# Supporting information S4

**DERRINGER TRANSFORMATION**

The biological responses were transformed into a dimensionless desirability (d) scale via the following linear desirability functions:

$d\left(Y\right)= \frac{0.9-0.1}{Y\_{max}- Y\_{min}} ×\left(Y\_{i}-Y\_{min}\right)+ 0.1$

or $d\left(Y\right)= \frac{0.1-0.9}{Y\_{max}- Y\_{min}} ×\left(Y\_{i}-Y\_{min}\right)+ 0.9$

for parameters to be maximized or minimized respectively. In the equations above, Yi is the experimental value for the respective response, whereas Ymin and Ymax are the minimum and maximum response values found, respectively. Ymin was arbitrarily set at the reporting threshold of 0.01 for every response; Ymax is the maximal value found for each methodology in our dataset and their values are reported in *Supporting information S5*. Compounds with a high Th1 desirability should have high IFN-γ values, combined with low IL-4 responses. This contrasts the Th2 desirability: high IL-4 responses combined with low IFN-γ values are required. After this linear d-transformation of each of the responses, all values range from 0.1 (undesirable) to 0.9 (most desirable). These standardized d-values were combined to calculate a global D-value:

$$D= \sqrt[n]{\prod\_{i=1}^{n}d\_{i}^{Pi}}$$

The compound with the highest D-value expresses the best combination of the different desired responses. In this equation, pi is the relative importance given to the respective response. Here, we weighted the responses equally, so pi = 1 for each of the 20 responses.