AI Driven Automation of Model Selection and Data Quality Control in SPR Production Screens

Pooja Sharma[1], Juan Florez[2], Aman Singh [1], Daniel Siegismund[2], Mario Wieser[2], Moritz Pfreundschuh[2], Qing Chen [1], Carolyn Ch'ng[1], Stephan Heyse[2] and Stephan Steigele[2]

[1]Discovery Attribute Sciences, Amgen Research, Amgen Inc., Thousand Oaks, CA; [2]Genedata AG, Basel, Switzerland

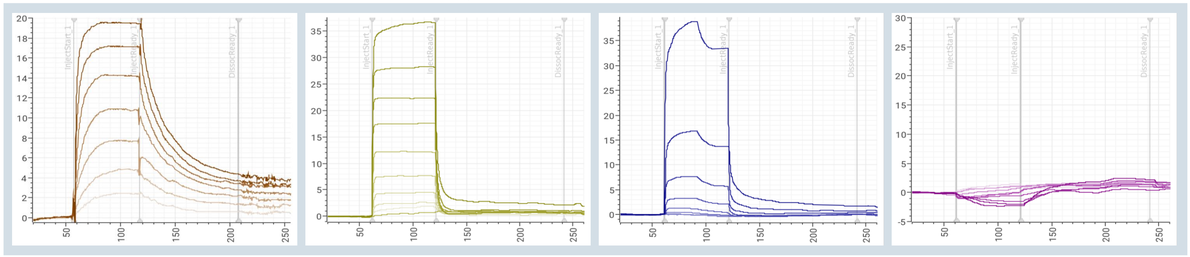
Background and Motivation

The advent of automated laboratory and data analytics has increased the success of drug discovery programs, by optimizing experimental workflows to better control result quality with less investment of resources and time. However, biophysical assays such as Surface Plasmon Resonance (SPR) still require tedious manual review for quality control, result review and decision-making, due to the complex nature of sensorgrams and possible outcomes. Genedata, together with Amgen, developed an AI-driven data analysis workflow for automation of complex biophysical analysis. In this workflow, sensorgrams are classified automatically into four categories before the appropriate 1:1 binding models (kinetic or steady-state) are applied. This workflow ensures that binding affinity and kinetic parameters are reproducibly and precisely determined for the multitude of outcomes typically observed in a compound screen, reducing the need for expert review to just a very few corner cases. We explain the basic elements of this new AI-driven solution for automating SPR data analysis, illustrate its use on two Amgen production screens, and conclude with a discussion on its potential impact for future drug screening programs.

Problem Statement

Four major binding profiles occur in a typical SPR binding experiment with 1:1 binding stoichiometry:

Resolved binding kinetics	**Binding too fast to resolve kinetics**	**Non-standard binding** (unclear process)	**Insufficient or no binding**
→ use kinetic fit	**→ use steady-state fit**	**→ Defer to review**	**→ label; no fit**

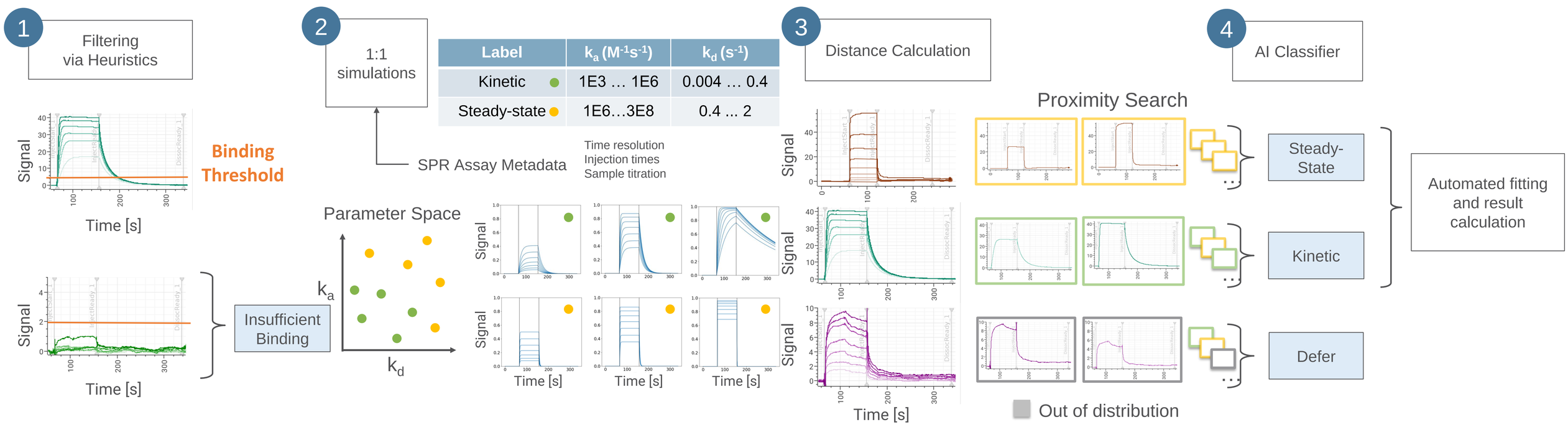


In current practice, to obtain precise results, a screening scientist must manually review traces and set the correct fit method for each.

To solve this, an *automated data analysis workflow* for SPR data must:

- Label tested analytes as following either a steady-state or kinetic binding model and determine their binding affinities and rate constants.
- Exclude insufficient binders from analysis and label them accordingly
- Identify corner-cases that might require further expert review.

Workflow and Solution



Identify Insufficient or Non-Binders: Analytes where most traces lie below the binding threshold are categorized as "Insufficient Binding" and excluded from further analysis.

Simulate Perfect Outcomes: Based on the assay specifics, 1:1 binding models are simulated within the defined parameter space for both Kinetic and Steady-State fits ("outcome distribution").

Calculate Distances: The distance of each measured trace to the simulated data is calculated. A proximity search determines the most likely model—Kinetic or Steady-State—and labels sensorgrams outside the expected outcome distribution.

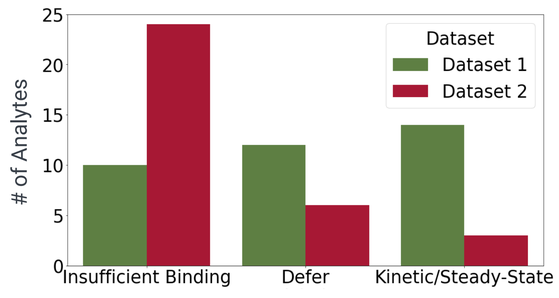
Classification with AI Classifier: The classifier pools the results from the individual traces and assigns the analyte to either the **Steady-State** or **Kinetic** category. The **Defer** category indicates a quality problem, for expert follow-up.

Data

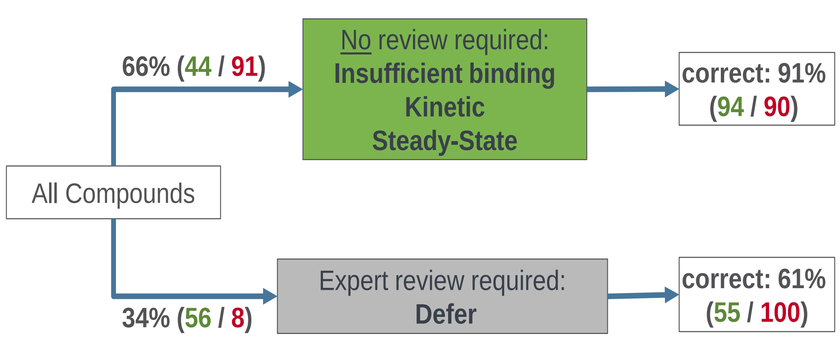
SPR datasets: Data from two production screens of 69 analytes in total, measured in 8 doses each, including 'challenging' outcomes (non- or weak binders).

	Dataset 1	Dataset 2
# of Analytes	36	33
Instrument	Biacore T200	
Objective (Screen)	Target Engagement or Binding Characterization	

As ground truth, the agreed-upon outcome from review by Amgen experts was used.



Results



Automated, Efficient Model Selection: The AI-driven workflow efficiently triaged analytes, automatically choosing the correct fit model for **66%** of analytes, while **34%** were deferred for review by an expert, who could then reject measurements whose binding profiles did not follow a 1:1 model. To guide reviewers in projects decision-making, each deferred analyte was also given a second label with the model (steady-state or kinetic) that could best describe it.

Reliable Classification: Accuracy of automated classification vs. expert-created classification was 0.72 and 0.91, respectively, for Dataset 1 (left) and Dataset 2 (middle). An automated misclassification to the 'Defer' category (green boxes) is reviewed by an expert, and thus does not negatively impact result quality. Thus, problematic false-predictions are rare (red boxes, with corresponding traces shown at the right), raising the accuracy in these data sets to 0.94 or 0.91, respectively.

The need for speed in drug discovery is being answered by earlier characterization of screening molecules using information-rich biophysical, mechanistic, phenotypic, and liability assays. Results from many approaches are then fed into increasingly digitalized decision-making and prediction processes, which depend on reliable, standardized, and rapid acquisition of experimental data on large sets of molecules. Here, we have demonstrated how AI-driven analytics for SPR screens produces instant, reliable results with minimal need for expert intervention. Such a workflow is amenable to other assay formats, such as Biolayer Interferometry. Through workflows like these, automated data analysis of information-rich assays will likely play a key role in the success of discovery projects.