Utility of Resazurin, Horseradish Peroxidase, and NMR Assays to Identify Redox-Related False-Positive Behavior in High-Throughput Screens

Matthew Tarnowski,[1,2] Amélie Barozet,[2,3] Carina Johansson,[2] Per-Olof Eriksson,[2] Ola Engkvist,[2] Jarrod Walsh,[4] and J. Willem M. Nissink[1]

[1]Oncology IMED, AstraZeneca, Cambridge, United Kingdom.
[2]Discovery Sciences, AstraZeneca R&D Gothenburg, Gothenburg, Sweden.
[3]Department of Biology, Lund University, Lund, Sweden.
[4]Discovery Sciences, AstraZeneca, Alderley Park, United Kingdom.

ABSTRACT

Discerning false positives from true actives in high-throughput screening (HTS) output is fraught with difficulty as the reason of anomalous activity seen for compounds is often not clear-cut. In this study, we introduce a novel medium-throughput NMR assay for the identification of redox-cycling compounds (RCCs), which is based on detection of oxidation of a reducing agent. We compare its outcomes to those from horseradish peroxidase (HRP)/phenol red and resazurin (RZ)-based assays that are more commonly used for triaging HTS outputs. Data from NMR, RZ, and HRP redox assay are shown to correlate, with the NMR assay showing the greatest accuracy. In addition, historical data analysis was used to identify compounds frequently active in assays for redox-susceptible targets. We provide examples of compound classes found and conclude that the NMR redox assay offers a novel and reliable way of identifying RCCs at a medium throughput. The HRP and RZ assays are reasonable higher-throughput alternatives, with both showing similar sensitivity to redox-cycling and false-positive compounds. The RZ assay has a higher hit rate, reflecting its ability to pick up multiple modes of action.

Keywords: redox false positives, HTS output deconvolution, horse-radish peroxidase/phenol red assay, resazurin-based redox assay, NMR

INTRODUCTION

With the advent of high-throughput screening (HTS), it has become possible to rapidly screen large libraries of compounds.[1,2] However it is now clear that with larger compound libraries, the risk of identifying and pursuing results that are artifacts of screening increases.[3] Pursuing such false positives results in wasted resource, time, and effort. False-positive results may arise from interference with signal readout through fluorescence or quenching, or from an undesirable mechanism of inhibition of the assay target. Examples of the latter include inhibition by aggregation[4,5] and chemical interference such as redox-mediated modification or reactive modification of the target.[6]

Redox-cycling compounds (RCCs) are known to cause false-positive hits in HTS against redox-susceptible targets such as cysteine proteases, protein tyrosine phosphatases, and metalloproteases.[7–10] RCCs catalyze a radical-mediated redox cycle between a strong reducing agent such as dithiothreitol or tris(2-carboxyethyl)phosphine (TCEP) (often added to protect protein from oxidation in assays)[11,12] and oxygen, ultimately resulting in the production of hydrogen peroxide, which damages the target through oxidation. Such damage may subsequently lead to loss of activity, and this results in an apparent inhibition of the target.

Redox-detecting assays have been put forward in the literature, and we have implemented two commonly used variants: a resazurin (RZ)-based generic redox assay[11] and a horseradish peroxidase-based assay,[13] specifically developed to detect RCCs. Other assays detecting reactivity have been reported in the literature, but are more generic in nature.[14] Redox effects remain a challenge in certain target classes, and application of redox-deconvoluting assays has been reported recently, for example, for lysine demethylases where half or more of the hits may suffer from such interference.[15] In this study, we compare results from two commonly used assays to those of a novel and orthogonal NMR-based approach, which

monitors the emergence of TCEP oxidation product as a proxy indicator of redox cycling.

To further analyze the nature of the compounds that are suggested by these assays to be false positives, we performed retrospective data mining of HTS results for a set of targets from classes known to be redox sensitive: cysteine proteases, metalloproteases, and protein tyrosine phosphatases.

MATERIALS AND METHODS

NMR Redox Assay

The NMR redox assay is based on detection of the ^{1}H NMR spectrum of the P-oxide of TCEP (TCEPO) in an aqueous buffer containing the compound of interest and TCEP (see *Supplementary Fig. S2*; Supplementary Data are available online at www.liebertpub.com/adt). The spectrum of TCEP at pH 7.5 is heavily broadened by chemical exchange effects,[16] whereas the spectrum of TCEPO consists of two well-resolved multiplets from the methylene groups (*Fig. 2* and *Supplementary Data*). By applying a transverse relaxation filter to the 1D NMR pulse sequence, the exchange-broadened TCEP signals are eliminated and the appearance of TCEPO in the sample is easy to follow.

The samples were prepared in an aqueous buffer with 50 mM TRIS-d_{11}, (98% D; Cambridge Isotope Laboratories, Inc., #DLM-1814-5), 200 mM NaCl, pH 7.4, 10% D_2O, and 22 μM TMSP, and compound of interest at 100 μM. The samples were prepared just in time before the NMR measurements with a TECAN EVO100 sample preparation robot running the Gemini software (ver 4.0) (Tecan AG Zurich, CH). Aliquots of 5 μL of cooled and pH-adjusted stock solution of TCEP (Sigma Aldrich, MA, #C4706) (0.3 M), and of stock solutions of compounds (10 mM in nondeuterated DMSO) were added to the NMR tube containing 490 μL. The samples were mixed mechanically by using a pipetting needle with a teflon plug attached, acting as a piston. The samples were then transported individually to the NMR spectrometer with a SampleRail™ (Bruker Biospin, GmbH, Rheinstetten, DE) followed by spectrum acquisition.

The 1D ^{1}H NMR spectra were acquired at 293 K using a Bruker Avance III HD 800 MHz spectrometer equipped with a cryogenically cooled probe head. The water magnetization and the signal from DMSO were simultaneously suppressed by excitation sculpting,[17,18] using biselective 3 ms shaped 180 degree flip-back pulses. The suppression of the DMSO peak also resulted in suppression of the low-field methylene doublet (2.4 ppm) of TCEPO (*Fig. 2* and *Supplementary Data*). The excitation sculpting sequence (Bruker Topspin pulse program zgesgp) had two pairs of 2.5 ms gradient pulses, whereby the entire pulse sequence contained a transverse relaxation filter of 16.9 ms, which effectively eliminated the exchange-broadened signals from TCEP (*Fig. 2* and *Supplementary Data*). Each spectrum was collected with 8 dummy scans and 128 scans, an acquisition time of 1.36 s, and a repetition delay of 3.2 s, resulting in a total experimental time of 10 min and 23 s. From the time point of the files in the NMR data directory, the start and finish of data acquisition could be determined. Together with the timestamp of the first mixing stroke in the sample preparation obtained from the Gemini logfile, the time between sample preparation and start of data collection was 8 min and 12 s with a standard deviation of 6 s. Thereby, the NMR data represent a ±5 min time average with a midpoint at 13 min and 24 s after sample preparation. The NMR data were processed with 2 Hz exponential line broadening and baseline correction around the high field TCEPO multiplet (2.1 ppm) (*Fig. 2* and *Supplementary Data*). Control samples containing TCEP in the absence of compound were run after every 10 samples. TCEP oxidation was quantified by integration of the high field peak multiplet of TCEPO (2.1 ppm) coming from the methylene protons next to the phosphorus atom. The peak integrals were extracted using a modified TopSpin™ (Bruker) multispectra integration script and normalized to a control sample run in absence of added compound. Compounds were regarded as redox active if they increased the amount of TCEPO by at least 10% compared to the control. Classification of the redox actives in categories corresponds to TCEPO integrals (relative to control) of 1–5 for weak, 5–15 for medium, and 15–25 for strong actives.

RZ-Based Redox Assay

We refer to Lor *et al.*[11] for a complete description of the assay. The original procedure was taken and optimized for assaying in the presence of the reducing agent TCEP (Sigma Aldrich, #C4706). Based on titration of RZ (Sigma Aldrich, #199303) and TCEP concentration (*Supplementary Data*), 5.0 μM TCEP and 5.0 μM RZ were selected for screening. These provided a compromise between an acceptable signal to background ratio and Z′ value, while maintaining assay sensitivity to detect RCCs. In summary, compounds (100 nL, 10 concentrations from 100 to 0.2 μM in DMSO) were loaded into black, low-volume 384-well Greiner 784076 plates using an Echo® (Labcyte, Inc.). On-board DMSO and NSC663284 (1.5 μM in DMSO; Sigma Aldrich, #N7537) control wells were used. RZ (5.0 mM in assay buffer) was sonicated for at least 10 min and then added to an assay buffer (50 mM HEPES [Sigma Aldrich, #H7006] and 50 mM NaCl [Sigma Aldrich, #S7653], pH 7.4 in milliQ water) to give a final concentration of 5.0 μM, and TCEP (5.0 mM in assay buffer, pH 7.4) was added to give a final concentration of 5.0 μM. Then, 10 μL of this solution was added to all wells using a Multidrop™

384 Reagent Dispenser (Thermo Scientific™). The wells were allowed to react for 60 min at room temperature after addition of RZ/TCEP solution and were then read on a PHERAstar (BMG LABTECH) with Ex = 560 nm/Em = 590 nm.

Horseradish Peroxidase/Phenol Red Based Assay

We refer to Johnston *et al.*[13] for a complete description of the assay. Their procedure was followed using 1 mM TCEP (Sigma Aldrich, #C4706) as the reducing agent. In summary, compounds (200 nL, 10 concentrations 100 to 0.2 μM in DMSO) were loaded into clear polystyrene 384-well Greiner 781101 plates using an Echo (Labcyte, Inc.). On-board DMSO and NSC663284 (1.5 μM in DMSO; Sigma Aldrich, #N7537) control wells were used. TCEP (0.5 M, pH 7.4) was diluted to 1 mM in HBSS (Sigma Aldrich, #H6648). The phenol red-HRP detection reagent was prepared (400 μg/mL phenol red [Sigma Aldrich, #P3532] and 25 U/mL HRP [Sigma Aldrich, #P2088] in HBSS). Twenty microliters of the TCEP solution was added to all wells and after 15 min of incubation at ambient temperature, 20 μL of the HRP-phenol red solution was added to all wells. After 5 min of incubation at ambient temperature, 10 μL of NaOH (0.5 M) was used to stop the reaction. All additions were completed using a Multidrop 384 Reagent Dispenser (Thermo Scientific). Each well was measured using a PHERAstar (BMG LABTECH), with 600 nm absorption.

EC50 curves. For both assays described above, EC50s were determined using 10 concentrations, with a maximum concentration of 100 μM, in half-step dilutions, and in triplicate. Original compound stock was kept in 10 mM DMSO solutions. Twelve wells of DMSO (–ve control) and 12 wells of NSC663284 (+ve control) were present on each assay plate and used to apply a plate-based normalization of assay signals in Genedata Screener® 12 (Genedata AG, Basel, Switzerland, 2015). The data were plotted utilizing Genedata Screener's "Smart Fit" option to apply a four-parameter fit. This enabled the software to automatically fix the asymptotes to 0% and 100% if required. A minimal nHill value of 0.5 and a maximal nHill value of 8 were permitted. All curves had to exhibit a span in percentage effect of greater than 50%, but less than 150% to be classed as a valid fit. Automated outlier masking was enabled.

Data Mining Methodology

Characterization of anomalous inhibitor behavior. We characterized both generic "frequent-hitter" behavior using a previously described method[19] and potential redox-cycling behavior by analyzing a subset of data derived from activity data for redox-susceptible targets. In summary, the method calculates a score (pBSF) for a compound by comparing the actual incidence of activity across the assays it has been tested in against the expected incidence of activity. The latter is determined from the larger set of data for all compounds. The pBSF score is the negative logarithm of a chance and reflects the probability that the observed tally of activity for the compound occurs randomly, given the expected incidence of activity. If this probability is very low, it suggests that the observed activity pattern is an outlier and goes beyond what is expected. In practice, such outliers tend to be compounds more colloquially known as "frequent-hitters," anomalous inhibitors that appear to give rise to inhibition readouts in a larger-than-expected number of assays. Typical threshold values for pBSF lie in the 2–3 range, corresponding to a false-positive chance of 1%–0.1% that the compound is *not* an outlier, indicating a large likelihood that the compound is an anomalous inhibitor.

Selection of a set of putative redox-active compounds. Results from historical high-throughput screens were used to identify compounds potentially acting through a redox-cycling mechanism by focusing on a subset of relevant assays. Data from a set of 21 screens run at AstraZeneca and targeting various redox-susceptible proteins, including cysteine proteases, metalloproteases, and protein tyrosine phosphatases, were gathered (*Supplementary Table S1*), and pBSF scores corresponding to results in these assays were calculated. Only compounds with a pBSF higher than or equal to three were included. These compounds show clear evidence of anomalous hitter behavior and will likely interfere through unknown mechanisms, including redox activity.

As non-redox mechanisms can raise issues in assays more generically, additional sets of "background" assays were collated and the same approach was applied to obtain pBSF scores based on these sets. Since all selected redox-susceptible assays, except one, are Fluorescence Intensity, Förster Resonance Energy Transfer, or spectrophotometric assays, sets of background assays for each of these categories were selected, as well as a generic set of diverse assays to assess general promiscuity of compounds (*Supplementary Data* and *Supplementary Table S2*). pBSF values were derived for each of these background sets. A set of putative redox-active compounds was then selected by retaining those tested at least once and with a pBSF lower than 2, that is, "clean" in each set of background assays. This set comprises 1,097 compounds and is referred to as ***S***.

S consists of compounds found to act as frequent hitters in redox-susceptible assays, but not in other assays, suggesting that these are either redox active (*e.g.*, RCCs) or compounds actually capable of binding to several proteins of one class (such as protein tyrosine phosphatases inhibitors containing

phosphate-mimetic groups). The latter were removed in a further step comprising visual inspection and checks of historical data.

Derivation of fragments enriched for redox behavior. Compounds in S were fragmented and an enrichment analysis was performed to identify overrepresented and underrepresented fragments (see *Supplementary Data*). This led to the manual selection of 42 representative enriched fragments. Binding properties of these fragments were assessed subsequently by investigating existing complexes in the Protein Data Bank (PDB)[20] (see *Supplementary Data* for details).

The group of compounds associated with an enriched fragment f was defined as the set of compounds tested in at least one of the selected redox-susceptible assays and containing f, but no other representative enriched fragment (unless latter fragment is a subfragment of f). To determine the frequent-hitter and redox-cycling tendencies of a given fragment, we tested for an increased proportion of frequent hitters among its associated group with regard to the redox target and background sets of assays. One-tailed binomial tests[21] were applied to frequent-hitter classifications using pBSF >3. A test was applied on each pair of fragment and set of assays. Each test was followed by Holm correction[22] and considered positive if $P < 0.001$. A positive test indicates an increased incidence of frequent hitters for the group of compounds associated to a fragment in a set of assays.

RESULTS AND DISCUSSION

Redox Assays

Two assays with a suitably high throughput have been reported in the literature for the detection of redox effects like reducing agent-mediated redox cycling. One uses resazurin (RZ) to detect compounds that are able to catalyze its radical-mediated conversion to resorufin (referred to as the RZ assay).[11] The other employs horseradish peroxidase (HRP) to detect formation of hydrogen peroxide by monitoring the oxidation of phenol red by HRP (referred to as the HRP assay).[13] The latter assay will detect effects of direct oxidation, as well as effects resulting from superoxide anion and peroxide formation. Triaging is generally done in two different ways—a worst-case scenario is checked where a compound is tested at relatively high concentration for any redox activity or the level of redox activity (*e.g.*, determined as EC50) is compared to the level of activity in the primary HTS target. Considering the level of redox activity in the context of the primary target activity is relevant, but can be ambiguous, for example, as a result of anomalous redox curve shapes. We focus here on the former approach, which is generally applicable when the range of the activity seen in the primary HTS is limited.

HTS Compound Set for Triage

A set of 4,921 compounds was identified based on actives from a MALT1 HTS campaign. MALT1 is a cysteine protease. In the course of the HTS workup, these actives were subjected to a hit triaging cascade to filter out, among others, compounds that appeared false positive as a result of redox-related interference. A subset of this set was processed using our NMR-based assay as part of the cascade.

Comparison of HRP and RZ Assays

A set of 4,921 compounds were screened in the HRP and RZ assays to assess their correlation. The HRP assay followed the Johnston *et al.*[13] protocol utilizing 1 mM TCEP for a 15-min reaction. This provided linear detection of H_2O_2 production up to 100 μM, although there was capacity in the system for nonlinear detection up to 400 μM. The RZ assay was optimized to use 5 μM, performed over 60 min, and capable of linear signal detection for full conversion of the dye to resorufin. Differences in the relative sensitivity of the respective assays to detect RCCs resulted in varying and bespoke criteria needing to be adopted when assigning activity. This was exacerbated by the anomalous (non-sigmoidal) fitting of a number of EC50 curves. The classification of actives was determined as follows and provides, in our opinion, an accurate representation of compound activity across the three assays: compounds with a pEC50 of four or less in the RZ assay were designated as inactive (N), thereby classifying as active (A) all compounds that gave rise to bell-shaped curves, and/or had a pEC50 > 4. EC50s derived using the HRP assays turned out to be too weak to be used consistently for classification, and therefore the compounds with %effect at the highest concentration (100 μM) >5 × standard deviation of the inhibition values of the null control wells were labeled as active.

Curves determined for our set of compounds using the RZ assay were commonly found to exhibit bell-shaped curves. This behavior has previously been noted by Lor *et al.*[11] In such cases, the curve-fitting software (Genedata Screener) was unable to assign a pEC50 value; these compounds were assigned as active when %effect at any concentration tested >5 × standard deviation of the null control wells.

RZ and HRP assays show a reasonable correlation in the designation of redox activity (*Fig. 1*). The majority of compounds flagged as active by the HRP assay is also flagged by the RZ assay (95%); a small fraction of HRP inactives (8%) is flagged

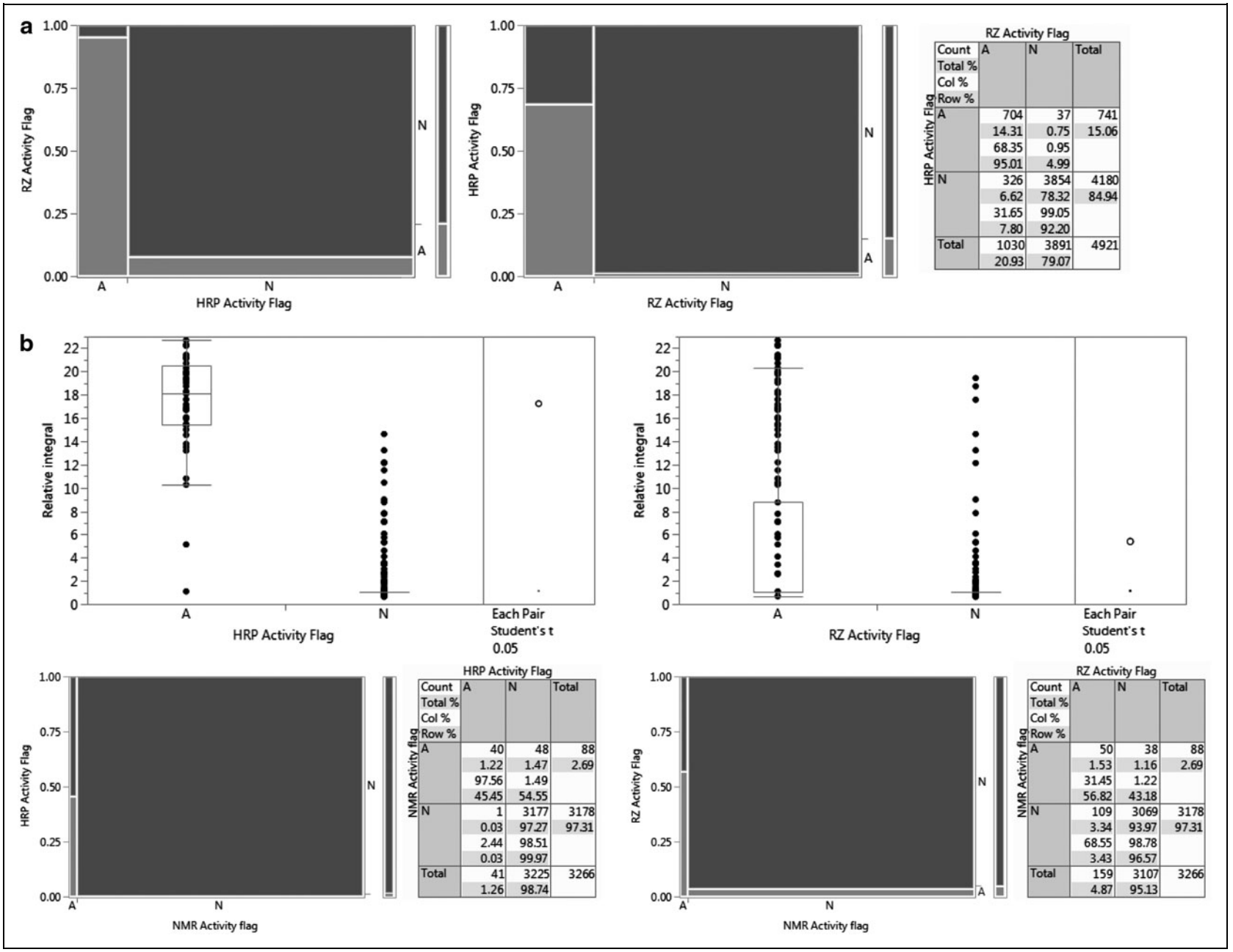

Fig. 1. HRP and RZ actives' counts in biochemical and NMR assays. A: active; N: not active. **(a)** Diagrams designating overlap between HRP and RZ actives (*left*, *middle*), and counts (on *right*); **(b)** Relative NMR integrals for compounds classed as active and inactive by HRP (on *left*) and RZ assays (on *right*). A compound is considered active in the NMR assay if its relative integral is greater than 1.1.

as active by the RZ assay. Of the RZ actives, approximately two thirds are designated active by the HRP assay; only 1% of the RZ inactives are labeled as active by the HRP assay.

The RZ assay shows a hit rate of 21% across the compounds tested, versus 15% in the HRP assay. This may either be a consequence of the difference in the way these assays detect redox behavior or be indicative of one assay having a greater sensitivity for false positives than the other. Because a positive readout in the HRP assay indicates formation of hydrogen peroxide, such a result strongly suggests a redox-cycling mechanism. Therefore, a compound active in the HRP assay through an H_2O_2-mediated mechanism must appear active in the RZ assay (which employs a measure of radical formation). Conversely, compounds that appear active in the RZ assay need not always show up in the HRP assay if their mode of action does not give rise to formation of H_2O_2. This seems to be supported by the data (*Fig. 1*). An interesting observation is that compounds are generally active in the RZ assay at 10- to 100-fold lower concentrations than in the HRP assay. This may reflect differences in assay setup as optimized in this study, for example, time in the presence of reducing agent and reducing agent concentration.

NMR Assay

To characterize the sensitivity and specificity of the RZ and HRP assays, we developed a novel 1D NMR assay with an accurate, oxidation-dependent readout orthogonal to the readouts of the biochemical assays. This assay detects the formation of the TCEPO, which arises for example, from the formation of hydrogen peroxide catalyzed by presence of the compound (*Fig. 2a*).[23]

The NMR redox assay takes advantage of the pH-dependent changes in the ^{1}H NMR spectra of TCEP and TCEPO. The assay is designed to be carried out at neutral pH, at which the ^{1}H spectrum of TCEP is highly broadened, while the TCEPO ^{1}H NMR spectrum presents two well-resolved multiplets (*Fig. 2* and *Supplementary Data*). Samples initially contain only TCEP and the compound of interest in the aqueous buffer. The assay requires no protein and little optimization is needed. However, since the redox mechanism is catalytic, it is important to minimize any variation in timing between preparation of the sample and NMR measurements. In our case, the samples were prepared by a sample preparation robot, which executes a pipetting script at a certain time point before start of the NMR experiment acquisition. After mixing, the sample was delivered

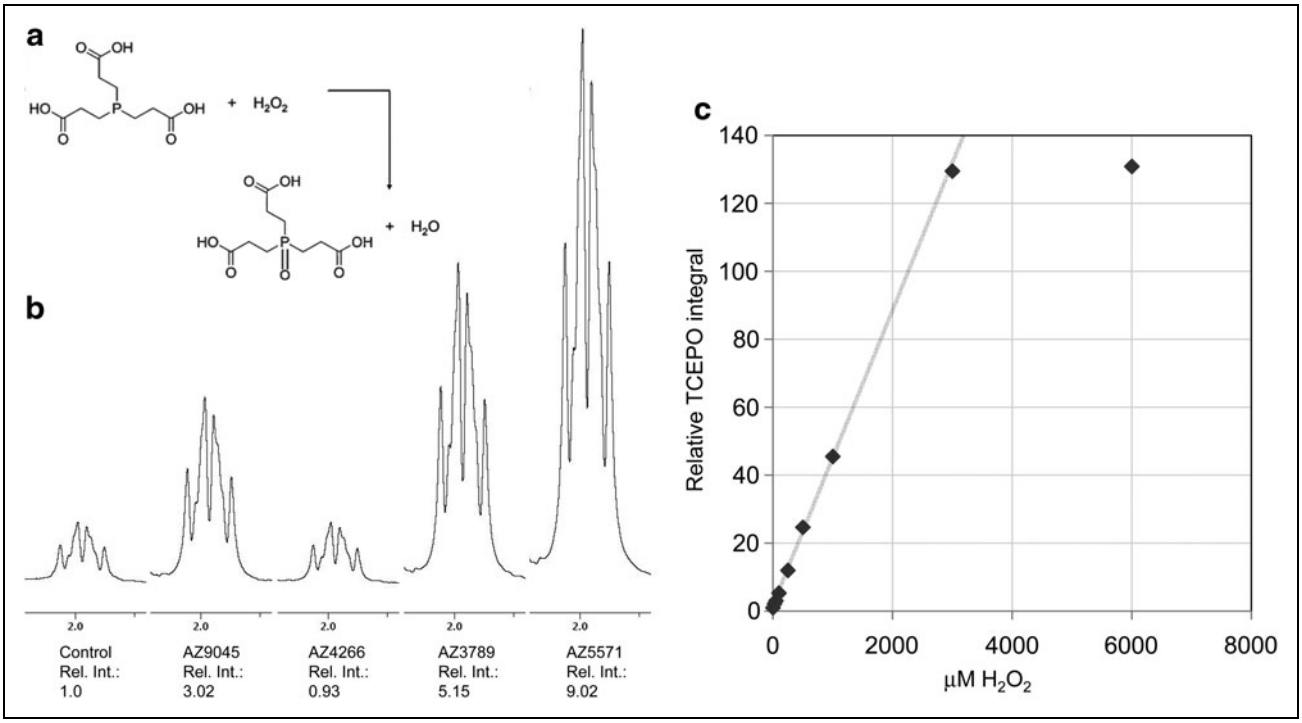

Fig. 2. **(a)** Conversion of TCEP to TCEPO by hydrogen peroxide. From Tan *et al.*[23] **(b)** Part of 1D ^{1}H NMR spectrum showing the high field multiplet signal from TCEPO. The TCEPO signal is integrated and normalized relative to a control sample lacking compound. Compound AZ4266 has no redox activity, whereas A9045, AZ3789, and AZ5571 are redox active and cause an increase in TCEP oxidation compared to the control. **(c)** Hydrogen peroxide calibration curve with peroxide concentration on *x*-axis and TCEPO integral relative to control on *y*-axis.

to the NMR spectrometer by a SampleRail sample delivery system. The automated setup eliminates timing variability between samples (see Materials and Methods section for details) and is superior to manual preparation and data acquisition in precision in timing.

The risk of false-positive results is low since the assay is based on direct observation of the TCEPO signal and there is little opportunity for interference by other components in the system. A false-positive risk could arise, however, from redox-active impurities in the compound samples that are not visible in the NMR spectrum, or from an overlap between compound and TCEPO multiplet signal, leading to overestimation of the integral. The presence of overlapping compound signals can easily be assessed by recording reference spectra of the compound of interest in the absence of TCEP.

The assay specifically detects redox-active compounds acting through the RCC mechanism since TCEP oxidation in the absence of a protein must proceed by production of hydrogen peroxide. The hydrogen peroxide calibration curve in *Figure 2c* shows that addition of up to 3 mM hydrogen peroxide to a sample results in a relative TCEPO integral, which is up to 130-fold larger than that in a sample containing no hydrogen peroxide. Further addition of hydrogen peroxide does not increase the TCEPO integral, indicating that this is the maximum relative integral, corresponding to full conversion of the 3 mM TCEP that was originally present in the sample. The largest value for the relative TCEPO integral observed for a compound in our experiments was 24, which suggests that 100 μM of added compound was able to generate around 500 μM hydrogen peroxide within the timeframe from sample mixing to the end of data acquisition.

Triaging of RZ and HRP Hits by NMR

A combined set of 3,266 compounds tested in RZ and HRP assays was tested in the NMR assay. Compounds were regarded as NMR redox active if they increased the amount of oxidized TCEP by at least 10% compared to the control (a relative integral fold difference of 1.1 or greater). Both the RZ and HRP assay were compared to the NMR redox assay (*Fig. 1b*).

The hit rates in the NMR, HRP, and RZ assays within the set of 3,266 amount to 3%, 1% and 5% respectively. Surprisingly, neither the RZ nor the HRP assay flags all the NMR actives—they flag 40 and 50 of 88 NMR-active compounds, respectively. Using the NMR relative integral signal as an indication of the degree of true redox-cycling behavior, we can assess whether this signal predicts the classification derived from the RZ and HRP assay data. *Figure 3a–d* contain receiver-operator curves (ROCs) depicting sensitivity (true-positive rate) and 1-specificity (1-true negative rate) for the RZ and HRP assays detecting NMR actives, as well as logistic plots for the same. The ROCs suggest that the classification of the HRP assay is much in line with the relative integral from the NMR assay (*Fig. 3a*), and the logistic regression graph (*Fig. 3c*) suggests that it flags a subset of nearly all NMR-active compounds over a relative integral threshold of approximately 12. The graphs for the RZ assay (*Fig. 3b, d*) paint a different picture and suggest that the RZ assay has a somewhat lesser ability than the HRP assay to discern true NMR positives from false positives (ROC *Fig. 3b*), apparently recognizing a subset of NMR actives across the lower range of relative integral values (regression plot *Fig. 2b*). The tables in *Figure 1b* suggest that the hit rate of the RZ assay could be about 100 times as large as that of the HRP assay (1 HRP active vs. 109 RZ actives that are not active in NMR), but it should be recognized that the RZ assay detects additional modes of action.

What we are really interested in is how the HRP and RZ assays can be used as high-throughput proxy to find those compounds that the NMR assay will flag as active. Using the HRP and RZ activity (% readout) at 100 μM as indication of activity, we construct the ROCs shown in *Figure 3e and f*, and logistic plots in *Figure 3g and h*. The ROCs suggest that neither assay is very predictive of compound activity in the NMR assay, and in particular for the RZ assay, false-positive activity affects the capability of ranking NMR-active compounds using the %-readout at 100 μM. In the case of the RZ assay, this results in a high hit rate. In summary, the HRP and RZ assay seem capable of picking up approximately half of the compounds that are active in the NMR assay, and the HRP assay appears to do this with the lowest incidence of false positives. Further work suggested that the hits picked up by the HRP and RZ assays, but not by the NMR assay, in part, arise from redox-mediated effects, as removal of reducing agent removes activity for some (results not shown). Neither the RZ nor HRP assay is suited for identifying redox cyclers exclusively.

Use of NMR Data in the MALT1 Hit Triaging Cascade

The NMR assay was trialed in a hit triaging cascade for a cysteine protease (MALT1). Before the start of the HTS campaign, NMR was used to evaluate MALT1 inhibitors observed in a 10,000-compound validation set. Thirty two of the hits resulting from this run were soluble enough to be tested in the NMR assay and 13 proved to be redox active. The large

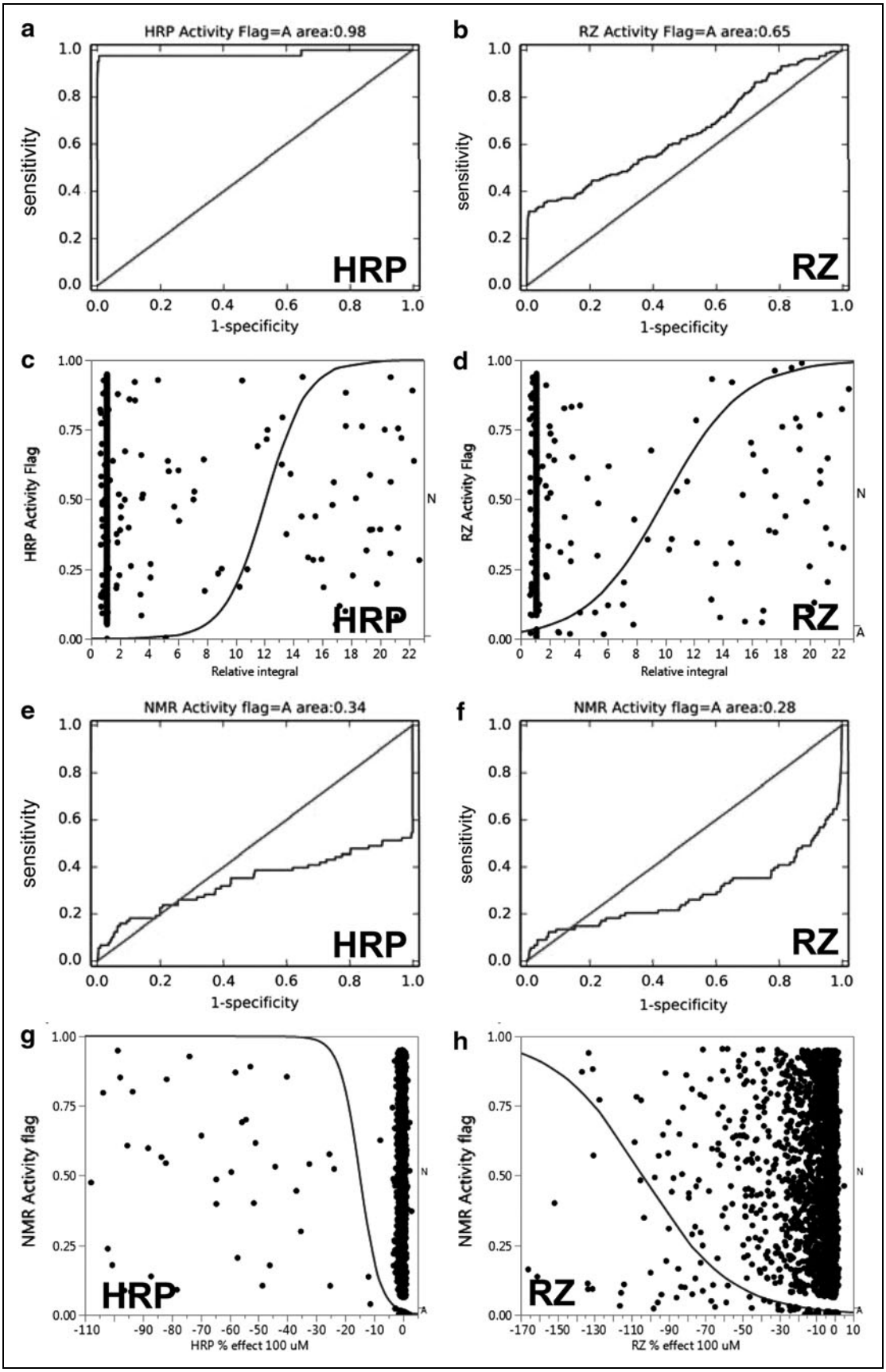

Fig. 3. Prediction of HRP and RZ assay classification by NMR signal. **(a, b)** Receiver-operator curves for relative NMR integral predicting the HRP and RZ activity flags; *red line* indicates random response. **(c, d)** Logistic regression plots for same. **(e, f)** Receiver-operator curves for %effect at 100 μM compound concentration predicting NMR activity, for HRP and RZ (maximum effect at –100%); **(g, h)** logistic regression plots for same.

proportion of redox actives among the hits from the validation set led to a change in the HTS strategy; a high-throughput RZ assay was included in the screening cascade. This work was performed before the large-scale comparison of results between the HRP and RZ assays. As such, at the time, the RZ assay was deemed the most suitable high-throughput approach for hit triage in that project. Of the MALT1 hits arising from the screen, 122 were measured in the NMR redox assay. Nine out of these 122 hits showed redox activity, that is, the proportion of false positives had decreased from 40% in the validation set to 7% in the final HTS output, which is, in part, explained by a relatively high content of "bad" compounds in the validation set. This demonstrates the value of screening validation sets for flagging potential problems, as well as the benefit of including redox artifact assays in the screening cascade.

Using Historical Data to Assess Anomalous Behavior

We derived a general frequent-hitter score (pBSF) to analyze hits for anomalous behavior other than redox. In short, the descriptor captures frequent-hitting behavior based on historical, in-house HTS data by comparing a compound's activity across many HTS campaigns to the expected activity for an "average" compound. Details are described elsewhere.[19] The score enables identification of compounds that may affect the assays through other modes of interference.

Further analysis of the HRP and RZ results suggests that a strong signal (% effect at 100 μM) in either assay is indicative of a compound showing frequent-hitter behavior across many HTS assays. Forty-six percent of HRP actives, 21% of RZ actives, and 50% of NMR actives were frequent hitters according to their generic pBSF scores. The high overall hit rate seen in the RZ assay combined with relatively low frequent-hitter incidence again suggests that this assay may have a high false-positive rate.

Comparison of Redox Activity in Subsets of Actives and Inactives

To compare the false-positive behavior of the HRP and RZ assays in more detail, we compared incidence of redox activity across a set of compounds confirmed active in the primary assay, and a set of compounds that showed no activity in the primary assay. Data in this case consisted of dose–response data, and were collated retrospectively across multiple assays following primary screening. Counts for these subsets can be found in the *Supplementary Data*. In the set of unconfirmed active compounds (*i.e.*, compounds more likely to be false positives), the incidence of RZ actives is 63%, and 51% for HRP actives, confirming that both assays identify similar numbers of compounds. The corresponding incidence of redox activity in the set of confirmed actives are 20% (RZ) and 11% (HRP). This makes sense as the confirmation of the target activity seen in the HTS for these compounds suggests they are better behaved and more likely to be true actives. Again, the hit rate of the RZ assay is shown to be higher compared with the HRP assay. This may be expected as the RZ assay will pick up multiple modes of action, whereas the HRP assay specifically detects formation of H_2O_2.

The remaining question is then around the nature of the compounds that are flagged as active by the HRP and RZ assays. Using our frequent-hitter descriptor, we assign all compounds as well or badly behaved. Latter compounds have shown activity in many HTS assays, and not necessarily as a result of redox activity. For the set of unconfirmed compounds, we see an enrichment of bad compounds in the RZ or HRP actives subset ($\sim$35% of the redox actives is classified as bad, vs. $\sim$10% of the compounds that are not redox active, and this is similar across HRP and RZ data). This suggests that both the RZ and HRP assays are sensitive to known frequent hitters. The fraction gives an indication of the false-positive rate being potentially as bad as 1 in 3 in this case, although it is interesting to note that this is similar across either HRP or RZ assay. Looking at the contingent of frequent-hitting compounds within this set, both assays can be seen to designate a large fraction of these as "redox-active" (HRP, 73%; RZ, 88%). This suggests a sensitivity of either assay to anomalous binders, as not all of these compounds will be redox-active.

Analysis of Historical Results for Redox-Susceptible Targets

To focus the historical data analysis further on modes of interference that are relevant for redox-sensitive proteins, we performed a data-mining analysis for a subset of targets from the following classes: cysteine proteases, metalloproteases, and protein tyrosine phosphatases. A set of 1,097 compounds found to be frequent hitters in redox-susceptible assays, but not in background assays was selected. The fragmentation and fragment enrichment analysis run on this set of likely redox-active compounds led to the selection of 42 representative enriched fragments. We tested for an increased incidence of frequent hitters for the groups of compounds associated to each fragment (*Table 1*; fragments in *Fig. 4*; see also *Supplementary Fig. S1*). A "Y" at the intersection of the row corresponding to group *i* and the column corresponding to the set of assays *j* indicate that an increased proportion of compounds from group *i* were found to be frequent hitters in set *j*. "N" indicates that no significant increase was detected.

Table 1. Results of the Binomial Tests Assessing Increased Proportion of Frequent Hitters in Groups of Compounds Representing Each Fragment

Fragment group	Number of compounds	Redox-susceptible assays				Background assays					Comments
		CP	MP	PTP	All RS	FI	FRET	Spec	Tech	Diverse	
1	1772	N	Y	N	N	N	N	N	N	Y	MP only
2	7252	N	Y	N	N	N	N	Y	Y	Y	MP only
3	151385	Y	N	N	Y	N	N	Y	Y	N	Ambiguous
4	6769	N	N	N	N	N	N	N	N	N	Not Redox
5	160085	N	N	N	Y	N	N	Y	N	N	Ambiguous
6	13665	N	N	N	N	N	N	N	N	N	Not Redox
7	9706	N	N	N	N	N	N	N	N	N	Not Redox
8	17548	N	Y	Y	Y	N	N	Y	Y	Y	Ambiguous
9	47	N	N	N	Y	N	N	N	N	N	Redox
10	259	N	N	N	N	N	N	N	N	N	Not Redox
11	86	N	N	N	Y	N	N	N	N	N	Redox
12	5844	N	N	N	N	N	N	N	N	N	Not Redox
13	141	N	N	N	N	N	N	N	N	Y	Not Redox
14	35	N	N	N	N	N	N	N	N	N	Not Redox
15	214	N	Y	N	Y	N	N	N	N	N	MP only
16	6	Y	N	N	Y	N	N	N	N	N	Redox
17	169	Y	N	N	Y	N	N	N	Y	Y	Ambiguous
18	39	Y	N	N	Y	N	N	N	N	N	Redox
19	4	N	N	N	N	N	N	N	N	N	Not Redox
20	2504	N	N	N	N	N	N	N	N	N	Not Redox
21	425	N	N	N	N	N	N	N	N	N	Not Redox
22	804	N	N	N	N	N	N	N	N	N	Not Redox
23	275	N	N	N	N	N	N	N	N	N	Not Redox
24	979	N	N	N	N	N	N	N	N	N	Not Redox
25	230	N	N	N	N	N	N	N	N	N	Not Redox
26	218	N	N	N	N	N	N	N	N	Y	Not Redox
27	183	N	Y	N	Y	Y	N	N	Y	Y	MP only
28	165	Y	N	N	N	N	N	N	N	N	Redox
29	2	N	N	N	N	N	N	N	N	N	Not Redox
30	2	N	N	N	N	N	N	N	N	N	Not Redox
31	6	N	N	N	N	N	N	N	N	N	Not Redox

(*continued*)

Table 1. *Continued*

Fragment group	Number of compounds	Redox-susceptible assays				Background assays					Comments
		CP	MP	PTP	All RS	FI	FRET	Spec	Tech	Diverse	
32	11	Y	N	Y	Y	N	N	N	N	N	Redox
33	3	Y	N	Y	Y	N	N	N	N	N	Redox
34	2	Y	N	N	N	N	N	N	N	N	Redox
35	10053	N	N	N	N	N	Y	N	N	N	Not Redox
36	1553	N	N	N	N	N	N	N	N	N	Not Redox
37	38	N	N	N	N	N	N	N	N	N	Not Redox
38	1971	N	N	N	N	N	Y	N	N	N	Not Redox
39	21	N	N	N	N	N	N	N	N	N	Not Redox
40	419	N	N	N	N	N	N	N	N	N	Not Redox
41	46	N	N	N	N	N	N	N	N	N	Not Redox
42	5	N	N	N	N	N	N	N	N	N	Not Redox

The table reports "Y" when the test is positive (increased proportion of frequent hitters) and "N" when the test is negative.

CP, cysteine proteases; MP, metalloproteases; PTP, protein tyrosine phosphatases; All RS, all redox-susceptible targets; FI, fluorescence intensity assays; FRET, Förster Resonance Energy Transfer; Spec., spectrophotometry; Tech., technology.

Surprisingly, most fragments whose corresponding group of compounds exhibited an increased incidence of frequent hitters for cysteine proteases did not show an increased incidence of frequent hitters for metalloproteases. A study of the assay protocols revealed that only one out of the nine metalloproteases assays used a reducing agent, while eight out of the nine cysteine proteases assays and all three of the protein tyrosine phosphatases assays did. In the absence of a reducing agent, RCCs are not expected to inhibit the redox-susceptible targets. Thus, the analysis of metalloproteases can be regarded as a negative control and it highlights that inclusion of a reducing agent in the assay is required to see redox-cycling frequent-hitter behavior.

Four groups showed an increased incidence of frequent hitters for metalloproteases, but not for other redox-susceptible assays. Given that eight of the nine metalloprotease assays did not contain a reducing agent, the corresponding fragments were not considered likely to give rise to redox-cycling properties. Twenty-six groups did not show a statistically significant increase in the incidence of frequent hitters for any set of redox-susceptible assays.

Eight groups exhibited an increased proportion of frequent hitters for redox-susceptible targets, but not for any background set of assays (fragments 9, 11, 16, 18, 28, 32, 33, and 34). These fragments are thus likely to convey redox-cycling properties and their compounds were kept for further experimental testing. The four remaining groups showed an ambiguous profile. They exhibited an increased incidence of frequent hitters both for sets of redox-susceptible assays and for background sets of assays, suggesting that these could act through redox cycling or another mechanism. Their corresponding compounds were also kept for experimental testing. It is important to emphasize that the presence of such fragments alone is not a sufficient condition to make a compound an RCC, but our analysis indicates that such compounds are more likely to display redox-cycling behavior.

For each of the 8 + 4 fragment groups selected for further experimental testing, a few representative compounds were tested in the HRP, RZ, and NMR redox assays. Selection took into account three criteria: pBSF value for all susceptible targets, diversity of structures, and purity. We prioritized compounds with sample purity over 80%, with diverse structures, and with a high pBSF in the set of redox-susceptible assays. All were tested in HRP and RZ assays, and a subset of compounds was tested in the NMR redox assay. Results can be found in *Table 2*.

The results from the *in silico* analysis are mostly confirmed experimentally. For five out of eight compound groups selected

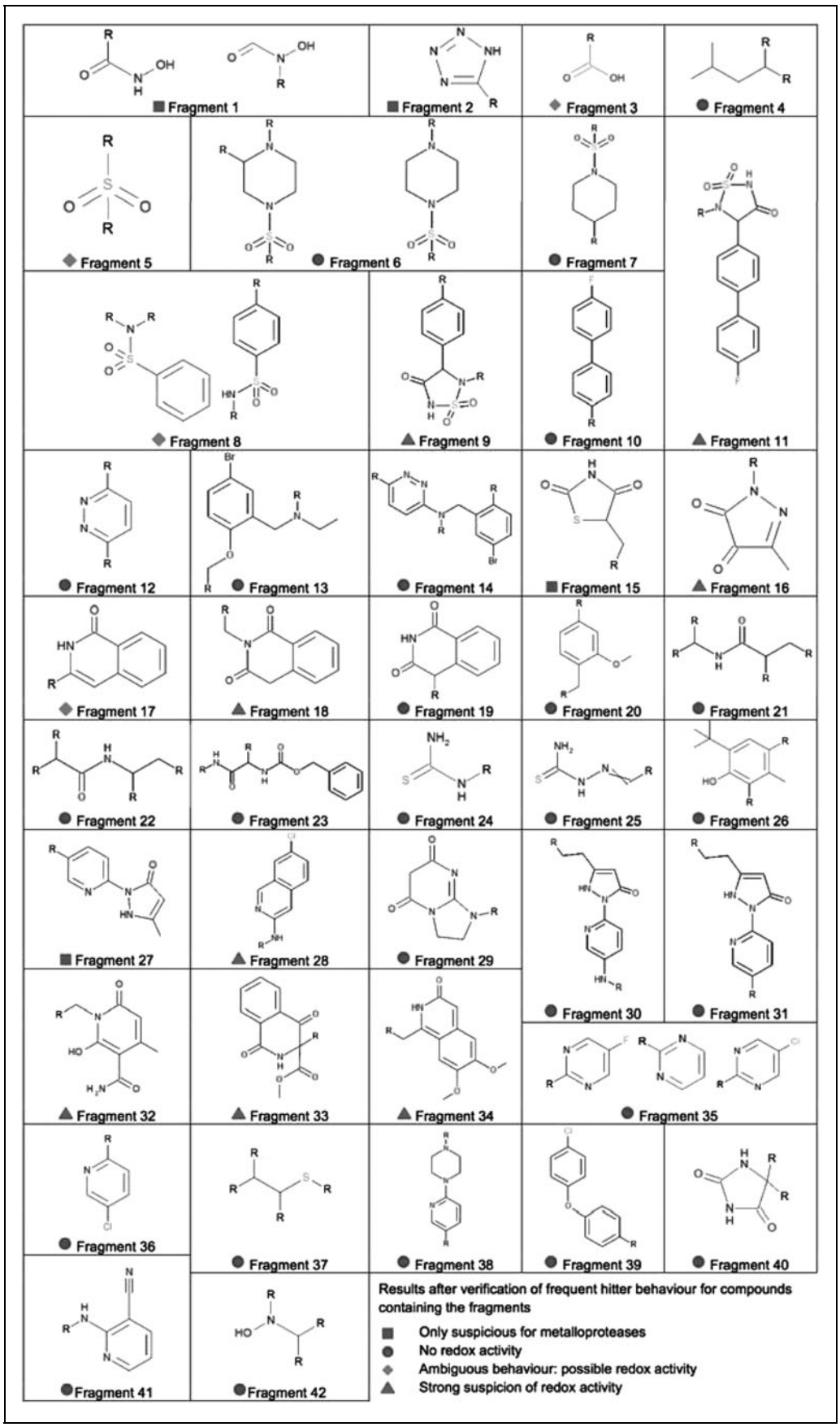

Fig. 4. Manually selected fragments representative of the enriched fragments in S. Numbers correspond to those in *Table 1*. Redox fragments are those assigned as redox in *Table 1*. See *Supplementary Data* section "Fragment selection" and *Figure S1* for further details.

Table 2. Experimental Results in the RZ, HRP, and NMR Redox Assays for the Fragments Identified Using Our *In Silico* Analysis

In silico results	Fragment	Number of compounds tested in		Active compounds in							
				RZ	HRP	NMR			At least one assay	At least two assays	All assays
		RZ and HRP	NMR	A	A	wA	mA	sA			
"Redox": Increased incidence of frequent hitters in redox susceptible but not in background	9	10	8	0	0	1	1	0	2	0	0
	11	10	10	1	0	1	1	0	3	0	0
	16	5	3	3	3	1	0	2	5	3	1
	18	10	8	10	10	0	0	8	10	10	8
	28	10	9	10	8	0	4	5	10	10	7
	32	6	4	0	2	0	2	1	4	1	0
	33	3	2	3	3	0	1	1	3	3	2
	34	1	1	1	1	0	0	1	1	1	1
"Ambiguous" Increased incidence of frequent hitters in redox susceptible and in background	3	10	7	3	0	3	3	0	6	3	0
	5	10	6	0	0	3	1	0	4	0	0
	8	10	8	1	0	2	1	0	3	1	0
	17	10	7	6	6	1	0	5	7	6	5
"Redox-clean"[a]	–	71	71	2	0	0	0	0	2	0	0

[a]Redox-clean compounds refer to compounds tested in all twelve redox-susceptible HTS assays employing reducing agent and inactive in all of them.
A, active; wA, weakly active; mA, medium active; sA, strongly active (activity categories are defined in the Results and Discussion for each assay).

to represent fragments that are suspected to cause redox effects by *in-silico* methods, at least one compound was active in all three redox assays. The "ambiguous" category is still ambiguous experimentally and suggests disagreement between the NMR, HRP, and RZ assay results. One fragment showed strong activity in all three assays, while the others only show activity in the RZ or NMR redox assays.

The "redox-clean" compounds included as control did not show any activity in redox-susceptible HTS assays. Any positive result in the redox assays can therefore be considered a false positive. False positives do not appear to be a problem in the HRP and NMR assays since these assays do not detect any activity in this set of compounds. The RZ assay flags two compounds from this set as active, which may be compounds containing impurities or compounds that are unstable and reactive.

The HRP assay only shows activity for groups 16, 18, 28, 32, 33, 34, and 17. Except for group 32, all these groups exhibit high activity rates and show activity on at least one occasion in all three redox assays, and these observations suggest strong redox-cycling properties for these compounds.

The remaining groups showed no HRP activity. At least 30% of compounds from groups 9, 11, 3, 5, and 8 were not active in any assay. Where possible, compounds selected for experimental testing had high pBSF for redox-susceptible targets. For groups with a lot of representative compounds such as groups 3, 5, and 8, this means that compounds with high pBSF have been chosen. The absence of activity in all redox assays thus suggests alternative frequent-hitter mechanisms for compounds harboring these motifs. For groups 3, 5, and 8, a few compounds may be redox cycling (as confirmed by the redox assays), but different frequent-hitter mechanisms may feature here, as suggested by the *in silico* analysis. For groups 9 and 11, the activity rates in the NMR redox assays are low, but not zero, and this assay did not seem prone to give false negatives. We conclude that each of these fragment sets contain RCCs that are not detected by the HRP assay. It is likely that these groups have very weak redox-cycling properties only detectable in the NMR assay. The absence of RZ-active compounds for cluster nine suggests that the RZ assay is prone to give false-negative results for this category.

The results confirm that the HRP assay is susceptible to false negatives when the redox-cycling activity is weak, but that it rarely gives false positives. The NMR and RZ assays both detect more RCCs than the HRP assay, and they both successfully identify compounds from clusters identified as redox active *in silico*, where the HRP assay fails. However, results for the RCCs still suggest substantial false-negative rates for the RZ assay.

In summary, we have analyzed the capacity of RZ, HRP, and a novel NMR-based redox assay to flag suspicious compounds. The former two biochemical assays are considered high throughput, whereas the NMR assay is medium throughput. Overall hit rates for these assays within a set of HTS actives are different, 1% for the HRP assay, 3% for the NMR assay, and 5% for the RZ assay. Surprisingly, neither the RZ nor the HRP assay flags all of the NMR actives—they identify 40 and 50 of 88 NMR-active compounds, respectively, and may be related to sensitivity differences arising from differently optimized assay readouts.

Computational analysis enabled assessment of the ability of each assay to identify redox-cycling frequent-hitter compounds (RCCs), and the results suggest that all three assays are capable of identifying compounds that frequently hit RCC-susceptible targets in HT screens. This suggests that even if the primary reason for activity is not pure redox-cycling behavior, the flagged compounds are generally likely to misbehave and of lesser interest. Historical HTS data proved to be a valuable source of information with regard to the detection of frequent-hitter compounds acting through a specific mechanism. In addition to identifying previously identified problematic RCCs, the computational analysis run on these data generalizes the results and reveals redox-cycling fragments that identify compound sets enriched in redox-active behavior. Hits from redox-susceptible screens that contain such fragments can thus be flagged as suspicious, regardless, when historical data for these compounds are limited.

Using observed behavior in historical HTS data of redox-sensitive and non-redox-sensitive target assays, we designated sets of "bad" and "suspicious" RCCs, as well as "clean" compounds. The HRP, NMR, and RZ assays identified 76%, 84%, and 80% of "bad" RCC frequent hitters, respectively, and 54%, 68% and 73% of "suspicious" RCC frequent hitters, respectively. They also identified 11%, 20%, and 17% of the "clean" set, respectively, which may be false-positive results, or due to these compounds not having been tested enough times to be identified as problematic by their data yet.

Analysis of anomalous-binder behavior of compounds in a wider set of HTS assays suggests that 46%, 50%, and 21% of HRP, NMR, and RZ actives are frequent hitters that can act by mechanisms other than redox, and that compounds flagged in the RZ and HRP assays may be undesirable still for reasons other than redox-cycling behavior. Closer inspection of a set of "compounds-of-interest" (confirmed-active and confirmed-inactive compounds for the primary target) yields numbers in the same range, but also suggests that both HRP and RZ assays behave similarly in terms of flagging frequent hitters. We conclude that the NMR redox assay offers a novel and reliable way of identifying RCCs at a medium throughput, with the HRP assay being a straightforward higher-throughput option with reasonable recognition and low false-positive rate. The RZ assay has a similar ability to recognize compounds active in the NMR assay as the HRP assay, but its higher overall hit rate necessitates further deconvolution of output.

DISCLOSURE STATEMENT

All authors were employees of AstraZeneca when working on this project.

REFERENCES

1. Drews J: Drug discovery: a historical perspective. *Science* 2000;287:1960–1964.
2. Wigglesworth MJ, Murray DC, Blackett CJ, Kossenjans M, Nissink JW: Increasing the delivery of next generation therapeutics from high throughput screening libraries. *Curr Opin Chem Biol* 2015;26:104–110.
3. Thorne N, Auld DS, Inglese J: Apparent activity in high-throughput screening: origins of compound-dependent assay interference. *Curr Opin Chem Biol* 2010; 14:315–324.
4. Feng BY, Simeonov A, Jadhav A, *et al.*: A high-throughput screen for aggregation-based inhibition in a large compound library. *J Med Chem* 2007; 50:2385–2390.
5. Roche O, Schneider P, Zuegge J, *et al.*: Development of a virtual screening method for identification of "frequent hitters" in compound libraries. *J Med Chem* 2002;45:137–142.
6. Baell J, Walters MA: Chemistry: chemical con artists foil drug discovery. *Nature* 2014;513:481–483.
7. Soares KM, Blackmon N, Shun TY, *et al.*: Profiling the NIH Small Molecule Repository for compounds that generate H2O2 by redox cycling in reducing environments. *Assay Drug Dev Technol* 2010;8:152–174.
8. Bova MP, Mattson MN, Vasile S, *et al.*: The oxidative mechanism of action of ortho-quinone inhibitors of protein-tyrosine phosphatase alpha is mediated by hydrogen peroxide. *Arch Biochem Biophys* 2004;429:30–41.
9. Brisson M, Nguyen T, Wipf P, *et al.* Redox regulation of Cdc25B by cell-active quinolinediones. *Mol Pharmacol* 2005;68:1810–1820.
10. Smith GK, Barrett DG, Blackburn K, *et al.*: Expression, preparation, and high-throughput screening of caspase-8: discovery of redox-based and steroid diacid inhibition. *Arch Biochem Biophys* 2002;399:195–205.
11. Lor LA, Schneck J, McNulty DE, *et al.*: A simple assay for detection of small-molecule redox activity. *J Biomol Screen* 2007;12:881–890.
12. Lal M, Rao R, Fang XW, Schuchmann HP, von Sonntag C: Radical-induced oxidation of dithiothreitol in acidic oxygenated aqueous solution: a chain reaction. *J Am Chem Soc* 1997;119:5735–5739.
13. Johnston PA, Soares KM, Shinde SN, *et al.*: Development of a 384-well colorimetric assay to quantify hydrogen peroxide generated by the redox cycling of compounds in the presence of reducing agents. *Assay Drug Dev Technol* 2008;6:505–518.

14. Sameshima T, Miyahisa I, Yamasaki S, Gotou M, Kobayashi T, Sakamoto J: High-throughput quantitative intrinsic thiol reactivity evaluation using a fluorescence-based competitive endpoint assay. *SLAS Discov* 2017:2472555217704654.
15. Wigle TJ, Swinger KK, Campbell JE, *et al.*: A high-throughput mass spectrometry assay coupled with redox activity testing reduces artifacts and false positives in lysine demethylase screening. *J Biomol Screen* 2015;20:810–820.
16. Krezel A, Latajka R, Bujacz GD, Bal W: Coordination properties of tris(2-carboxyethyl)phosphine, a newly introduced thiol reductant, and its oxide. *Inorg Chem* 2003;42:1994–2003.
17. Hwang TL, Shaka AJ: Water suppression that works—excitation sculpting using arbitrary wave-forms and pulsed-field gradients. *J Magn Reson Ser A* 1995;112: 275–279.
18. Dalvit C: Efficient multiple-solvent suppression for the study of the interactions of organic solvents with biomolecules. *J Biomol Nmr* 1998;11:437–444.
19. Nissink JWM, Blackburn S: Quantification of frequent-hitter behavior based on historical high-throughput screening data. *Future Med Chem* 2014;6: 1113–1126.
20. Berman HM, Westbrook J, Feng Z, *et al.*: The protein data bank. *Nucleic Acids Res* 2000;28:235–242.
21. Hollander M, Wolfe, DA, Chicken, E: *Nonparametric Statistical Methods*. John Wiley & Sons: New York, 1973, pp. 15–22.
22. Holm S: A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics* 1979;6:65–70.
23. Tan Z, Ihnat PM, Nayak VS, Russell RJ: Quantitative analysis of tris(2-carboxyethyl)phosphine by anion-exchange chromatography and evaporative light-scattering detection. *J Pharm Biomed Anal* 2012;59:167–172.

Address correspondence to:
J. Willem M. Nissink, PhD
Oncology IMED
AstraZeneca
Darwin Building
310 Cambridge Science Park
Milton Road
Cambridge CB4 0WG
United Kingdom

E-mail: willem.nissink@astrazeneca.com

Abbreviations Used

FI = fluorescence intensity
FRET = Förster resonance energy transfer
HRP = horseradish peroxidase
HTS = high-throughput screening
RCC = redox-cycling compound
ROC = receiver-operator curves
RZ = resazurin
TCEP = tris(2-carboxyethyl)phosphine
TCEPO = P-oxide of TCEP