Assessing HTS Performance Using BioAssay Ontology: Screening and Analysis of a Bacterial Phospho-N-Acetylmuramoyl-Pentapeptide Translocase Campaign

Andreas Moberg,^{1,*} Linda Zander Balderud,^{2,*}
Eva Hansson,¹ and Helen Boyd³

¹Screening Sciences, Discovery Sciences, AstraZeneca, Mölndal, Sweden.

²Chemistry Innovation Center, Discovery Sciences, AstraZeneca, Mölndal, Sweden.

³Reagents and Assay Development, Discovery Sciences, AstraZeneca, Mölndal, Sweden.

*Co-first authors.

ABSTRACT

With the public availability of biochemical assays and screening data constantly increasing, new applications for data mining and method analysis are evolving in parallel. One example is BioAssay Ontology (BAO) for systematic classification of assays based on screening setup and metadata annotations. In this article we report a high-throughput screening (HTS) against phospho-N-acetylmuramoyl-pentapeptide translocase (MraY), an attractive antibacterial drug target involved in peptidoglycan synthesis. The screen resulted in novel chemistry identification using a fluorescence resonance energy transfer assay. To address a subset of the false positive hits, a frequent hitter analysis was performed using an approach in which MraY hits were compared with hits from similar assays, previously used for HTS. The MraY assay was annotated according to BAO and three internal reference assays, using a similar assay design and detection technology, were identified. Analyzing the assays retrospectively, it was clear that both MraY and the three reference assays all showed a high false positive rate in the primary HTS assays. In the case of MraY, false positives were efficiently identified by applying a method to correct for compound interference at the hit-confirmation stage. Frequent hitter analysis based on the three reference assays with similar assay method identified additional false actives in the primary MraY assay as frequent hitters. This article demonstrates how assays annotated using BAO terms can be used to identify closely related reference assays, and that analysis based on these assays clearly can provide useful data to influence assay design, technology, and screening strategy.

INTRODUCTION

Within any primary assay there are a number of compounds that are false actives. These compounds are usually excluded downstream in the screening cascade through hit evaluation assays, counter screens, and orthogonal assays. New methods are constantly being developed in an attempt to better predict false positives, such as frequent hitters and promiscuous compounds, to be able to exclude them early on in the screening cascade.

One initiative to enhance assay setup and the use of technologies within AstraZeneca is the utilization of BioAssay Ontology (BAO) for annotation of high-throughput screening (HTS) assays and meta-data.¹ The in-house initiative is a part of the Innovative Medicines Initiative (IMI) OpenPHACTS project with the aim to develop a standardized vocabulary for assay design and technology to enhance access and exchange of data within both the pharmaceutical industry and the public domain. Therefore, HTS assays from in-house drug discovery programs have been annotated retrospectively using BAO terms, to be used, for example, in frequent hitter analysis and to support questions around specificity and selectivity in early drug discovery programs.

Frequent hitters are compounds with a high activity rate in several unrelated assays. Numerous compounds have shown to be frequent hitters within a target class or due to interference with a specific detection technology or assay format.² The use of BAO for annotation of assays for screening facilitates the classification and analysis of diverse HTS data sets. This methodology exploits the flexibility of a semantic data set by introducing well-defined relationships between classes and the possibility of using the data at different levels of granularity.¹ By the use of annotated data a more customized frequent hitter analysis can be performed, including a combination of parameters such as assay design, detection technology, target class, reagents, or other parameters.³

There is still a large unmet medical need for novel antibacterial therapies; phospho-N-acetylmuramoyl-pentapeptide translocase (MraY) is an attractive target involved in peptidoglycan synthesis.^{4–7} Peptidoglycan synthesis is an essential process in bacteria and thus makes it attractive for small molecule intervention. MraY has been a target of interest for a number of years and a number of natural

product inhibitors have been identified⁸ using different assay technologies.^{9–12} In this report we describe the 1,536-well miniaturization of a fluorescence resonance energy transfer (FRET)-based MraY assay and screening campaign, which led to the identification of novel chemical matter. A Compound Interference Correction (CIC) method¹³ was applied to compensate for compound interference with the FRET signal in the compound activity assessment. In addition to the Core Compound Library (CCL), a High Content Recognition Library (HCRL) and Fragment Library (FL) were screened to open up for alternative lead discovery approaches.

In this article the active compounds from the MraY assay were analyzed to identify frequent hitters using both BAO annotated assays and by using simple assay descriptions in an ordinary frequent hitter method, that is, what is usually captured without a thorough annotation or by reading the protocols. The MraY assay protocol was manually annotated using BAO terms and analyzed together with both internal and external assays annotated according to BAO. Three reference assays, previously used for in-house HTS of other targets, were identified with a similar assay design and detection technology, as well as other specifications like wavelength, as the MraY assay. The assays all had high false positive rates, deduced from their low hit confirmation rates. A cheminformatics analysis using the CCL screening data identified common false positives and frequent hitters with 14% of the MraY hits, being active in at least one of the three reference assays. This retrospective analysis highlights the need for powerful tools to analyze screening outputs to influence decision making in assay development and the design of screening cascades. With increasing awareness of the utility of assay ontology and ability to access public databases, this offers a new tool for both screening and assay development across the pharmaceutical industry and the public domain.

MATERIALS AND METHODS

Compound Libraries

Compounds of three different libraries were screened in the MraY FRET assay (Table 3): 1.4 million compounds of a core compound collection (CCL),¹⁴ 18,000 compounds of a FL (molecular weight <300 Da, AlogP <3), and 50,000 compounds of a HCRL, comprising a diverse set of medium-sized, nonlipophilic compounds.

High-Throughput Screening

The FRET donor BODIPY-FL-UDP-N-acetylmuramoyl pentapeptide (B-UNAM-pp) and *Escherichia coli* membranes containing MraY were produced in-house.¹⁰ The FRET acceptor lissamine rhodamine B-dipalmitoyl phosphatidylethanolamine (LRPE) was purchased from Avanti Polar Lipids (Alabaster, AL) and undecaprenyl phosphate (C₅₅P) was obtained from Larodan Fine Chemicals (Malmö, Sweden). Tris-HCl was from Invitrogen (Stockholm, Sweden), Triton X-100 was from Thermo Fisher Scientific

(Gothenburg, Sweden), Tunicamycin was from Apollo (LabKemi AB, Stockholm, Sweden), ethylenediaminetetraacetic acid (EDTA)-NaOH, magnesium chloride, potassium chloride, dithiothreitol, and trehalose were all obtained from Sigma-Aldrich (Stockholm, Sweden). The screening assay was run using black polystyrene 1,536-well microtiter plates (Greiner Bio-One, Frickenhausen, Germany). Assay ready screening plates were provided by in-house Compound Management using Echo[®] acoustic dispensing (Labcyte, Sunnyvale, CA).^{15,16} *E. coli* membranes (33 µg/mL), LRPE (24 µM), and C₅₅P (15 µM) were mixed in a buffer solution (50 mM Tris-HCl pH 7.5, 500 mM Trehalose, 150 mM potassium chloride, 1 mM magnesium chloride, 1 mM dithiothreitol, 0.02% w/v Triton X-100) and preincubated for 30 min at room temperature before use (Tables 1 and 2). Four microliters of this membrane mix was added to the assay plates and preincubated with test compounds for 15 min before starting the reaction with 2 µL B-UNAM-pp (0.3 µM) made in the buffer solution. The reaction was run at room temperature for 130 min and stopped by the addition of 2 µL EDTA (8 mM) in buffer solution. Assay performance was monitored by Z'-factor observations.¹⁷ Tunicamycin at 1 µM was used for inhibition

Table 1. High-Throughput Screening Assay Protocol Table: Live Assay

Step	Parameter	Value	Description
1	Deseal assay ready plates		Compounds/controls in DMSO
2	Enzyme addition	4 µL	MraY <i>Escherichia coli</i> membranes, LRPE, C ₅₅ P in Assay Buffer
3	Incubation time	15 min	Ambient temperature
4	Substrate addition to start reaction	2 µL	B-UNAM-pp in Assay Buffer
5	Incubation time	130 min	Ambient temperature
6	Endpoint quench	2 µL	EDTA in Assay Buffer
7	Centrifuge	1 min	Spin plates at 2,900 rpm
8	Assay Readout	520 and 590 nm	PheraStar (BMG), dual emission read. Excitation wavelength 485 nm

Step Notes

1. Black polystyrene 1,536-well plates (Greiner). Library compound, DMSO or tunicamycin dispensed by internal compound management facility using Echo[®] (Labcyte) acoustic liquid handling.
2. Eight-tip noncontact liquid handling.
4. Eight-tip noncontact liquid handling.
6. Eight-tip noncontact liquid handling.
8. Tip noncontact liquid handling.

Assay Buffer: 50 mM Tris-HCl pH 7.5, 500 mM Trehalose, 150 mM Potassium chloride, 1 mM Magnesium chloride, 1 mM Dithiothreitol, 0.02% w/v Triton X-100.

B-UNAM-pp, BODIPY-FL-UDP-N-acetylmuramoyl pentapeptide; C₅₅P, undecaprenyl phosphate; DMSO, dimethyl sulfoxide; EDTA, ethylenediaminetetraacetic acid; LRPE, lissamine rhodamine B-dipalmitoyl phosphatidylethanolamine; MraY, phospho-N-acetylmuramoyl-pentapeptide translocase.

Table 2. High-Throughput Screening Assay Protocol Table: Artifact Assay

Step	Parameter	Value	Description
1	Mix assay reagents to start reaction		MraY <i>E. coli</i> membranes, LRPE, C ₅₅ P, B-UNAM-pp in Assay Buffer
2	Incubation time	130 min	Ambient temperature
3	Endpoint quench		EDTA in Assay Buffer
4	Dispense	8 μ L	Quenched bulk reaction to assay ready plates containing compounds/controls in DMSO
5	Centrifuge	1 min	Spin plates at 2,900 rpm
6	Assay Readout	520 and 590 nm	PheraStar (BMG), dual emission read. Excitation wavelength 485 nm

Step Notes

1. Run reaction in glass bottle.
 4. Eight-tip noncontact liquid handling. Black polystyrene 1,536-well plates (Greiner). Library compound, DMSO or tunicamycin dispensed by internal compound management facility using Echo (Labcyte) acoustic liquid handling.
- Assay Buffer: 50 mM Tris-HCl pH 7.5, 500 mM Trehalose, 150 mM Potassium chloride, 1 mM Magnesium chloride, 1 mM Dithiothreitol, 0.02% w/v Triton X-100.

controls ($n=512$) and max signal ($n=512$) was obtained with dimethyl sulfoxide (DMSO) controls. The HTS was run on a fully automated BioCel system (Agilent, Santa Clara, CA). All reagents were dispensed using a MultiDrop™ Combi (ThermoScientific, Waltham, MA) and plates were centrifuged with a VSpin (Agilent) for 1 min at 2,900 rpm, before reading on a PheraStar (BMG, Ortenberg, Germany). Fluorescence was read using 485 nm excitation and simultaneous detection of 520 and 590 nm emission. The artifact assay was run at the hit confirmation and IC₅₀ stage in the screen cascade. It was run as a bulk reaction of enzyme plus substrate over a fixed time period, followed by adding EDTA, with reagent concentrations as described above (Tables 1 and 2). This was added to the compounds, plates were centrifuged, and read immediately on the PheraStar. Data were analyzed using both wavelengths according to the method described by Shapiro *et al.*,¹³ using in-house software. The two emission wavelengths were corrected individually to allow for the calculation of a corrected FRET ratio. This method will be referred to as CIC. The cut-off defining an active compound in primary screening of the CCL was set as the third quartile plus two interquartile ranges ($Q3 + 2 \times IQR$) per plate, which equated to 30% effect. At the hit evaluation stage, compounds were screened using a seven-point concentration-response (CR) curve for the determination of IC₅₀ values. Four-parameter logistic equations were fitted to CR data and IC₅₀-values were determined from the sigmoidal dose-response curves. Assay development data were analyzed using the Spotfire Decision Site (Gothenburg, Sweden) and IDBS XLfit (Guildford, United Kingdom). HTS data were analyzed using an in-house database in conjunction with the Spotfire Decision Site.

BAO and Frequent Hitter Analysis

Enzyme activity and fluorescent frequent hitters have been evaluated, in what has been defined as an ordinary frequent hitter analysis, using over 100 in-house primary enzyme target HTS assays or HTS assays detected by fluorescence. The active compounds in the primary MraY CCL assay, that were tested in at least 20 reference assays, were defined as frequent hitters when active in more than 20% of the assays.

The MraY assay was annotated according to BAO and three reference assays with similar assay design and detection technology were identified using the ontology classification. The reference assays were enzyme activity assays; also using CIC in the hit evaluation stage and with fluorescence read out in a similar wavelength range, and were used in a frequent hitter analysis based on BAO (Fig. 3A). The number of compounds active in one or more of these reference assays along with compounds active in all three reference assays was evaluated both in the primary and in the hit evaluation tests.

RESULTS**High-Throughput Screening**

A 384-well FRET-based assay for MraY¹⁰ was miniaturized to run in a 1,536-well format. The FRET assay readout reflects MraY-mediated attachment of UDP-N-acetylmuramoyl-pentapeptide to C₅₅P. MraY, C₅₅P, and the FRET acceptor, LRPE are present in detergent/lipid mixed micelles. UDP-N-acetylmuramoyl pentapeptide is labeled with a BODIPY FRET donor and FRET occurs as MraY puts B-UNAM-pp in close proximity with LRPE upon the attachment of B-UNAM-pp to C₅₅P. Essential steps in miniaturization and HTS adaptation were taken by introducing liquid handling equipment, robotics, and by studies of reagent and signal stability over time under careful quality control of assay performance. Assay window and the reaction time course were also studied (Fig. 1A, B), showing no difference between the 384- and 1,536-well assay. Tunicamycin, a natural product known from literature to inhibit MraY, was used as a positive control. The IC₅₀ of Tunicamycin was 4.1 nM (Fig. 1C, D), which is in agreement with previously published data using this assay.¹⁰ IC₅₀ values reported using other assays in the range of 360 nM to 2 μ M.^{8,11,12} Z' values for the 1,536-well assay during the HTS were equal to or greater than 0.6 and the signal-to-background ratio was 1.4–1.6.

Primary MraY HTS and associated confirmatory assays were screened using the three compound libraries described in the “Materials and Methods” section. The dose-response data were of high quality with >80% of the coefficients of correlation, $R^2 \geq 0.99$. Confirmation rates and hit rates for the three compound libraries are shown in Table 3. Around 90% of the primary screening hits from the CCL were removed as false positives at the hit confirmation stage with the use of

Fig. 1. Time course showing normalized fluorescence resonance energy transfer (FRET) signal, reflecting MraY catalyzed B-UNAM-pp attachment to C₅₅P. **(A)** The 384-well format in the presence (●) or absence (○) of 22 µg/mL MraY membranes (average of three wells, error bars represent SD). **(B)** The 1536-well format in the presence (■) or absence (□) of 1 µM Tunicamycin (average of 768-wells, error bars represent SD). A straight line has been fitted to the normalized FRET ratio data by linear regression, giving slopes of **(A)** 0.011 min⁻¹ and **(B)** 0.0089 min⁻¹, respectively. **(C)** Normalized FRET signal in presence of various concentrations of Tunicamycin (Δ). IC₅₀ of Tunicamycin was determined to 4.1 nM (*n* = 3, error bars represent SD, four-parameter logistic curve fit). **(D)** Structure of Tunicamycin.

CIC. The frequency distribution of the percent effect data of compounds in primary screening of the CCL showed a normal distribution around zero, indicating that the screen was robust and most compounds were inactive (Fig. 2A). The CCL had an IC₅₀ distribution around ~15 µM, whereas the IC₅₀ values of the HCRL and FL were distributed around 90 and 400 µM, as expected, due to the nature of these libraries (Fig. 2B–D). The overall screen hit rates after hit evaluation and CIC were 0.35% for the CCL, 0.75% for the HCRL, and 3% for the fragments (see hit rate relative total number of compounds in respective library, Table 3 and Fig. 3B).

BAO and Frequent Hitter Analysis

In the MraY hit evaluation assay 11% of the compounds from the primary assay were confirmed actives (Table 3 and Fig. 3B). To explore the quality of the high number of false positive compounds the data were retrospectively analyzed for identification of assay design and technology frequent hitters using both ordinary frequent hitter analysis and frequent hitter analysis based on BAO annotated reference assays (described in the “Materials and Methods” section). Among the actives in the primary MraY assay 3.2% of the compounds were identified by the ordinary frequent hitter analysis, that is, in more than one out of five in-house enzyme activity assays (described in the “Materials and Methods” section), whereas 3.3% were identified in the analysis using assays detected by fluorescence. From these compounds, defined as frequent hitters using the ordinary frequent hitter analysis, 30.2% respective 27.3% of were confirmed as actives in the hit evaluation assay, and thus not filtered out by the CIC tool.

To further evaluate the false positives in the primary CCL MraY screen, the assay was annotated according to BAO. The main parameters anticipated to influence false activity readouts in the MraY assay were defined and analyzed together with other similar in-house HTS assays and assays published in the PubChem BioAssay database,¹⁸ annotated using BAO terms. Optimal reference assays would be other enzyme activity as-

says in a micelle assay format detected by FRET using labels with excitation wavelength 485 nm and emission 520 and 590 nm. Furthermore, to be able to explore the usefulness of the CIC tool, the hit evaluation assays were to be analyzed using this method. The study resulted in three in-house assay cascades using CIC tool in the confirmatory assay, where enzyme activity were measured in a protein-based format using fluorescence intensity (FI) with excitation 485 nm and emission 520 nm (Fig. 3A). In addition, two reference assays were identified in the PubChem BioAssay; AID 1385,¹⁹ a human target protein–protein interaction assay using FRET detection and AID

Table 3. Screening Conditions, Hit Rates, and Confirmation Rates for the Three Compound Libraries Screened in the Mray Fluorescence Resonance Energy Transfer Assay

	Screening concentration (μ M)	Number of replicates	CIC applied	Hit rate relative total number of compounds in respective library (%)	Confirmation rate from previous step (%)
CCL					
Primary screening	10	1	No	3.6	—
Hit confirmation	10	2	Yes	0.39	11
Hit evaluation	0.069–50	2	Yes	0.35	88
HCRL					
Primary screening	100	1	Yes	0.75	—
Hit evaluation	7.8–500	1	Yes	0.75	97
FL					
Primary screening	1,000	1	Yes	3.0	—
Hit evaluation	31–2,000	1	Yes	3.0	100

CIC, Compound Interference Correction; CCL, Core Compound Library; FL, Fragment Library; HCRL, High Content Recognition Library.

898,²⁰ a bacterial enzyme activity assay with FI detection, similar to the in-house reference assays. Both PubChem assays had an excitation wavelength of 485 nm and emission 520 nm. To evaluate the possibility of using external assays in frequent hitter analysis the compound overlap between the in-house and the PubChem assays were analyzed. Of the active compounds in the primary Mray assay, only 200 were tested in the public reference assays. Hence, only the in-house reference assays were used for frequent hitter analysis.

The three in-house reference assays represent targets from two different enzyme classes, and neither of them are part of the Mray protein family. The screening cascades and hit rates for Mray and the reference assays are presented in Figure 3B and described in Supplementary Table S1 (Supplementary Data are available online at

www.liebertpub.com/adt). Similar to the Mray assay the reference assays had a high false positive rate in the primary CCL assay (84%, 88%, and 90% compared with 89% in the Mray primary assay; Fig. 3B). Fourteen percent of the active compounds in the primary Mray assay were also active in at least one of the other three primary reference assays and 10% of these were active in all three primary reference assays (Fig. 4). Of the frequent hitters in the primary Mray assay that were active in all three reference assays, 25% of the compounds were still active in the evaluation assay, that is, not removed by the confirmatory assays where CIC were applied. The confirmatory CR data were also used to identify unspecific compounds in the Mray assay. Out of the active compounds in the Mray CR assay, 561 compounds were tested in any of the three confirmatory reference assays. As many as 178 (32%) of these were active in at least one of the reference assays.

Furthermore, we evaluated the two frequent hitter analysis methods; the ordinary frequent hitter method and the method using BAO annotated reference assays were compared. As many as two out of three frequent hitters in the primary Mray assay could only be identified using the three BAO annotated reference

Fig. 2. Frequency distribution of the percent effect of compounds in primary screening (A) and frequency distributions of pIC₅₀ values observed at hit evaluation stage for the Core Compound Library (CCL) (B), High Content Recognition Library (HCRL) (C), and Fragment Library (FL) (D), respectively.

Fig. 3. (A) Schematic picture of the BioAssay Ontology (BAO) annotation of MraY (in gray) and the three reference assays (thick border). **(B)** Percentage hit rate of the MraY cascade and the three reference screening cascades in the different compound libraries: CCL, HCRL, and FL.

assays (Fig. 4). Of the compounds active in all four primary HTS assays 28% could not be identified as frequent hitters by the ordinary enzyme activity and fluorescent frequent hitter analysis.

DISCUSSION

This report describes a HTS for MraY inhibitors followed by a retrospective analysis of the HTS output. The screen identified putative MraY inhibitors in all the three compound libraries; CCL, FL, and HCRL. However, the number of false positives using the MraY FRET assay was high—for the CCL almost 90% of the hits from primary screening were removed as false positives at the hit confirmation stage. The final output from the screen produced novel chemical entity and hence a step forward in the important quest for novel antibiotics.

Experienced researchers might have anticipated a high false positive rate in the MraY HTS even before the first experiment was made, based on the nature of the assay. Such information is, however, traditionally not shared or easily accessed due to the lack of an appropriate infrastructure. BAO allows for classification of assays globally, between companies and the public domain, to build an open access inventory of assays and their classifications. By comparing the assay to be assessed with similar assays, it is possible to make proactive decisions based on the setup of previous screens. In the MraY study the assay was retrospectively annotated and the screening output was subsequently compared with data from three similar enzyme activity assays. These used a fluorophor with 485 nm excitation and emission wavelength close to 520 nm in conjunction with the CIC tool. Unlike the MraY FRET assay, the reference assays used recombinant soluble proteins and were detected by FI. The FI reference assays had a high rate of false positives in the primary screen (Fig. 3B) similar to the MraY screen. In the ideal case of perfect matches FRET assays measuring translocase activity in a micelle format that used the same liquid handling equipment, etc. would be preferred for the proactive BAO assessment. Through a broad annotation using defined and detailed HTS assay terminology,

more specified assay design and technology issues can be analyzed. The need for complex assay annotations should, however, not be exaggerated. The false positive rate information obtained from mining the similar, but less complex FI assays would be a strong argument for caution with hit annotation in any screen. A prospective analysis for MraY would most likely have resulted in a different assay technology or screening cascade. In addition the importance of the CIC would have been more apparent at the assay development stage. This is only one level at which the annotation of information according to BAO has potential applications. In this study we show that a restricted number of tests with closely related assay methods can, with benefit, be used for frequent hitter analysis. Frequent hitters are compounds showing a high activity rate in a number of different assays. In the MraY assay the predicted frequent hitters are

Fig. 4. Frequent hitter analysis of compounds active in the primary Mray high-throughput screening assay. Percentage frequent hitters of (A) the compounds active in the primary Mray assay, but inactive in the following confirmatory Mray assays using Compound Interference Correction (CIC), (B) compounds confirmed active in the confirmatory concentration–response assay. Frequent hitters were identified using both ordinary frequent hitter analysis and by the three assay design and technology-specific reference assays identified using BAO (see “Materials and Methods” section). The main circles show the percentage of compounds not identified as frequent hitters, that is, real actives (dark gray), the compounds identified as frequent hitters by both frequent hitter analysis methods (medium gray), and the fraction of compound detected only by either the ordinary method or by the BAO reference assays (seen in black respective light gray). The secondary circles show all compounds detected by the method using three reference assays identified using BAO. The percentage of frequent hitters active in all three reference assays (dark gray) and the compounds only active in one or two reference assays (light gray). The percentage frequent hitters also identified using the ordinary frequent hitter method is shown by the black outline.

compounds interfering with the fluorescent signal or micelles, and promiscuous compounds activating several unrelated enzyme targets. The CIC tool, used in the Mray confirmatory assay, is designed to compensate for false activity due to compounds interacting with any of the two fluorescence wavelengths in the FRET detection signal. Most frequent hitters (more than 80%) were, in fact, identified in the confirmation assay where the CIC tool was used. However, from the frequent hitters identified using the three reference assays, over 1,000 compounds, about one out of four of the actives in the CR remained, which stresses the importance of an accurate frequent hitter analysis early in the screening cascade. Valuable information can be gained by exploring assay design and technology from in-house data and public domain information to influence decision making around the assay development and frequent hitter analysis. The increasing number of PubChem assays annotated using BAO classification opens up a wealth of information to aid assay design

and frequent hitter analysis. In Mray assays with similar assay technology, for example, the PubChem AID 1385 and 898^{19,20} and target class-specific assays such as the PubChem AID 594268,²¹ an NMR assay evaluating Mray inhibitors, can be easily identified and evaluated as a support during assay development. Due to a low number of overlapping compounds, that is, compounds tested in both the in-house Mray HTS and the PubChem assays, the external assays could not be included in this analysis.

In conclusion, we show that by the use of only three closely related assays a significant number of compounds with frequent hitter properties were identified. The more unspecific reference assays, restricted by either assay design or detection technology, often used (probably due to the lack of detailed annotation) in our study only did identify about 30% of the frequent hitters found in the primary Mray assay, analyzed using BAO annotated reference assays. A retrospective analysis of both in-house and external data is also shown to

be useful as a support during assay development and screening cascade design, for example, influencing HTS method decisions and the design of orthogonal assays and counter screens. With a growing number of both internal and external assay annotations using BAO and with extended use of the information to explore its full potential, we envisage this type of assay annotation evolving into an impactful tool for pre-HTS decision making.

ACKNOWLEDGMENTS

The authors wish to thank Per-Erik Strömstedt, Ola Engkvist, Silvio DiCastro, Adam Shapiro, Helen Plant, and Ulf Börjesson in addition to the AstraZeneca Global Compound Management Team. The authors would also like to thank Stephan Schürer and Uma Vempati from the BioAssay Ontology team, University of Miami, for sharing their annotation of PubChem assays. The work was supported by Open PHACTS (<http://openphacts.org>), funded by the Innovative Medicines Initiative of the EU and EFPIA (www.imi.europa.eu).

DISCLOSURE STATEMENT

No competing financial interests exist.

REFERENCES

- Vempati UD, Przydzial MJ, Chung C, et al.: Formalization, annotation and analysis of diverse drug and probe screening assay datasets using the BioAssay Ontology (BAO). *PLoS One* 2012;7:e49198.
- Bocker A, Bonneau PR, Edwards PJ: HTS promiscuity analyses for accelerating decision making. *J Biomol Screen* 2011;16:765–774.
- Schurer SC, Vempati U, Smith R, Southern M, Lemmon V: BioAssay ontology annotations facilitate cross-analysis of diverse high-throughput screening data sets. *J Biomol Screen* 2011;16:415–426.
- Boyle DS, Donachie WD: mraY is an essential gene for cell growth in *Escherichia coli*. *J Bacteriol* 1998;180:6429–6432.
- Dini C: MraY inhibitors as novel antibacterial agents. *Curr Top Med Chem* 2005;5:1221–1236.
- Ikedo M, Wachi M, Jung HK, Ishino F, Matsuhashi M: The *Escherichia coli* mraY gene encoding UDP-N-acetylmuramoyl-pentapeptide: undecaprenyl-phosphate phospho-N-acetylmuramoyl-pentapeptide transferase. *J Bacteriol* 1991;173:1021–1026.
- Zawadzke LE, Wu P, Cook L, et al.: Targeting the MraY and MurG bacterial enzymes for antimicrobial therapeutic intervention. *Anal Biochem* 2003;314:243–252.
- Brandish PE, Kimura KI, Inukai M, et al.: Modes of action of tunicamycin, liposidomycin B, and mureidomycin A: inhibition of phospho-N-acetylmuramoyl-pentapeptide translocase from *Escherichia coli*. *Antimicrob Agents Chemother* 1996;40:1640–1644.
- Branstrom AA, Midha S, Longley CB, et al.: Assay for identification of inhibitors for bacterial MraY translocase or MurG transferase. *Anal Biochem* 2000;280:315–319.
- Shapiro AB, Jahic H, Gao N, Hajec L, Rivin O: A high-throughput, homogeneous, fluorescence resonance energy transfer-based assay for phospho-N-acetylmuramoyl-pentapeptide translocase (MraY). *J Biomol Screen* 2012;17:662–672.
- Solapure SM, Raphael P, Gayathri CN, et al.: Development of a microplate-based scintillation proximity assay for MraY using a modified substrate. *J Biomol Screen* 2005;10:149–156.
- Stachyra T, Dini C, Ferrari P, et al.: Fluorescence detection-based functional assay for high-throughput screening for MraY. *Antimicrob Agents Chemother* 2004;48:897–902.
- Shapiro AB, Walkup GK, Keating TA: Correction for interference by test samples in high-throughput assays. *J Biomol Screen* 2009;14:1008–1016.
- Kogej T, Blomberg N, Greasley PJ, et al.: Big pharma screening collections: more of the same or unique libraries? The AstraZeneca-Bayer Pharma AG case. *Drug Discov Today* 2013;18:1014–1024.
- Zaragoza-Sundqvist M, Eriksson H, Rohman M, Greasley PJ: High-quality cost-effective compound management support for HTS. *J Biomol Screen* 2009;14:509–514.
- Ellson R, Mutz M, Browning B, et al.: Transfer of low nanoliter volumes between microplates using focused acoustics—automation considerations. *J Lab Autom* 2003;8:29–34.
- Zhang JH, Chung TD, Oldenburg KR: A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J Biomol Screen* 1999;4:67–73.
- Wang Y, Xiao J, Suzek TO, et al.: PubChem's BioAssay Database. *Nucleic Acids Res* 2012;40:D400–D412.
- National Center for Biotechnology Information. PubChem BioAssay Database AID:1385 (RAD51), Penn Center for Molecular Discovery. Available at: <https://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=1385>. (Last accessed on June 19, 2014).
- National Center for Biotechnology Information. PubChem BioAssay Database AID:898 (YopH), University of Pittsburgh Molecular Library Screening Center. Available at: <https://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=898>. (Last Accessed on June 19, 2014).
- National Center for Biotechnology Information. PubChem BioAssay Database AID:594268 (MraY), ChEMBL (744419). Available at: <https://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=594268>. (Last accessed June 19, 2014).

Address correspondence to:
Linda Zander Balderud, PhD
Chemistry Innovation Center
Discovery Sciences
AstraZeneca
Pepparedsleden 1
Mölndal 431 83
Sweden

E-mail: linda.zander-balderud@astrazeneca.com

Abbreviations Used

B-UNAM-pp = BODIPY-FL-UDP-N-acetylmuramoyl pentapeptide
BAO = BioAssay Ontology
C₅₅P = undecaprenyl phosphate
CCL = Core Compound Library
CIC = Compound Interference Correction
CR = Concentration–response
DMSO = dimethyl sulfoxide
EDTA = ethylenediaminetetraacetic acid
FI = fluorescence intensity
FL = Fragment Library
FRET = fluorescence resonance energy transfer
HCRL = High Content Recognition Library
HTS = high-throughput screening
IMI = Innovative Medicines Initiative
IQR = interquartile range
LRPE = lissamine rhodamine B-dipalmitoyl phosphatidylethanolamine
MraY = phospho-N-acetylmuramoyl-pentapeptide translocase
NMR = nuclear magnetic resonance