Supporting Information

Elucidating the Interdependence of Drug Resistance from Combinations of Mutations

Debra A. Ragland[1], Troy W. Whitfield[2,3], Sook-Kyung Lee[4], Ronald Swanstrom[4], Konstantin B. Zeldovich[3], Nese Kurt-Yilmaz[1] and Celia A. Schiffer[1,*]

[1]Department of Biochemistry and Molecular Pharmacology, [2]Department of Medicine, [3]Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, MA 01605, USA
[4]Department of Biochemistry and Biophysics, and the UNC Center for AIDS Research, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

*Corresponding Author

Table S1. EC_{50} values for clinically-derived HIV-1 protease variants determined by cell-based assays at the Swanstrom laboratory at UNC-Chapel Hill (unpublished data), and fold changes relative to wild type NL4-3. Note that EC_{50} value for variant ATA_{21} could not be determined.

Figure S1. Sequence identity matrix for all 15 HIV-1 protease variants in the panel. Values are colored from blue to red for high to low % identities.

Figure S2. Root mean square deviation (RMSD) values for MD simulations of HIV-1 protease variants. **A**. DRV accessory RAMs L76V, V32I, L33F, and V32I+L33F compared to SF-2 WT. **B**. Viral passaging variants I84V, I93L, DRV^r8 and DRV^r10 compared to NL4-3 WT. **C**. Clinically-derived variants compared to SF-2 WT.

Figure S3. The variance of hydrogen bond occupancies among the variants is largely accounted for by the first principal component (eigenvalue spectrum, proportion of variance). The first principal component for the mean hydrogen bond occupancy dataset accounts for 89% of the variance.

Figure S4. The variance of mean van der Waals energy is accounted for by the first principal component (eigenvalue spectrum, proportion of variance). The first principal component for the mean van der Waals energy dataset accounts for 97.6% of the variance.

Figure S5. Mean hydrogen bond occupancies share similarities with sequence and dynamics. **A**. Fifteen sequence variants projected onto the first two principal components for the correlation matrix of 143 dynamic hydrogen bond occupancies. Variants are partitioned into three groups, those variants possessing K41 (black), those

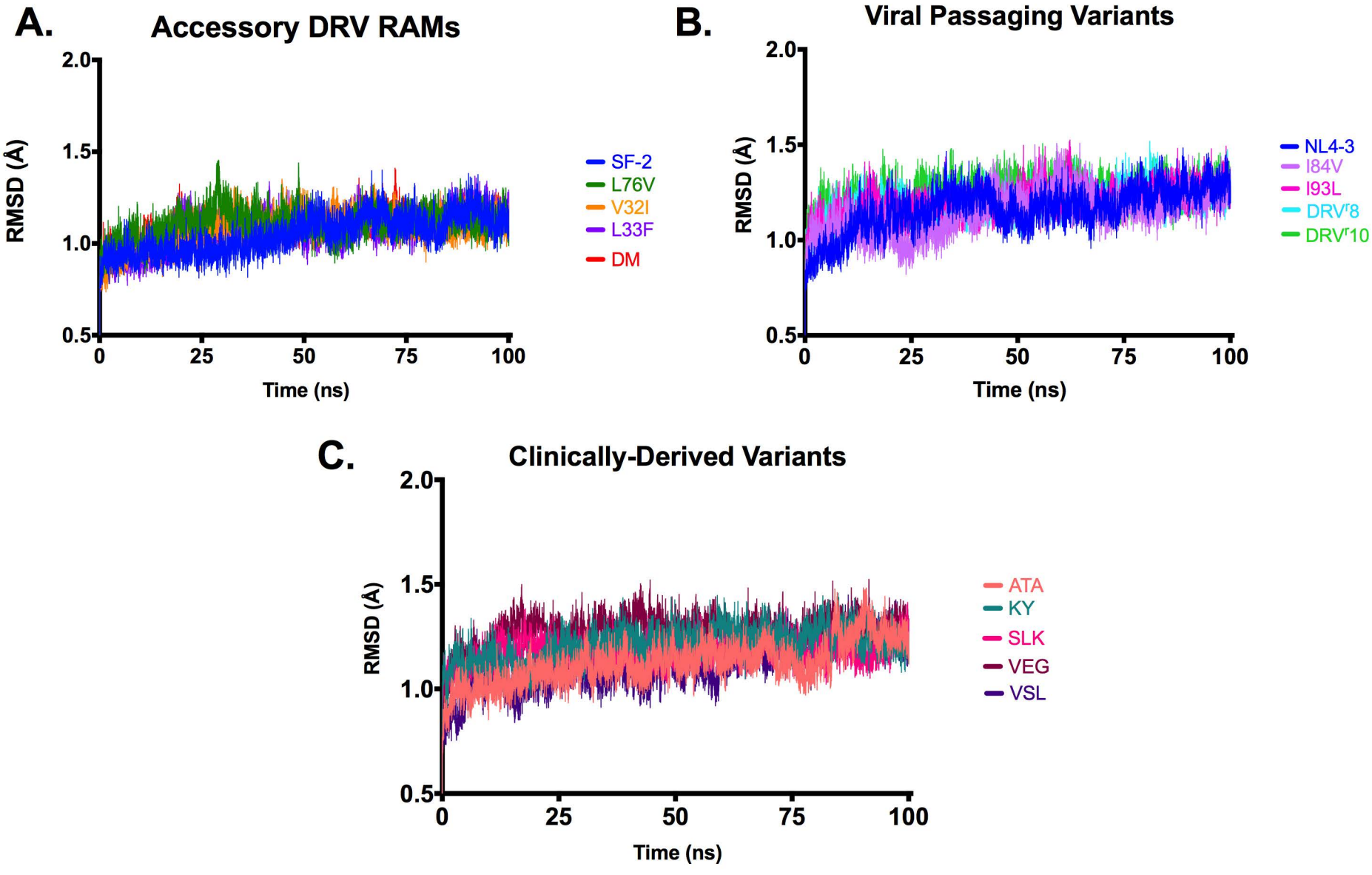
possessing R41, (bold, purple) and those possessing a combination of substitutions at positions 41, 10, and 54 simultaneously (bold, blue). **B**. List of top five single positions found to most likely underlie alterations in hydrogen bond occupancy patterns. **C**. List of top five position pairs found to most likely underlie hydrogen bond occupancy changes.

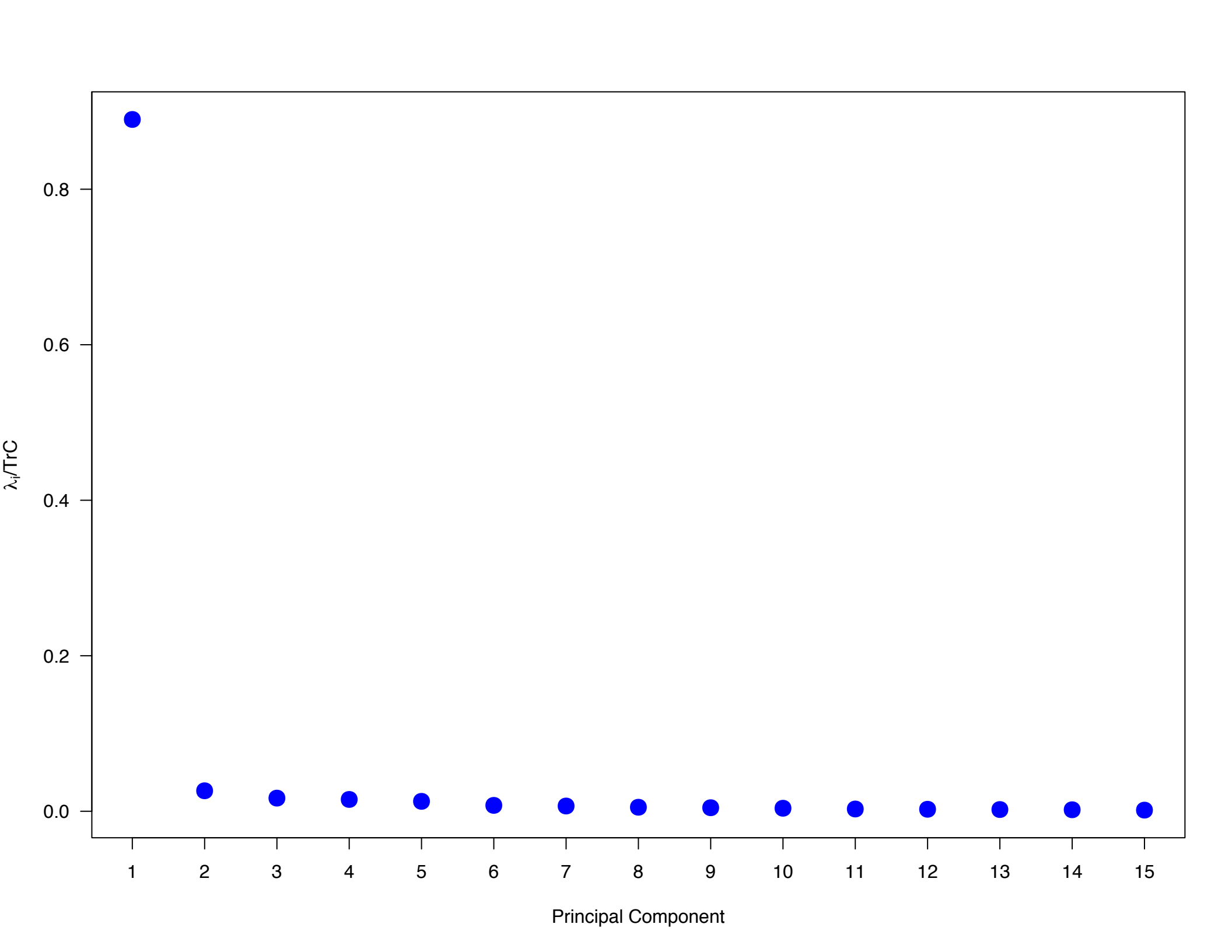
Figure S6. Density distribution of 15 variants along the first principal component $\rho(u_1)$ for the 111 mean hydrogen bond occupancies seen in Figure 3A.

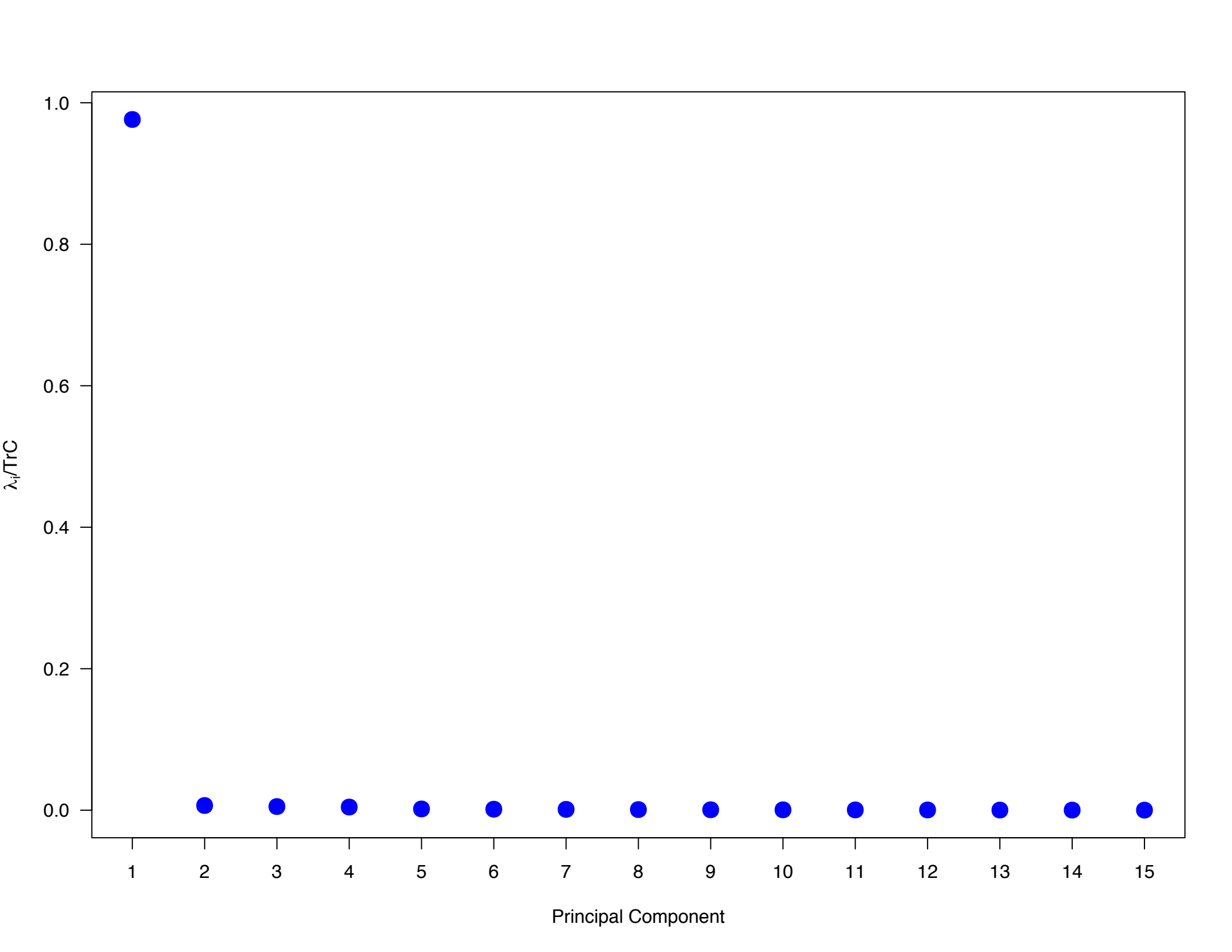
Figure S7. Density distribution of 15 variants along the first principal component $\rho(u_1)$ for the 64 mean van der Waals contact energies seen in Figure 5A.

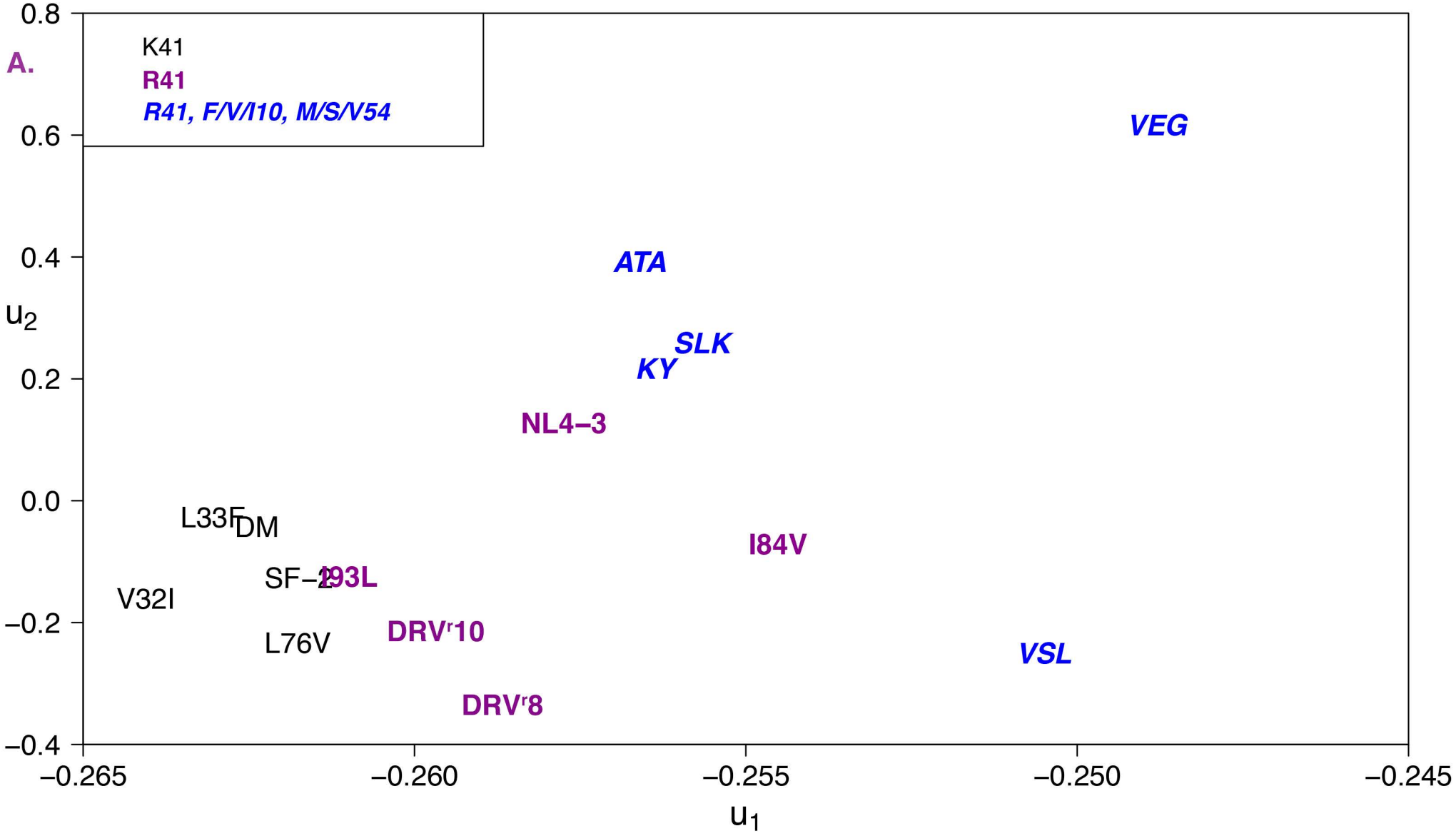
EC_{50} Values for Clinically-Derived Variants		
	EC_{50} in nM	Fold Change EC_{50}
NL4-3	3.98	
ATA_{21}	-	-
KY_{26}	1160	291
SLK_{19}	32.5	8
VEG_{23}	7800	1960
VSL_{23}	320	80

	SF-2	NL4-3	L76V	L33F	V32I	DM	I84V	I93L	DRV'8	DRV'10	ATA	KY	SLK	VEG	VSL
SF-2	100	95	99	99	99	98	94	94	87	85	79	74	81	77	77
NL4-3	95	100	94	94	94	93	99	99	92	90	78	75	80	76	78
L76V	99	94	100	98	98	97	93	93	86	86	78	73	80	76	76
L33F	99	94	98	100	98	99	93	93	88	86	80	75	80	78	78
V32I	99	94	98	98	100	99	93	93	88	86	80	75	80	78	76
DM	98	93	97	99	99	100	92	92	89	87	81	76	79	79	77
I84V	94	99	93	93	93	92	100	98	93	91	79	74	79	77	77
I93L	94	99	93	93	93	92	98	100	91	89	77	74	79	75	79
DRV'8	87	92	86	88	88	89	93	91	100	98	80	76	75	78	75
DRV'10	85	90	86	86	86	87	91	89	98	100	78	75	75	78	75
ATA	79	78	78	80	80	81	79	77	80	78	100	79	73	88	74
KY	74	75	73	75	75	76	74	74	76	75	79	100	68	77	74
SLK	81	80	80	80	80	79	79	79	75	75	73	68	100	73	80
VEG	77	76	76	78	78	79	77	75	78	78	88	77	73	100	75
VSL	77	78	76	78	76	77	77	79	75	75	74	74	80	75	100


A.
Accessory DRV RAMs
RMSD (Å)
2.0
1.5
1.0
0.5
0
25
50
75
100
Time (ns)
SF-2
L76V
V32I
L33F
DM
B.
Viral Passaging Variants
RMSD (Å)
2.0
1.5
1.0
0.5
0
25
50
75
100
Time (ns)
NL4-3
I84V
I93L
DRVr8
DRVr10
C.
Clinically-Derived Variants
RMSD (Å)
2.0
1.5
1.0
0.5
0
25
50
75
100
Time (ns)
ATA
KY
SLK
VEG
VSL


0.8
0.6
0.4
0.2
0.0
λi/TrC
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Principal Component


1.0
0.8
0.6
0.4
0.2
0.0
λi/TrC
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Principal Component

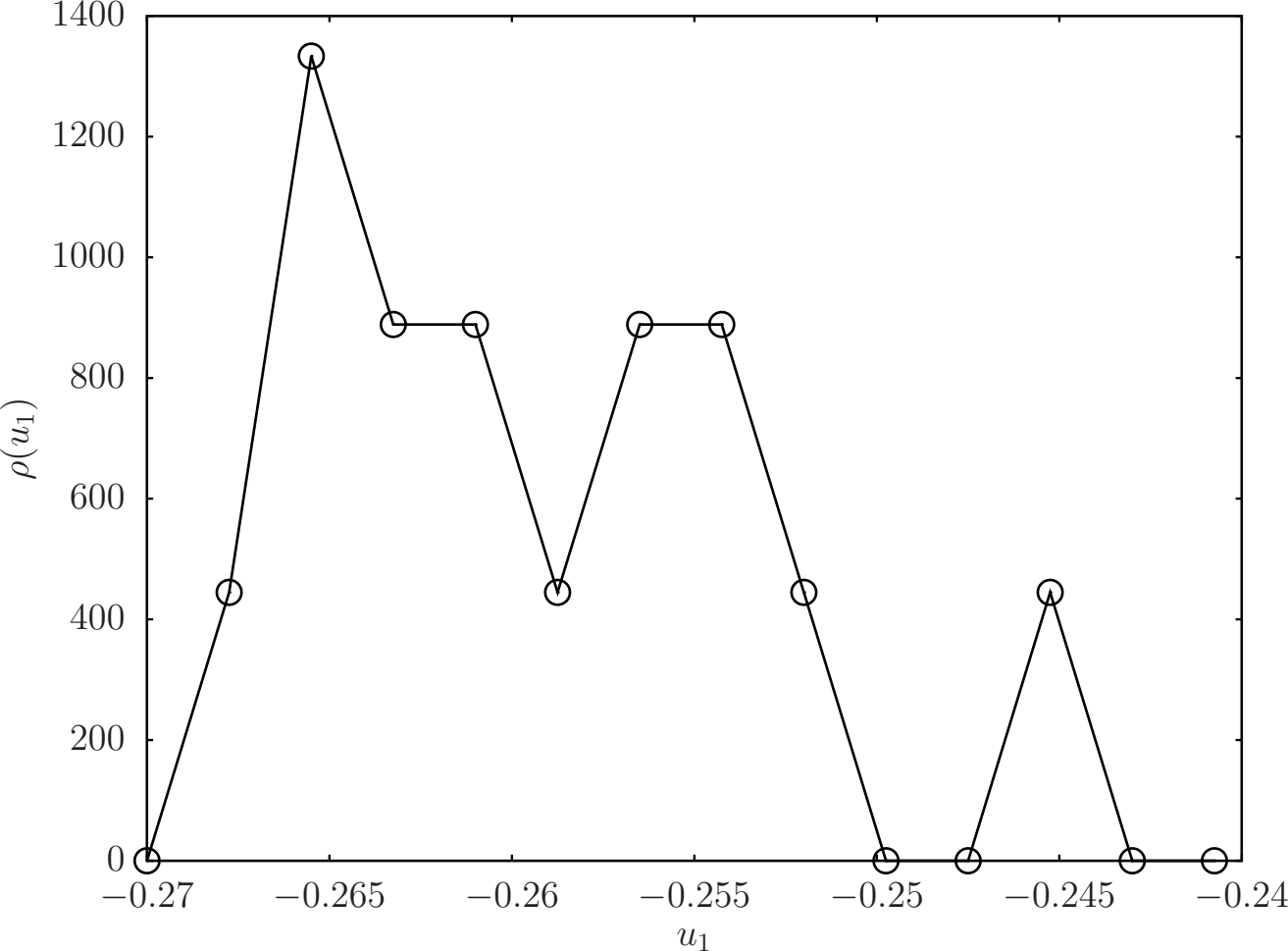


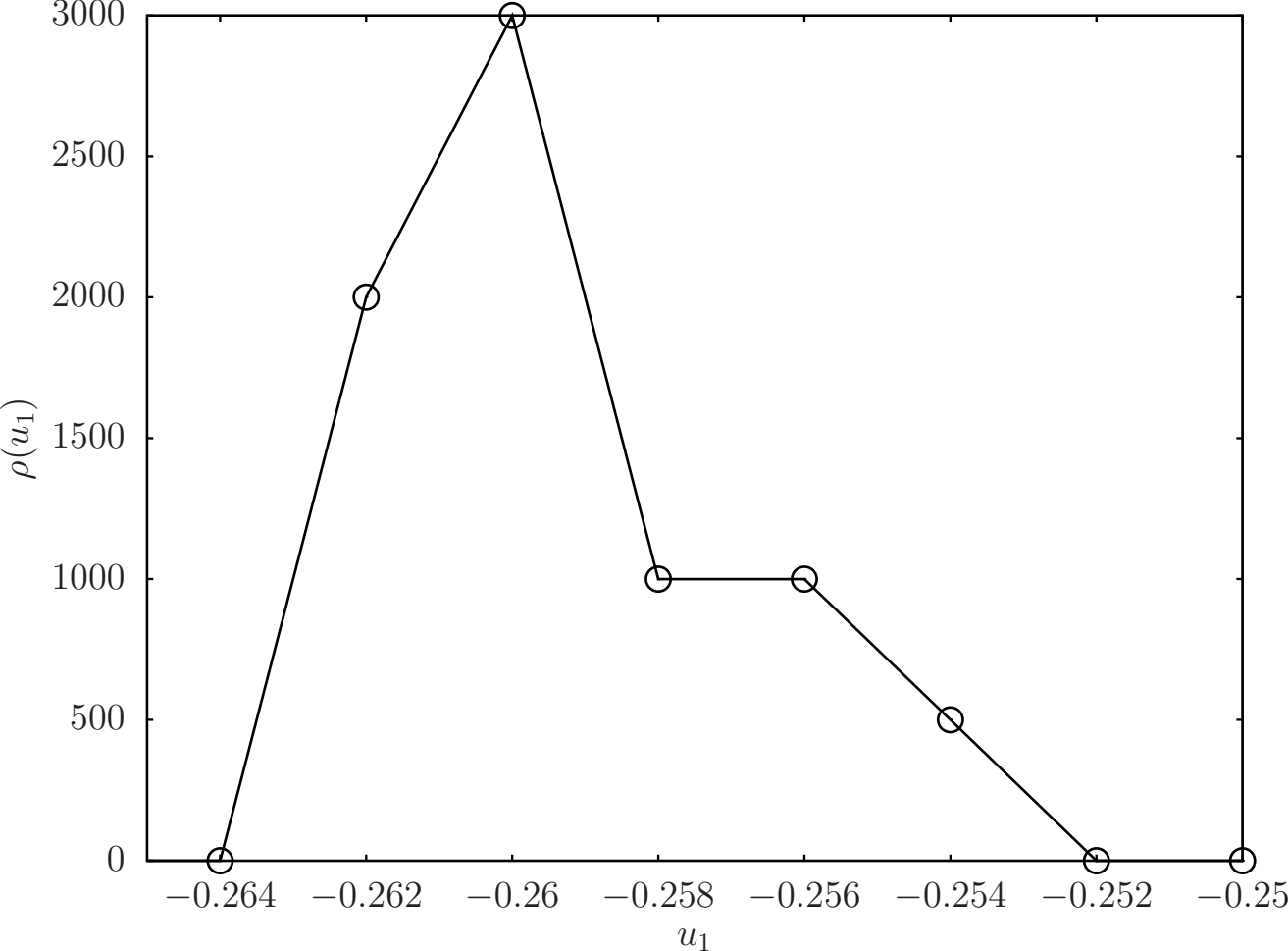
B.

Residue Position	p–value
41	0.000666
10	0.00466
54	0.00466
37	0.00879
35	0.0102

C.

Residue Position 1	Residue Position 2	p–value
10	41	0.00466
10	54	0.00466
41	54	0.00466
10	71	0.00586
54	71	0.00586


1400
1200
1000
800
600
400
200
0
$\rho(u_1)$
−0.27
−0.265
−0.26
−0.255
−0.25
−0.245
−0.24
u_1


3000
2500
2000
1500
1000
500
0
$\rho(u_1)$
-0.264
-0.262
-0.26
-0.258
-0.256
-0.254
-0.252
-0.25
u_1