quadrupole time-of-flight mass spectrometer (UPLC-Q-TOF-MS). In this study, we used tyrosine phosphatase 1B (PTP1B, 37.5 kDa) and low-density lipoprotein receptor (LDLR, 11.5 kDa) to observe the distribution of single proteins in the sucrose density gradient solution after centrifugation. And bovine serum albumin (BSA, 66.4 kDa) was used as the model of interacted proteins to observe the redistribution after centrifugation. It was supposed that BSA would deposit to the lower part of the density gradient solution than PTP1B while LDLR would redistribute in the most top part of the solution because of their different sedimentation coefficients, namely \(S_{20,w}\). Additionally, the respective binding ligands of each protein should show similar redistribution to their targets after centrifugation. The compounds, which were found to show identical redistribution of BSA after centrifugation, represented the ligands that would interact with multiple targets. The compounds, which were redistributed identically to either PTP1B or LDLR after centrifugation, represented the ligands that would only modulate a single target. In this study, we limited representative compound tools to be either tight binding ligands or non-binder for the development and validation of this ultracentrifugation-based multi-target affinity selection mass spectrometry method. However, this method can also be applied to sort binding events like the method Harlan et al. introduced \(^8\).

Materials and Methods

Materials and Instrument Information

PTP1B and LDLR were purified and provided by Viva Biotech Ltd. (Shanghai, China). Heat shock isolated BSA (pH 7.0) was purchased from Sangon Biotech Corp. (Shanghai, China). Warfarin analytical standard and sucrose (\( \geq 99.5\% \)) were purchased from Sigma-Aldrich (St. Louis, MO, USA). The PTP1B inhibitor named FL \(^{10}\) and a negative control (((S)-3-(2-(2-bromo-4,5-difluorophenyl)benzo[d]oxazol-5-yl)-2-((1-ethoxyvinyl)amino)propanoic acid) named NC were provided by Fu Lab in the School of Pharmacy, Shanghai Jiao Tong University (Shanghai, China). 45 compounds for screening were provided by Viva Biotech Ltd. (Shanghai, China). The Bradford protein assay kit, which contained 1 mL of 5-mg/mL BSA standard solution and 200 mL Coomassie brilliant blue G-250 solution, was purchased from Beytime Biotech (Shanghai, China).

Fig. 1 Scheme of the ultracentrifugation-based simultaneous multi-target affinity selection experiment. A discontinuous gradient of sucrose solution is prepared. The mixed proteins and compounds are laid on top of the sucrose solution. The tube is centrifuged after incubation. The solution in the centrifugal tube is fractionated with equal volumes into 5 fractions from the top to the bottom of the tube, and the released compounds within each fraction are analyzed by MS. The molecular weight of the proteins is assumed to be B>A>C, and the sedimentation coefficient difference of each two proteins is > 30%.
Water was purified in-house using a Milli-Q water purification system (Milford, MA). Dissolving buffer (10 mM Tris at pH 7.5, 25 mM NaCl, 1 mM EDTA), all other chemicals and solvents otherwise not mentioned were provided by Viva Biotech Ltd. (Shanghai, China).

Sample Preparation

Standard compounds including warfarin, FL and NC were respectively made as 40 mM solutions in DMSO and diluted into either protein(s) or buffer solutions so that the final concentration of DMSO was < 10% (vol/vol). 45 compounds with no monoisotopic mass redundancy for screening, ranging in monoisotopic mass from 169.0972 to 344.1427 (m/z) were prepared by starting with single compounds dissolved to 40 mM in DMSO. Successive steps of mixing resulted in stocks of 800 μM each compound for dilution in screening samples. We restricted the total concentration of compounds to be lower than each individual protein. This restriction afforded each ligand equal access to binding site(s) on the target protein(s).

Each protein stock was centrifuged by a MicroCL 17 microcentrifuge (Thermo Scientific™, Rockford, IL, USA) at room temperature for 10 min at 12,000 rpm to exclude possible hybrid proteins before affinity screening.

Centrifugation-based affinity selection

Concentrations of 5%, 10%, 15%, 20%, 25%, 30%, 35% and 40% (W/V) sucrose solutions were initially diluted with dH2O to prepare a discontinuous sucrose density gradient. In a 0.7× 2.1 (φ × L cm) centrifuge tube (Hitachi Koki Co. Ltd., Japan), the needle of a 100-μL gastight syringe (Hamilton Co., Reno, Nevada) was carefully inserted into the bottom of the tube to inject 25 μL of the 5% sucrose solution. Once the 5% sucrose solution had been placed in the tube, 25 μL of the 10% sucrose solution could be injected into the bottom of the tube. The procedure was continued with the injection of 15%, 20%, 25%, 30%, 35% and 40% sucrose solution, respectively. A total of 10 μL solution containing well mixed 1 μM per compound and 50 μM per protein was carefully placed on the top of the sucrose medium by pipetting after 20 min of equilibrium. After 30 min of incubation at room temperature, the samples in the respective centrifuge tubes were centrifuged using a CX-150FX micro-ultracentrifuge (Hitachi Koki Co. Ltd., Japan) with a fixed angle rotor at 90,000 rpm. A control experiment with no protein but compounds in buffer laid directly on the top of the discontinuous sucrose medium was also performed for each screening experiment.

Fractionation

The samples after centrifugation were fractionated manually into 5 equal parts, respectively. For this purpose, a pipette was carefully inserted into the liquid column just below the meniscus. As the fluid volume was removed, the pipette was moved slowly down into the liquid so that it remained close to the meniscus without breaking the surface and allowing air to enter the pipette tip. The collected fractions were sequentially dispensed for MS analysis and/or quantitative determination.

Bradford protein assay

The methodology has been described by Bradford 11. Firstly, 10 μL of the 5 mg/mL BSA standard solution was diluted 10 times with the dissolving buffer. A series of volumes (i.e., 0, 1, 2, 4, 8, 12, 16, 20 μL) of 10X diluted BSA standard was added into individual wells in a 96-well plate. A corresponding series of volumes (i.e., 20, 19, 18, 16, 12, 8, 4, 0 μL) of the dissolving buffer was then added into the respective wells to produce equal volumes of 20 μL. Subsequently, 20 μL of each sample or the dilutions of samples was added to a separate well in the plate. Then, 200 μL of Coomassie brilliant blue G-250 solution was quickly added to each well. The absorbance values of the samples were measured immediately using a Multiskan™ MK3 96-well microplate spectrophotometer (Thermo scientific, USA) at 595 nm after 5 min of incubation. The different quantities of BSA were plotted against the corresponding absorbance, resulting in a standard curve (r^2 > 0.99) used to calculate the quantities of proteins after centrifugation. The quantification of each protein and the standard curve were derived from three independent experiments.

Compound release and analysis

For MS analysis, the standards and the mixtures of 45 compounds must be firstly released from the protein-ligand complexes in the sucrose solution after fractionation. For this purpose, 20 μL solution of each fraction was well mixed with 80 μL methanol (MeOH). Then, 20 μL acetonitrile (ACN) was added and mixed using a SI Vortex-Genie 2T (Scientific Industries, New York, USA). Next, 60 μL dH2O was added and mixed again. Compounds in the organic phase were collected for MS analysis after 10 min of centrifugation using a MicroCL 17 microcentrifuge at 13,300 rpm.

MS Analysis and Data Processing

Ultra-performance liquid chromatography (UPLC) was performed using a Waters ACQUITY UPLC system (waters, Milford, MA, USA) with an ACQUITY UPLC BEH C18 column (2.1 × 50 mm, 1.7 μm) and a HypURITY C6 column (2.1 × 50 mm, 5 μm) for compounds and proteins analysis, respectively. The solvents used were as follows: A, 0.1% diluted aqueous formic acid, and B, 100% ACN. The gradient conditions for compound analysis were as follows: 0 → 1 min, 2→5% B; 1 → 9 min, 5→90% B; and 9 → 10 min, 95% B. The gradient conditions for protein analysis were as follows: 0 → 1 min, 2→10% B; 1 → 2.5 min, 10→90% B; 2.5 → 3 min, 90% B; and 3 → 5 min, 90→10% B. The injection volume was 5 μL, and the column and sample temperatures were maintained at 40 °C and 5 °C, respectively.

Mass spectrometric detection was coupled with UPLC and performed using a Synapt™ quadrupole time-of-flight (Q-TOF) High Definition Mass Spectrometer (Waters, Milford, MA, USA) equipped with an electrospray ionization (ESI) source.
operating in positive ionization mode. The optimized mass spectrometric parameters were as follows: capillary voltage, 2.5 kV; sample cone, 25 V; extraction cone, 4.0 V; source temperature, 120 °C; and desolvation temperature, 400 °C. Nitrogen was used as a desolvation and a cone gas at flow rates of 600 and 50 L/h, respectively. Argon was used as a collision gas. A lock mass of leucine-enkephalin at a concentration of 200 pg/mL in 50% ACN-water solution (including 0.1% formic acid) was employed as the external reference to generate a [M+H]+ ion in positive mode at m/z 556.2771 via a lock spray interface at a flow rate of 5 μL/min to acquire accurate mass during the analysis. The samples of compounds only were scanned in full-scan mode from m/z 80 to 800 in 1 sec scan intervals. The samples of proteins were scanned in full-scan mode from m/z 500 to 2200 in 0.5 sec scan intervals. Masslynx software (Waters, Milford, MA) was used to integrate and visually inspect the peaks of each compound and protein. A potential ligand was estimated by comparing its mass peak heights with the amount of its target proteins in the 5 parallel fractions after centrifugation.

Results and Discussion

Protein Distribution

Fundamentally, the centrifugation time (t), which is also called the clearing time, is calculated from S_{20,w} of the molecules and the clearing factor (k) of the rotor (Eq. 1). The k value of a rotor can be estimated using Eq. 2, where ω is the angular velocity in radians per second (0.105 × rpm), R_{max} is the maximum radius (bottom of tube), and R_{min} is the minimum radius (meniscus). S_{20,w} contains information about the size, shape, and density of the molecules and is estimated from a standard protein of known molecular weight and S_{20,w}^{12}.

\[ t = \frac{k}{S_{20,w}} \quad \text{Eq. 1} \]
\[ k = \left(\frac{\ln(R_{\text{max}}/R_{\text{min}})}{\omega^2}\right) \times \left(10^{13}/3600\right) \quad \text{Eq. 2} \]

However, as the aim of our experiment was to separate the mixed proteins distinctly in the sucrose medium by centrifugal force, in this case, the respective bound ligands would have similar redistribution to their target proteins. Therefore, we had to first determine the most workable time for obviously separating and redistributing the mixed proteins instead of using the calculated centrifugation time.

To our knowledge, proteins can be separated by rate-zonal density gradient centrifugation if the difference of S_{20,w} between every two proteins is higher than 30%9. The approximate S_{20,w} values of BSA, PTP1B and LDLR were 4.67, 3.27 and 1.53, respectively, which indicated that the three proteins were usable as representative protein tools to validate the distinct separation of mixed proteins by rate-zonal density gradient centrifugation. By considering the largest protein, 210 μL BSA, for which the centrifugation time is 100 min using a CX-150FX micro-ultracentrifuge with a fixed angle rotor at 90,000 rpm, while the respective theoretical centrifugation times of 210 μL PTP1B and LDLR are 142 and 304 min, we began with 60 min of centrifugation and measured the distribution of individual proteins and the mixed proteins every 20 min until 220 min. This investigation range was expected to show us the redistribution process of individual proteins and the mixture of them, and to indicate the optimal time for separating the protein mixture.

After 60 min of centrifugation, the relative concentration distribution of individual proteins including BSA, PTP1B and LDLR in fractions after centrifugation was determined by Bradford protein assay, as shown as Fig. 2a. The fractions contained the 3 mixed proteins were analyzed using the integration of peak height from
LC-MS chromatography (Fig. 2b). All the protein samples, including the individual proteins and the mixed proteins, were clearly distributed mostly in the top 3 fractions (i.e., a, b, c) after 60 min centrifugation, although less than 10% of the total BSA content was found in the d fraction according to the Bradford protein assay, and a small amount of PTP1B was also detected in the same fraction by MS. We determined the distributions of individual proteins after centrifugation by Bradford protein assay to ensure that the distribution tendency of the mixed proteins provided by LC-MS chromatograms was not caused by protein interaction. In other words, if any protein in the mixture bind with any other protein in the solution, the molecular weight of the interacted proteins would increase, leading to much faster deposition than expected. After 80 min of centrifugation, the distribution of individual proteins and mixed proteins was altered (Fig. 2c-d). Although the individual LDLR was still concentrated in the top 2 fractions (i.e., a, b), 58.51±0.86% of individual BSA and 62.71±2.56% of individual PTP1B were concentrated in the d and c fractions, respectively. LC-MS chromatogram indicated a similar tendency of protein distribution (Fig. S1). By increasing the centrifugation time to 100 min, LDLR was still most concentrated in the b fraction, although 37.67±3.76% was deposited in the c fraction (Fig. 2e). Additionally, 22% more of BSA was deposited in the e fraction, and PTP1B was still majorly concentrated in the c fraction (Fig. 2e-f). After 140 min, almost no proteins were measured in the top fraction, and the proteins were further concentrated in the d and e fractions after 220 min of centrifugation (Fig. 2g-j). Based on our results, 80 min of centrifugation was most suitable for screening the ligands, especially the multi-target directed ligands, of these mixed proteins because each protein (i.e., BSA, PTP1B and LDLR) was predominantly (> 50%) distributed in a distinct fraction.

Binding behaviour of individual compounds as monitored by MS
Initially, we selected 3 representative compounds to validate the feasibility of the multi-target affinity selection based on rate-zonal density gradient centrifugation. Warfarin ($K_D = 4.0±2.8 \mu M$) is a commonly used anticoagulant that binds with strong affinity to BSA $^{13-15}$. The PTP1B inhibitor (FL) has an IC$_{50}$ value of 2.4 $\mu M$ $^{10}$, while the negative control (NC) showed no binding with BSA, LDLR or PTP1B according to our preliminary experiments. Hence, it was assumed that the redistribution of warfarin by centrifugation would be similar to BSA while FL would exhibit a similar redistribution to PTP1B after an appropriate equilibrium between proteins and compounds. On the other hand, there should be no compound showing a similar distribution to LDLR because none of the three compounds have affinity binding with LDLR according to our preliminary experiments. We mixed the 3 compounds with individual proteins and the mixed proteins, respectively, for an 80 min ultracentrifugation-based affinity selection (Fig. S2). As shown in Fig. 3, 1 $\mu M$ warfarin was mostly concentrated in the c and d fractions of the sample, which treated with 50 $\mu M$ BSA and mixed proteins (50 $\mu M$ each), respectively, after 80 min of centrifugation. Although the concentration distribution of warfarin when treated with either single protein of BSA or the mixed protein was slightly different to the individual protein distribution of BSA (Fig. 2c), it was considered that warfarin exhibited specific affinity binding to BSA because its concentration distribution when treated with single protein of BSA was identical to its concentration distribution when treated with the mixed proteins (Fig. 3a). Additionally, the concentration distribution of warfarin with respect to PTP1B and LDLR showed identical binding behavior to its redistribution in the solution without protein, which also indicated
that warfarin is specific bound to BSA among the three types of proteins. A decreasing gradient of the concentration distribution of warfarin was shown when the warfarin was treated with respective LDLR, PTP1B and buffer. This might be because of diffusion, a spontaneous movement of particles from an area of high concentration to an area of low concentration, which is led by the mechanism of Brownian motion \(^\text{16}\). Second, the concentration distribution of FL in the mixture, when treated with different proteins and subjected to 80 min centrifugation, was similar to our expectations (Fig. 3b). That is, PTP1B was predominantly in the c fraction, where PTP1B was likewise majorly concentrated after 80 min centrifugation (Fig. 2c). Moreover, NC showed a nearly identical decreased concentration distribution in all samples (Fig. 3c). Thus, a distinct separation of each compound was presented with the redistribution of mixed proteins by centrifugation. And it is suggested that the potential ligands of target protein(s) would be apparently deposit with the respective bound target(s) with regard to the molecular weights of protein(s).

An initial study of the binding behavior of compound mixture as monitored by MS

The results discussed above suggest that rate-zonal density gradient centrifugation-based AS-MS method would be able to screen binding ligands of both single target and multi-targets. It was demonstrated that a specific ligand would essentially distribute identically with its targeted protein(s). In the next, we randomly selected 45 synthetic compounds, which were initially identified by comparing the mass spectra with the exact masses of respective compounds, to treat with the mixed proteins for screening the potential ligands to BSA, or LDLR, or PTP1B in a single run. Figure 4 shows the molecular distribution of the 45 compounds treated with the mixed proteins after centrifugation. It was found that 2 compounds showed similar concentration distribution to BSA after centrifugation, 1 compound had a similar tendency to PTP1B, and 1 compound showed a tendency that suggested predominantly binding with LDLR (Fig. 4). By MS detection, we also found a compound that potentially bound to both LDLR and PTP1B, as its distribution tendency was unlike other non-bound compounds showing a significantly decreased gradient of molecular distribution, but displayed a combinational distribution of LDLR and PTP1B. It is considered that such kind of compounds should be non-specific binder because it bound to more than one target but exhibited distinct redistribution to the interacted proteins after centrifugation. The in vitro and in vivo activities of these identified potential ligands to the single target (BSA, PTP1B, or LDLR) should be further confirmed.

Finally, although we majorly aimed to use representative proteins and compounds to validate the feasibility of rate-zonal density gradient centrifugation-based AS-MS approach for the screening of ligands, especially MTDLs, we also speculated that this method can be used to screen the protein, which can be unpurified and would be altered during isolation procedures. Lack of compound solubility will limit the application of this method, all the other solution based AS-MS approaches, NMR and X-ray crystallography ways. In this case, the rate-zonal centrifugation-based AS-MS approach has great potential for screening ligands, especially MTDLs.

Fig. 4 The concentration distribution of compound mixture including 45 compounds treated with mixed LDLR, PTP1B and BSA after 80 min centrifugation. The red line indicates a potential ligand that binds specifically to PTP1B, the light and dark blue lines indicate compounds that bound to BSA, the green line indicates a compound that might bind to LDLR, and the brown line may indicate a compound that can bind to both LDLR and PTP1B. The unit of Y-axis is the integration of MS peak height.

References
A rate-zonal density gradient ultracentrifugation-based affinity selection mass spectrometry approach has been developed for simultaneous multi-targets screen.