Supplemental information

Directed mutational scanning reveals a balance between acidic and hydrophobic residues in strong human activation domains

Max V. Staller, Eddie Ramirez, Sanjana R. Kotha, Alex S. Holehouse, Rohit V. Pappu, and Barak A. Cohen
Figure S1: Validating the high-throughput activation domain (AD) assay

A) Reproducibility of AD measurements (n=525). Diagonal: histograms of AD activity measurements. Upper panels: reproducibility of ADs between biological replicate measurements. Lower panels: Pearson (R_p) and Spearman (R_S) correlation coefficients. Sort 1-4 are independent biological replicates with 4 sorted pools. Sort 5 is an 8 way sort that combined all 4 replicates, so it is not independent but has a larger dynamic range.

B) Variants of VP16, CITED2 and Hif1α that have previously been shown to reduce activity reduce activity in our assay (Berlow et al., 2017; Cress and Triezenberg, 1991; Freedman et al., 2003; Regier et al., 1993). *, significant at 5% FDR.

C) The activities of published variants of p53 are correlated with the activities we measure (Chang et al., 1995; Lin et al., 1994). WT p53 AD1, dark blue dot; SEM, light blue box. WT p53 AD2, dark green dot; SEM, green box (cutoff at the top).
Figure S2: Both activation domains of p53 require acidic residues and hydrophobic motifs for function

A) The motif and alpha helix locations in both p53 ADs.

B) Both ADs of p53 contain aromatic and leucine rich motifs that are important for AD function (Dyson and Wright, 2016; Raj and Attardi, 2017). The activity of No AD TF is plotted with a green line. Note for the p53 data, we use the normalized Z score as the activity because the data was collected on multiple cell sorters.

C) Removing negative charge and/or adding positive charge decreased the activity of both p53 ADs. In p53 AD1, adding acidity (more negative net charge) increases activity in about half of variants. Note that p53 AD1 is 40 residues and p53 AD2 is 20 residues, so AD2 starts out with a more negative net charge per residues. WT is indicated with a large black dot. WT means and SEMs are shown in all plots with a black line and gray box. Activity is average Z score. A linear regression with a confidence interval summarizes the general dependence on net charge.

D) For both ADs, removing aromatic residues generally decreases activity. Adding aromatic residues can increase activity. The least squares regression and 95% confidence interval were calculated with seaborn.regplot().
Figure S3: Hand designed mutations to remove hydrophobic motifs or break alpha helices

For each AD, we mutated hydrophobic motifs to alanine. We identified motifs from the literature and also predicted additional motifs by looking for clusters of aromatic and leucine residues. In an attempt to break each alpha helix we introduced 2 prolines or 2 glycine residues. We used two substitutions because a single proline or glycine does not always reliably break a helix. For each AD, we divided the sequence into 2 or 3 regions, marked with the purple lines. For each of these regions we made 3 Shuffle variants that randomly permuted the sequence of this region. These Shuffle variants disrupt motif spacing and helix formation.
Figure S4: Net charge, but not aromatic content is correlated with AD activity

In the main text, we regressed the average activity of the variants designed to perturb charge against net charge. Here, we plotted all four biological replicates for all variants and repeated the regression with 95% CI. Mean WT activity, black dot. WT standard deviations shown with gray lines. There are 468, 448, 292, 444, 448 points on each plot, respectively. A) Net charge is a statistically significant determinant of the activity of CITED2, Hif1α and VP16 under ordinary least squares regression. B) The relationship between the number of aromatic residues and activity is not significant. C) For VP16, regressing activity against the number of W,F,Y,L residues yields a significant relationship, further evidence for how leucine residues make large contributions to activity in this AD. D) For each variant we calculated the Omega statistic for mixture between [W,F,Y,L] and [D,E] residues. Well-mixed variants (low Omega) tended to have higher activity while variants that segregated negative charge and W,F,Y,L residues tended to have lower activity. Pearson correlations shown. The 3 ADs panel combines CITED2, Hif1α and VP16. We excluded variants for which Omega was undefined.
Figure S5: Activity depends on net charge, not the identity of the substitution
A) Removing acidic residues (removing negative charge, light blue) has a similar effect as adding basic residues (adding positive charge, dark blue), or mutating all charged residues to alanine (green).
B) The identity of added basic residues does not matter. Adding K’s (dark blue) has a similar effect on activity as adding R’s (light blue).
C) Adding acidic residues (light blue) has similar effects to replacing basic residues with alanine (green) or replacing basic residues with acidic residues (dark blue). In Hif1α, removing positives always increases activity.
D) The identities of added acidic residues frequently does not matter. Adding D’s (dark blue) has a similar effect on activity as adding E’s (light blue). In Hif1α, adding E’s is more likely to increase activity than adding D’s.
Figure S6: Systematically adding and removing aromatic residues has complex effects on AD activity

For each AD, we plotted mean activity vs the number of aromatic residues. The standard deviations are shown with gray error bars. Points colors are: fewer aromatic residues and more active than WT (brown), fewer aromatic residues and less active than WT (dark green), aromatic exchanges that increase activity (light green), aromatic exchanges that decrease activity (yellow), more aromatic residues and more active than WT (purple), and more aromatic residues and less active than WT (pink). WT mean is the black line and standard deviation is the gray box. A linear regression (ordinary least squares) is plotted as a gray line. Middle: the sequences of the variants are shown with the mutated residues colored to match the top panel. In each set, the sequences are sorted from highest activity to lowest activity. The colors match the points in the top panel.

A) For CITED2, variants that replace aromatic residues with leucine have small increases in activity and variants that replace aromatic residues with alanine have small decreases in activity. For this AD, aromatic residues are not essential for activity.

B) For Hif1α, adding and removing aromatic residues had smaller effects on activity than adding or removing charged residues (See Figure 3).

C) For VP16, all mutations of F442 decreased activity. In 6/10 cases, adding up to 4 aromatic residues increased activity.
Figure S7: All atom Monte Carlo simulations of conformational ensembles are consistent with the Acidic Exposure Model

A) For each variant, we ran ten simulations starting in a random coil and computed the radius of gyration, a parameter that captures the size of the conformational ensemble. Each simulation is plotted as a separate point and all simulations are shown. Because the X values are discrete, x_jitter was added with the Seaborn plotting package to reduce the number of overlapping points. WT mean and standard deviation are shown in red. Variants with a more net negative charge have expanded ensembles. The blue line is a ordinary least squares linear regression and the shading is a 95% CI generated by Seaborn regplot. CITED2 n = 1060, VP16 n = 730.

B) Variants with more aromatic residues have more collapsed ensembles. For A and B, we excluded variants with more than 10 aromatic residues, because these simulations did not converge well. We also excluded variants with a positive net charge, because these variants start to expand, as expected from the positive charges repelling each other.

C) The supercharge variants that added both acidic and aromatic residues show similar trends. Note, we only simulated 21Hif1α variants, including WT.
Figure S8: Leucine and acidic residues contribute to activity of VP16 and Hif1α

A-B) For each position, the activities of all variants that introduced a substitution at that position are summarized. The 4 biological replicates are included separately. Most variants changed multiple positions. Each residue was mutated a different number of times (1-22 variants, 4-88 measurements) and mutated to different amino acids. The shuffle mutants are not included.

A) In VP16, F442 (purple) is critical for activity. The leucine (green) and acidic (red) residues also contribute to activity.

B) For Hif1α, WT activity was low, reducing our ability to resolve decreases in activity. Leucine and acidic residues make large contributions to activity.

C) In the structure of the Hif1α-TAZ1 interaction (1L8C), positions with the lowest average activities panel B (purple) in Hif1α (orange) point towards the surface of the TAZ1 (white). Three rotations of the same snapshot are shown. The mutagenesis can detect positions that contact the coactivator.
Figure S9: Mutations designed to disrupt alpha helix formation in CITEDs and VP16 frequently reduce helicity in the all-atom Monte Carlo simulations

A) Example of predicted helicity at each residue for simulations for one CITED2 shuffle variant. The fraction helicity on the y-axis is the fraction of conformations where each residue (x-axis) is in a helix. For each variant, we ran 10 simulations. We ran 30 simulations for WT. For clarity, simulations starting in a helix are shown. The arrow indicates the alpha helix observed in the NMR interaction structure with TAZ1. The orange box indicates the shuffled region.

B, C) Shuffling CITED2 disrupts activity as much as removing key motifs. Disrupting the helix with 2xProline also decreased activity.

C) Shuffling regions of VP16 that overlap the alpha helix disrupt activity as much as removing key motifs. Disrupting the helix with 2xGlycine or 2xProline also decreased activity.

D, E) For each simulation, we integrated the fraction helicity over the full sequence (e.g. integral of dark blue curve in A). Sets of shuffle variants are shaded.

F) Schematic of motif locations, mutants, alpha helices (arrows) and shuffled regions.
Figure S10: Snapshots of CITED2 bound to the TAZ1 domain of CBP.

In 18/20 structures (1R8U), D224 (red) of CITED2 (orange) sits between the narrow, positively charged rims (blue) of the binding canyon on TAZ1 (white). CITED2 D224 is closest to R439 (upper blue) and K365 (lower blue) of TAZ1. The canyon is widest in structures 1 and 10, and D224 interacts primarily with R439. Structure 16 is shown in Figure 4C. Most variants that exchanged charged residues had nearly WT activity, but two variants with the D221E, D224E and D238 substitutions had decreased activity. D221 and D238 are not sterically constrained.

| Sequence          | Activity | Description                        |
|-------------------|----------|-------------------------------------|
| TDFI0EVLMLVINGNLR | R, D, K  | (C-terminal half)                   |
| TDFI0EVLMLVINGNLR | R, D, D  | (N-terminal half)                   |
| TDFI0EVLMLVINGNLR | R, D, D  | (D, D)                             |
| TDFI0EVLMLVINGNLR | R, D, D  | (D, D)                             |
| TDFI0EVLMLVINGNLR | R, D, D  | (D, D)                             |
| TDFI0EVLMLVINGNLR | R, D, D  | (D, D)                             |

WT
Figure S11: Strong ADs require acidic and hydrophobic residues

A) The mean activity of all variants of VP16 (X), CITED2 (†) and Hif1α (+) are shown. The location of the point indicates the net charge (x-axis) and number of W,F,Y,L residues (y-axis). The color indicates activity (AU). To prevent overlap, ADs with the same net charge and W,F,Y,L count are arrayed in a grid and ordered by activity. Figure 6A summarizes the same data with the median of each set. N = 302.

B) Similar to 6A but replacing W,F,Y,L count with aromatic count (W,F,Y).

C) Similar to 6A but replacing W,F,Y,L count with Kyte Doolittle Hydropathy. All variants are shown.

D-E) Composition based machine learning models can separate strong and weak ADs. The 5 parameter model (E) performs better than the 2 parameter model (D).

F) To simplify the learning problem and balance the data, we split the data into high and low activity sets. The black dot is the performance of the AD predictor that uses Net Charge and W,F,Y,L count (Figure 6A). The cyan dot is the performance of a predictor that uses Next Charge and aromatic (W,F,Y) count (Figure S11B). Including leucine residues improves performance in all cases.
Figure S12. Example of the FACS sorting gates (Replicate 2 of 5 AD experiment)

A) Gating for live cells with Forward Scatter Area and Side Scatter Area
B) Gating for single cells with Forward Scatter Area and Forward Scatter Width
C) Gating for single cells with Side Scatter Area and Side Scatter Width
D) Unused plot.
E) The R4 gate is the GFP positive population. This boundary is set as the 90th percentile of the uninduced cell population (cell without β-estradiol). R8 is the GFP negative population and ‘Bin1’ for low activity. It is set based on the bottom 50% of the uninduced population.
F) R5, R6, R7 are the high activity gates, Bin2, Bin3 and Bin4. They split the GFP positive population into three populations of similar size.
| Model Parameters          | Support Vector Avg. AUC | delta | Logistic Regression Avg. AUC | delta | Random Forest Avg. AUC | delta |
|---------------------------|-------------------------|-------|----------------------------|-------|------------------------|-------|
| Charge, WFYL              | 0.8511                  | 0     | 0.8524                     | 0     | 0.8976                 | 0     |
| Charge                    | 0.8498                  | -0.001| 0.8498                     | -0.003| 0.8752                 | -0.022|
| WFYL                      | 0.5184                  | -0.333| 0.4836                     | -0.369| 0.5301                 | -0.368|
| Charge, W, F, Y, L        | 0.8523                  | 0     | 0.865                      | 0     | 0.9221                 | 0     |
| W, F, Y, L                | 0.5574                  | -0.295| 0.6685                     | -0.197| 0.7891                 | -0.133|
| Charge, F, Y, L           | 0.8498                  | -0.002| 0.8632                     | -0.002| 0.9284                 | 0.0063|
| Charge, W, Y, L           | 0.8402                  | -0.012| 0.8584                     | -0.007| 0.9306                 | 0.0085|
| Charge, W, F, L           | 0.8429                  | -0.009| 0.8714                     | 0.0064| 0.9392                 | 0.0171|
| Charge, W, F, Y           | 0.8385                  | -0.014| 0.839                      | -0.026| 0.8778                 | -0.044|
| W, F, Y, L, E, D, K, R   | 0.8553                  | 0     | 0.8589                     | 0     | 0.9219                 | 0     |
| F, Y, L, E, D, K, R      | 0.8488                  | -0.006| 0.8589                     | 0     | 0.9201                 | -0.002|
| W, Y, L, E, D, K, R      | 0.8539                  | -0.001| 0.8593                     | 0.0004| 0.9013                 | -0.021|
| W, F, L, E, D, K, R      | 0.8515                  | -0.004| 0.8682                     | 0.0093| 0.9199                 | -0.002|
| W, F, Y, E, D, K, R      | 0.8362                  | -0.019| 0.8511                     | -0.008| 0.8839                 | -0.038|
| W, F, Y, L, D, K, R      | 0.7946                  | -0.061| 0.7947                     | -0.064| 0.9176                 | -0.004|
| W, F, Y, L, E, K, R      | 0.8089                  | -0.046| 0.859                      | 1E-04 | 0.8642                 | -0.058|
| W, F, Y, L, E, D, R      | 0.855                   | -3E-04| 0.8576                     | -0.001| 0.9056                 | -0.016|
| W, F, Y, L, E, D, K      | 0.8264                  | -0.029| 0.8291                     | -0.03  | 0.889                  | -0.033|

**Table S1: Average AUC values for Machine Learning Model variants**

For each of the three base models, we left out each parameter, retrained the model and computed the average AUC from 5-fold cross validation (see also Figure S11D,E). The delta column shows the change in model performance after removing each parameter. Net charge contributed more to model performance than aromatic and leucine residues. Leaving out leucine residues generally caused a larger drop in model performance than leaving out any one aromatic residue (bolded delta values). The WFYL parameter combines the counts of these four amino acids. W, tryptophan; F, phenylalanine; T, tyrosine; L, leucine.
| Experiment      | Replicate | Bin1  | Bin2  | Bin3  | Bin4  | Bin5  | Bin6  | Bin7  | Bin8  |
|-----------------|-----------|-------|-------|-------|-------|-------|-------|-------|-------|
| 5 acidic ADs    | Sort 1    | 82.3  | 604   | 2037  | 11215 |       |       |       |       |
|                 | Sort 2    | 56.5  | 489   | 1755  | 9456  |       |       |       |       |
|                 | Sort 3    | 43.7  | 471   | 1750  | 9431  |       |       |       |       |
|                 | Sort 4    | 41.2  | 386   | 1618  | 9716  |       |       |       |       |
| 8 bin Sort      |           | 23.4  | 49.4  | 84.1  | 176   | 464   | 1292  | 5184  | 18964 |
| p53             | Sort 1    | 6.59  | 25.08 | 53.95 | 141.3 |       |       |       |       |
|                 | Sort 2    | 3.54  | 5.46  | 8.12  | 25.81 |       |       |       |       |
|                 | Sort 3    | 31.31 | 46.05 | 56.87 | 66.25 | 79.55 | 99.71 | 146.61| 329.44|
| Predicted ADs   | Sort 1    | 35.19 | 62.33 | 89.43 | 258.7 |       |       |       |       |
|                 | Sort 2    | 39.92 | 72.05 | 109.73| 273.49|       |       |       |       |
|                 | Sort 3    | 40.64 | 70.58 | 109.92| 337.19|       |       |       |       |

Table S2: GFP Fluorescence values from FACS sorting
The median GFP fluorescence values from each FACS sort. These values are used to compute the activities of each AD variant (see methods).