Supplemental Results: Improved HIV-1 Neutralization Breadth and Potency of V2-Apex Antibodies by *In Silico* Design

Graham T. Holt[1,2,†], Jason Gorman[3,†], Siyu Wang[2], Anna U. Lowegard[1,2], Baoshan Zhang[3], Tracy Liu[3], Bob C. Lin[3], Mark K. Louder[3], Marcel S. Frenkel[4], Krisha McKee[3], Sijy O'Dell[3], Reda Rawi[3], Chen-Hsiang Shen[3], Nicole A. Doria-Rose[3], Peter D. Kwong[3,*], and Bruce R. Donald[1,4,5,6,*,‡]

[1] *Department of Computer Science, Duke University, Durham NC, USA*
[2] *Program in Computational Biology & Bioinformatics, Duke University, Durham NC, USA*
[3] *Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA*
[4] *Department of Biochemistry, Duke University, Durham NC, USA*
[5] *Department of Mathematics, Duke University, Durham NC, USA*
[6] *Department of Chemistry, Duke University, Durham NC, USA*

Table SR1: Predicted K* scores for PGT145 WT and alanine mutants at residues of interest using OSPREY vs. corresponding IC_{50} data[1], related to Results.

	IC_{50} (μg/mL)[1]		K* score ($\log_{10}$K*)		Summary	
Residue	WT	Ala	WT	Ala	$-\Delta \log_{10}IC_{50}$	$\Delta \log_{10}$K*
100d	0.010	2.380	3.438	2.483	−2.377	−0.955
100e	0.010	9.991	7.995	6.278	−3.000	−1.717
100l	0.010	0.004	15.454	15.155	0.398	−0.299
100m	0.010	>10	4.909	3.952	< -3	−0.957

[1] IC_{50} for BG505.Env.C2 pseudovirus.

Table SR2: Upper and lower bounds on the K* score and partition functions for PGT145 designs at residues 100e and 100m predicted using OSPREY, related to Results. H_δ and H_ε refer to the δ and ε protonation states of histidine.

		$\log_{10}$K*		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100e	100m	UB	LB	UB	LB	UB	LB	UB	LB
E	E	18.239	16.797	29.182	28.687	6.963	6.468	4.971	4.476
E	D	18.202	16.802	29.035	28.540	6.862	6.367	4.961	4.466
E	Y	16.646	15.292	31.986	31.491	11.369	10.874	4.961	4.466
E	M	16.412	15.047	30.716[1]	30.221	10.213[1]	9.825	4.961[1]	4.479
E	H_ϵ	16.341	14.876	30.823	30.328	10.511	10.016	4.961	4.466
D	E	16.336	14.908	27.329	26.834	7.028	6.533	4.955	4.460
E	H_δ	16.327	14.900	30.937[1]	30.442	10.581[1]	10.143	4.961[1]	4.468
M	E	16.221	14.766	30.752[1]	30.257	10.536[1]	10.061	4.955[1]	4.470
D	D	16.161	14.718	27.003	26.508	6.870	6.375	4.961	4.466
E	Q	16.159	14.718	29.664	29.169	9.533	9.039	4.961	4.466
H_ϵ	E	16.148	14.671	30.558[1]	30.063	10.430[1]	9.941	4.962[1]	4.469
M	D	16.088	14.608	30.438[1]	29.943	10.372[1]	9.881	4.963[1]	4.469
H_ϵ	D	16.077	14.612	30.286[1]	29.792	10.222[1]	9.745	4.957[1]	4.465
L	E	15.976	14.515	30.573	30.078	10.627	10.132	4.960	4.465
Q	D	15.948	14.496	29.749[1]	29.254	9.801[1]	9.331	4.957[1]	4.470
Q	E	15.902	14.428	29.884[1]	29.389	9.997[1]	9.511	4.964[1]	4.471
L	D	15.896	14.467	30.340	29.845	10.478	9.983	4.956	4.461
E	I	15.824	14.569	30.803	30.308	10.999	10.505	4.969	4.474
E	T	15.689	14.222	29.558	29.063	9.897	9.403	4.961	4.466
E	V	15.647	14.217	29.620	29.126	10.003	9.508	4.961	4.466
E	S	15.525	14.056	29.006	28.511	9.514	9.019	4.958	4.463
E	C	15.510	14.096	28.813	28.318	9.332	8.837	4.961	4.466
E	A	15.467	14.041	28.794	28.299	9.355	8.860	4.961	4.466
E	G	15.352	13.938	28.250	27.755	8.926	8.431	4.961	4.466
H_δ	E	15.264	13.854	30.129[1]	29.634	10.821[1]	10.398	4.960[1]	4.468
H_δ	D	15.247	13.818	29.935[1]	29.441	10.660[1]	10.220	4.962[1]	4.469
T	E	15.090	13.609	29.368[1]	28.873	10.307[1]	9.812	4.958[1]	4.466
T	D	15.075	13.596	29.173[1]	28.678	10.128[1]	9.635	4.954[1]	4.462
V	D	15.032	13.594	28.938	28.443	9.935	9.440	4.961	4.466
V	E	15.021	13.682	29.194	28.699	10.202	9.707	4.961	4.466
D	Y	14.333	12.893	29.594	29.099	11.290	10.795	4.961	4.466
M	Y	14.053	12.579	32.213[1]	31.718	14.178[1]	13.688	4.960[1]	4.471
M	M	14.040	12.563	31.158[1]	30.663	13.138[1]	12.644	4.963[1]	4.474
D	H_ϵ	13.952	12.489	28.437	27.942	10.503	10.008	4.972	4.477
D	M	13.918	12.448	28.261	27.766	10.368	9.873	4.965	4.470
M	H_ϵ	13.909	12.432	31.375[1]	30.880	13.494[1]	13.004	4.954[1]	4.462
D	H_δ	13.881	12.400	28.497	28.002	10.648	10.153	4.957	4.462
H_ϵ	M	13.852	12.372	30.886[1]	30.391	13.057[1]	12.563	4.961[1]	4.470
Q	Y	13.836	12.389	31.446[1]	30.951	13.605[1]	13.127	4.957[1]	4.483
Q	M	13.830	12.348	30.382[1]	29.887	12.577[1]	12.083	4.961[1]	4.469
H_ϵ	Y	13.821	12.372	31.884[1]	31.389	14.058[1]	13.578	4.958[1]	4.485

Upper and lower bounds on the K* score and partition functions for PGT145 designs at residues 100e and 100m predicted using OSPREY (continued).

		$\log_{10}$K*		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100e	100m	UB	LB	UB	LB	UB	LB	UB	LB
D	Q	13.788	12.340	27.322	26.827	9.562	9.068	4.961	4.466
M	Q	13.768	12.293	30.170[1]	29.676	12.420[1]	11.934	4.962[1]	4.469
M	H_δ	13.768	12.312	31.141[1]	30.646	13.383[1]	12.911	4.952[1]	4.462
L[2]	Y[2]	13.763	12.296	32.039	31.544	14.309	13.814	4.956	4.461
L	M	13.732	12.275	31.009	30.514	13.306	12.811	4.961	4.466
H_ϵ	H_ϵ	13.710	12.260	31.091[1]	30.596	13.382[1]	12.915	4.954[1]	4.466
Q	H_ϵ	13.688	12.245	30.610[1]	30.115	12.915[1]	12.442	4.955[1]	4.480
H_ϵ	Q	13.628	12.180	29.958[1]	29.463	12.322[1]	11.862	4.961[1]	4.467
Q	H_δ	13.600	12.131	30.410[1]	29.915	12.828[1]	12.342	4.957[1]	4.469
Q	Q	13.571	12.102	29.412[1]	28.917	11.862[1]	11.377	4.953[1]	4.463
L	H_ϵ	13.571	12.196	31.225	30.730	13.683	13.188	4.961	4.466
H_ϵ	H_δ	13.546	12.076	30.835[1]	30.341	13.304[1]	12.823	4.960[1]	4.466
L	Q	13.503	12.037	30.019	29.525	12.551	12.056	4.955	4.460
L	H_δ	13.503	12.024	30.964	30.469	13.492	12.997	4.959	4.464
D	I	13.356	11.935	28.198	27.703	10.860	10.365	4.972	4.477
D	T	13.305	11.829	27.164	26.669	9.892	9.397	4.957	4.462
M	T	13.295	11.815	30.043[1]	29.548	12.772[1]	12.280	4.961[1]	4.467
H_δ	M	13.204	11.777	30.605[1]	30.110	13.378[1]	12.941	4.956[1]	4.461
M	I	13.177	11.728	30.990[1]	30.495	13.807[1]	13.347	4.960[1]	4.465
H_ϵ	T	13.142	11.666	29.815[1]	29.320	12.693[1]	12.201	4.961[1]	4.472
D	V	13.132	11.699	27.154	26.659	10.051	9.556	4.961	4.466
M	C	13.100	11.626	29.174[1]	28.679	12.093[1]	11.604	4.960[1]	4.469
M	S	13.097	11.620	29.423[1]	28.928	12.352[1]	11.863	4.955[1]	4.464
Q	T	13.084	11.614	29.303[1]	28.808	12.235[1]	11.751	4.959[1]	4.468
H_δ	Y	13.081	11.641	31.472[1]	30.977	14.373[1]	13.915	4.963[1]	4.475
D	C	13.075	11.648	26.330	25.835	9.283	8.788	4.961	4.466
H_δ	H_ϵ	13.054	11.586	30.820[1]	30.325	13.779[1]	13.300	4.960[1]	4.465
D	S	13.040	11.588	26.588	26.093	9.573	9.078	4.965	4.470
Q	I	13.026	11.620	30.246[1]	29.751	13.175[1]	12.757	4.956[1]	4.462
L	T	13.018	11.541	29.892	29.397	12.905	12.410	4.958	4.464
H_ϵ	C	12.998	11.528	28.930[1]	28.436	11.951[1]	11.465	4.957[1]	4.467
M	V	12.986	11.543	29.841[1]	29.346	12.844[1]	12.390	4.958[1]	4.464
T	M	12.969	11.497	29.849[1]	29.354	12.898[1]	12.408	4.959[1]	4.472
M	A	12.964	11.499	29.149[1]	28.654	12.187[1]	11.708	4.968[1]	4.477
T	Y	12.955	11.476	30.892[1]	30.397	13.958[1]	13.466	4.962[1]	4.471
E	R	12.953	11.495	30.116	29.622	13.182	12.687	4.972	4.477
Q	C	12.951	11.474	28.451[1]	27.956	11.527[1]	11.038	4.955[1]	4.462
H_ϵ	S	12.945	11.464	29.188[1]	28.693	12.276[1]	11.782	4.954[1]	4.462
D	A	12.934	11.566	26.339	25.844	9.434	8.939	4.961	4.466
V	Y	12.925	11.468	30.651	30.157	13.755	13.260	4.961	4.466
H_δ	Q	12.888	11.412	29.553[1]	29.058	12.687[1]	12.198	4.960[1]	4.466
L	I	12.885	11.475	30.824	30.329	13.964	13.469	4.965	4.470
H_ϵ	I	12.871	11.445	30.569[1]	30.074	13.668[1]	13.229	4.961[1]	4.469

Upper and lower bounds on the K* score and partition functions for PGT145 designs at residues 100e and 100m predicted using OSPREY (continued).

		$\log_{10}$K*		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100e	100m	UB	LB	UB	LB	UB	LB	UB	LB
Q	S	12.859	11.401	28.679[1]	28.184	11.820[1]	11.348	4.963[1]	4.472
D	G	12.834	11.444	25.793	25.298	8.987	8.492	4.961	4.466
L	C	12.830	11.370	29.041	28.546	12.232	11.737	4.969	4.474
L	S	12.824	11.351	29.274	28.779	12.478	11.984	4.961	4.466
M	G	12.815	11.402	28.560[1]	28.066	11.707[1]	11.280	4.957[1]	4.465
T	H_ϵ	12.785	11.313	30.043[1]	29.549	13.275[1]	12.781	4.960[1]	4.477
V	M	12.782	11.418	29.605	29.110	12.852	12.357	4.961	4.466
H_δ	H_δ	12.776	11.418	30.460[1]	29.965	13.587[1]	13.215	4.960[1]	4.469
H_ϵ	V	12.758	11.390	29.549[1]	29.054	12.707[1]	12.328	4.958[1]	4.464
Q	A	12.756	11.308	28.400[1]	27.905	11.638[1]	11.167	4.958[1]	4.477
Q	V	12.756	11.340	29.097[1]	28.602	12.304[1]	11.860	4.958[1]	4.481
V	H_ϵ	12.750	11.299	29.806	29.311	13.089	12.595	4.956	4.461
H_ϵ	A	12.673	11.362	28.899[1]	28.405	12.089[1]	11.762	4.954[1]	4.464
L	V	12.657	11.286	29.669	29.174	13.034	12.540	4.967	4.472
T	Q	12.634	11.162	28.834[1]	28.340	12.219[1]	11.728	4.959[1]	4.472
T	H_δ	12.623	11.176	29.808[1]	29.313	13.182[1]	12.707	4.954[1]	4.477
L	A	12.584	11.262	28.996	28.501	12.447	11.952	4.955	4.460
V	H_δ	12.582	11.158	29.556	29.061	13.002	12.507	4.961	4.466
Q	G	12.581	11.180	27.815[1]	27.320	11.180[1]	10.761	4.960[1]	4.472
V	Q	12.549	11.175	28.609	28.114	12.086	11.591	4.964	4.469
L	G	12.514	11.110	28.402	27.907	11.916	11.421	4.961	4.466
H_ϵ	G	12.473	11.001	28.303[1]	27.809	11.846[1]	11.362	4.962[1]	4.469
H_δ	T	12.406	10.939	29.446[1]	28.951	13.052[1]	12.573	4.961[1]	4.467
H_δ	I	12.405	10.934	30.324[1]	29.829	13.942[1]	13.454	4.953[1]	4.466
H_δ	S	12.223	10.742	28.828[1]	28.333	12.633[1]	12.140	4.959[1]	4.465
H_δ	C	12.208	10.735	28.573[1]	28.078	12.386[1]	11.897	4.957[1]	4.468
T	T	12.158	10.682	28.704[1]	28.209	12.570[1]	12.077	4.957[1]	4.469
V	T	12.128	10.664	28.467	27.972	12.368	11.873	4.961	4.466
T	I	12.119	10.644	29.643[1]	29.148	13.543[1]	13.056	4.961[1]	4.468
H_δ	G	12.047	10.568	28.103[1]	27.608	12.079[1]	11.590	4.960[1]	4.466
H_δ	A	12.023	10.651	28.534[1]	28.039	12.436[1]	12.053	4.952[1]	4.458
T	C	11.968	10.501	27.838[1]	27.344	11.882[1]	11.404	4.961[1]	4.467
T	S	11.953	10.471	28.090[1]	27.595	12.167[1]	11.674	4.957[1]	4.463
V	S	11.926	10.450	27.840	27.345	11.942	11.447	4.961	4.466
V	I	11.921	10.641	29.419	28.924	13.527	13.033	4.961	4.466
V	C	11.897	10.467	27.575	27.080	11.703	11.208	4.964	4.469
T	V	11.857	10.408	28.513[1]	28.018	12.649[1]	12.188	4.962[1]	4.468
T	A	11.835	10.372	27.819[1]	27.324	12.006[1]	11.513	4.946[1]	4.471
H_δ	V	11.835	10.536	29.093[1]	28.598	13.100[1]	12.789	4.962[1]	4.469
M	R	11.784	10.304	29.923[1]	29.428	14.160[1]	13.668	4.964[1]	4.471
V	V	11.739	10.375	28.272	27.777	12.568	12.073	4.955	4.460
T	G	11.704	10.248	27.234[1]	26.739	11.537[1]	11.046	4.954[1]	4.484
V	A	11.692	10.343	27.564	27.069	11.901	11.406	4.961	4.466

Upper and lower bounds on the K* score and partition functions for PGT145 designs at residues 100e and 100m predicted using OSPREY (continued).

		$\log_{10}$K*		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100e	100m	UB	LB	UB	LB	UB	LB	UB	LB
Q	R	11.579	10.104	29.258[1]	28.763	13.705[1]	13.216	4.954[1]	4.463
L	R	11.473	10.010	29.745	29.250	14.301	13.806	4.961	4.466
V	G	11.456	10.215	26.956	26.461	11.533	11.038	4.957	4.462
H_ϵ	R	11.434	9.963	29.676[1]	29.181	14.263[1]	13.780	4.956[1]	4.463
H_δ	R	10.948	9.473	29.260[1]	28.765	14.333[1]	13.848	4.959[1]	4.464
V	R	10.608	9.132	28.326	27.832	13.748	13.253	4.961	4.466
T	R	10.593	9.118	28.603[1]	28.109	14.035[1]	13.540	4.956[1]	4.470
D	R	10.578	9.151	27.473	26.979	12.914	12.419	4.972	4.477

[1] Upper bounds computed using lower bound and $\epsilon = 0.68$.
[2] Wild-type amino-acid.

Table SR3: Root mean-squared deviation between existing structures of PGT145, related to Results.

	RMSD (Å)[1,2]					
PDB ID	8FLW[3]	6NIJ	58VL	8DVD	3U1S	5U1F
8FLW[3]	0.00 (140)					
6NIJ	1.95 (140)	0.00 (140)				
5V8L	1.32 (140)	1.69 (140)	0.00 (140)			
8DVD	1.40 (140)	1.86 (140)	1.03 (140)	0.00 (140)		
3U1S	1.35 (140)	1.66 (140)	0.76 (140)	0.76 (140)	0.00 (240)	
5U1F	1.37 (138)	1.69 (138)	0.80 (138)	0.84 (138)	1.16 (235)	0.00 (235)

[1] RMSD computed for all resolved PGT145 antibody heavy chain C_α atoms using the PyMOL `align` command with the `cycles=0` option.
[2] Numbers in parentheses represent the number of atoms included in the RMSD calculation.
[3] Cryo-EM structure of PGT145 DU303.

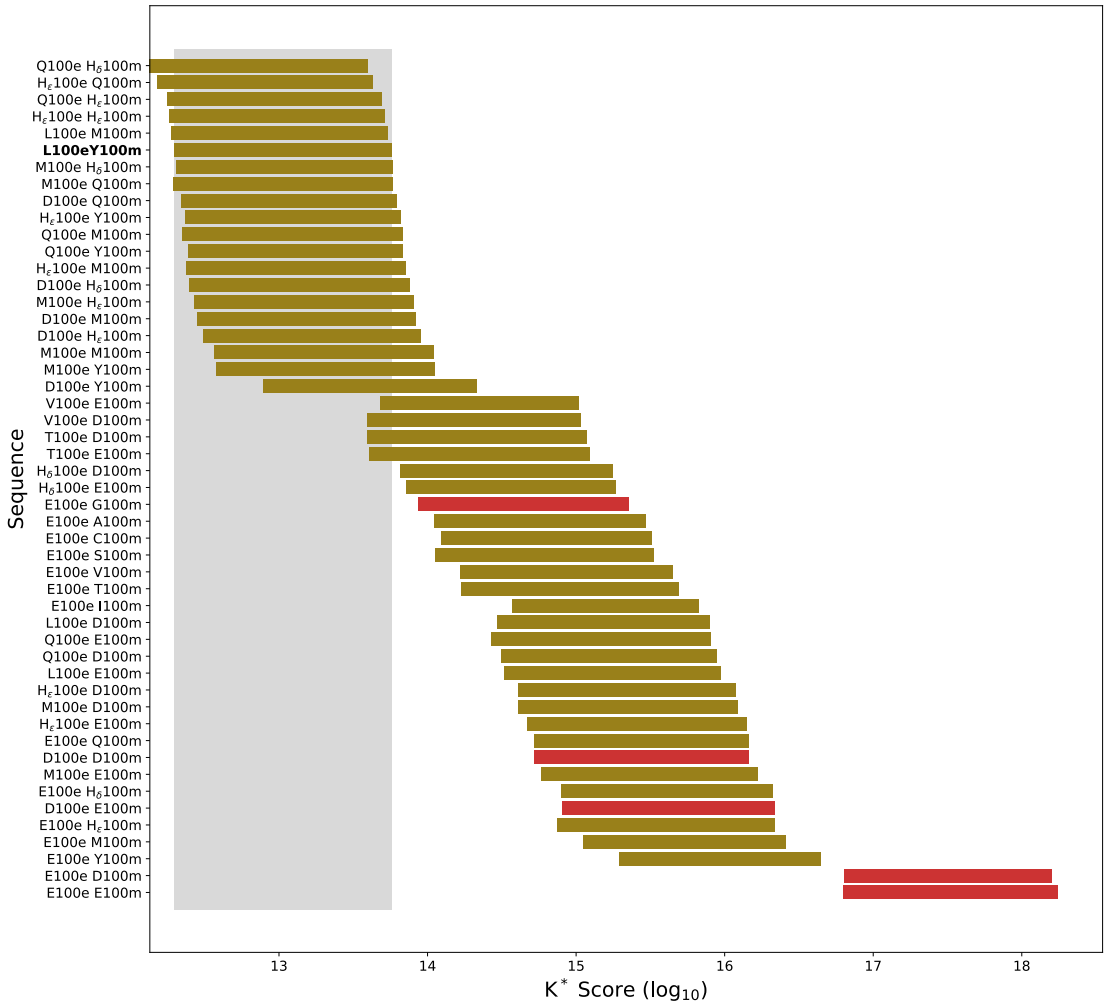


Figure SR1: K* scores for a double-mutation design of PGT145 at residues 100e and 100m, related to Results. Bounds on the K* score for single mutations predicted using OSPREY are shown as horizontal bars. Pictured results are limited to the top 50 design predictions. Designs for which the unbound antibody is predicted to be more stable, approximately equally-stable, or less stable than wild-type are indicated by green, tan, or red coloring, respectively. Relative stability was estimated using the lower-bound on the partition-function value for the unbound antibody state. Wild-type antibody residue labeled in bold. The bounds on the K* score for the wild-type design are indicated by the gray, shaded box.

Table SR4: Root mean-squared deviation between existing structures of PG9, related to Results.

	RMSD (Å)[1,2]							
PDB ID	8FK5[3]	8FL1[4]	7T77	7T76	3U4E	3U2S	4YAQ	5VJ6
8FK5[3]	0.00 (136)							
8FL1[4]	0.57 (136)	0.00 (136)						
7T77	0.61 (133)	0.85 (133)	0.00 (133)					
7T76	0.81 (135)	1.00 (135)	0.61 (133)	0.00 (136)				
3U4E	0.62 (136)	0.83 (136)	0.62 (133)	0.84 (135)	0.00 (239)			
3U2S	0.62 (136)	0.87 (136)	0.44 (133)	0.77 (135)	3.38 (230)	0.00 (230)		
4YAQ	0.69 (136)	0.92 (136)	0.53 (133)	0.85 (135)	5.14 (239)	1.94 (230)	0.00 (239)	
5VJ6	2.18 (135)	2.34 (135)	2.06 (133)	2.02 (135)	2.35 (229)	3.59 (229)	5.16 (229)	0.00 (229)

[1] RMSD computed for all resolved PG9 antibody heavy chain C_α atoms using the PyMOL `align` command with the `cycles=0` option.
[2] Numbers in parentheses represent the number of atoms included in the RMSD calculation.
[3] Cryo-EM structure of PG9RSH N(100f)Y.
[4] Cryo-EM structure of DU025.

Table SR6: Flexible residues for designs PGT145 with AMC011 or SIVmac239, related to Results.

PDB ID	**Name**	**Target**	**Mutable res**	**Flexible res**
6NIJ	PGT145:Env (AMC011)	100d, 100l	H100D, H100L	H100A, H100N, A168, A169, C166, C167
8DVD	PGT145:Env (SIVmac239)	100d, 100l	H100D, H100L	H100A, H100N, E168, E169, G166, G167

Table SR5: Predicted K* scores for PGT145 WT and variants using OSPREY vs. corresponding small-panel neutralization data, related to Results. In general, K* scores do not correlate quantitatively with binding affinity[2] except in some cases in which multiple high-resolution structures are used[3]. For these reasons, in this study, the K* scores were used as a filter, to select a small number (10) of candidate PGT145 designs (out of thousands of possible sequences) for experimental testing, with the hypothesis that some would increase breath or potency as predicted.

	IC_{50} (μg/mL)[1]		K* score ($\log_{10}$K*)[2]		Summary	
Ab Variant	WT	Variant	WT	Variant	$-\Delta \log_{10}IC_{50}$	$\Delta\log_{10}$K*
F(100d)E	0.080	0.698	5.288	6.911	-0.094	1.623
L(100e)E	0.080	>50	13.763	16.646	< -2.80	2.883
N(100l)D	0.080	0.017	5.288	8.891	0.398	3.603
Y(100m)E	0.080	>50	13.763	15.976	< -2.80	2.213

[1] IC_{50} for BG505.Env.C2 pseudovirus. Geometric mean over several assays reported for wild-type value.
[2] Upper bound on the K* score reported.

Table SR7: Upper and lower bounds on the K* score and partition functions for designs of PGT145 in complex with AMC011 at residues 100d and 100l predicted using OSPREY, related to Results. H_δ and H_ε refer to the δ and ε protonation states of histidine.

		$\log_{10}$K*		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100d	100l	UB	LB	UB	LB	UB	LB	UB	LB
W	D	9.845	8.816	46.390	46.149	20.009	19.527	17.324	17.018
Y	D	9.355	8.493	46.396	46.151	20.334	20.023	17.324	17.018
F	D	9.325	8.506	45.941	45.731	19.901	19.598	17.324	17.018
W	Y	9.105	8.096	46.037	45.849	20.429	19.914	17.324	17.018
H_δ	Y	8.857	7.749	44.974	44.759	19.686	19.099	17.324	17.018
D	H_ϵ	8.798	7.841	41.889	41.644	16.479	16.073	17.324	17.018
W	T	8.794	7.720	48.532	48.321	23.277	22.721	17.324	17.018
W	H_ϵ	8.762	7.936	48.123	47.939	22.679	22.342	17.324	17.018
F	H_ϵ	8.737	7.767	47.691	47.491	22.400	21.936	17.324	17.018

Upper and lower bounds on the K* score and partition functions for designs of PGT145 in complex with AMC011 at residues 100d and 100l predicted using OSPREY (continued).

		$\log_{10}$K*		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100d	100l	UB	LB	UB	LB	UB	LB	UB	LB
Y	Y	8.681	7.796	46.114	45.908	20.789	20.415	17.324	17.018
W	Q	8.628	7.512	44.841	44.596	19.760	19.194	17.324	17.018
H_δ	H_ϵ	8.587	7.498	47.009	46.765	21.944	21.404	17.324	17.018
W	S	8.551	7.389	48.141	47.898	23.185	22.572	17.324	17.018
F	Y	8.530	7.681	45.545	45.307	20.302	19.997	17.324	17.018
D	M	8.511	7.389	43.109	42.863	18.151	17.580	17.324	17.018
W	M	8.417	7.427	49.377	49.139	24.388	23.942	17.324	17.018
Y	H_ϵ	8.414	7.547	48.161	47.922	23.051	22.729	17.324	17.018
D	N	8.405	7.380	42.413	42.178	17.475	16.991	17.324	17.018
W	N	8.353	7.418	48.753	48.515	23.774	23.382	17.324	17.018
H_ϵ	H_ϵ	8.253	7.353	46.331	46.092	21.415	21.060	17.324	17.018
F[1]	N[1]	7.941	7.088	48.420	48.174	23.762	23.461	17.324	17.018

[1] Wild-type amino-acid.

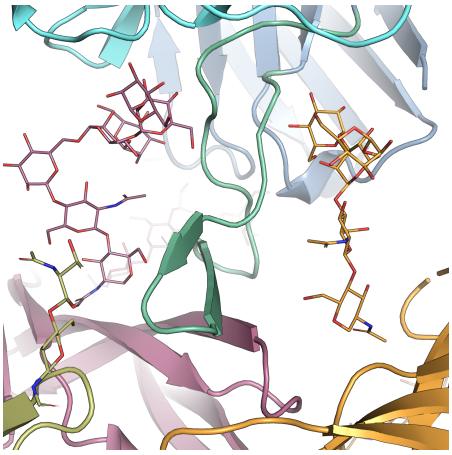
Table SR8: Upper and lower bounds on the K* score and partition functions for designs of PGT145 in complex with SIVmac239 at residues 100d and 100l predicted using OSPREY, related to Results. H_δ and H_ε refer to the δ and ε protonation states of histidine.

		$\log_{10}$K*		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100d	100l	UB	LB	UB	LB	UB	LB	UB	LB
H_δ	E	12.040	10.673	44.537	44.292	21.204	20.585	12.415	11.912
F	E	11.704	10.418	45.280	45.034	22.202	21.664	12.415	11.912
E	M	11.685	10.329	46.220	45.975	23.232	22.624	12.415	11.912
H_ϵ	E	11.504	10.172	45.117	44.872	22.285	21.701	12.415	11.912
E	H_ϵ	11.479	10.210	46.581	46.336	23.712	23.190	12.415	11.912
E	Q	11.418	10.080	45.046	44.801	22.306	21.716	12.415	11.912
E	V	11.311	9.959	44.408	44.163	21.789	21.185	12.415	11.912
E	T	11.250	9.890	44.781	44.536	22.232	21.620	12.415	11.912
E	N	11.243	9.938	45.009	44.764	22.411	21.855	12.415	11.912
E	S	11.153	9.790	44.176	43.931	21.726	21.112	12.415	11.912
E	C	11.099	9.829	44.518	44.273	22.029	21.508	12.415	11.912
H_δ	M	10.970	9.620	47.177	46.932	24.897	24.296	12.415	11.912
F	M	10.946	9.635	48.032	47.786	25.737	25.174	12.415	11.912
F	H_ϵ	10.915	9.600	49.488	49.242	27.228	26.660	12.415	11.912
W	H_ϵ	10.881	9.573	47.387	47.141	25.153	24.593	12.415	11.912
W	N	10.802	9.447	45.429	45.184	23.323	22.716	12.415	11.912
H_δ	H_ϵ	10.773	9.449	48.280	48.035	26.171	25.596	12.415	11.912
Y	R	10.761	9.414	47.990	47.745	25.916	25.318	12.415	11.912
H_δ	Q	10.639	9.286	46.724	46.479	24.778	24.173	12.415	11.912
F	Q	10.539	9.289	47.731	47.486	25.782	25.281	12.415	11.912

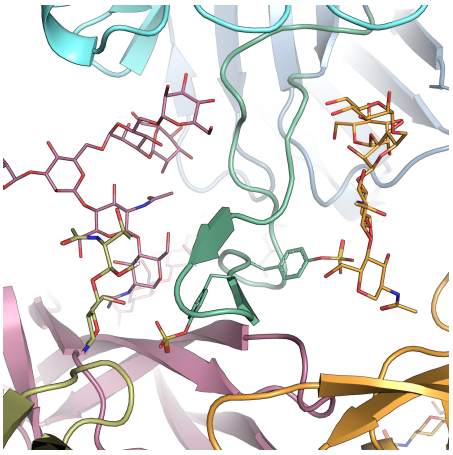
Upper and lower bounds on the K* score and partition functions for designs of PGT145 in complex with SIVmac239 at residues 100d and 100l predicted using OSPREY (continued).

		$\log_{10}$K*		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100d	100l	UB	LB	UB	LB	UB	LB	UB	LB
F[1]	N[1]	10.288	9.235	47.248	47.003	25.353	25.048	12.415	11.912

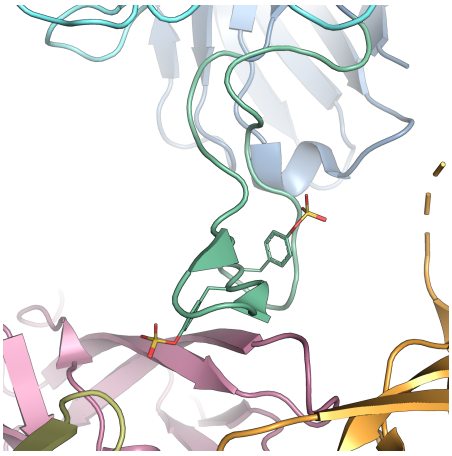
[1] Wild-type amino-acid.



(A) The CDRH3 loop for PG9RSH N(100f)Y exhibits a twisted conformation.



(B) The CDRH3 loop for DU025 exhibits a twisted conformation.



(C) The CDRH3 loop for PG9 (PDB ID: 5VJ6) exhibits an untwisted conformation.

Figure SR2: The CDHR3 loop conformation differs between PG9RSH variants and PG9, related to Results. Backbone shown in ribbon representation with glycans and sulfated tyrosines shown as lines with Env subunits colored with warm colors, antibody in cool colors, and antibody CDRH3 loops (residues 95-102) shown in green. A pronounced twist in the CDRH3 loop of PG9RSH variants N(100f)Y (A) and DU025 (B) can be seen, contrasted with the untwisted conformation in PG9 (C). As a result, residues 100i - 100n are displaced by up to 3.6 Å.

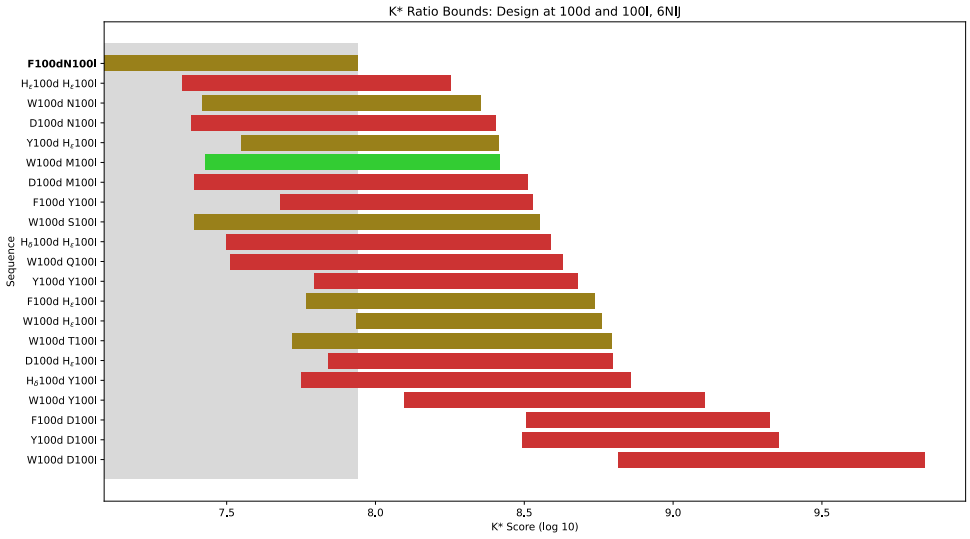


Figure SR3: K* scores for a double-mutation design of PGT145 in complex with AMC011 at residues 100d and 100l, related to Results. Bounds on the K* score for single mutations predicted using OSPREY are shown as horizontal bars. Pictured results are limited to the top 20 design predictions. Designs for which the unbound antibody is predicted to be more stable, approximately equally-stable, or less stable than wild-type are indicated by green, tan, or red coloring, respectively. Relative stability was estimated using the lower-bound on the partition-function value for the unbound antibody state. Wild-type antibody residue labeled in bold. The bounds on the K* score for the wild-type design are indicated by the gray, shaded box.

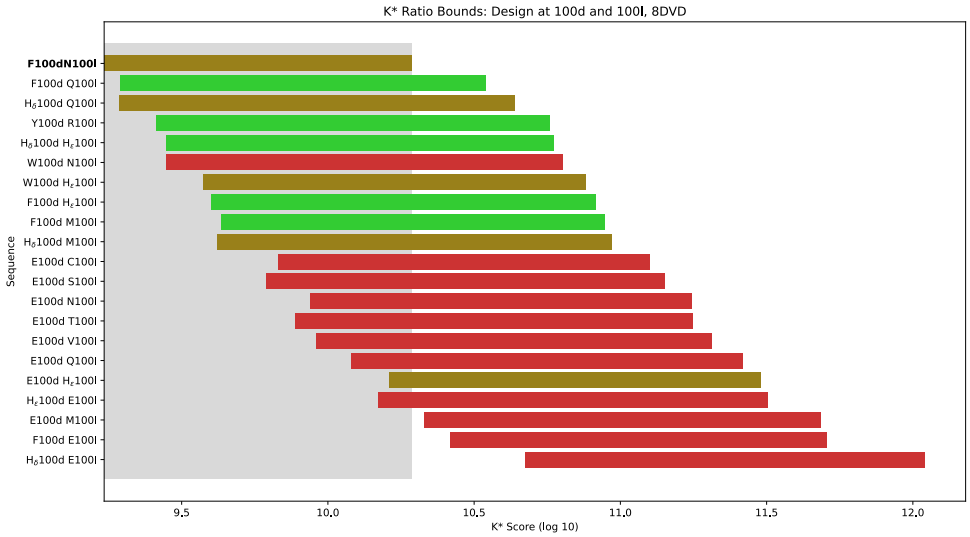
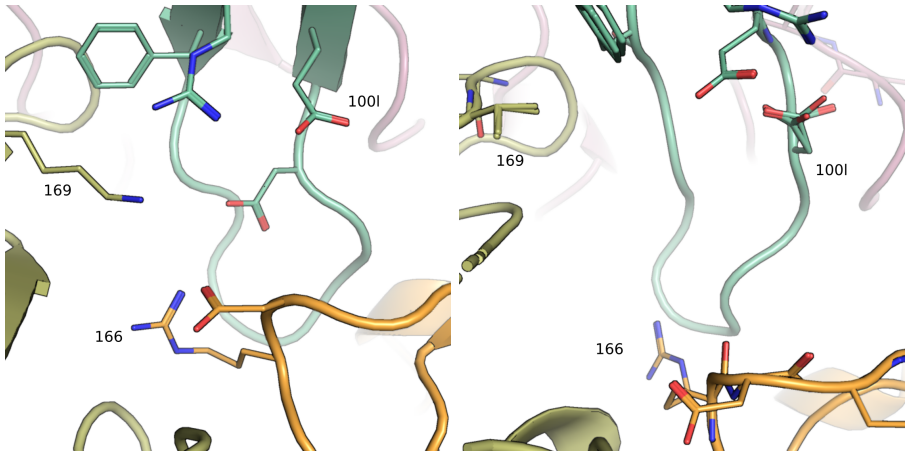
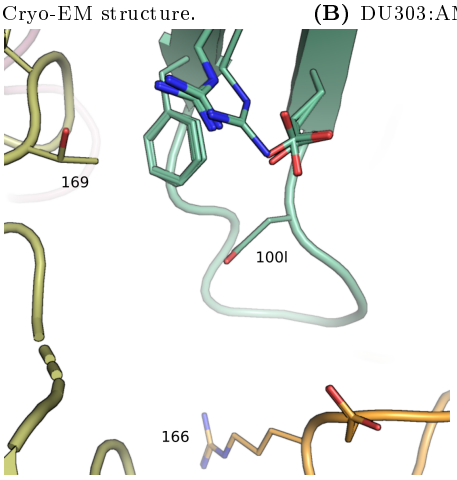


Figure SR4: K* scores for a double-mutation design of PGT145 in complex with SIVmac239 at residues 100d and 100l, related to Results. Bounds on the K* score for single mutations predicted using OSPREY are shown as horizontal bars. Pictured results are limited to the top 20 design predictions. Designs for which the unbound antibody is predicted to be more stable, approximately equally-stable, or less stable than wild-type are indicated by green, tan, or red coloring, respectively. Relative stability was estimated using the lower-bound on the partition-function value for the unbound antibody state. Wild-type antibody residue labeled in bold. The bounds on the K* score for the wild-type design are indicated by the gray, shaded box.

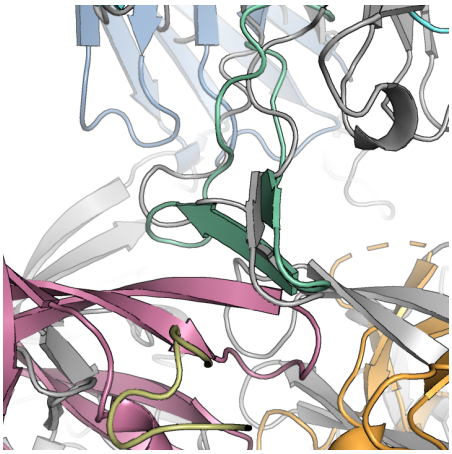


(A) DU303:BG505 Cryo-EM structure. **(B)** DU303:AMC011 model.

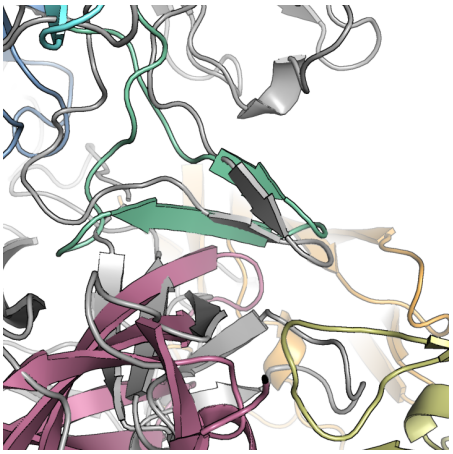


(C) DU303:SIVmac239 model.

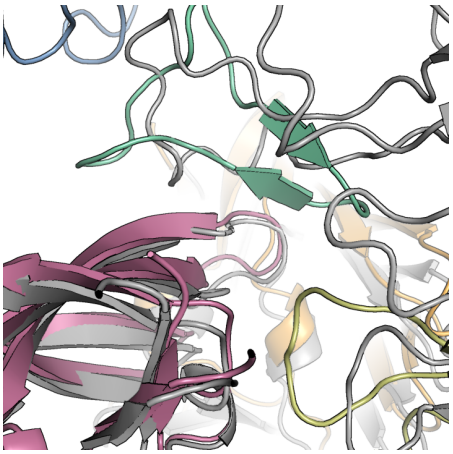
Figure SR5: Conservative homology models predict that the N(100l)D mutation will not create intermolecular contacts with AMC011 or SIVmac239, related to Results. Backbone shown in ribbon representation with flexible residues shown as lines with Env subunits colored with warm colors and antibody CDRH3 loops (residues 95-102) shown in green. Models predict that antibody residue D(100l) does not interact with position 169 due to residue substitution, and does not interact with residue 166 due to CDRH3 loop and V2 loop conformational shifts.



(A) AF2 models of DU011 bound to JRFL.JB Env



(B) AF2 models of DU011 bound to ZM233.6 Env



(C) AF2 models of DU025 bound to JRFL.JB Env

Figure SR6: Alphafold Multimer predicted interactions between antibodies and Env are incorrect, related to Results. Backbone shown in ribbon representation. The Cryo-EM structures are shown in color, with Env subunits in warm colors and antibody in cool colors, and antibody CDRH3 loops (residues 95-102) shown in green. Alphafold Multimer models are shown in gray. Interestingly, in some cases Alphafold Multimer does predict that PG9 binds to the Env apex, but incorrectly predicts A) an antiparallel β-strand interaction, B) a "cross strand" interaction, or C) an interaction in which the antibody lies parallel to the Env apex plane. In most cases, including all predicted models with BG505 Env and predicted models of DU025 in complex with TRO.11, physically reasonable predicted structures did not place the antibody at the Env apex (not pictured).

SR1 Extended design results

SR1.1 Validation of design pipeline on published experimental data

To validate our computational protein design methods on the antibody structure input models we first attempted to compare point-mutation designs to published experimental data. We performed four single-residue designs on the PGT145 CDRH3 loop (Residues 100d, 100e, 100l, and 100m) using OSPREY and computed K* scores for both the wild-type amino-acid and substitution with alanine. We compared the resulting change in K* score with previously determined IC_{50} data from alanine-scanning of PGT145[1]. The changes in the $\log_{10}$K* scores qualitatively recapitulated trends seen in the experimental data (See Table SR1).

SR1.2 Additional designs for PGT145

We performed double-mutation designs of PGT145 at residues L(100e) and Y(100m) and predicted mutations to improve binding (Figure SR1 and Table SR2). The $\log_{10}$K* score for the wild-type sequence was bounded by the interval $[12.3, 13.8]$ and the $\log_{10} Z_{\mathrm{Ab}}$ was bounded by the interval $[13.8, 14.3]$. Sequences for which the lower-bound on $\log_{10} Z_{\mathrm{Ab}}$ was less than 9.70 were excluded from consideration. Combinations of L(100e)E, L(100e)M, and L(100e)H with Y(100m), Y(100m)M, Y(100m)H, and Y(100m)E were predicted to improve binding without a large reduction in antibody stability.

SR1.3 Additional designs for PG9RSH

Designs of PG9RSH based on a high-resolution, monomeric input model of PG9 binding to the gp120 V1 and V2 loops (PDB ID: 3U2S) returned mutations at various residues predicted to improve binding. A three-residue design of residues R(100b), N(100c), and Y(100e) suggested that positions 100b and 100e are not amenable to mutation, while position 100c exhibits many mutations that are predicted to be slightly favorable, including Met, His, Arg, and Lys. Most mutations are predicted to result in only a slight increase in $\log_{10}$K* score over the wild-type ($\log_{10}\mathrm{K}^* = 13.0$)

SR2 Small-panel neutralization results

We assayed selected designs for neutralization on a small panel of 10 pseudovirus strains. Designs that performed well on these small assays were selected for large-panel (208 pseudoviruses) neutralization assays. Variants were selected for further characterization based on the number of pseudoviruses neutralized with an $IC_{80} \leq 50\mu$g/mL, the median IC_{80} value, the number of pseudoviruses neutralized with an $IC_{50} \leq 50\mu$g/mL, and the median IC_{50} value (listed in order of importance). Additionally, the sequence and neutralization diversity of the set of variants to be characterized was considered. For example, although several variants (N(100f)Y, DU012-DU017) achieved good neutralization breadth and median potency, only N(100f)Y and DU017 were chosen for further characterization due to their good performance at the IC_{80} cutoff and their incorporation of two different mutations at the 100f position.

Based on the OSPREY designs of PGT145 we selected 10 variants to test in small-panel pseudovirus neutralization assays (See Table S5). These PGT145 variants were assayed for neutralization against a panel of 10 pseudovirus strains from clades A, B, and C (See Data S1). Variant DU303 (containing the N(100l)D mutation) achieved a significant increase in breadth against this small panel, neutralizing 7 of the 10 tested strains with both IC_{50} and $IC_{80} \leq 50$

μg/mL, while the wild-type controls neutralized only 4 strains (median) with $IC_{50} \leq 50$ μg/mL and only 3 strains (median) with $IC_{80} \leq 50$ μg/mL. The median IC_{50} (0.218 μg/mL) over all neutralized pseudoviruses for DU303 was comparable to that for the wild-type controls, while the median IC_{80} (1.83 μg/mL) was greater than that for the controls.

Based on the OSPREY PG9RSH design results we selected 34 variants to test in small-panel pseudovirus neutralization assays (See Table S6). These variants were assayed for neutralization against a panel of 10 pseudovirus strains from clades A, B, and C (See Data S1). Variant PG9RSH N(100f)Y (containing the N(100f)Y mutation) achieved a clear increase in potency against this small panel, achieving median IC_{50} values of 0.073 and 0.185 μg/mL and median IC_{80} values of 0.821 and 3.87 μg/mL. PG9RSH N(100f)Y also showed a slight increase in breadth as one of the few tested variants to achieve measurable neutralization against the 3873.v1.c24 pseudovirus (clade C). Similarly, variant DU025 (containing the Y(100k)D mutation) achieved an increase in potency against this small panel, with median IC_{50} (0.185 μg/mL) and IC_{80} (2.21 μg/mL) improving relative to the wild-type Ab controls.

Overall, the small-panel neutralization panels indicated that DU303 resulted in increased breadth, PG9RSH N(100f)Y resulted in improved potency with perhaps a slight increase in breadth, and DU025 resulted in improved potency. In summary, the small panels were used as a screening tool to select computationally designed antibodies for further analysis using a large panel. Antibodies with exceptional biological properties, typically neutralization potency or breadth as per the large panel, were subsequently selected for further characterization, including high resolution cryo-EM structures.

SR3 Extended structural analysis of antibody variants

SR3.1 Structural analysis of DU303

The EM structure of PGT145 variant DU303 reveals a binding mode consistent with previous structures of PGT145. We focus our detailed structural analysis around DU303 residue 100l. Applying different masking parameters to the image dataset resulted in two electron density maps with 3.58 Å resolution that differ near residue 100l. The atomic model fit to the second density map shows the antibody interacting with the Env trimer apex through a long, negatively charged CDRH3 (Figure 2C).

D(100l) interacts with residues from two distinct gp120 subunits. Both electron density maps show well-resolved electron density for gp120 residues R166 and K169 (Figure S14). In contrast, density for residue D(100l) and for neighboring F(100d) is more ambiguous. Both maps show significant density between the modeled side-chain locations of D(100l) and F(100d). The first map localizes the density more closely to F(100d), a feature that suggests multiple occupied rotamers for F(100d) (Figure S14A). Density isosurfaces (1.2 and 3.0σ) for residue D(100l) show little density associated with the modeled location of the side-chain carboxyl group, perhaps indicative of a mobile side-chain. The second map, in contrast, localizes the density more closely to residue D(100l) (Figure S14B). This does result in density localized to the modeled location of the side-chain carboxyl of D(100l), but the peak extends further toward the modeled locations of F(100d) and gp120 K169.

The atomic model of DU303 fit to the second density map indicates that D(100l) could form electrostatic interactions with gp120 residues R166 and K169 (Figure 3D). The side-chain nitrogen of K169 lies 5.1 Å from a side-chain carboxyl oxygen of D(100l). Similarly, one of the side-chain nitrogens of R166 lies 4.0 Å from a side-chain carboxyl oxygen of D(100l). The position of these side-chains suggests that the negatively charged D(100l) forms favorable interactions with positively charged residues on gp120 to improve breadth and potency of neutralization.

SR3.2 Comparative analysis with DU303 design ensemble

To determine the extent to which design predictions corresponded with experimental structures we compared 10 members of the low-energy ensemble (LEE) generated during OSPREY design with the new structure of DU303 (Figure 3C,F). We aligned structures by the backbone heavy-atoms of the V2 region of a single gp120 subunit (Chain G, residues 154-177). Members of the LEE aligned to the Cryo-EM model with an RMSD of 0.79 Å calculated using 93 of 96 backbone heavy-atoms.

The CDRH3 loop in the experimental structure model is shifted laterally by 2.1 Å compared to the structures in the LEE, measured at the C_α carbon of residue 100l. Despite this, the distances between the side-chains of D(100l), R166, and K169 are accurately predicted in the LEE. The closest side-chain nitrogen of K169 in the LEE lies 4.2 Å from a side-chain carboxyl oxygen of D(100l). Similarly, one of the side-chain nitrogens of R166 lies 5.0 Å from a side-chain carboxyl oxygen of D(100l). While the LEE predicts that D(100l) interacts more closely with K169 than R166 (a relationship that is inverted in the experimental structure model), these predicted interaction distances differ by at most one angstrom from those in the experimental model. The side-chain orientations of this system are qualitatively similar between the LEE and the experimental model, indicating that OSPREY correctly predicted the structural consequences of the N(100l)D substitution.

SR3.3 Structural analysis of PG9RSH N(100f)Y

The EM structure of PG9RSH N(100f)Y shows the antibody interacting with the Env trimer apex through a hammer-like, negatively charged CDRH3 similar to previous bound structures of PG9RSH (Figure 2A). However, in contrast to a previous electron microscopy structure (PDB ID: 5VJ6)[4], the CDRH3 loop appears in a twisted conformation (Figure SR2). We focus our detailed structural analysis around PG9RSH N(100f)Y residue 100f.

Y(100f) interacts with side-chains and glycans from two distinct gp120 subunits. Electron density maps show well-resolved density for gp120 residues D167, K168, and K169 (Figure S15A). The first two N-acetylglucosamine (GlcNAc) sugars of gp120 glycan N160 are also resolved. Density corresponding to bNAb residues is more ambiguous. Both 1.2 σ and 3.0 σ isosurfaces show density peaks between the modeled side-chain locations of residues Y(100f) and Y(100a), perhaps suggesting the presence of alternate rotamer configurations. These isosurfaces do not contain the modeled locations of terminal hydroxyls for either tyrosine. Examination of low-density peaks (0.5 σ) reveals a small peak in density of the second GlcNAc of glycan N160 that is oriented toward the modeled location of Y(100f), suggesting possible interactions between Y(100f) and the glycan shield (Figure S15B).

The atomic model of PG9RSH N(100f)Y fit to the density map indicates that the primary interaction between Y(100f) and gp120 is a π-cation interaction with residue K168 (Figure 3A). The ammonium nitrogen of K168 lies 4.7 Å from the center of the Y(100f) π system, and the angle between the distance vector and the ring normal vector is approximately 20°, which is representative of typical π-cation geometry[5]. Y(100f) also forms van der Waals interactions with antibody residues P99, Y(100a), and TYS(100h). The side-chain geometry suggests that the aromatic Y(100f) side-chain participates in a π-cation interaction with the positively charged K168 to improve breadth and potency of neutralization.

SR3.4 Comparative analysis with PG9RSH N(100f)Y design ensemble

To determine the extent to which design predictions corresponded with experimental structures we compared 10 members of the low-energy ensemble (LEE) generated during OSPREY design

with the new structure of PG9RSH N(100f)Y (Figure 3A, D). We aligned structures by the backbone heavy-atoms of the V2 region of a single gp120 subunit (Chain I, residues 154-177). Members of the LEE aligned to the Cryo-EM model with an RMSD of 0.54 Å calculated using 92 of 96 backbone heavy-atoms.

The CDRH3 loops for the LEE structures and the experimental model near residue 100f are nearly identical in backbone conformation, despite the global difference between the twisted and untwisted conformations. The V2 loop of one gp120 subunit (chain G) exhibits a slight lateral shift relative to the experimental model. The side-chain locations in the LEE are qualitatively similar to those in the experimental model. The ammonium nitrogen lies as close as 5.1 Å from the center of the Y(100f) π system. Interestingly, multiple rotamers of Y(100f) and Y(100a) appear in the LEE, resulting in a conformation in which these side-chains have rotated and stacked. While for the majority of the LEE Y(100f) presents the face of the π ring toward K168, for a few members K168 forms van der Waals interactions with the edge of the aromatic ring. Notably, the design ensemble indicates D167 may interact with Y(100f), but these predicted interactions are not supported by the experimental model. However, the overall correspondence between LEE and experimental model suggests that OSPREY correctly predicted the structural consequences of the N(100f)Y substitution.

SR3.5 Structural analysis of DU025

The EM structure of PG9RSH variant DU025 shows the antibody interacting with the Env trimer apex through a hammer-like, negatively charged CDRH3 similar to previous bound structures of PG9RSH (Figure 2A). However, in contrast to a previous electron microscopy structure (PDB ID: 5VJ6)[4], the CDRH3 loop appears in a twisted conformation (Figure SR2).

We focus our detailed structural analysis around DU025 residue 100k. The electron density around the side chain of D(100k) is well-resolved at 1.2σ and maintains a small but noticeable peak at 3.0σ (Figure S16A). Similarly, clear density is observed at both 1.2 and 3.0σ for the side-chains of Env residues Q170, K305, and Y173, along with the first two core GlcNAc monomers of glycan N156. Nearly the entire core of glycan N156 is resolved at 1.2σ. Interestingly, three unassigned density peaks arise in the groove between the V2 and V3 loops at both 1.2 and 3.0σ, which could indicate the presence of solvent at this interface (Figure S16B). Furthermore, a bridge of density at 1.0σ arises between the modeled locations of Env residues Q170 and R308, hinting at long-range or solvent-mediated interactions between the two. Multiple strong density peaks around the modeled side-chain location of R308 provide evidence for multiple side-chain conformations at this residue.

These data suggest that residue D(100k) may form long-range or solvent-mediated interactions with residues Q170 and K305, which lie at distances of 4.1, and 6.8 Å, respectively. Additionally, D(100k) may interact with residue Y173 and glycan N156. The multiple unsatisfied densities at both low and high σ suggest that solvent may play a critical role at this interface.

SR3.6 Comparative analysis with DU025 design ensemble

To determine the extent to which design predictions corresponded with experimental structures we compared 10 members of the low-energy ensemble (LEE) generated during OSPREY design with the new structure of DU025 (Figure 3B, E). We aligned structures by the backbone heavy-atoms of the V2 region of a single gp120 subunit (Chain I, residues 154-177). Members of the LEE aligned to the Cryo-EM model with an RMSD of 0.50 Å calculated using 90 of 96 backbone heavy-atoms.

Members of the LEE indicate that improvement in binding is predicted to result from favorable electrostatic interactions between D(100k) and Q170 or R308. The lowest-energy member of the design ensemble places one carboxyl oxygen of D(100k) 3.2 Å from the side chain nitrogen of Q170 and 5.6 Å from R308. K305, on the other hand, lies distal (9 Å) from D(100k).

The design ensemble (based on PDB: 5VJ6) displays the CDRH3 loop in the untwisted conformation, whereas the cryo-EM structure of DU025 exhibits a twisted CDRH3 loop. Although this does not cause large changes for most PG9RSH residues that contact Env directly, residues 100i-100n are displaced by up to 3.6 Å measured at the C_α. Residue D(100k) in particular is displaced by 3.5 Å, resulting in a significant change in its environment. As a result, although the design ensemble correctly predicts that D(100k) creates long-range electrostatic interactions, and correctly predicts a favorable interaction with Q170, the predicted interaction between D(100k) and R308 is not supported by the cryo-EM structure. Instead, shifts in the backbone create interactions between D(100k) and K305. However, the overall quality and type of interactions formed by D(100k) in the design ensemble is in fact consistent with the experimental structure.

SR3.7 Inconsistent CDRH3 loop conformation in design inputs of PG9

Interestingly, we noticed differences between the CDRH3 conformations of PG9 of our design input model (based on PDB ID: 5VJ6) and the experimentally determined structures of both PG9RSH N(100f)Y and DU025. Cryo-EM structures placed the PG9RSH CDRH3 loop in a twisted conformation, similar to the conformations displayed in previous structures (PDB IDs: 3MME, 3U2S, 3U4E, 4YAQ), but the model represented by 5VJ6 displays an "untwisted" conformation (Figure SR2). The consensus between all but one structure suggests that the twisted conformation is in fact the correct CDRH3 conformation. The aberrant structure, 5VJ6, was generated by fitting crystal structures of PG9 bound to a V1V2 scaffold (PDB ID: 3U4E) and unliganded gp120 core (PDB ID: 3DNN) to density obtained via electron microscopy[6], but surprisingly the loop conformations in 5VJ6 and 3U4E are different. This suggests that fitting or processing of 5VJ6 inappropriately modified the CDRH3 loop, resulting in a hallucinated conformation.

SR4 Comparison to putative germline antibodies

The mutation defining PGT145 DU303, N(100l)D, does not appear in the putative germline antibody ([7] via IMGT/V-QUEST[8]) nor the clonally-related PGT141-PGT144 sequences[9]. As such, this mutation appears to represent a higher bar for elicitation by vaccines, albeit by only one mutation. It is important to note that, due to the long CDRH3 insertion, the CDRH3 sequence of the putative germline antibody was taken from the mature PGT145 antibody when the D gene did not provide sufficient coverage[7].

Similarly, the mutations PG9RSH N(100f)Y and PG9RSH DU025 (Y(100k)D) do not appear in the putative germline antibodies for PG9 ([7] via IMGT/V-QUEST[8]) nor the somatically-related PG16[10]. Nor do either of these mutations appear in the mature PG16 antibody. These observations suggest that, like DU303, these mutations represent a higher bar for elicitation. However, position 100k does appear contain a mutation away from germline in both PG9 and PG16, suggesting that, at least for DU025, the bar for vaccine elicitation is no higher than for the wild-type PG9 antibody, ignoring the RSH chimeric context.

SR5 Investigation of additional Env genotypes

We are interested in additional questions and extensions to this work, but unfortunately many of these are beyond the scope of the current manuscript. Can this method be applied to *any* antigen genotype to successfully improve breadth? We suspect that the answer is no. We have in fact been careful to avoid claims that improved breadth can result from design against *any* antigen in this manuscript. How, then, can we differentiate between antigens that are more "canonical" (i.e., antigens for which design for potency can result in breadth), versus those that are less? In the discussion section of this manuscript we provide several hypotheses for how one could begin to address this question, but we leave the evaluation of these hypotheses for future work. Designs on additional genotypes could provide information to help answer these questions, although structural information for these antibodies complexed with other Env genotypes is sparse. The results of such designs would not themselves be definitive, and design results would need to be characterized experimentally.

Nonetheless, to satisfy our curiosity we additionally peformed OSPREY designs on PGT145 in complex with the AMC011 (PDB ID: 6NIJ) and SIVmac239 (PDB ID: 8DVD) Env trimers, focusing on CDRH3 residues 100d and 100l. The primary differences between these trimers and the BG505 trimer at this site lie at Env residue 169, which is a lysine, valine, or threonine for the BG505, AMC011, and SIVmac239 trimers, respectively. We used the unminimized input PDB structures (after adding hydrogens using MolProbity) as the input structure, with flexible and mutable residues described in Table SR6, and approximated the top 20 sequences by K* score (and wild-type sequence) for each model.

The results of these designs can be found in the Supplementary Information (Table SR7 and SR8, Figure SR3 and SR4). Overall, we do see sequences with high K* scores within the top 20 sequences of both designs that appear similar to those in the original BG505 design, especially when focusing on the mutations at position 100d. Differences can be seen with respect to postion 100l, which is expected given the sequence variation at residue 169. Similarly to the BG505 designs, both the 6NIJ and 8DVD designs assign high K* scores to sequences with either the N(100l)D or N(100l)E mutations, and, like the BG505 design, the AMC011 design assigns a high K* score to the F(100d) D(100l) sequence. However, the low antibody partition function Z_{Ab} and low complex partition function for the PGT145-F(100d)D(100l):AMC011 sequence indicates that this is not a viable mutation for binding to the AMC011 Env trimer.

Most substitutions of N(100l) with a negatively charged residue result in a decrease in the unbound antibody partition function Z_{Ab} to some extent, indicating that the additional negative charge may decrease the antibody CDRH3 stability. However, examining the complex and antibody partition functions for these sequences reveals a critical difference: The complex partition function for the BG505:PGT145-D100l complex remains approximately equal to that for the BG505:PGT145, but the complex partition function for the AMC011:PGT145-D100l complex is significantly lower than that for the AMC011:PGT145 complex. Thus, our models predict that the the negatively charged D(100l) is compensated for in the BG505 complex state by interaction with K169, but not in the AMC011 complex state, where no productive interactions are formed with V169. The high K* score of F(100d)-D(100l) in the AMC011 design is a case of a bad complex partition function divided by a worse unbound partition function – a phenomenon that we select against using the unbound partition function stability cutoffs. As a result, we would not have chosen to test the N(100l)D mutation from this design, but would have identified position 100l as a site of interest.

Overall, these additional designs suggest that OSPREY is correctly modeling the environment at this site. Similar mutations were found by OSPREY ranked by K* score alone, but there were significant differences in results after filtering by stability. It is impossible to assess the effects

of many of these mutations (e.g., Y(100d)W, N(100l)M) without experimental characterization, but these results provide a direction for possible future work.

SR6 Structure modeling of additional Env genotypes

To gain further insight into the mechanisms of improved breadth for these variant antibodies we constructed homology models of each antibody in complex with Env trimers from other strains of HIV or SIV. We first leveraged existing structures of PGT145 in complex with the AMC011 (PDB ID: 6NIJ) and SIVmac239 (PDB ID: 8DVD) trimers to construct conservative homology models. Unfortunately, to our knowledge there are still no structures of PG9 in complex with Env trimers from strains other than BG505. The primary difference between AMC011 or SIVmac239 trimers and the BG505 trimer near residue 100l is the identity of Env residue 169, which is a lysine, valine, or threonine in BG505, AMC011, and SIVmac239, respectively. We used the unminimized input PDB structures (after adding hydrogens using MolProbity) as the input structure, with flexible and mutable residues described in Table SR6. We used OSPREY to mutate PGT145 residue N(100l) to aspartate in both structures and generated ensembles of predicted low energy conformations in the same manner as described in the methods section of the manuscript. We then compared the resulting ensembles to the Cryo-EM structure of DU303 to model the effect of Env sequence and structural variation on DU303.

Broad structural features match well between the models of DU303 bound to AMC011 or SIVmac239 and the DU303:BG505 experimental structure, but small differences in loop conformations are evident (See Figure SR5) We aligned structures by the 23 C_α atoms of the CDRH3 loops, which aligned to the cryo-EM structure of DU303:BG505 with an RMSD of 1.62 Å or 0.86 Å for the models of DU303:AMC011 or DU303:SIVmac239, respectively. Slight differences are evident in the CDRH3 loop of DU303:AMC011 relative to DU303:BG505, which significantly change the orientation of the C_α-C_β vector of residue 100l relative to the DU303:BG505 structure. The CDRH3 loop of DU303:SIVmac239 appears quite similar to that of DU303:BG505. The V2 loops in both models are roughly consistent with those in the experimental structure, but shifts of up to 3 Å or 7 Å relative to the BG505 V2 conformations can be seen. These differences could reflect real conformational differences, conformational heterogeneity, model fitting error, or experimental uncertainty.

The intermolecular interactions of residue D(100l) in both models are significantly reduced relative to the experimental structure. Not only is V169 too short to interact with the relatively distant D(100l) in the DU303:AMC011 model, but the reorientation of residue 100l due to CDRH3 conformation differences seems to preclude positive interactions with V169. Additionally, this reorientation increases the distance between residue 100l and R166, and, as a result, the model predicts no significant interactions between these residues. T169 is also too short to interact with D(100l) in the DU303:SIVmac239 model. R166 appears to interact primarily with the sulfated tyrosine at residue 100f, possibly due to shifts in the V2 loop position relative to the DU303 experimental structure. Overall, these models suggest that DU303 would not gain neutralization potency for the AMC011 HIV strain and for the SIVmac239 SIV strain due to lack of interaction with the important K169 residue. Similar observations were also previously gleaned from the analysis of large-panel neutralization data (See Figure S10).

Due to the lack of experimental structure information for PG9 binding to Env trimers other than BG505, to model the interactions of P9RSH N(100f)Y and DU025 we also pursued less conservative modeling approaches. We used AlphaFold Multimer[11,12] to predict structures of PG9RSH N(100f)Y in complex with Envs from strains ZM233.6, and JRFL.JB, and predicted structures of DU025 in complex with JRFL.JB and TRO.11. Additionally, we predicted struc-

tures of all antibodies in complex with BG505.W6M.C2 as a positive control. We examined 25 predicted structures for each complex. In general, although the antibody and Env structures individually showed good homology to existing structures, the antibody binding epitope was poorly predicted. Only a few predicted structures placed antibodies near the Env apex (Figure SR6), and no structures successfully recapitulated the parallel β-strand interaction between PG9RSH and the Env V2 loop. Interestingly, no predicted structures for the BG505 control placed these antibodies near the apex of the Env trimer.

Overall, our modeling yielded good results when based on conservative modification of existing experimental structures, but was less successful when such structures were unavailable. Due to the scarcity of experimental structure data, the difficulty of modeling protein loops, and the difficulty of accurately predicting binding interfaces, further modeling of the interactions between these variant antibodies and diverse Env trimers is deferred to future work.

References

[1] Jeong Hyun Lee et al. "A Broadly Neutralizing Antibody Targets the Dynamic HIV Envelope Trimer Apex via a Long, Rigidified, and Anionic β-Hairpin Structure". In: *Immunity* 46.4 (Apr. 2017), pp. 690–702. ISSN: 10974180. DOI: 10.1016/j.immuni.2017.03.017. URL: https://pubmed.ncbi.nlm.nih.gov/28423342/.

[2] Mark A. Hallen et al. "OSPREY 3.0: Open-Source Protein Redesign for You, with Powerful New Features". In: *J. Comput. Chem.* 9999 (2018), pp. 1–14. DOI: 10.1002/jcc.25522. URL: https://www.biorxiv.org/content/early/2018/04/23/306324.

[3] Anna Ulrika Lowegard. "Novel Algorithms and Tools for Computational Protein Design with Applications to Drug Resistance Prediction, Antibody Design, Peptide Inhibitor Design, and Protein Stability Prediction". PhD thesis. Duke University, 2022.

[4] Haoqing Wang et al. "Asymmetric recognition of HIV-1 Envelope trimer by V1V2 loop-targeting antibodies". In: *eLife* 6 (2017).

[5] Hervé Minoux and Christophe Chipot. "Cation-π interactions in proteins: Can simple models provide an accurate description?" In: *Journal of the American Chemical Society* 121.44 (Nov. 1999), pp. 10366–10372. ISSN: 00027863. DOI: 10.1021/ja990914p. URL: http://www.sander.embl-heidelberg.de/.

[6] Jean Philippe Julien et al. "Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9". In: *Proceedings of the National Academy of Sciences of the United States of America* 110.11 (Mar. 2013), pp. 4351–4356. ISSN: 00278424. DOI: 10.1073/pnas.1217537110. URL: https://www.pnas.org/content/110/11/4351%20https://www.pnas.org/content/110/11/4351.abstract.

[7] Kwinten Sliepen et al. "Binding of inferred germline precursors of broadly neutralizing HIV-1 antibodies to native-like envelope trimers". In: *Virology* 486 (Dec. 2015), pp. 116–120. DOI: 10.1016/j.virol.2015.08.002. URL: https://doi.org/10.1016%2Fj.virol.2015.08.002.

[8] Véronique Giudicelli, Xavier Brochet, and Marie-Paule Lefranc. "IMGT/V-QUEST: IMGT Standardized Analysis of the Immunoglobulin (IG) and T Cell Receptor (TR) Nucleotide Sequences". In: *Cold Spring Harb Protoc* 2011.6 (June 2011), pdb.prot5633. DOI: 10.1101/pdb.prot5633. URL: https://doi.org/10.1101%2Fpdb.prot5633.

[9] Laura M. Walker et al. "Broad neutralization coverage of HIV by multiple highly potent antibodies". In: *Nature* 477.7365 (Sept. 2011), pp. 466–470. ISSN: 00280836. DOI: 10.1038/nature10373. URL: https://www.nature.com/articles/nature10373.

[10] Marie Pancera et al. "Crystal Structure of PG16 and Chimeric Dissection with Somatically Related PG9: Structure-Function Analysis of Two Quaternary-Specific Antibodies That Effectively Neutralize HIV-1". In: *Journal of Virology* 84.16 (Aug. 2010), pp. 8098–8110. ISSN: 0022-538X. DOI: 10.1128/jvi.00966-10. URL: http://jvi.asm.org/.

[11] John Jumper et al. "Highly accurate protein structure prediction with AlphaFold". In: *Nature* 596.7873 (Aug. 2021), pp. 583–589. ISSN: 14764687. DOI: 10.1038/s41586-021-03819-2. URL: https://doi.org/10.1038/s41586-021-03819-2.

[12] Richard Evans et al. "Protein complex prediction with AlphaFold-Multimer". In: *bioRxiv* (Mar. 2022), p. 2021.10.04.463034. DOI: 10.1101/2021.10.04.463034. URL: https://doi.org/10.1101/2021.10.04.463034.