ISGF3 and Antiproliferative Activities of Type I Interferons

Francis Lin - 804457464 - W18 BE188 Final Report

Introduction and Problem Definition

Interferons (IFN) are proteins that are secreted by cells in the immune system in response to the presence of pathogens. Specifically, Type I IFNs are released when cells recognizes a viral infection. By upregulating molecules that prevent viruses from replicating RNA and DNA, IFNs are able to increase host defenses, increasing antiviral, antiproliferative, and immunomodulatory activities. One interesting characteristic about the IFN complex is that there are redundancies for the IFN ligands which bind to one IFN receptor, even though they exhibit different levels of activities. Due to the highly sensitive environment of IFNs, characterizing differential signal activation is challenging for researchers and remains a topic of research among scientists.

In the Jaks et al. paper, the authors concluded that different IFNs recruit signaling complexes with similar architectures, which likely cannot account for the differential signal activation. Instead, the authors measured the kinetic rate constants and equilibrium dissociation constants of the IFN receptor subunits ifnar1 and ifnar2 using solid phase binding assays and discovered a correlation between different types of cellular signal activation and the IFN's affinity towards the receptor subunits. Specifically, they found that affinity towards ifnar1 correlated with antiproliferative activity and that relative affinity towards ifnar2 correlated with ISGF3 activity. Therefore, the authors suggested that differences in IFN affinity towards the individual IFN receptor subunits, specifically ifnar-1 and ifnar-2, are the main reason for differential signal activation.

While a correlation can be observed by the Jaks et al. paper, the authors did not offer any insight on whether the correlation is strong to have predictive power. The goal of this exercise is to confirm the correlations observed by Jaks et al. and determine whether the rate and affinity constants can reliably predict ISGF3 and antiproliferative activities. A better characterization may lend researchers and scientists a better understanding of IFN binding mechanics which can lead to improved treatments for patients.

Methods

First, major figures of the paper were re-implemented using information given from Table 1 in Jaks et al. Table 1 contains the affinity constants, dissociation rate constants, and equilibrium dissociation constants to each of the two IFNs receptor subunits (ifnar1 and ifnar2) for 10 IFNs, 3 of which are mutant. The 7 wild-type IFNs are IFNα1, IFNα2, IFNα8, IFNα21, IFNβ, IFNω, IFNτ2. The 3 mutant IFNs are E58A, R144A, and E58A+R114A. The table also includes the ISGF3, antiproliferative, and antiviral activities for each of the 10 IFNs. Table 1 is shown below for reference. Figures 1g, 1h, 2e, 2f, 5c-5f, 6a-6d, and 7a-7c were re-implemented.

	Ifnar2-EC			Ifnar1-EC			Activities (EC_{50})		
IFN	k_a ($M^{-1}s^{-1}$)	k_d (s^{-1})	K_d (nM)	k_a ($M^{-1}s^{-1}$)	k_d (s^{-1})	K_d (μM)	ISGF3 (pM)	Antiproliferative (nM)	Antiviral (pM)
IFNα1	2×10^6	0.12	100	~ 2×10^5	0.5	2.5	47	2	400
IFNα2	3×10^6	0.015	5	~ 2×10^5	1	5.0	8	1.1	40
IFNα8	6×10^6	0.02	3	~ 2×10^5	0.5	2.2	5	1.1	20
IFNα21	3×10^6	0.08	25	~ 2×10^5	0.5	2.5	20	1.5	50
IFNβ	1×10^7	0.001	~ 0.1	~ 2×10^5	0.025	0.05	1.5	0.065	–
IFNω	6×10^6	0.008	1	~ 2×10^5	0.08	0.40	3	0.2	–
IFNτ2	5×10^5	0.8	1200	~ 2×10^5	2	10.0	3800	60	4000
E58A	3×10^6	0.015	5	~ 2×10^5	0.1	0.5	2	0.14	–
R144A	5×10^5	0.05	150	~ 2×10^5	1	5.0	575	10.0	–
E58A, R144A	5×10^5	0.05	150	~ 2×10^5	0.1	0.5	390	3.2	–
Error (%)	20	10	20		40	30	20	30	30

Table 1: Table of binding affinities and signal activation for different IFNs.

Second, Principal Component Analysis (PCA) was used to identify the principal components contained in the data. Furthermore, if clear clusters were visually observed, K-Nearest Neighbors would be used to more clearly define the clusters.

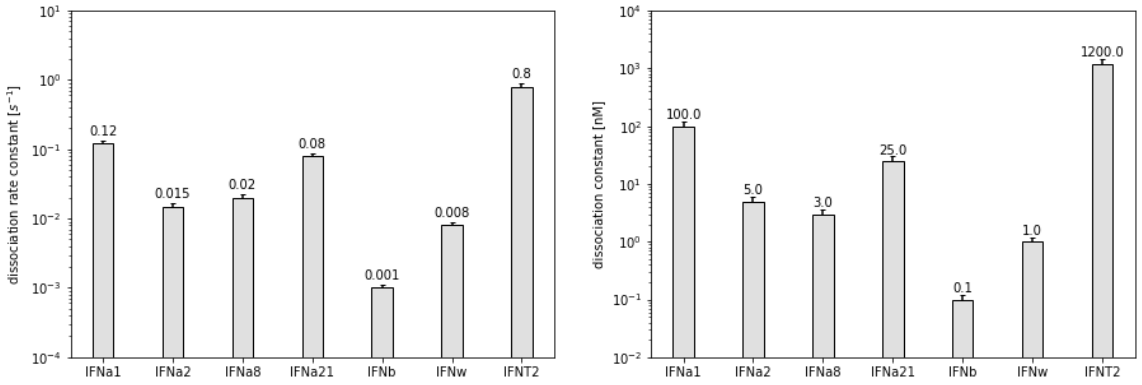
Third, Partial Least Squares Regression (PLSR) was used to predict the activities of the interferons. PLSR was chosen because its principal components rely on the covariance of the data and covariance is often present in biological systems. In order to evaluate the performance of PLSR, Leave-One-Out Cross Validation (LOOCV) will be used. LOOCV was chosen rather than K-fold cross-validation because the data provided by Table 1 contained few observations. If K-fold cross-validation was used, the model would be trained on an even smaller dataset, resulting in an inconsistent model.

Fourth, linear regression was used to find how strong the correlation is between certain pairs of relative potencies and relative binding affinities as described in the paper. The goodness of fit was calculated to describe how much variability in the data can be accounted for by the model.

Results

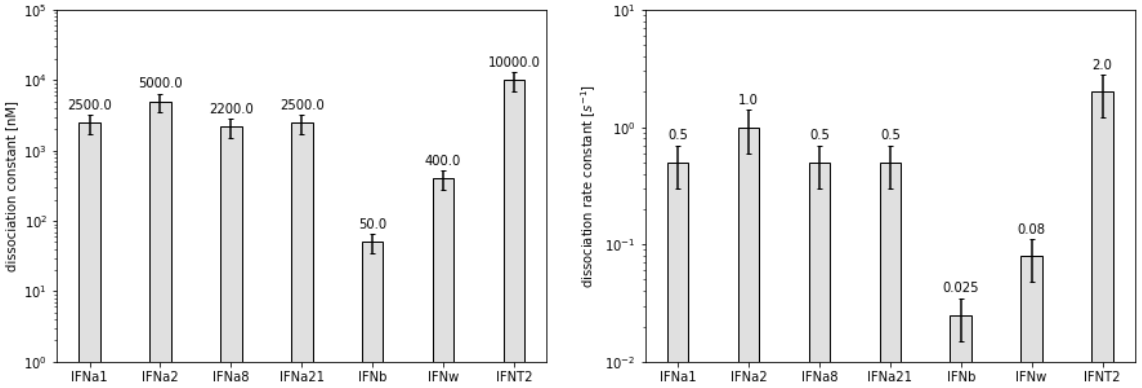
Re-implementation

The re-implementations are shown in Figures A through E.



(i) (ii)

Figure A: (i) Dissociation rate constants and (ii) equilibrium dissociation constants of different wild-type IFNs to immobilized ifnar2-EC.



(i) (ii)

Figure B: (i) Equilibrium dissociation constants and (ii) dissociation rate constants of wild-type IFNs to immobilized ifnar1-EC.

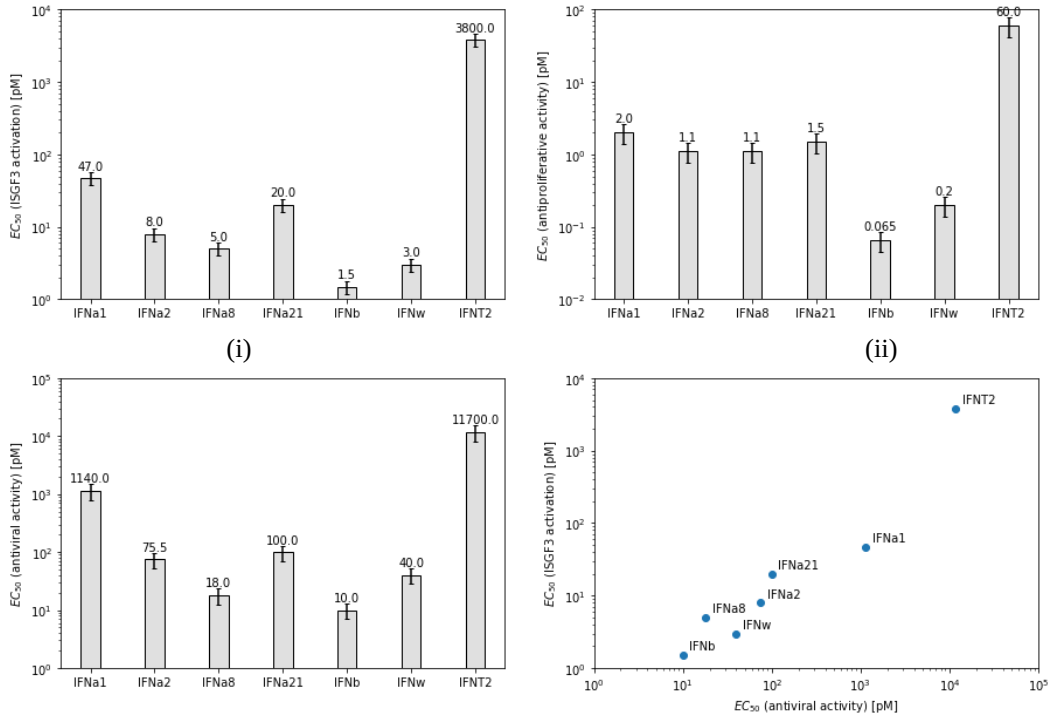


Figure C: (i) ISGF3 activities, (ii) antiproliferative activities, and (iii) antiviral activities of wild-type IFNs. (iv) The ISGF3 activities were graphed as a function of antiviral activities for the wild-type IFNs

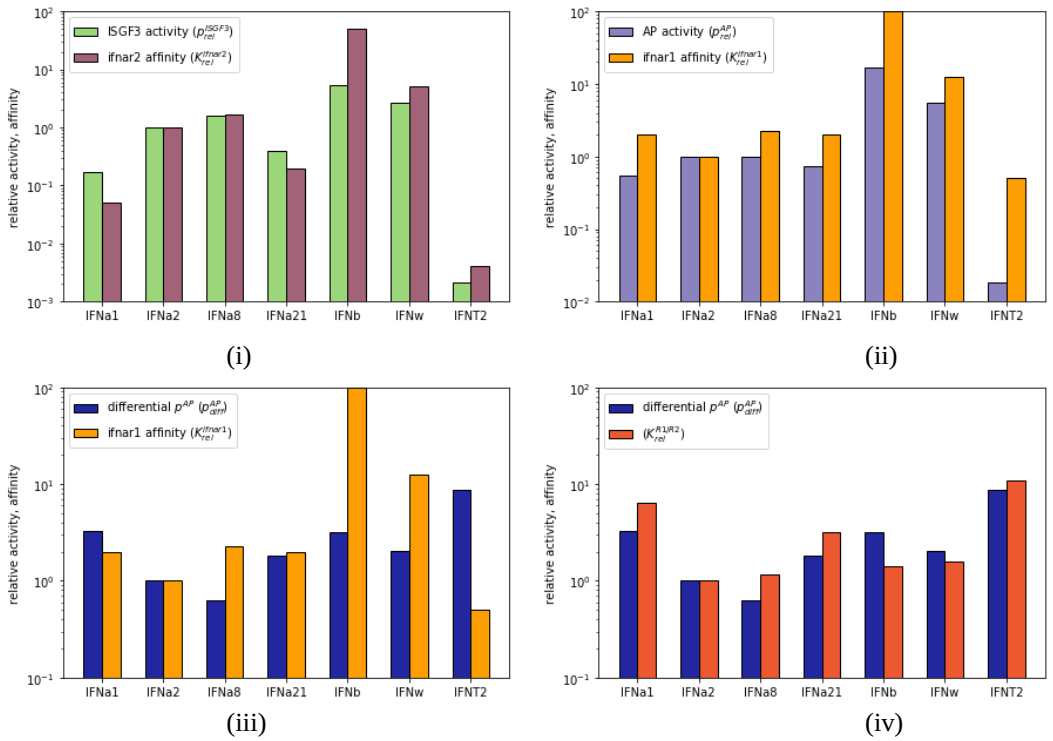
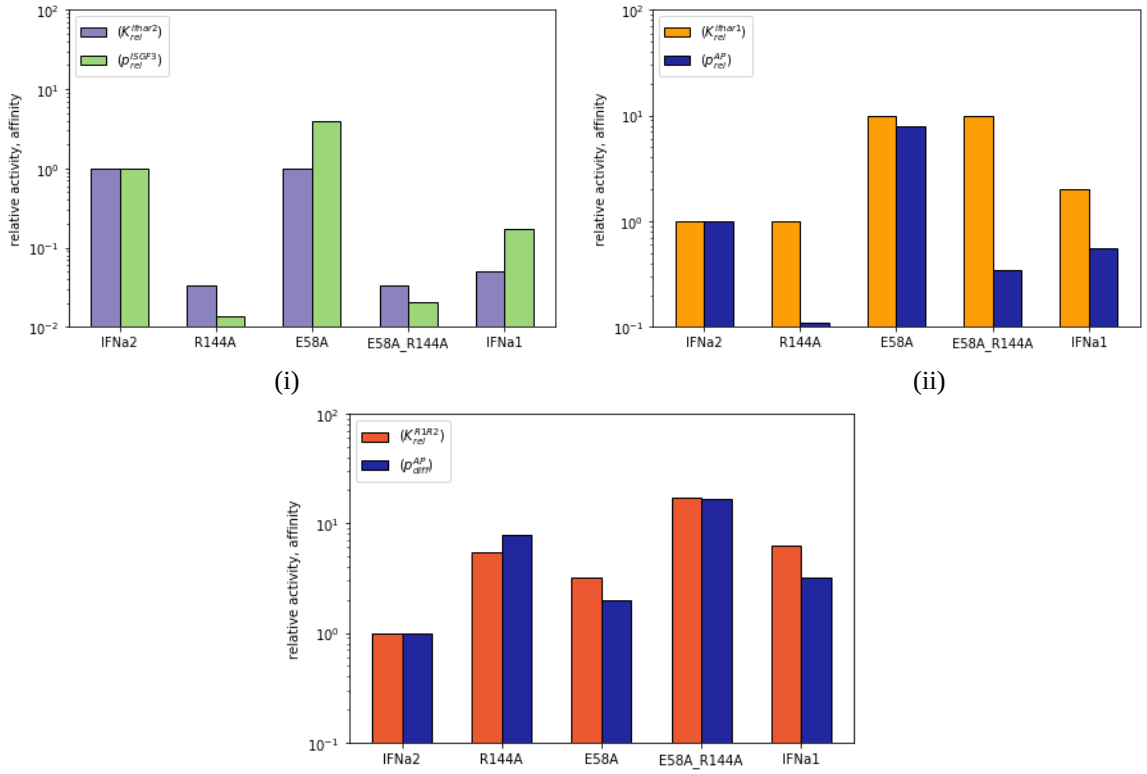


Figure D: (i) Relative ISGF3 potencies and relative binding affinities to ifnar2. (ii) Relative antiproliferative potencies and relative binding affinities to ifnar1. (iii) Differential antiproliferative potencies and relative binding affinities to ifnar1. (iv) Differential antiproliferative potencies and affinity towards ifnar1 relative to the affinity towards ifnar2.

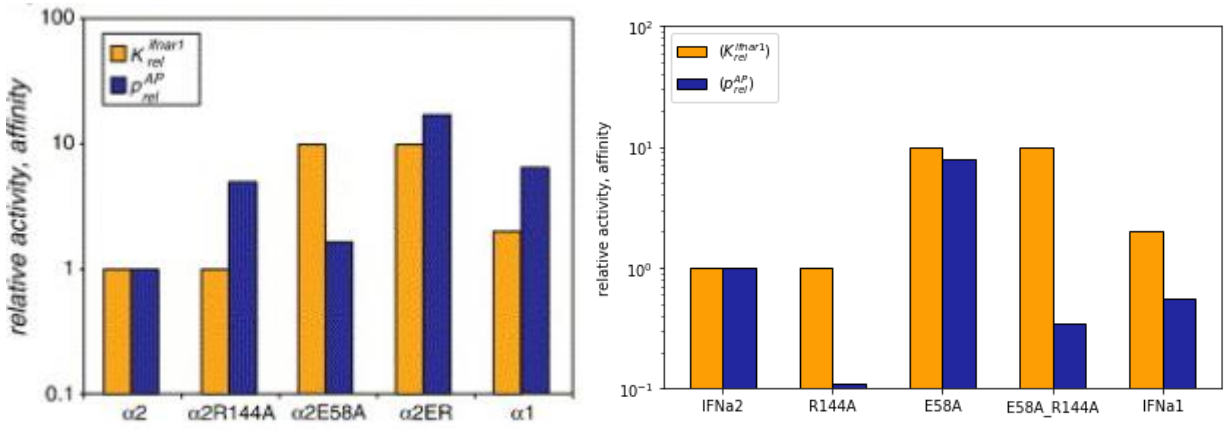


(i) (ii)

(iii)

Figure E: (i) Relative binding affinities to ifnar2 and relative ISGF3 potencies. (ii) Relative binding affinities to ifnar1 and relative antiproliferative potencies. (iii) Affinity towards ifnar1 relative to the affinity towards ifnar2 and differential antiproliferative potencies.

While most of the figures in Jaks et al. were re-implemented without issues, discrepancies for relative antiproliferative potencies and differential antiproliferative potencies were found. For example, shown in Figure F is a comparison of Figure 7b in Jaks et al. Since correlation was determined qualitatively in Jaks et al., this discrepancy may cause the authors to reevaluate their findings for correlations for analyses involving relative antiproliferative potencies.



(i) (ii)

Figure F: (i) The original Figure 7b in Jaks et al. (ii) The re-implementation of 7b from Jaks et al.

Principal Component Analysis

The results of the PCA is shown in Figure G. By visual inspection, there are no clusters in the data, indicating that no IFNs are more similar or different than the others in terms of their characteristics. Since there are no clusters, no clustering algorithms such as K-Nearest Neighbors will be implemented.

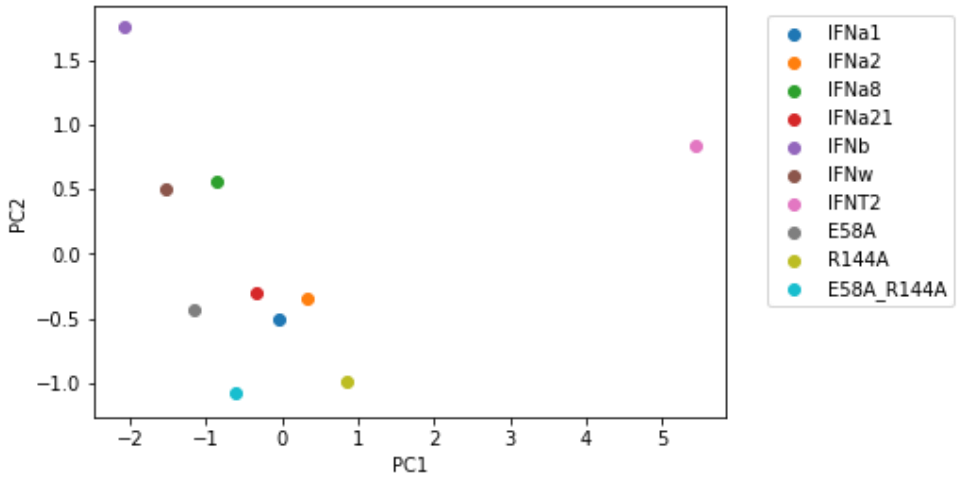


Figure G: Scores calculated using PCA performed on Table 1. By visual inspection, no clusters exist.

PLSR and LOOCV

Figure H shows the results of running PLSR on the data.

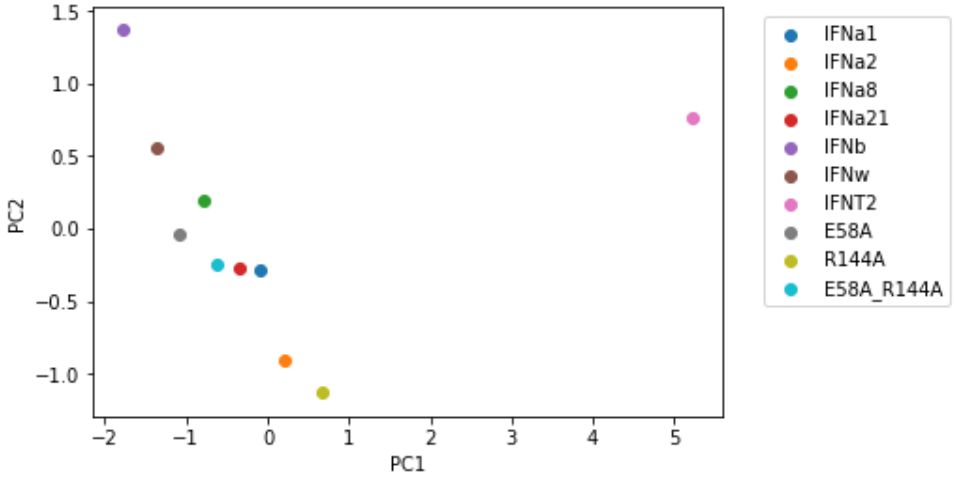


Figure H: Scores calculated using PLSR performed on Table 1.

Using the PLSR model, LOOCV was used to quantify how accurately we can predict signal activation from the affinity constants. Figure I and Figure J show the observed and predicted ISGF3 and antiproliferative activities for the wild-type IFNs, respectively. Additionally, Table 2 and 3 shows which signal activation were predicted within a certain margin of error provided by Table 1. The number of components included for this analysis was 4, which was determined because it was the number of components that offered the highest accuracy for the model.

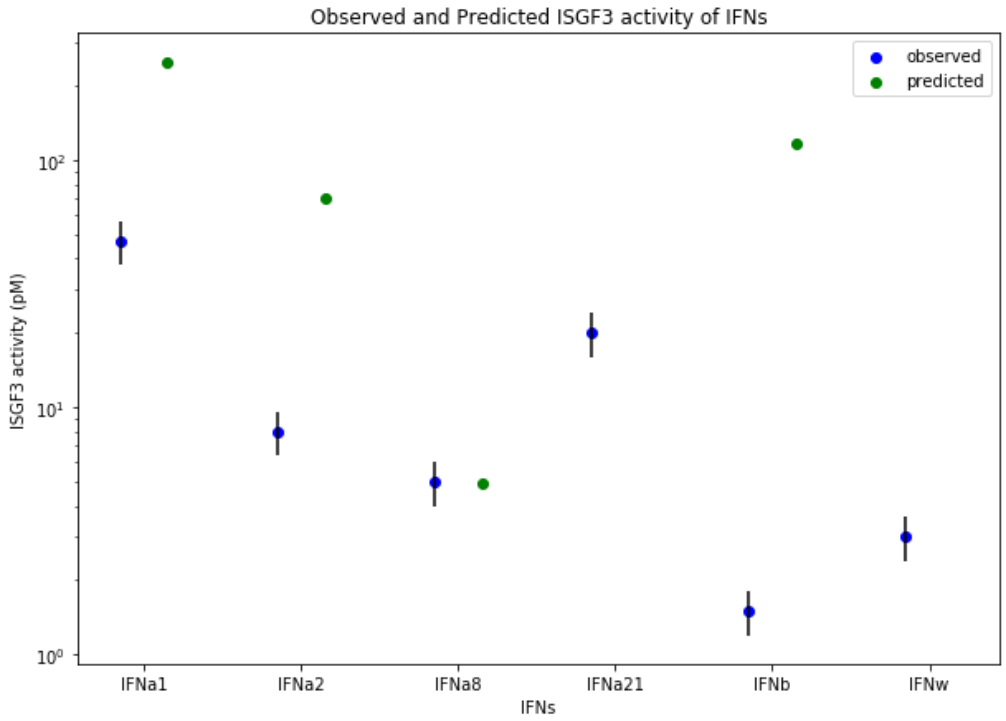


Figure I: Observed and predicted values for ISGF3 activities for wild-type IFNs.

	ISGF3	Acceptable Range	Predicted ISGF3	Prediction within range
IFNa1	47.0	37.6 - 56.4	250.403051	False
IFNa2	8.0	6.4 - 9.6	70.638494	False
IFNa8	5.0	4.0 - 6.0	4.941827	True
IFNa21	20.0	16.0 - 24.0	-115.270477	False
IFNb	1.5	1.2 - 1.8	116.092758	False
IFNw	3.0	2.4 - 3.6	-43.410168	False
IFNT2	3800.0	3040.0 - 4560.0	2525.935296	False
E58A	2.0	1.6 - 2.4	-103.617038	False
R144A	575.0	460.0 - 690.0	506.222188	True
E58A_R144A	390.0	312.0 - 468.0	530.224506	False

Table 2: Observed and predicted values for ISGF3 activities for wild-type IFNs.

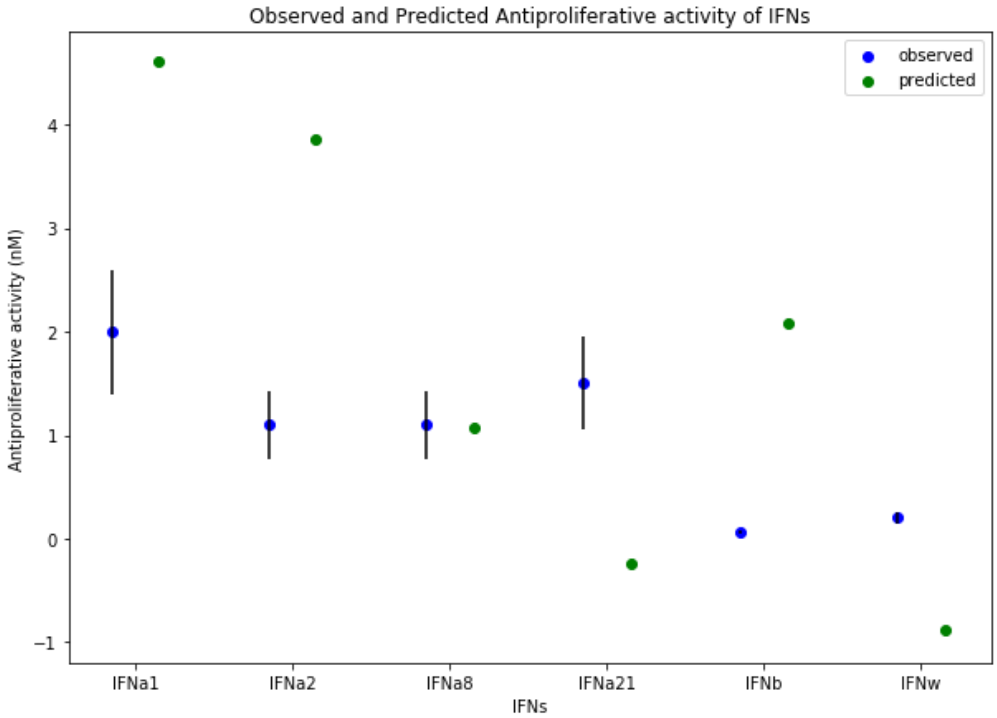


Figure J: Observed and predicted values for antiproliferative activities for wild-type IFNs.

	antiproliferative	Acceptable Range	Predicted antiproliferative	Prediction within range
IFNa1	2.000	1.4 - 2.6	4.606070	False
IFNa2	1.100	0.77 - 1.43	3.851453	False
IFNa8	1.100	0.77 - 1.43	1.078108	True
IFNa21	1.500	1.05 - 1.95	-0.244630	False
IFNb	0.065	0.0455 - 0.0845	2.086526	False
IFNw	0.200	0.14 - 0.26	-0.878390	False
IFNT2	60.000	42.0 - 78.0	43.750093	True
E58A	0.140	0.098 - 0.182	-2.154072	False
R144A	10.000	7.0 - 13.0	6.642939	False
E58A_R144A	3.200	2.24 - 4.16	8.136540	False

Table 3: Observed and predicted values for ISGF3 activities for wild-type IFNs.

Applying PLSR on Table 1 is not effective in predicting the ISGF3 and antiproliferative activities of the IFNs. For both ISGF3 and antiproliferative activities, only two IFNs were predicted within the margin of error. The results suggest the absolute values of affinity constants cannot be utilized to predict signal activation.

Linear Regression

Figure K displays the results of linear regression on the relative potencies and binding affinities discussed in the paper.

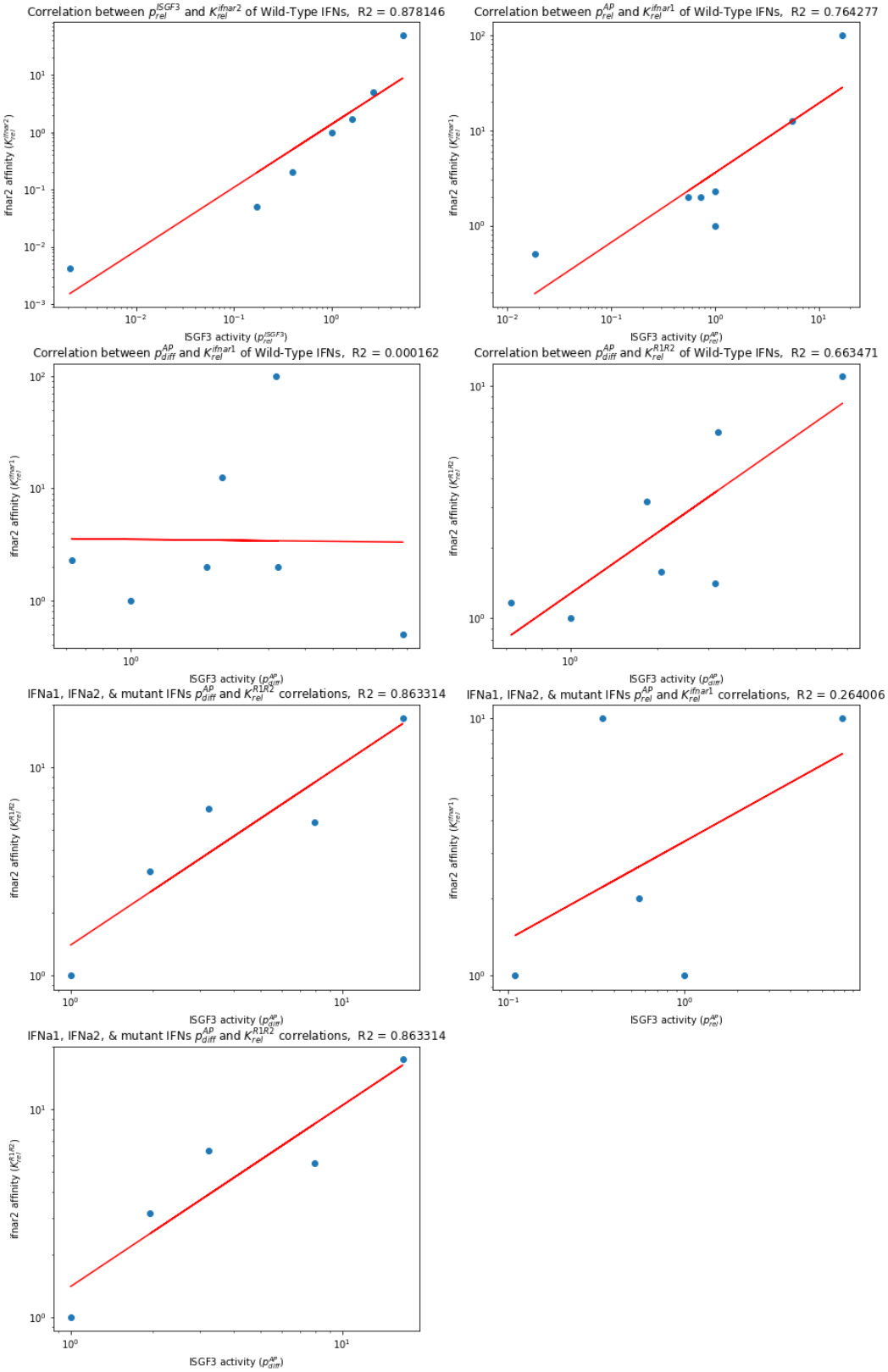


Figure K: Correlations between relative potencies and binding affinities

In Figure K, the goodness of fit stayed below 90% for all analyzed correlations, indicating that the model cannot account for more than 90% of the variability in the data. For correlations involving relative antiproliferative activity, the goodness of fit was lower than 30%. This may be a direct result of the discrepancy between our re-implementation and the Jaks et al. paper.

Discussion

The conclusion from this exercise is that while some correlations were seen between binding affinities and signal activation, it is not high enough to accurately predict signal activation from binding affinities. This could be seen in the lack of accurate predictions in PLSR and the poorness of fit in linear regression. These results could be validated by creating different mutant IFNs with varying affinity constants and activity to increase the number of data points.

There were several limitations of the analysis performed here. Firstly, the error percentages in the initial dataset were quite large, ranging from 10% to 40%. Therefore, the model, trained on inaccurate data, could deviate widely from the actual values. Secondly, from visual inspection, it seems as if the PLSR scores graph could have a line drawn through the data with high goodness of fit. However, the PLSR was not adjusted to logarithmic scale. This portion could be re-implemented so that PLRS can more accurately determine the eliminations. Finally, linear regression was performed on logarithmic data, which means that the R2 value may be heavily skewed.

To validate these results, different mutant IFNs should be created with varying affinity constants and activity. This could add more information for the algorithm to choose itself, resulting in a better model.

References

Jaks, Eva, et al. “Differential Receptor Subunit Affinities of Type I Interferons Govern Differential Signal Activation.” Journal of Molecular Biology, vol. 366, no. 2, 2007, pp. 525–539., doi:10.1016/j.jmb.2006.11.053.