HIGH-THROUGHTPUT SCREENING QUALITY CONTROL GENERAL GUIDELINES

EU-OPENSCREEN ERIC (EU-OS), the European Research Infrastructure Consortium of Open Screening Platforms for Chemical Biology, builds a distributed organization of national screening and chemistry facilities, a common database, and a central headquarter that manages a joint compound collection and coordinates project flow and training. It provides world-class services to academia and industry in the fields of small-molecule screening and medicinal chemistry.

To ensure that data produced at different sites are comparable and reproducible, a set of common operational standards for the EU-OPENSCREEN screening sites have been defined. While the operational standards need to cover several aspects of the screening, they are here kept to a minimum in order to avoid becoming a burden to the screening sites.

The following guidelines are based on the "Report on screening standards to be used" (Deliverable 11.7, October 2013, Grant Agreement number: 261861) and regulate the following aspects:

1. **DMSO Tolerance**

   DMSO related assay sensitivity should be determined before starting the assay. Minimum DMSO tolerance is defined as the DMSO concentration that does not negatively affect assay readout (i.e. enzyme activity or cell viability) in terms of signal variation or standard deviation. As a general rule, the assay system should have a minimum DMSO tolerance of 1% for biochemical screens and of 0.1% for cell-based screens.

2. **Z’-factor**

   Minimum Z´ factor cut-off (Z´ > 0.5) (*Zhang JH, J Biomol Screen 1999,* DOI: 10.1177/108705719900400206) is applied per default for reader-based

biochemical and cellular screens. A Z’ of <0.5 can be accepted for cellular assays if total number of hits can be validated in follow-up assays, i.e demonstrated by screening the pilot library. When Z´ cut-off is not applicable, an alternative method to assess whether the response in an assay is significant can be defined. The alternative parameter has to be defined up front and approved by the screening site. Plates with a Z' <0.5 during a screen will be rejected. However, if well-argued and accepted by the screening site, plates with Z’ <0.5 could be used as long as this is registered in the experimental report.

3. **Coefficient of variation**

When Z’-factor cannot be applied, data variability will be estimated using the coefficient of variation (CV%). CV% refers to the sample population of a screening plate. CV% lower than 20% is suggested.

4. **Acceptance criteria**

Priority will be given to Z’-factor and CV%. When a sufficient number and range of XC50 values can be measured, Minimum Significant Difference (MSD) will be used as additional parameter for assessing assay reliability. MSD represents the smallest efficacy difference between two compounds that is statistically significant (*https://www.ncbi.nlm.nih.gov/books/NBK83783/*). MSD lower than 20 is suggested. If the use of MSD is not applicable, inter-day variation CV may be used instead.

5. **Pilot screen**

A pilot screen with a sub-library comprising 5000 compounds is performed to validate assay performance and robustness. The pilot screen serves to establish the assay protocol, data evaluation and the test for robustness of the assay against compound-induced measurement artifacts (as frequently caused by cytotoxicity, auto-fluorescence or aggregation). In the case of very low hit rates the screening campaign is stopped (assay sensitivity too low). In the case of high hit rates above 1%, (assay specificity too low), the compound concentration can be reduced or the activity threshold for a hit can be changed.

6. **Positive and negative controls for primary screening**

When available, the following number of controls will be used:

- 4 positive + 4 negative / 96-well plate
- 16 positive + 16 negative controls / 384-well plate
- 32 positive + 32 negative controls / 1536-well plate

Negative controls must consist of the reactions with no compound (DMSO only) and must be present on each testing plate. Ideally, control compounds are used for generating the positive control. Only in the absence of any known compound standards or if the effect on assay readout is unstable, other measures for generating the positive controls are acceptable (e.g. knock-out cell lines for cellular assays, or wells without enzyme for biochemical assays).

7. **Positive and negative controls for IC50 determination**

For complete evaluation of IC50 data, maximum and minimum signal controls must be performed along with the inhibitor titration. For example, in inhibitor-based studies, if control compounds are not available, maximum controls (100% enzyme activity, 0% inhibition) should consist of enzyme reactions with no inhibitor (DMSO only). Minimum signal controls should be performed by using DMSO only (no inhibitor) and leaving enzyme out of the reaction (adding just buffer instead) to represent 0% enzyme activity, 100% inhibition.

8. **Failure rate**

After initial quality control, no more than 20% of the measured plates shall have to be remeasured. An initially failed plate can be repeated, but to a maximum of 3 repetitions.

9. **Reagent stability**

For biochemical assays, reagents (e.g. proteins, enzymes, buffers) shall be stable in the experimental conditions tested for the entire duration of the run using these reagents. Reagents stability should be proven before the beginning of the experiment.

10. Cell line quality control

Quality and integrity of cell lines used for drug screening should be proven through short tandem repeat (STR) profiling (https://www.atcc.org/en/Products/Cells_and_Microorganisms/Testing_and_Characterization/STR_Profiling_Analysis.aspx).

11. Correction of plate patterns and outlier correction

- Intra-plate normalisation to eliminate instrument and edge effects in standard assays (like b-scores) will not be applied. Hardware-based correction for spatial patterns might be acceptable. Intra-plate normalisation to reduce plate patterns will be allowed only in special types of assays/readouts and it should be justified before the screening campaign begins.
- Outlier correction will not be allowed for single dose assays. All data have to be uploaded on the ECBD and if necessary flagged as "inconclusive".

Regarding control wells, outliers (3 SD) can be removed as follow: up to 1 positive OR 1 negative control in the 96 well plate format, up to 2 positive AND 2 negative controls in the 384-well plate format and up to 4 positive and 4 negative controls in the 1536-well format (or in general, 1/8 of the control wells). However, reasonable activity ranges can be applied for hit definitions (e.g. no activities below negative controls) when creating the activity-call column for the ECBD database upload. Outlier correction is allowed for dose response assays. Outlier values removed for XC50 calculation should be uploaded together with the accepted data points, and correspondingly be flagged as outliers.

12. Hit rate

The acceptable range for the hit rate will be defined at the beginning of each screening campaign in agreement between user and screening site. Whenever possible a counter-screen should be identified and in case of elevated hit rates, the screening site may insist in the establishment of a counter-screening protocol.

The following **formulas** and **definitions** will be used.

a. The parameters for Z prime calculation are the standard deviations of the positive (σp) and negative (σn), and means of the positive (μp) and negative (μn) controls.
   The equation: **Z´ factor** = $1 - (3(\sigma^p + \sigma^n))/(|\mu^p - \mu^n|)$

   The equation for **standard deviation**: $\sigma = \sqrt{\frac{\sum(x-\mu)2}{N}}$, where σ = standard deviation, x = each value in the population, μ = mean of the values, N = number of values

b. **Coefficient of variation:** CV= $\frac{\sigma}{\mu}$

c. **Single dose active compound**: average plus/minus three standard deviations. Special assays might benefit from deviating criteria. Alternative criteria have to be defined up front and approved by the screening site.

d. **XC50**= half maximal response concentration (IC50, EC50 or AC50)

e. **Dose response active compound**: an XC50 value lower than the highest tested dose

f. **XC50 values will be calculated according to the following definitions:**
   - Minimum of 8 data points (concentrations) are used for calculating an XC50.
   - Minimum of 2 replicas for each experimental data will be used.
   - Dose response curve fitting is done using the four-parameter logistic model as default.

     Equation:

$$y = Bot + \frac{Top - Bot}{1 + 10^{(\log_{10}(IC50) - \log_{10}(x)) * slope}}$$

     Bot: minimal response

     Top: maximal response

IC50: concentration of drug at the symetric inflection point

slope: slope of the tangent to the curve when the concentration is IC50 (Hill-coefficient).

x: concentration

y: response

If the positions of IC50 and x are interchanged in the formula, the sign of the slope is changed.

For each assay, limitations of parameters are specified.

- An XC50 must be within the tested range - no extrapolation is allowed. XC50s outside the tested range are specified as: >[highest tested dose] (inactive) or "<[lowest tested dose]"
- XC50 error will be reported as error range (95% confidence interval) and fitting quality using R square.
- An XC50 must have a minimum visible activity change (activity difference between highest and lowest tested dose, determined using the fitted curve) of at least 25% in case of IC50. In case of EC50, this value has to be defined in the biological context of the assay. 25% is a minimal visible activity change, but other values >25% can be set for specific experiments and reported along with EC50 data.

General Definitions

EU-OPENSCREEN ERIC COMPOUND COLLECTION (EU-OS ERIC COMPOUND COLLECTION) is a screening collection comprising COMMERCIAL COMPOUNDS and/or ACADEMIC COMPOUNDS.

SCREENING PARTNER SITES (SPS) are the research institutes that provide the respective experimental facilities for compound screening and bioprofiling of compounds contained in the EU-OPENSCREEN COMPOUND COLLECTION. There are two (2) types of SPS, namely Screening (high capacity screening or specialized screening) and Bioprofiling Partner Sites (which are also Screening Partner Sites at the same time) which provide these facilities and the required staff.

The **EUROPEAN CHEMICAL BIOLOGY DATABASE (ECBD)** is EU-OPENSCREEN ERIC's open access database, in which structural information of commercial and proprietary compounds, BIOPROFILING results and primary screening data will be published.

USER shall mean any individual or its authorized representative, any legal entity or any organization utilizing COMPOUNDS for screening purposes through the SCREENING PARTNER SITE(S).

ASSAY(s) shall mean bioassay(s) in which a procedure is carried out containing experiments for determining the biological activity of COMPOUND(s) by measuring one or multiple effect(s) on a biomolecule, an organism, a tissue, a cell line or a biological model compared to control compounds.

PRIMARY ASSAY is the first assay performed in the screening campaign. The purpose of the primary assay is to identify primary hits, which are potentially biologically active chemical entities.

SECONDARY ASSAYS are the additional assays following the hit validation stage to confirm the biological activity of chemical entities via a different type of assay or to

eliminate certain active compounds based on their mechanism of action, toxicity or activity profile. SECONDARY ASSAYS can also include selectivity and specificity assays.

CONFIRMED HITS are compounds which were initially identified as hits in the primary assay and then retested at the same concentration (ideally in at least duplicates) in the same assay in order to exclude technical false positives.

COUNTER ASSAY is the assay run to eliminate those hits from the primary and confirmatory assay stages that are not of interest due to their artificial or non-selective activity

ASSAY READY PLATE is the screening plate containing a small aliquot of the compound to be screened, sufficient for a single ASSAY.

VALIDATED HIT shall mean term for a putative biological activity of a compound, validated by a concentration response curve and usually expressed as an EC50 (for stabilizers and activators) or an IC50 (for inhibitors) value. This is accompanied by lack of activity in assay-relevant counter screens and by independent confirmation of compound purity and identity.

DATA shall mean information output by any sensing device. In this present contract, it defines all information pertaining the physical and chemical property of a compound (for instance, mass spectra trace, molecular weight, etc.) and all physical-chemical or biological information originates from a) bioprofiling assessment and b) screening assays.