Affinity-Selection Mass Spectrometry Screening of Macrocycles

Zane R. Thistleford[1], Srinivasa Karra[2], Kevin J. Lumb[2] and Can C. Ozbal[1]

[1] Momentum Biotechnologies, 3 Federal Street, Suite 300, Billerica, MA 01821 USA

[2] Avilar Therapeutics, 400 Totten Pond Road, Waltham, MA 02451, USA

Abstract

The Affinity Ligand Identification System (ALIS) is a well-established form of affinity-selection mass spectrometry (ASMS) for identifying and characterizing small-molecule binders of proteins and RNA. The ability to pool many test compounds using 2D HPLC-mass spectrometry enables ALIS to be an extremely efficient approach for screening large numbers of compounds to identify binders. Typically, rule of five compliant compounds with molecular weights below 480 g/mol are interrogated as potential ligands in ASMS studies. However, larger compounds are problematic using established protocols. Here we describe methods to address such larger compounds with molecular weights ranging from 450 to 865 g/mol. For typical rule of five compliant test compound libraries, pools of 100-600 test compounds are fractionated with a 50 mm x 2 mm polyhydroxyethyl A size-exclusion chromatography column containing 3 micron beads with 60 Å porosity. With macrocyclic compounds, the 60 Å porosity column resulted in the breakthrough of an unacceptable number of test compounds even in the absence of target protein, resulting in an impractical number of false positives. After experimentation with different columns and compound pooling densities, a 200 Å porosity size-exclusion column run against pools of 50 test compounds was found to essentially eliminate compound breakthrough and false positives. A high-throughput ALIS screen of a 15,200-compound macrocyclic library was performed in pools of 50 compounds at 1 micromolar each (304 pools in total) against a target protein concentration of 5 micromolar which resulted in 72 hits (0.49% hit rate). All hits were tested as singletons in a secondary assay run at 10 micromolar test compound concentration against 2.5 micromolar target protein, resulting in 67 of 75 hits confirming as binders (89% confirmation rate). In summary, optimization of ASMS methods has enabled the affinity screening of macrocycles against a protein target, thus expanding the capability of ASMS to higher molecular weight molecules.

Introduction

ASMS is a well-established biophysical assay platform where one or more small molecule test compounds are mixed with a target of interest the target/test compound mixture is fractionated by size-exclusion chromatography (SEC) and the fraction containing the target is interrogated by HPLC-MS[1]. Identification of one or more test compounds in the target fraction is indicative of a binding event (Figure 1).

Targets interrogated by ASMS can be RNA, proteins and protein/protein or protein/oligonucleotide complexes while typical test compounds consist of traditional drug-like small molecules that are ideally rule-of-five compliant with molecular weights below 480 g/mol. The goal in this project was to screen a diverse set of macrocyclic compounds with molecular weights ranging from 450 to 865 g/mol. Macrocycles with medium- and large-sized rings are considered to be a distinct class of molecules, not only because of their interesting biological properties but also due to the specific features of their synthesis.

The principle of SEC separation depends primarily on the molecular cross-section of the analytes which is dictated by the molecular weight, hydration shell and intrinsic structure of the molecule. Molecules with molecular cross sections small enough to fit through the pores of the SEC resin have longer column residency while those molecules that are too large for the pores simply flow around them and elute much faster. Macrocycles are often composed at least in part of peptides and therefore are generally quite polar, resulting in large hydration shells. The cyclic nature of macrocycles also can limit their flexibility leading to larger molecular cross sections than molecular weight alone would suggest.

Initial experiments using macrocycles with methods and workflows optimized for rule-of-five compliant test compounds resulted in an unacceptable number of test compounds breaking through the SEC column. The macrocycles elute in the target peak even in the absence of target protein and appear as false positives. In this project methods and workflows were optimized to screen a library of macrocycles while eliminating false positives.

ASMS Methods

The ALIS instrumentation consists of an Agilent 1260 HPLC pump for the size-exclusion chromatography coupled to an Agilent 1290 UHPLC pump for the reversed phase (RP) chromatography with a high-pressure switching valve interfaced to an Agilent 6230B Time-of-Flight Mass Spectrometer. All MS acquisitions were done in positive ion acquisition mode. Data analysis was performed with a combination of Agilent Mass Hunter and proprietary custom software.

SEC conditions: Buffer A: 700 mM Ammonium Acetate

Buffer B: 70% Acetonitrile

Column: 50 x 2.1 mm Polyhydroxyethyl A, 3 µm 200 Å porosity (PolyLC)

RP conditions: Buffer A: Water + 0.1% Formic Acid

Buffer B: 90% Acetonitrile + 0.1% Formic Acid

Column: 50 x 2.1 mm Kinetex 2.6 um C18 100 Å (Phenomenex)

ALIS-ASMS Workflow

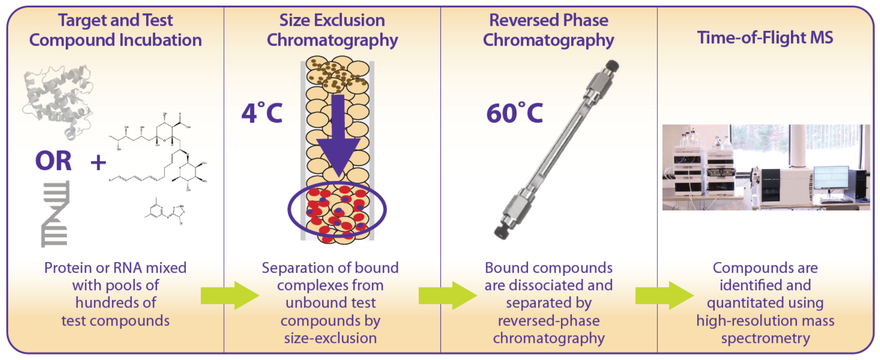

Figure 1: Schematic of the Affinity Ligand Identification System (ALIS). A mixture of one or more small molecule test compounds and a target are fractionated by size-exclusion chromatography. The target fraction is selectively interrogated by HPLC-MS to identify potential ligands

SEC-UV Chromatogram at 280 nm

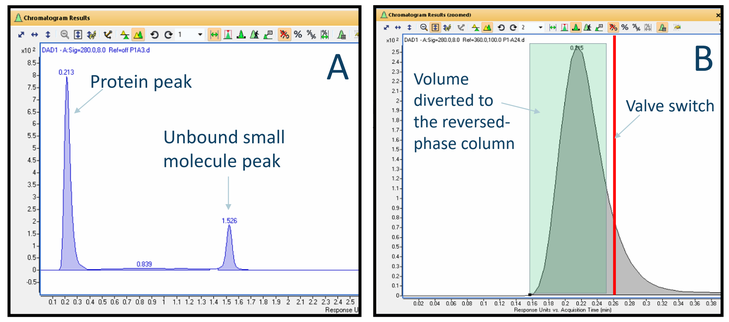

Figure 2: (A) UV chromatogram of the SEC column elute. The early eluting peak at 0.21 min contains the target consisting of molecules that are too large to fit through the pores of the SEC column packing material while the later eluting peak at 1.53 min are small molecules trapped in the SEC column pores and subsequently washed out with 70% acetonitrile. **(B)** Closeup of the protein peak in panel A. A 50 µL high pressure switching valve is used to divert the green shaded portion from the SEC dimension to the HPCL dimension of the system. Any test compounds identified by MS in this region are inferred to be ligands to the target

Isotope Pattern Matching for Hit Identification

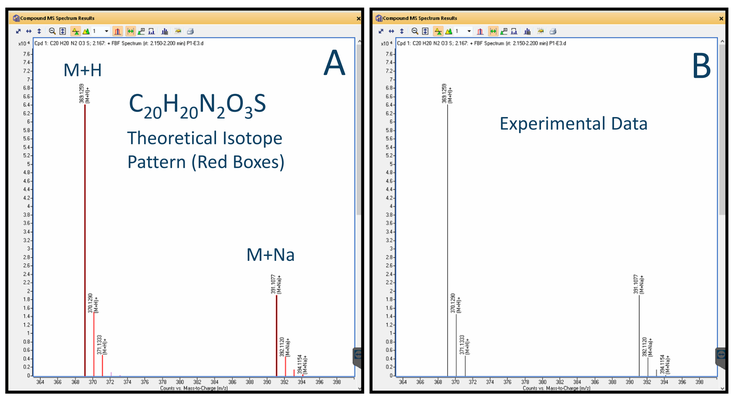

Figure 3: Hit identification is performed by 2 sets of filters. The first filter is an isotope pattern match of the theoretical isotope distribution for the protonated and/or sodiated adducts for the molecular formula of a given test compound (red lines in panel A) with the experimentally determined data (shown in panel B). The match is scored on a scale from 1 (no match) to 100 (perfect match) and any cutoff can be user selected in SW (default is 70).

Method Optimization

The aim of the experiment was to identify ligands to a protein target from a chemical library of 15,200 macrocycles. 16 pools of 20 macrocycles (320 compounds in total) at 40 µM were prepared in DMSO. Multiple copies of single use assay plates were prepared by dispensing 250 nL of each pool into a 384-well plate. The test compounds in DMSO were diluted to 10 µL in a buffer consisting of 50 mM Tris pH 7.5, 150 mM NaCl, 0.01% Tween 20. One copy of the 16 pools was run in a conventional HPLC-MS experiment to verify that the macrocycles could be detected. 302 of 320 macrocycles (94.4%) were detected with strong signal fidelity that would enable detection of a hit. Another 17 test macrocycles (5.3%) were detected but at a sensitivity levels that may result in a false negative due to low intensity. A single compound was not detected at all. The same 16 pools were tested using a 50 mm x 2 mm polyhydroxyethyl A SEC column (PolyLC, Columbia, MD) containing 3 micron beads with 60 Å porosity commonly used for rule-of-five compliant test compounds. These conditions resulted in over 20% of test compounds being detected in the early eluting fraction even in the absence of any target indicating many of the macrocycles are too large to fit through the pores and instead break through the column. To avoid the issue of SEC column breakthrough, a column with the same geometry containing 3 micron beads with 200 Å porosity was tested. None of the 320 test compounds were detected in the early eluting fraction indicating that the larger pore size is successful in trapping the macrocycles. All data analysis was done using Agilent Mass Hunter software.

References

1) Annis DA et al, (2004) International Journal of Mass Spectrometry **128**(2), 77-83

Pilot Screen

A 300-compound subset of the library was selected to run as a pilot. Mixtures of the 300 test compounds at 3 different pooling densities were prepared. To test the ASMS system test pools with densities of 25, 50 and 100 test compounds per pool were made such that each test compound in the pool was at a concentration of 40 µM. A 250 nL aliquot of this mixture was spotted in a polypropylene 384-well plates with an Echo acoustic liquid dispenser (Beckman Coulter). 10 µL of a 5 µM protein target in 25 mM Tris pH 7.5 and 100 mM NaCl was added to each well resulting in a 40x dilution of each test compound to a final concentration of 1 µM. The pools were interrogated using ASMS with a 200 Å porosity SEC column as described. Data analysis was performed using proprietary software with the results shown below

Compounds/Pool	# of Pools	# of Hits	Hit Rate
100	3	5	1.7%
50	6	10	3.3%
25	12	21	7.0%

All 5 compounds detected at the highest pooling density of 100 macrocycles per pool were also detected at the lower densities. A lower hit rate at higher pooling densities is likely the result of competition between test compounds for a limited amount of target protein such that weaker binders are displaced by stronger ones. Based on these results a density of 50 test compounds per pool was selected for the larger screen.

High Throughput Screen

The 15,200 macrocycles in the screening library were multiplexed at a density of 50 compounds per pool resulting in 304 pools. Each pool was interrogated using the same buffer and conditions as in the pilot screen. At a throughput of 6.5 minutes per pool the entire screen was completed in 33 hours (~8 seconds per compound). Data was analyzed using proprietary ASMS software. A total of 72 hits were identified for a final hit rate of 0.5%. The hits were fairly evenly distributed throughout the 304 pools.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
A	0	0	0	0	0	0	0	0	0	0	0	3	0	4	3	0	1	0	0
B	1	1	0	1	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0
E	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	5	0	0	1	0	1	0	0	0	0	0	1	0	0
G	0	0	1	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0
H	1	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0
I	0	0	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
J	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
K	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
L	1	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0
M	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
N	0	0	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
O	0	0	1	0	0	0	0	0	0	0	1	0	0	3	0	0	0	2	0
P	0	0	0	0	0	0	2	0	0	0	0	0	0	2	0	0	0	2	0

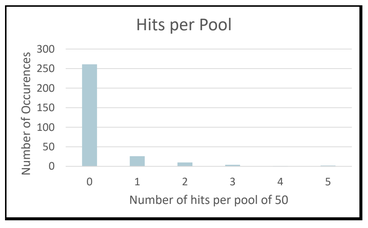

Figure 4: Distribution of the 72 hits across the 304 pools of 50 macrocycles (total 15,200 test compounds)

Hit Confirmation

The 75 hits identified in the HTS campaign together with 3 close structural analogs of interest were interrogated in singleton as part of a secondary screen to confirm binding. The confirmation screen was run at 10 µM test compound against 2.5 µM target. The macrocycle concentration was increased relative to the primary screen to maximize chances that weak binders would be identified. 67 of the 75 primary hits tested were confirmed as binders for a confirmation rate of 89%. The hits are being further interrogated in orthogonal assays

Summary of ASMS

- Label-free, solution-based binding with native target: target or test compound does not need a tag or modification or immobilization on a surface
- Ability to interrogate very large targets and target complexes
- Straightforward assay development, generally limited to buffer optimization
- High throughput: 100K test compound in pools of 250 can be screened in 36 hours
- Low protein usage: Typical screens requires 50 pmol of target per injection (2.5 µg of a 50 kDa target protein or 1 mg/100K test compounds)
- Single method for HTS and hit characterization minimizes assay development
- Compound quantification is dependent on molecular formula eliminating false results due to impurities or test compound breakdown/oxidation