Experimental data evaluation

Step 1: Primary data processing, elimination of primary outliers

For each column of the two-dimensional dataset $\mathbf{d}_{ij}$, a box-and-whiskers plot is plotted as shown in Figure 1.

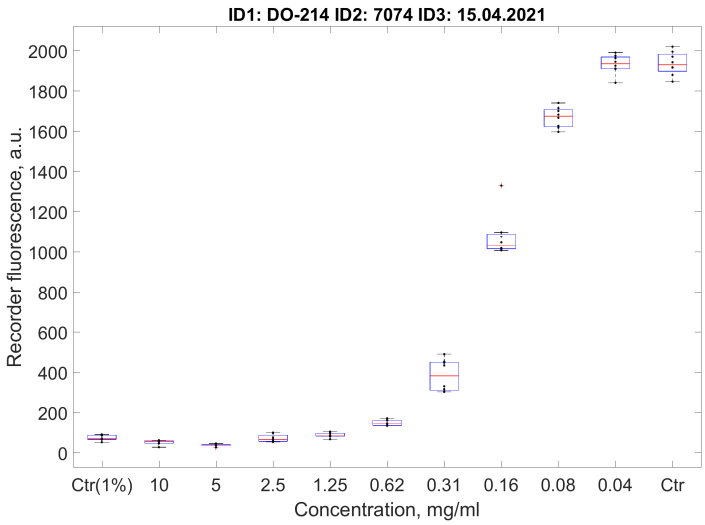


Figure 1: Experimental data for different concentrations (black dots) and corresponding boxplots (see parameters in the text); red crosses mark primary outliers.

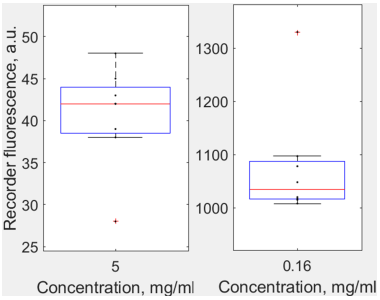
Box parameters are standard for graphical nonparametric methods of statistical data visualisation[1]: the top and bottom bars of the box mark the 1st and 3rd quartiles of the data (i.e. 0.25 and 0.75 quantiles when normalising the full probability to one) q_1 and q_3, respectively; the red line marks the median M, the limits of whiskers w_{up} and w_{down} are determined by the standard relations

$$w_{up} = q_3 + 1.5(q_3 - q_1), \tag{1}$$

$$w_{down} = q_1 - 1.5(q_3 - q_1) \tag{2}$$

between quartiles and interquartile intervals. Points outside the whiskers (checking $\mathbf{d}_{ij} > w_{up}$ or $\mathbf{d}_{ij} < w_{down}$ for each concentration column indicated by index j) are considered as outliers marked with red crosses in the graph and removed from the corrected dataset $\tilde{\mathbf{d}}_{\{i\}j}$ for further processing[2]. In the plot in Fig. 1, according to standard convention, the whiskers are trimmed at the corresponding minimum and maximum value of the remaining ones that are not outliers. Thus, the scatter plot allows for estimating the interval of the experimental values $\tilde{\mathbf{d}}_{\{i\}j}$ used in the subsequent analysis.

[1] J. M. Chambers, W. S. Cleveland, B. Kleiner, P. A. Tukey, Graphical Methods for Data Analysis, CRC Press (Boca Raton), 2018



[2] Here the curly braces of index $\{i\}$ denote a possible different value of the maximal index in different columns j.

Step 2: Determination of the discrete minimum inhibitory concentration (MIC) for the cleared sample

For the cleaned sample, the quartiles, the median, and the whiskers positions are estimated again by (1)–(2), but with new quartile values[3] The resulting "box-and-whiskers" diagram is shown in Fig. 2.

In addition, a confidence interval (*notches*) is calculated for each median using the formula [4].

$$\Delta_{up_j} = M_j + 1.57\left(q_{3j} - q_{1j}\right), \tag{3}$$

$$\Delta_{down_j} = M_j - 1.57\left(q_{3j} - q_{1j}\right) \tag{4}$$

The positions of the respective values are indicated in Fig. 2 as the centres of the green triangles with the vertices pointing downwards and upwards respectively. For the 1% control, these boundaries are extended over the whole study interval.

[3] Moreover, after removing large outliers from such a small sample, new outliers may appear due to shifting quartile boundaries, however, again due to the small sample size, there is no further redefinition of the median and the bounds.

[4] The multiplier's value equal to 1.57 corresponds to a 95 % probability of statistical insignificance of the median difference in normally distributed data; it is not interpreted directly for arbitrary samples, but the same value is accepted as the standard of significance for differences in this case too

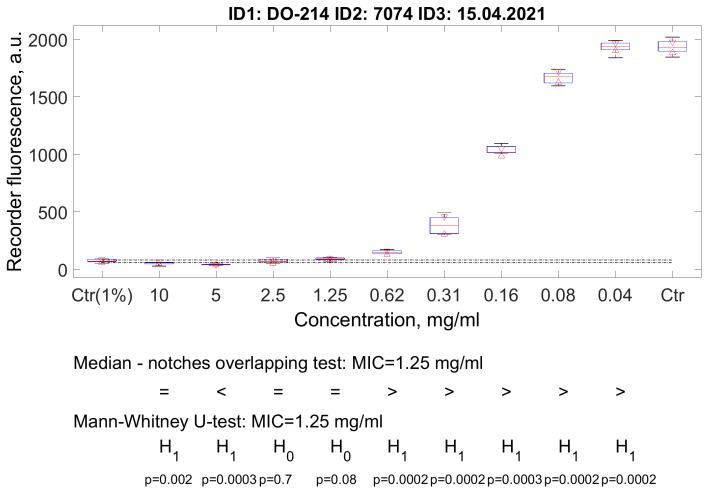


Figure 2: The boxplot of the experimental data with eliminated primary outliers showing the limits of the 1% control confidence interval (black dashed lines), below which for each concentration the significant coincidences or differences in medians compared with 1% control and the Mann-Whitney test comparison of these samples with the level of significance are indicated. For both methods, the corresponding MIC values are indicated.

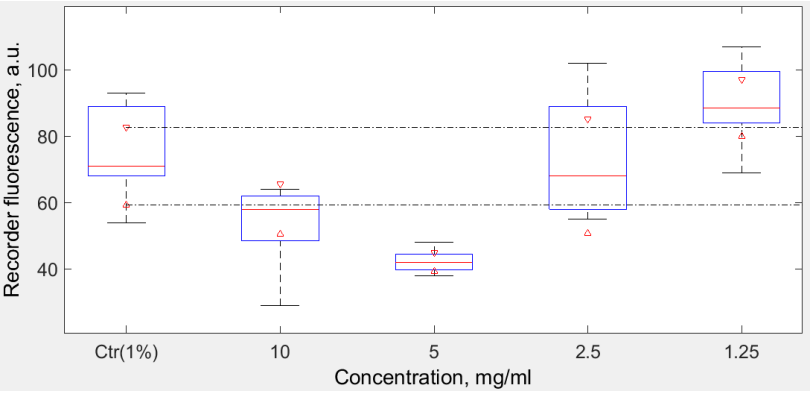


Figure 3: Enlarged range of large concentrations for a clearer demonstration of the overlapping ranges of the intervals for 1% control (dashed lines) and different concentrations (within the range bounded by the triangles).

To find the MIC, fluorescence intervals are compared with 1% control, checking for two conditions:

(i) whether there is a statistically significant difference between the fluorescence medians at this concentration and the 1% control;

(ii) if such statistically significant difference between them exist, which median is larger.

For (i), the automated calculation is carried out as follows: the lower limits of the confidence intervals for 1% control (Δ_{down1}) and the current concentration (Δ_{down1}, $j \leq 2$) are compared in pairs and the maximal of these is found,

$$\Delta_{maxj} = \max\left\{\Delta_{down1},\, \Delta_{down1}\right\}, \tag{5}$$

the upper limits of the confidence intervals for 1% control (Δ_{up_1}) and current concentration (Δ_{up_j}, $j \leq 2$) and the minimal of these is found,

$$\Delta_{minj} = \min\left\{\Delta_{up1},\, \Delta_{up_j}\right\}, \tag{6}$$

then the results are compared between (5) and (6). If $\Delta_{minj} > \Delta_{maxj}$, then the difference in medians is treated as statistically insignificant, with a "=" notation under the corresponding concentration in Fig. 2. Otherwise, the ratio of values of statistically different medians is checked: if $M_j < M_1$, that is, the median of 1% control is larger, then the mark "<" is indicated, otherwise the mark ">" is indicated. The MIC for a discrete set of concentrations is defined as the concentration corresponding to the highest index for which "=" or "<" are defined, that is, the median is either statistically the same as or less than the median of the 1% control.

As an additional test, the Mann-Whitney U-test was used to compare distributions of experimental data for 1% control and different concentrations as a nonparametric test applicable for randomly distributed comparing statistical differences of mediansciteConover1998book that replaces the two-sided t-test for means, which is defined only for normally distributed data. The corresponding labels are shown in Fig. 2: H_0 means that the data compared correspond to samples with equal medians at the 5% confidence level ($p > 0.05$), and H_1 indicates their statistically significant difference[5]; the corresponding significance levels found in the U-test are also indicated under each conclusion. In this case the MIC is defined as the maximal of the indices at which there is denoted either H_0 or H_1 with the median at this concentration being less than the median of the 1% control. In Fig. 2, both MICs are the same.

[5] The difference in the findings of the two tests for concentrations of 10 mg/ml is due to the fact that in the confidence interval intersection method the median deviations from the latter are symmetrical, but the real sample is biased downwards, as seen in Fig. 3: although the upper triangle falls between the dashed lines, the "boxes" themselves do not overlap at all and extend in opposite directions to the median.

Approximation of the fluorescence response curve to the active ingredient with a continuous curve

Biophysical background

The discrete values shown in Fig. 2 follow the dependence of the fluorescence level on the logarithm of concentration, which can be approximated by a sigmoidal function. It is convenient to go to dimensionless normalised values

$$m = \frac{M - M_1}{M_{contr} - M_1}, \tag{7}$$

where M_1 is the fluorescence value of 1% of control (i.e. the minimum detectable value), and M_{contr} is the fluorescence value of pure culture (the maximum detectable value).

Thus, $m(\ln(C))$ varies from 0 to 1, and we take the shifted hyperbolic tangent as its approximation[6].

$$m = \frac{1}{2}\left[1 - \tanh\left(\frac{1}{2}\alpha\left(\ln(C) - \ln(C_{0.5})\right)\right)\right], \tag{8}$$

where α and $C_{0.5}$ are constants to be found.

Eq. (8) can be transformed[7] to the form

$$m(C) = \frac{C_{0.5}^{\alpha}}{C_{0.5}^{\alpha} + C^{\alpha}}, \tag{9}$$

which, in turn, can written in the equivalent form

$$m(C) = 1 - \frac{C^{\alpha}}{C_{0.5}^{\alpha} + C^{\alpha}}. \tag{10}$$

Since the fluorescence measurements are made quite shortly after introduction of the substance, i.e. the influence of the medium's carrying capacity can be neglected, the growth dynamics of the bacterial population N satisfies the Malthus equation

$$\frac{dN}{dt} = r(C)N, \tag{11}$$

where $r(C)$ is the growth rate depending on the acting substance's concentration; in the case of the control, $r = r_0 =$ const.

For a small period of time Δt the derivative can be replaced by a finite difference

$$\frac{N(t_0 + \Delta t) - N(t_0)}{r_0 N(t_0)\Delta t} \approx \frac{r(C)}{r_0} \tag{12}$$

and, due to the fact that the level of fluorescence depends on the level of population growth, and comparing (12) with

[6] The multiplier $1/2$ is introduced for convenience of the kinetic interpretation, see below.

[7] Step-by-step calculations:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x - e^{-x}} \equiv \frac{e^{2x} - 1}{e^{2x} + 1},$$

$$x = \ln(z),\ 2x = \ln\left(z^2\right),$$

$$1 - \tanh(z) = 1 - \frac{z^2 - 1}{z^2 + 1} = \frac{2}{z^2 + 1}.$$

Since the quantity

$$\frac{1}{2}\alpha\left(\ln(C) - \ln(C_{0.5})\right) = \ln\left(\frac{C^{\frac{\alpha}{2}}}{C_{0.5}^{\frac{\alpha}{2}}}\right),$$

then its substitution as z leads to (9).

(7), we conclude that $m(C) \approx r(C)/r_0$, that is, the growth equation (11) can be written taking into account (10), as[8]

$$\frac{dN}{dt} = r_0 N - r_0 \frac{C^\alpha N}{C_{0.5}^\alpha + C^\alpha}. \qquad (13)$$

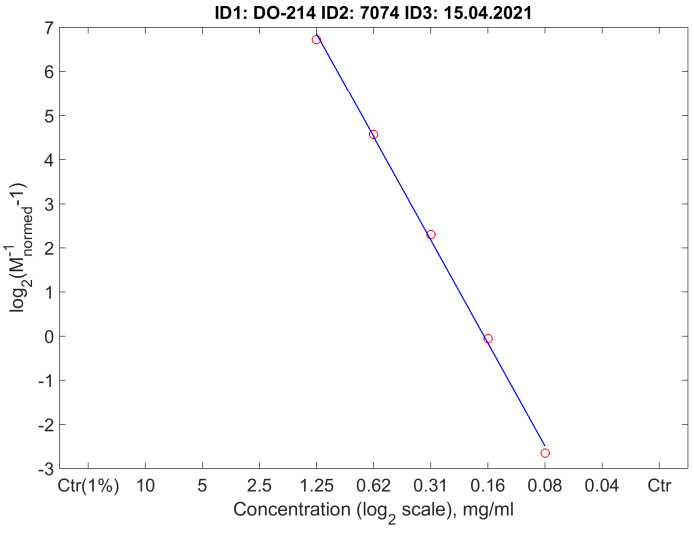
In Eq. (13), the first term corresponds to the unperturbed growth and the second term to the growth retardation due to inhibition, which has the exact form of an enzyme-substrate inhibitory Hill function with the active substance concentration as the inhibitory substrate[9].

A practical test of whether the change in concentration is indeed following Hill's kinetics can most easily be carried out by linearising the dependence (9) by using a suitable combination of variables[10] denoted by square brackets in the formula below[11]:

$$\left[\log\left(m^{-1} - 1\right)\right] = -\alpha \log(C_{0.5}^\alpha) + \alpha \left[\log(C)\right], \qquad (14)$$

where the logarithm can be taken on any basis.

Fig. 4 demonstrates the fit of the data studied to the Hill kinetics, as the medians found in the transition region (points corresponding to median concentrations equal or less than both controls are excluded as leading to divergence or imaginary logarithm values) in coordinate representation (14) with logarithms at base two due to the binary dilution sequence, are approximated with good accuracy by the straigh line. The tangent of its slope determines the power factor $\alpha = 2.33$ and the shift corresponds to the concentration of the half effect, close to 0.16 mg/ml.



The dependence of the normalised dimensionless fluorescence medians (markers) on the active substance concentration in logarithmic coordinates and their fitting by a straight line, indicating correspondence of the response to Hill's kinetics.

[8] J. Zhi, C. H. Nightingale, R. Quintiliani, A pharmacodynamic model for the activity of antibiotics against microorganisms under nonsaturable conditions, Journal of Pharmaceutical Sciences 75 (1986) 1063–1067. doi:10.1002/jps.2600751108

[9] S. Goutelle, M. Maurin, F. Rougier, X. Barbaut, L. Bourguignon, M. Ducher, P. Maire, The Hill equation: a review of its capabilities in pharmacological modelling, Fundamental & Clinical Pharmacology 22 (2008) 633–648. doi:10.1111/j.1472-8206.2008.00633.x; and R. Gesztelyi, J. Zsuga, A. Kemeny-Beke, B. Varga, B. Juhasz, A. Tosaki, The Hill equation and the origin of quantitative pharmacology, Archive for History of Exact Sciences 66 (2012) 427–438. doi:10.1007/s00407-012-0098-5

[10] Step-by-step it can by done by writing (9) as

$$m(C) = \frac{1^\alpha}{1 + \frac{C^{\frac{\alpha}{2}}}{C_{0.5}^{\frac{\alpha}{2}}}},$$

turning it over, $1 + C^{\frac{\alpha}{2}} \Big/ C_{0.5}^{\frac{\alpha}{2}} = 1/m$, and by moving one to the right-hand side and taking the logarithm from both parts.

[11] R. Barlow, J. F. Blake, Hill coefficients and the logistic equation, Trends in Pharmacological Sciences 10 (1989) 440–441. doi:10.1016/S0165-6147(89)80006-9

Figure 4: The dependence of the normalised dimensionless fluorescence medians (markers) on the active substance concentration in logarithmic coordinates and their fitting by a straight line, indicating correspondence of the response to Hill's kinetics.

Practical approximation

Since a number of active substances have intrinsic fluorescence properties, which results in a shift of the recorded steady-state value at high concentrations from the 1% control value, it is more convenient from a practical point of view to consider dimensionless quantities not as in the ideal case (7) but simply by the value of the normal control[12]

$$m = \frac{M}{M_{contr}} \tag{15}$$

From a practical point of view, the hyperbolic tangent (8 can be replaced by an equivalent form of the logistic curve, explicitly expressed as using an exponent, as shown above and, in addition, the upper m_{max} and lower m_{min} values of its asymptotes can be considered as explicit uncertain parameters, which correspond to standard methodology for fitting response data to the actiing dose [13]:

$$m(\ln(C)) = m_{min} + \frac{m_{max} - m_{min}}{1 + e^{\alpha(\log(C) - \log(C_{0.5}))}}. \tag{16}$$

The use of the logarithm of concentration, although there is an explicit formula for the dependence on concentration itself, (9), is motivated by the fact that the hyperbolic function (9) is rapidly decreasing, and the sequence of experimental concentration values varies uniformly exactly on a logarithmic scale[14].

The concentration at which $m = M_1/M_{contr}$ is used as the MIC within the continuous approximation; in addition, Fig. 5 shows the confidence intervals of the approximation, which, in principle, allows the MIC uncertainty interval to be estimated by their intersection with the upper limit line of the 1% control confidence interval as the concentration increases.

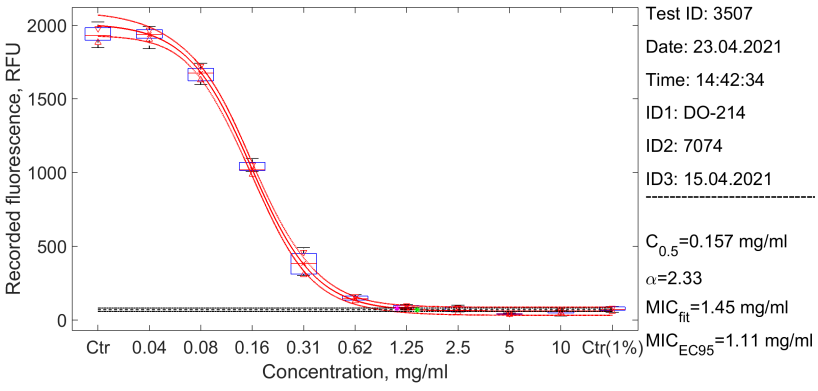


Figure 5: Experimental fluorescence data plotted as a boxplot, similar to Fig. 2; the red × are the points on the lines denoting the median used in the fitting. Solid red line is the fitting function (16), dashed line correspond to 95% confidence interval of approximation; green and purple asterisks denote MIC positions for equality of response curve and median of control and 5% exceeding the value of the lower stationary of the response curve, respectively.

Besides, the properties of the response to the active substance, not requiring comparison with the fluorescence value

[12] In addition, this allows working with a fluorescence recording situation without a diluted culture as such. In addition, in the case of small growth for 1% control, the condition of using 0.1 values for normal growth is implemented in the software.

[13] H. Motulsky, A. Christopoulos, Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting, Oxford University Press, 2004

[14] See comments on the stability of the approximation algorithm in the book[13].

of 1% of the control, but given by the properties of the approximating curve itself, the output also indicates: (i) the half-maximum effective concentration $C_{0.5} \equiv EC_{50}$ corresponding to an effect equal to half the possible for this substance; (ii) MIC_{EC95} equal to the concentration at which the response function, which approximates the data, reaches $005m_{min}$ that is considered as the criterion for reaching a steady-state value; at further increasing the concentration the fluorescence change is not significant considering the statistical scatter of the measurements and corresponds to the effective concentration standard EC_{95}.

Historical notes

A technique, which uses the quantitative comparison between fluorescence fluorescence level at presence of antimycobacterial compounds and the control level in the resazurin test was proposed in 1997[15] (using the EC_{90} criterion for the fluorescence level compared to the control mycobacterial culture). By 2007, such a quantitative method was already regarded as a standard protocol both with binary dilutions in the plate rows and with more detailed concentration resolution[16,17].

The usage of the fitting function (16) is the standard recommended method for drug sensitivity studies (*dose-response curve*), see book[15]; in works on drug sensitivity of mycobacteria in the resusurin test, such kind of interpolation started to be actively used in 2012-2015.[18,19].

Moreover, it is the usage of the four-parameter function (16) is recommended in the data processing step of modern microplate fluorescence readers operating with the ressarurin test; moreover, the output of the corresponding parameters (including EC_{50}) is included in the accompanying software, for example, see page 48 of the manual:

- `FLUOstar OPTIMA: Software Manual - Part IIIa: MARS Data Analysis`

However, the discussion of non-parametric estimations of raw datasets in microplate data seems to be not investigated properly in literature.

[15] L. A. Collins, S. G. Franzblau, Microplate alamar blue assay versus BACTEC 460 system for high-throughput screening of compounds against *Mycobacterium tuberculosis* and *Mycobacterium avium*, Antimicrobial Agents and Chemotherapy 41 (1997) 1004–1009. doi:10.1128/AAC.41.5.1004

[16] N. K. Taneja, J. S. Tyagi, Resazurin reduction assays for screening of antitubercular compounds against dormant and actively growing *Mycobacterium tuberculosis*, *Mycobacterium bovis* BCG and *Mycobacterium smegmatis*, Journal of Antimicrobial Chemotherapy 60 (2) (2007) 288–293. doi:10.1093/jac/dkm207

[17] T. P. Primm, S. G. Franzblau, Recent advances in methodologies for the discovery of antimycobacterial drugs, Current Bioactive Compounds 3 (2007) 201–208

[18] M. Zhang, C. Sala, R. C. Hartkoorn, N. Dhar, A. Mendoza-Losana, S. T. Cole, Streptomycin-starved *Mycobacterium tuberculosis* 18b, a drug discovery tool for latent tuberculosis, Antimicrobial Agents and Chemotherapy 56 (2012) 5782–5789. doi:10.1128/AAC.01125-12

[19] B. Gold, J. Roberts, Y. Ling, L. L. Quezada, J. Glasheen, E. Ballinger, S. Somersan-Karakaya, T. Warrier, J. D. Warren, C. Nathan, Rapid, semiquantitative assay to discriminate among compounds with activity against replicating or nonreplicating ycobacterium tuberculosis, Antimicrobial Agents and Chemotherapy 59 (2015) 6521–6538. doi:10.1128/AAC.00803-15