Supplementary data

Data pre-processing
Raw data from the nanoDSF instrument (F330, F350, A350, F330/F350 ratio, and their respective derivatives) were interpolated using InterpolatedUnivariateSpline from the scipy.interpolate python module in order to ensure the same temperature alignment for all data.

Machine Learning
The Machine Learning code developed for this work uses the scikit-learn toolbox with : (1) LogisticRegression from the linear_model module with parameter max_iter sets to 1000; (2) SVC from the svm module with the following combination of parameters: kernel = "poly", gamma = "auto", C = 1, degree = 3; (3) RandomForestClassifier from the ensemble module with parameter n_estimators fixed to 500; (4) AdaBoostClassifier from the module ensemble with the default parameter values except n_estimators = 100. The code and the data used for EGFR amplification detection is publicly available at **https://github.com/RemiEyraud/EGFR_amplification_detection**

Additional Plots

Figure S1: Means of nanoDSF output.