Biophysical Journal, Volume *121*

Supplemental information

Comparative evaluation of spin-label modeling methods for protein structural studies

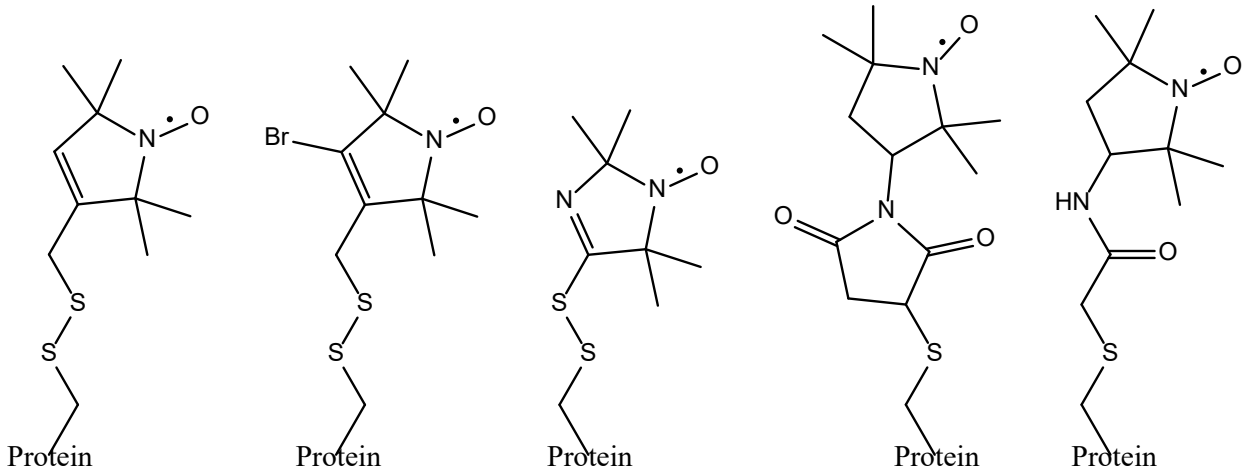
Maxx H. Tessmer, Elizabeth R. Canarie, and Stefan Stoll

Comparative Evaluation of Spin Label Modeling Methods for Protein Structural Studies

Maxx Tessmer, Elizabeth R. Canarie, Stefan Stoll

Department of Chemistry, University of Washington, Seattle, Washington 98195, USA

Supporting Information



R1 R7 V1 M1 I1

Figure S1: Spin label skeletal diagrams.

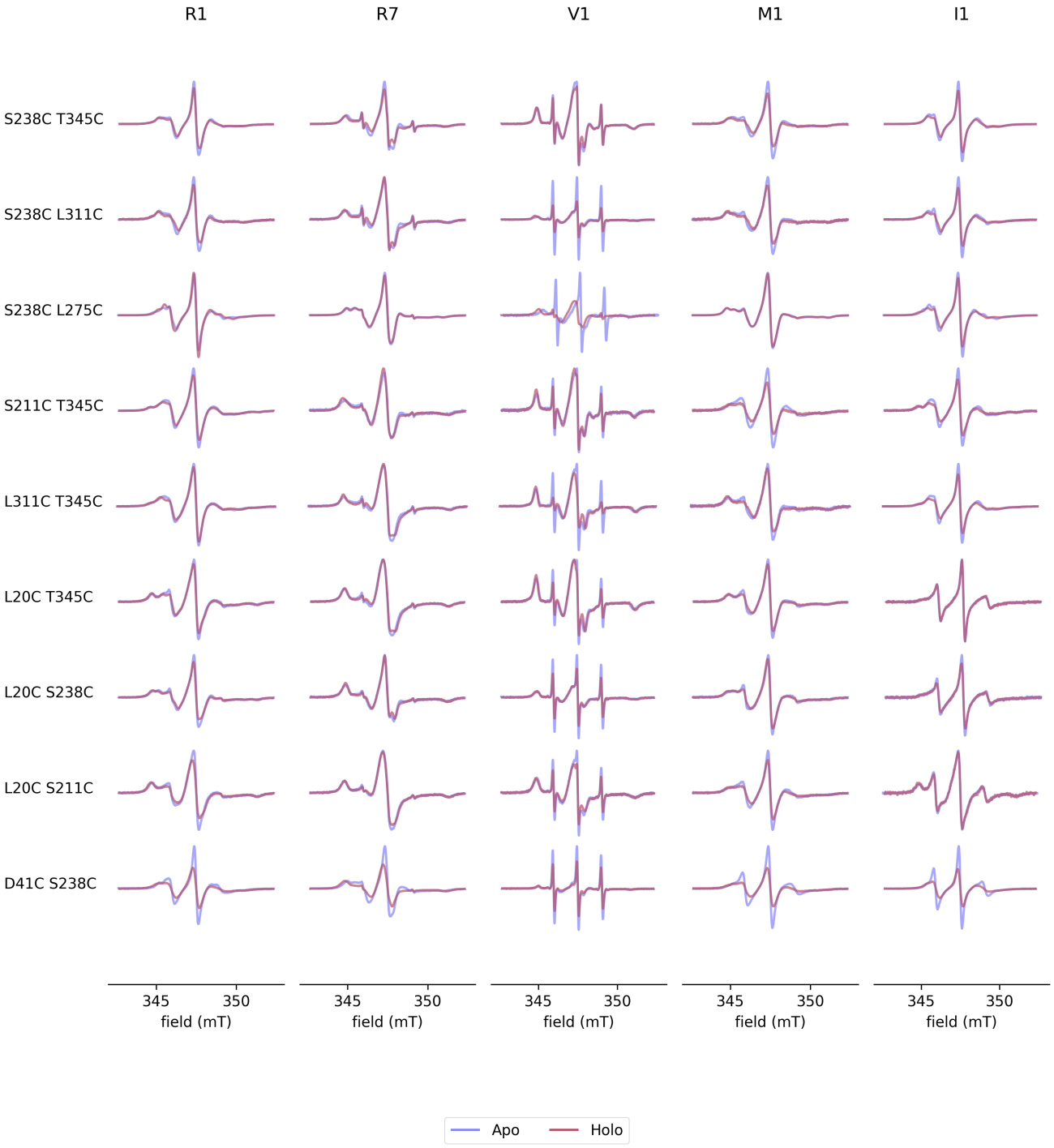


Figure S2. X-band CW EPR of all nine site pairs (rows) and all five labels (columns) in the presence (red) and absence (blue) of maltose.

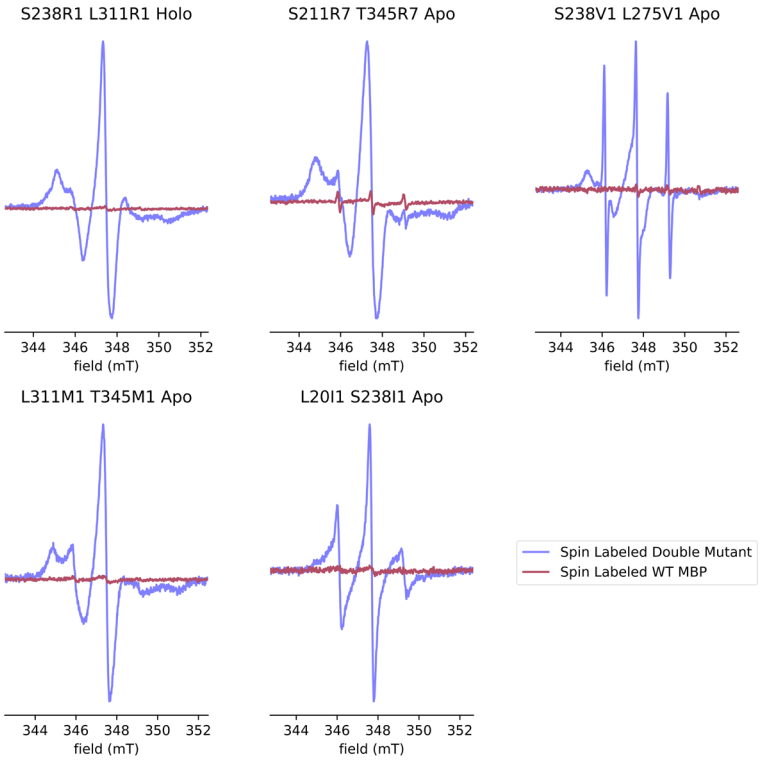
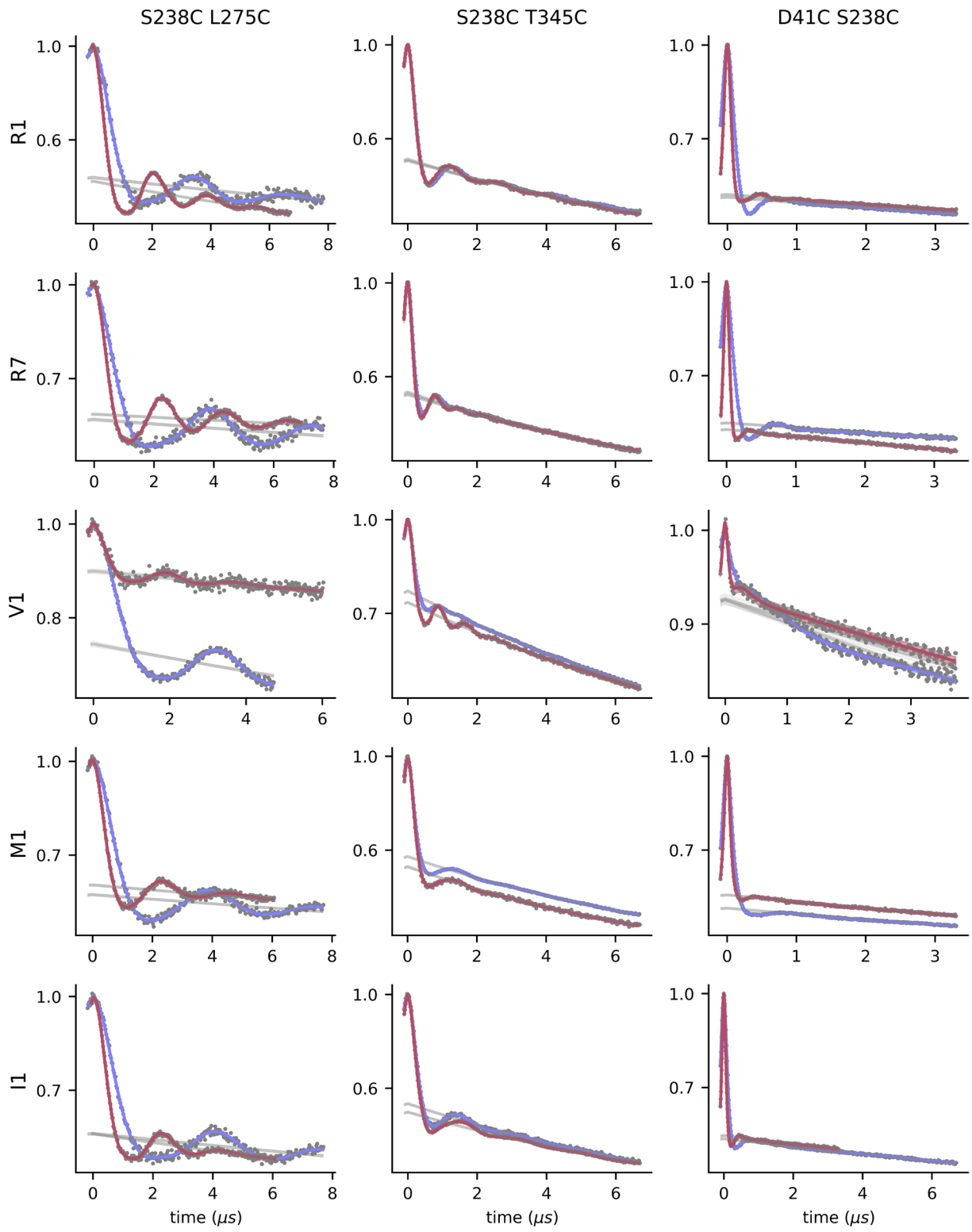
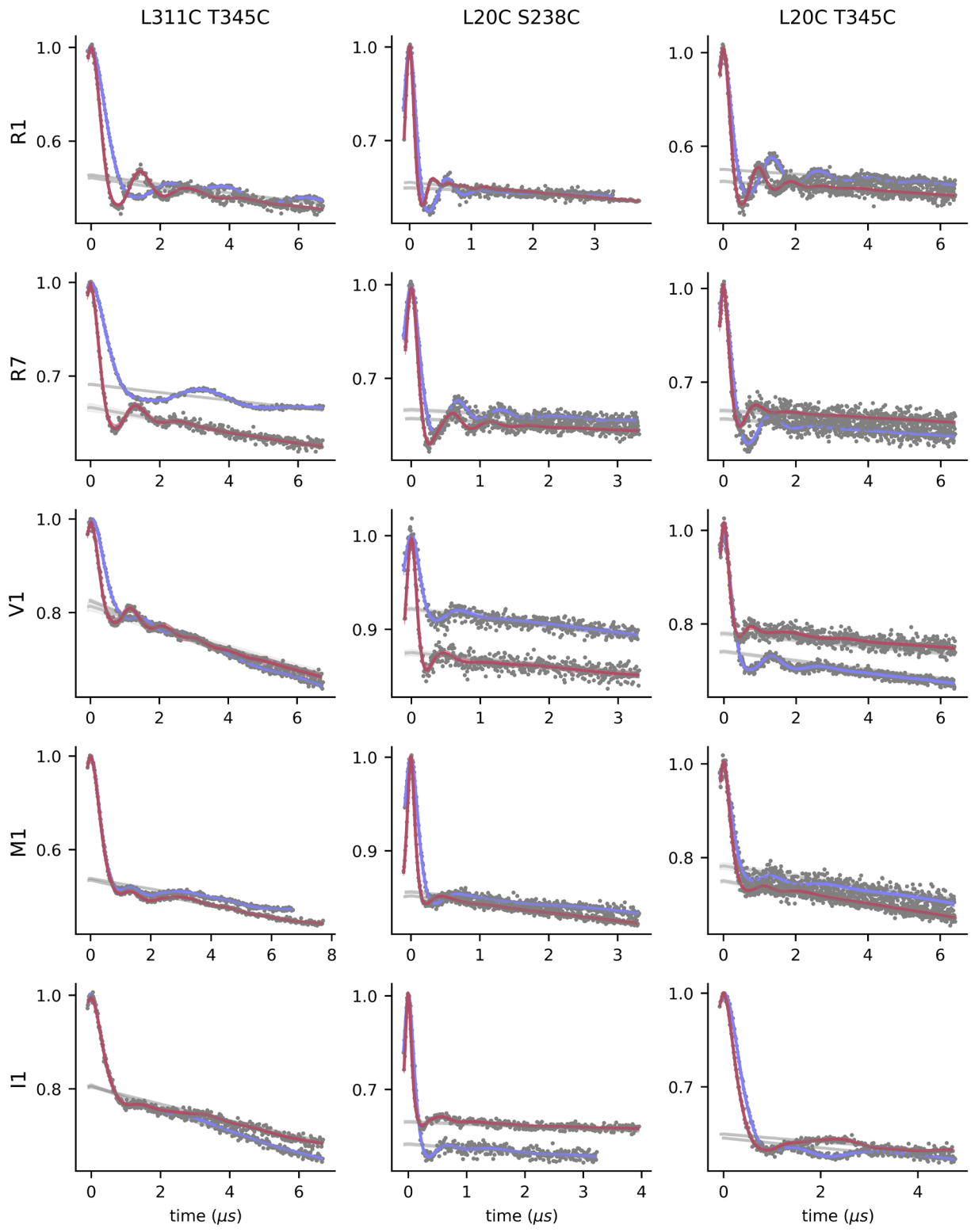


Figure S3. X-band CW EPR of WT MBP spin labeled compared to the CW spectra of the double mutants with the lowest overall signal (double integral) for each label.


S238C L275C
S238C T345C
D41C S238C
R1
R7
V1
M1
I1
time (μs)
time (μs)
time (μs)


L311C T345C
L20C S238C
L20C T345C
R1
R7
V1
M1
I1
time (μs)
time (μs)
time (μs)

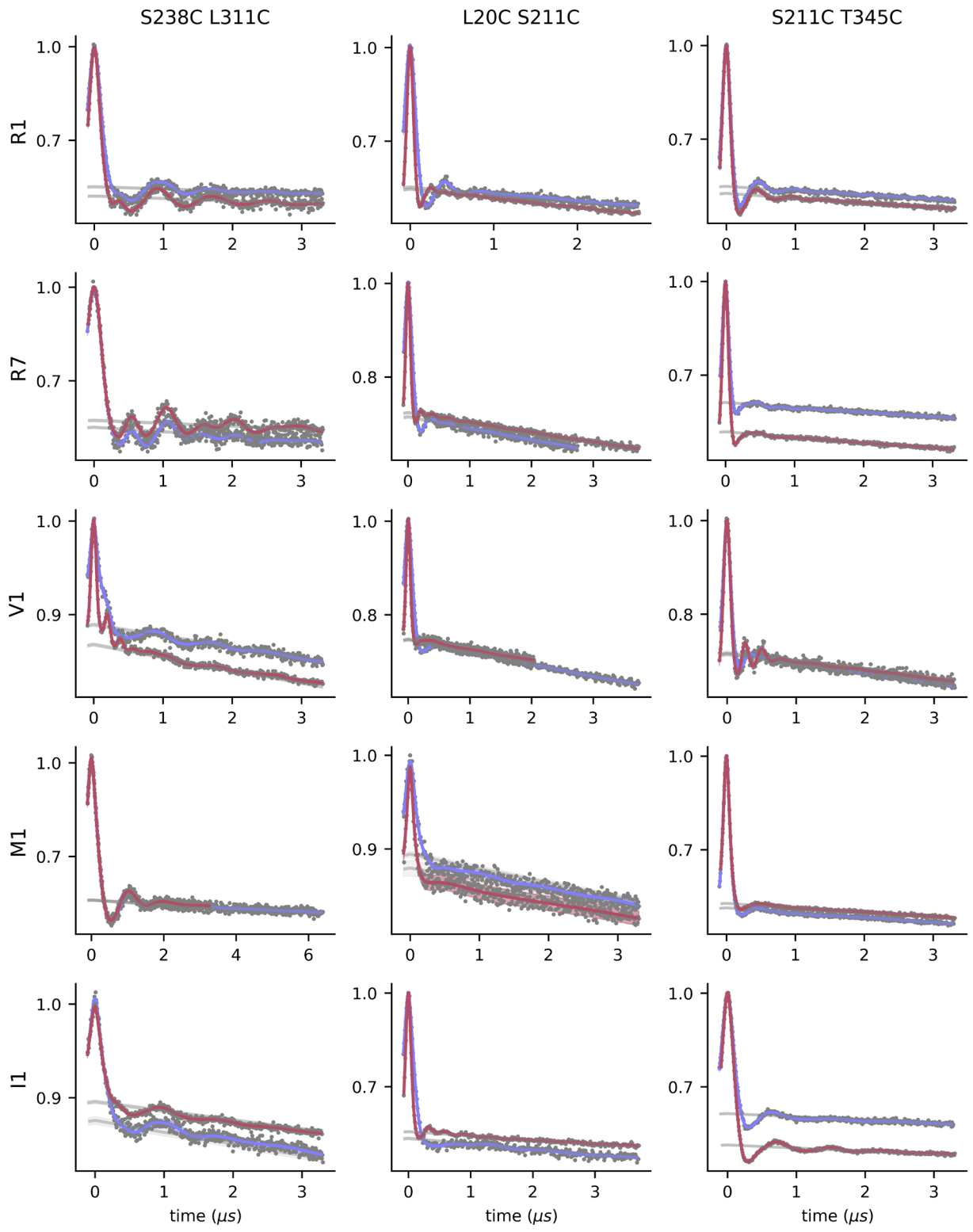


Figure S4. DEER time domain signals and fits. Raw data is shown as gray dots. Apo fits are shown in blue and holo fits are shown in red with 95% confidence intervals shown as transparent bands of the same. Background fits are shown as gray lines with 95 % confidence intervals shown as gray transparent bands

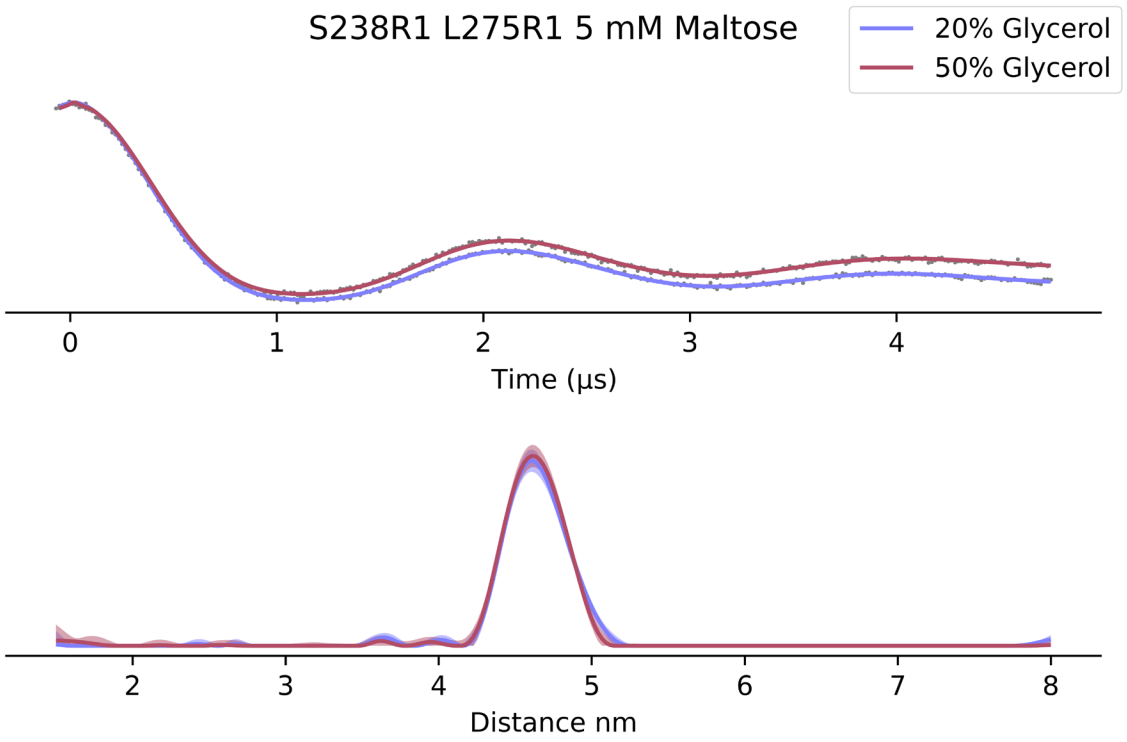


Figure S5. Effect of glycerol concentration on MBP conformation. Top. DEER Traces of MBP S238R1 L275R1 5 mM Maltose, the construct with the largest dynamic range, in 20% and 50% glycerol. Bottom. Comparison of the distance distributions of the same, showing nearly identical distributions.

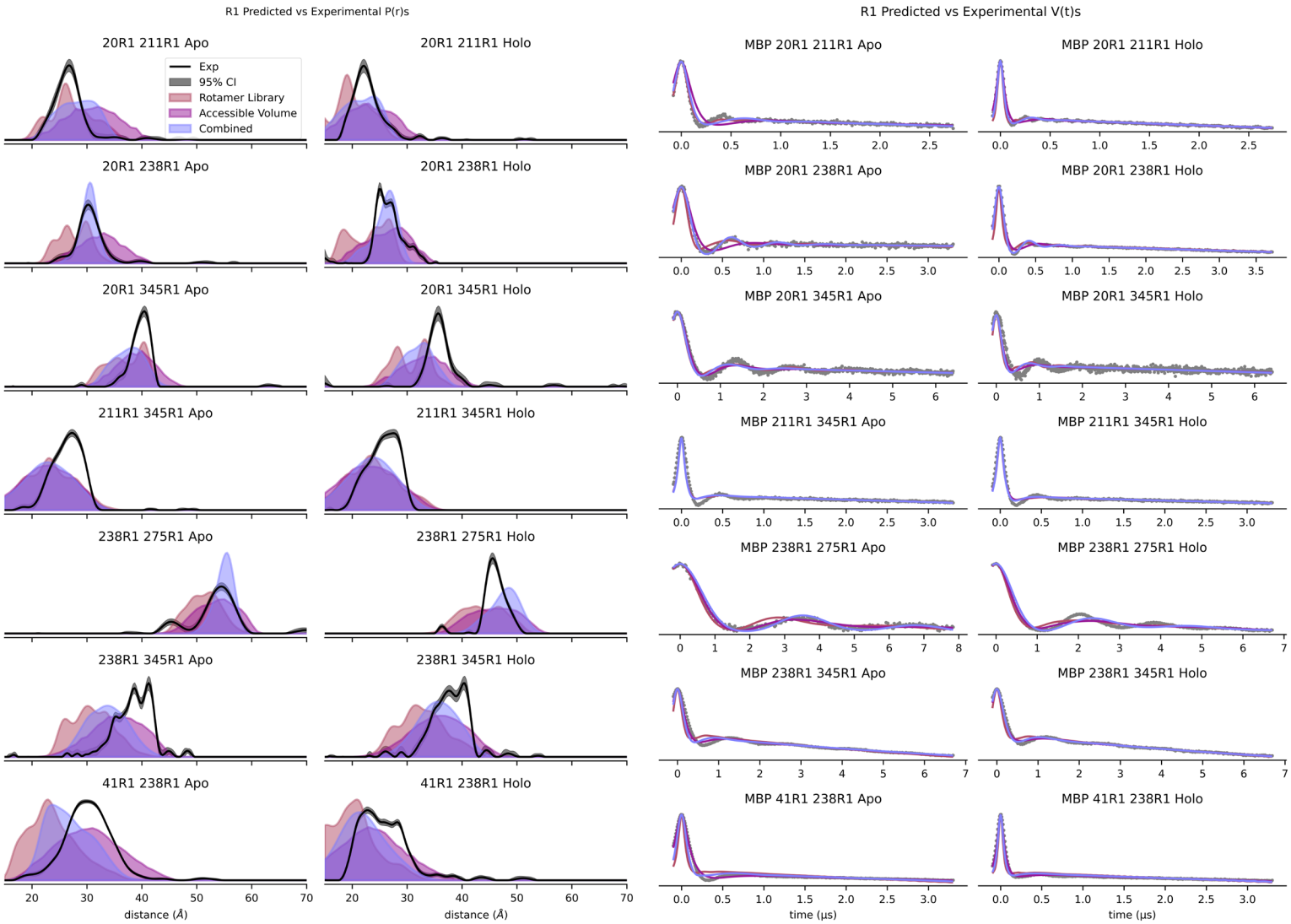


Figure S6. DEER distance distribution predictions of spin label modeling methods for R1 labeled site pairs (left). Simulated time domain signals of distance distributions predictions overlayed on experimental data (right). Time domain signal modulation depth and background were fit to optimize the root mean squared deviation of the simulated and experimental data.

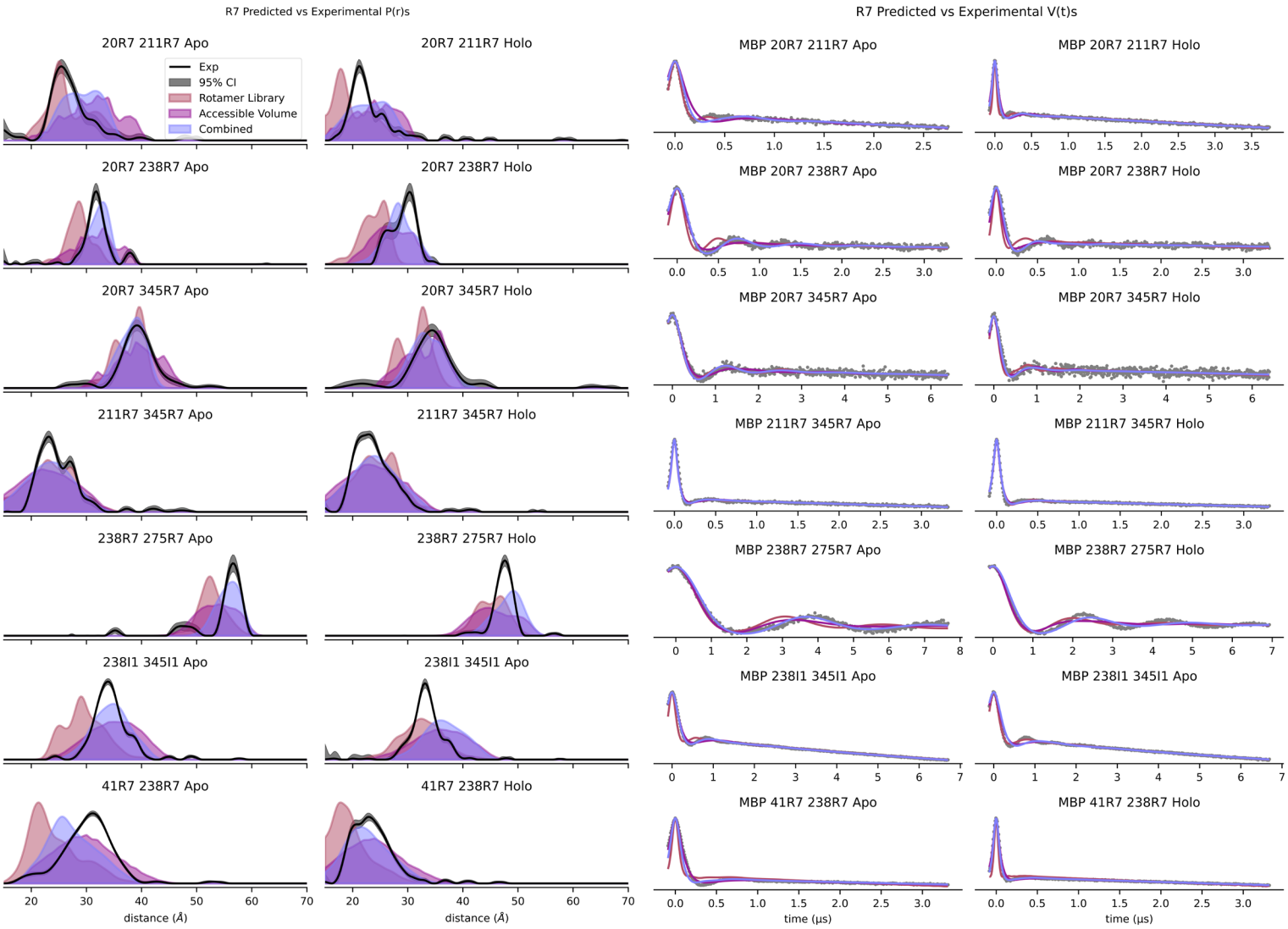


Figure S7. DEER distance distribution predictions of spin label modeling methods for R7 labeled site pairs (left). Simulated time domain signals of distance distributions predictions overlayed on experimental data (right). Time domain signal modulation depth and background were fit to optimize the root mean squared deviation of the simulated and experimental data.

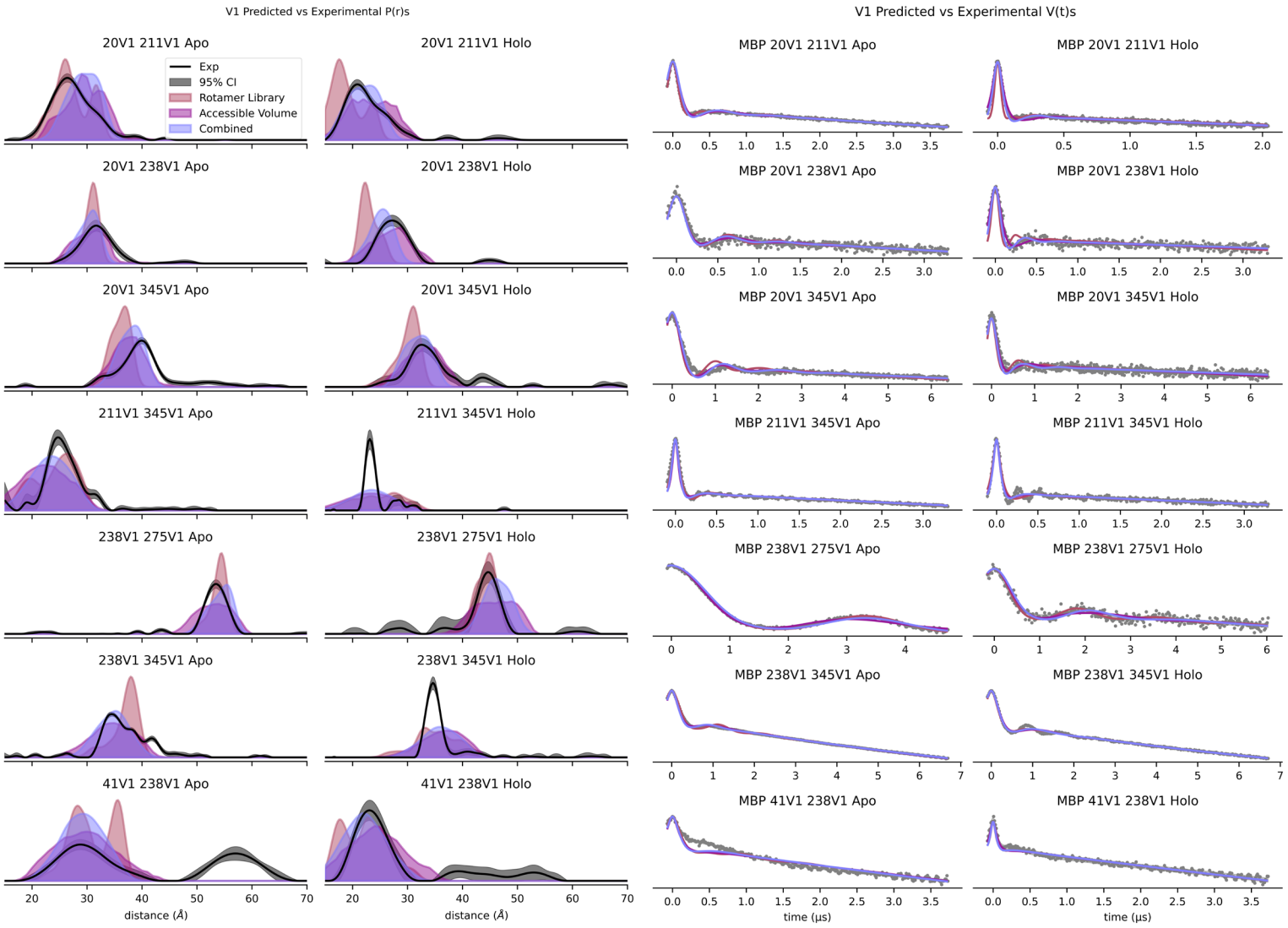


Figure S8. DEER distance distribution predictions of spin label modeling methods for V1 labeled site pairs (left). Simulated time domain signals of distance distributions predictions overlayed on experimental data (right). Time domain signal modulation depth and background were fit to optimize the root mean squared deviation of the simulated and experimental data.

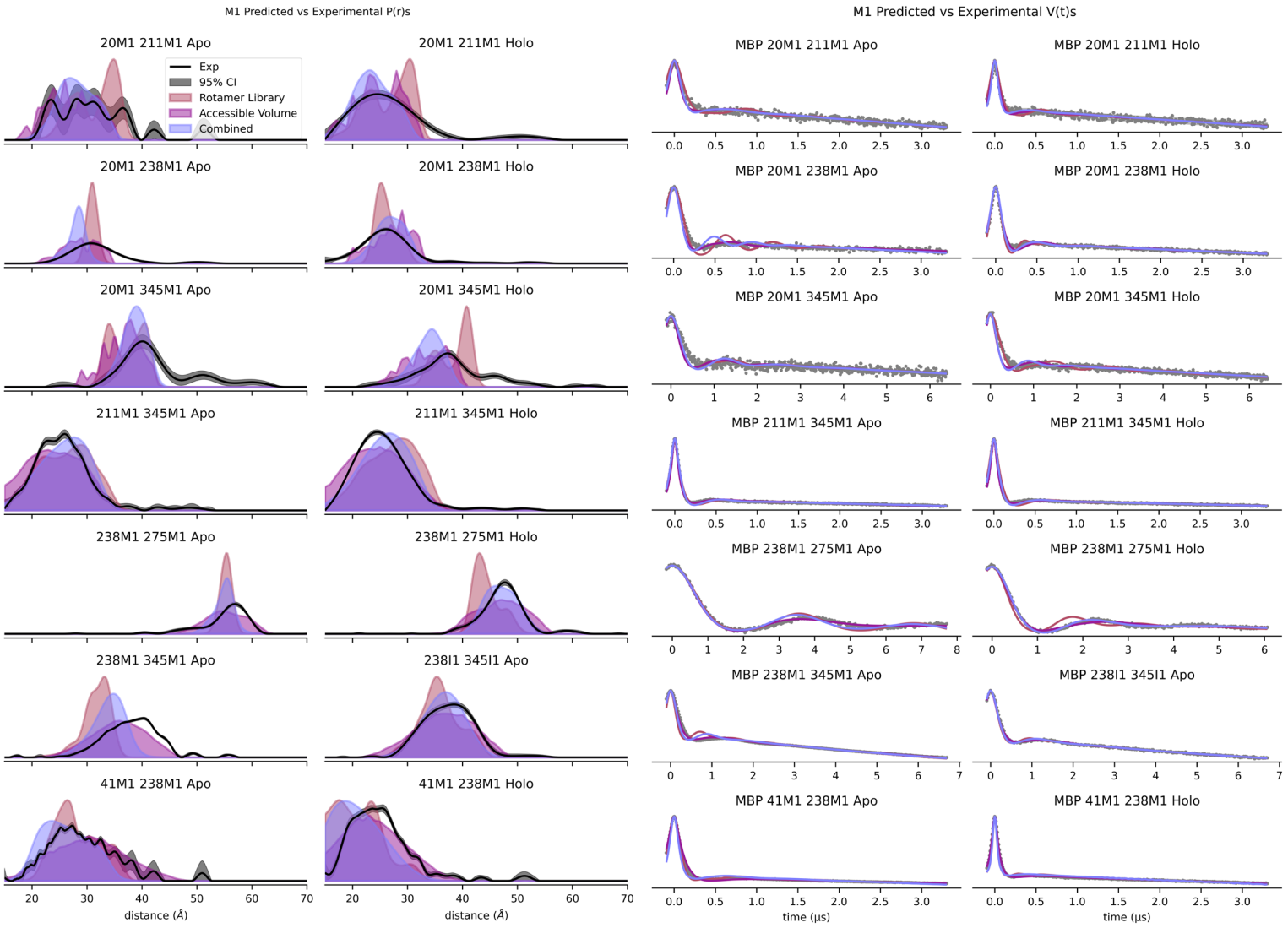


Figure S9. DEER distance distribution predictions of spin label modeling methods for M1 labeled site pairs (left). Simulated time domain signals of distance distributions predictions overlayed on experimental data (right). Time domain signal modulation depth and background were fit to optimize the root mean squared deviation of the simulated and experimental data.

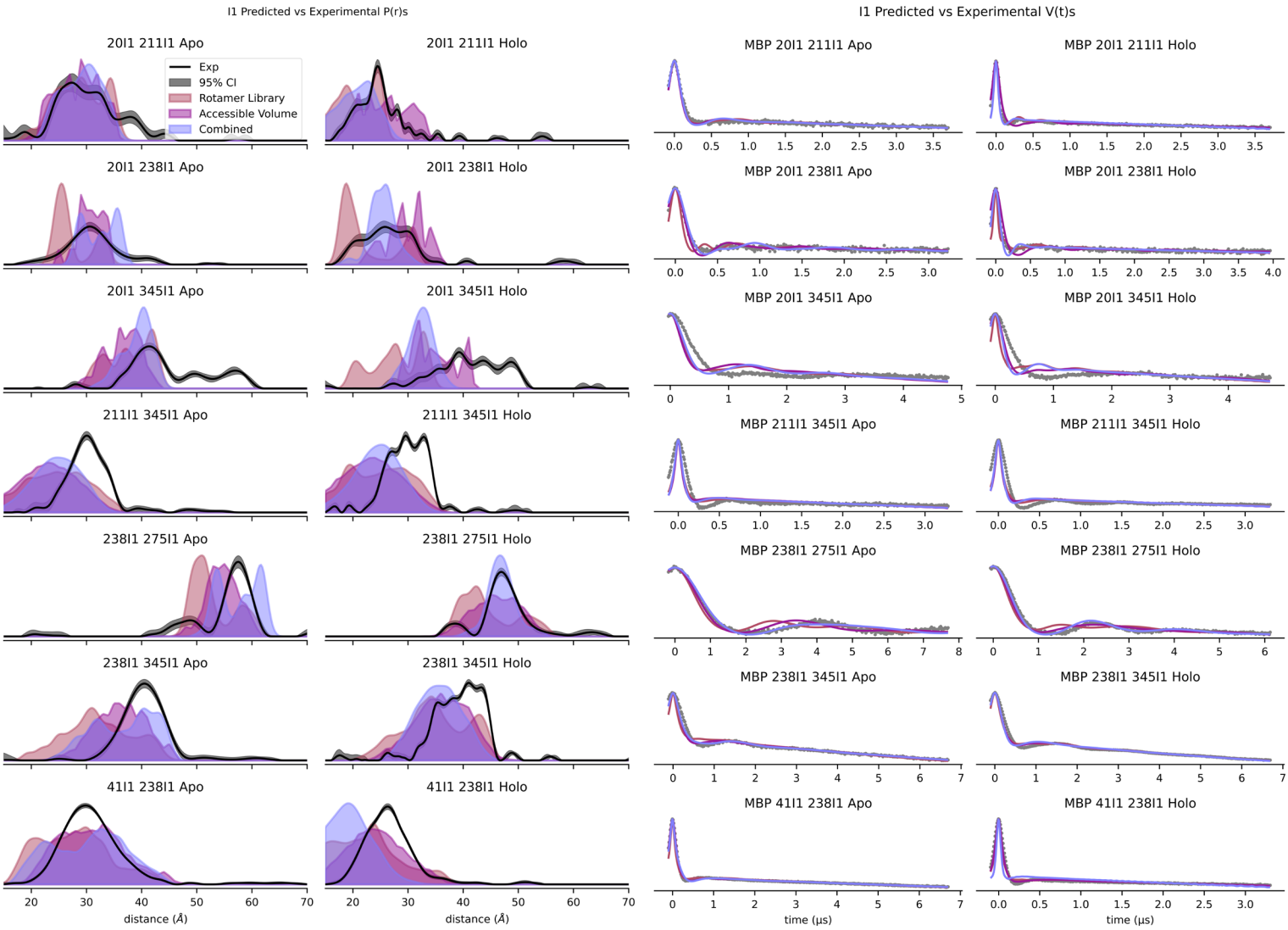


Figure S10. DEER distance distribution predictions of spin label modeling methods for I1 labeled site pairs (left). Simulated time domain signals of distance distributions predictions overlayed on experimental data (right). Time domain signal modulation depth and background were fit to optimize the root mean squared deviation of the simulated and experimental data.

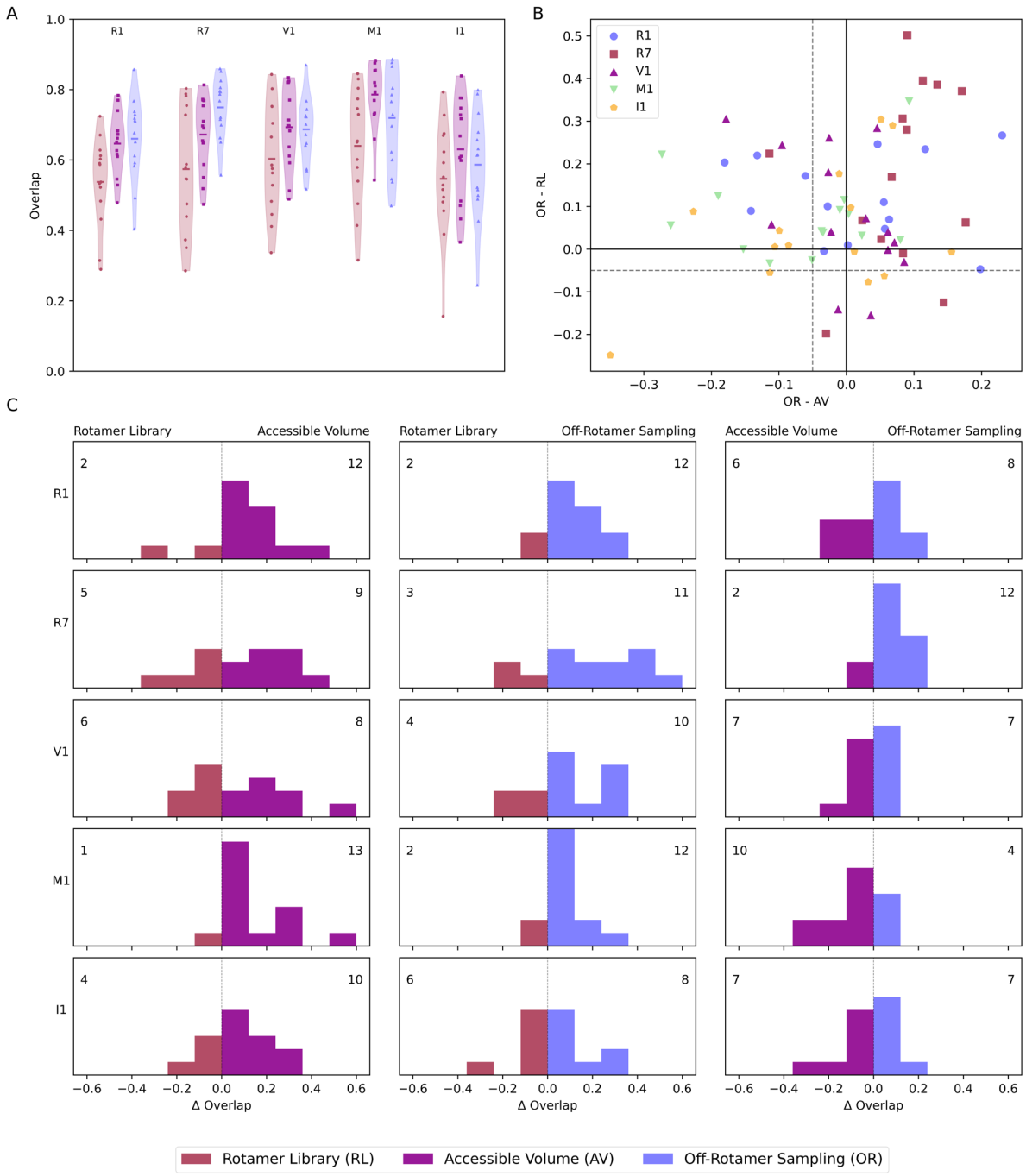


Figure S11. Extended comparison of prediction overlap. A) Violin plots comparing all methods and all spin labels. Dots indicate individual overlap scores, and horizontal lines indicate the mean overlap score of the group. Shaded bands are a density estimate of all data points in the group, indicating the characteristics of the distribution of the data. B) Scatterplot shows difference between off-rotamer sampling and other methods. Solid lines indicate the points where methods are equivalent and dashed lines indicate where points are within 5% overlap. C) Histograms of the change in overlap between modeling methods for each site pair. Colors indicate which method performed better for the bin. Numbers indicate the number of site pairs that showed improvement over the compared method.

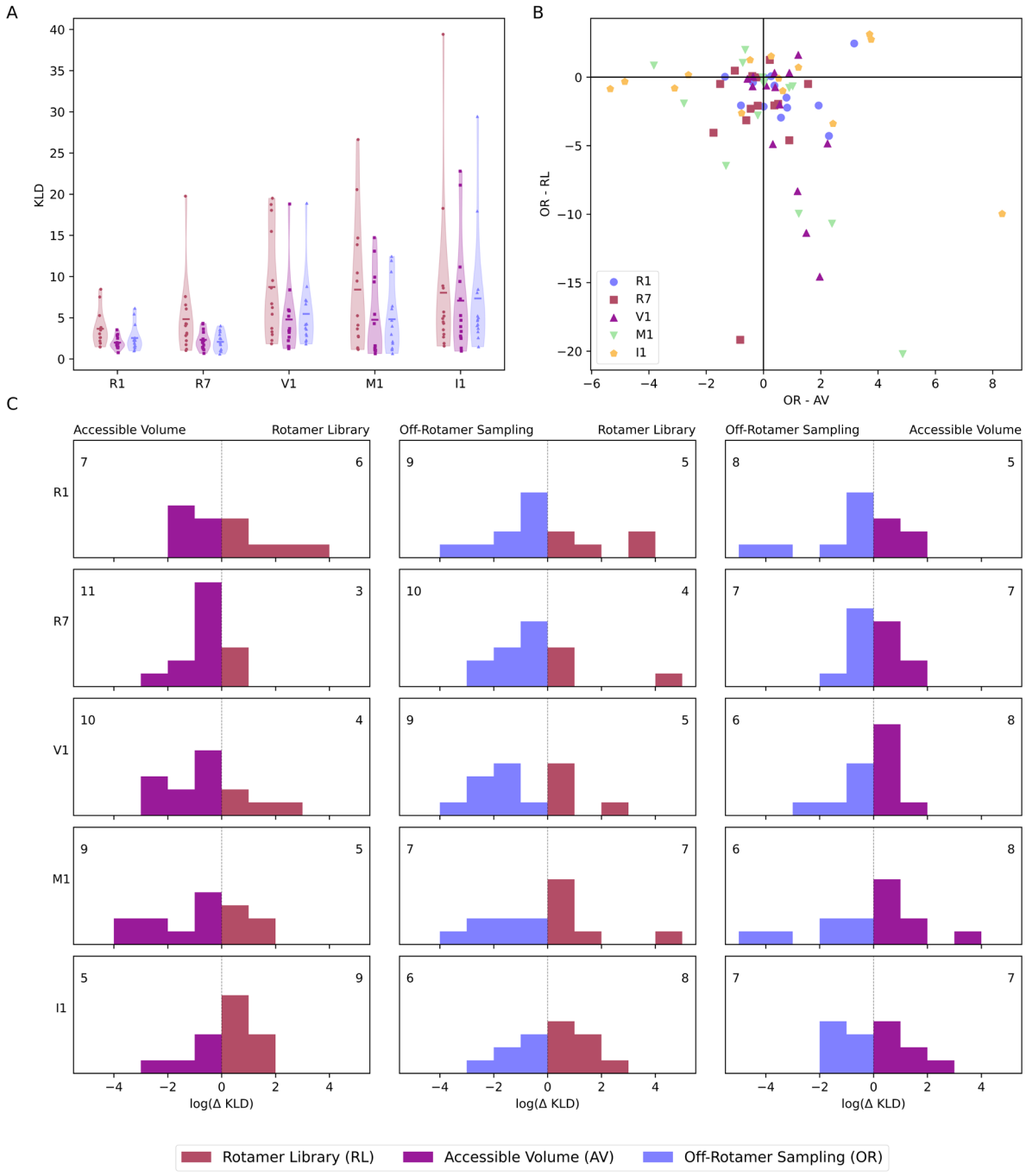


Figure S12. Extended comparison of prediction Kullback–Leibler divergence (KLD, Simulated||Experimental) . A) Violin plots comparing all methods and all spin labels. Dots indicate individual KLDs, and horizontal lines indicate the mean KLD of the group. B) Scatterplot shows difference between off-rotamer sampling and other methods. Shaded bands are a density estimate of all data points in the group, indicating the characteristics of the distribution of the data. C) Histograms of the change in log(KLD) between modeling methods for each site pair. Colors indicate which method performed better for the bin. Numbers indicate the number of site pairs that showed improvement over the compared method.

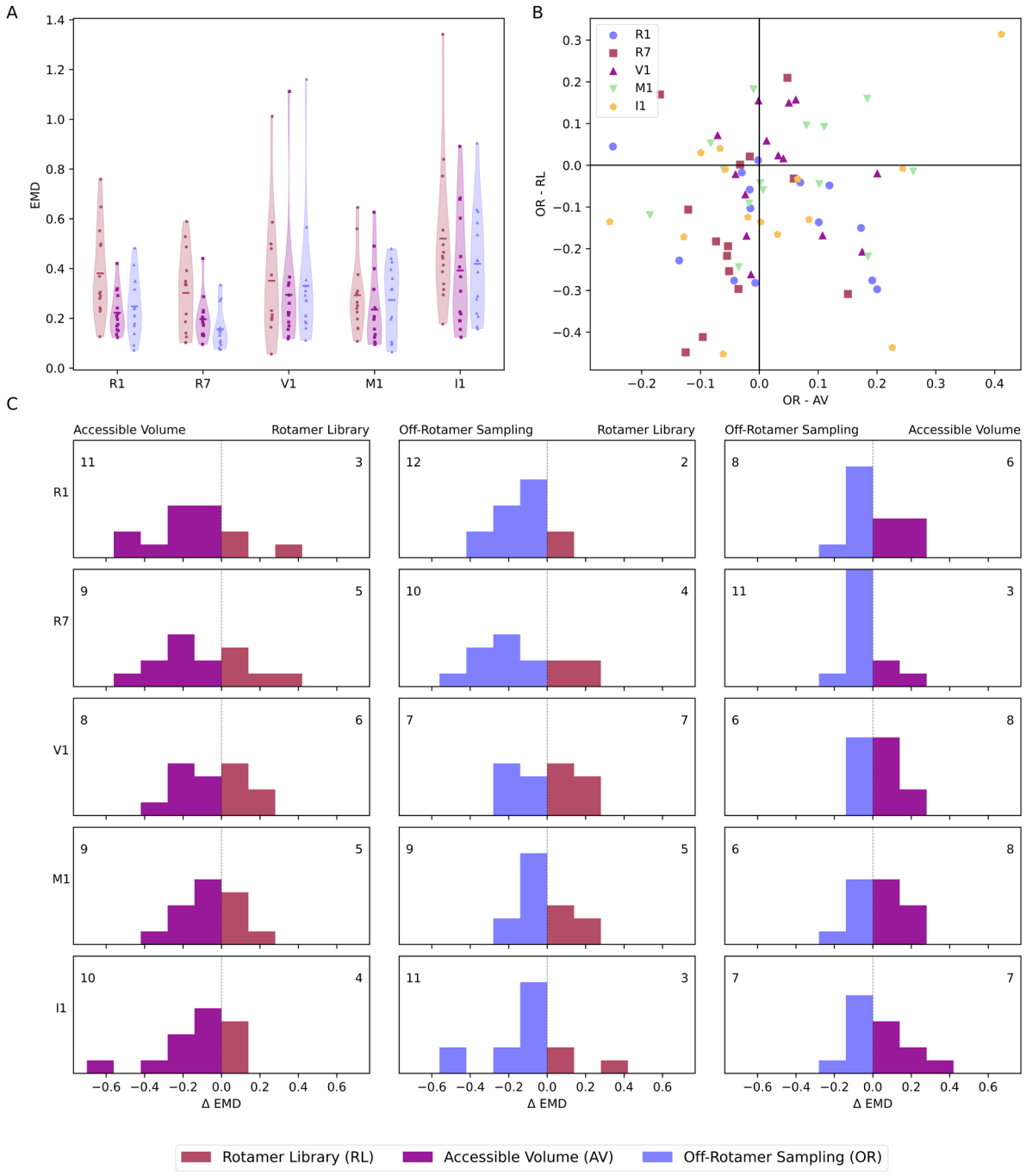


Figure S13. Extended comparison of prediction earth-movers distance (EMD) . A) Violin plots comparing all methods and all spin labels. Dots indicate individual EMDs, and horizontal lines indicate the mean EMD of the group. Shaded bands are a density estimate of all data points in the group, indicating the characteristics of the distribution of the data. B) Scatterplot shows difference between off-rotamer sampling and other methods. C) Histograms of the change in EMDs between modeling methods for each site pair. Colors indicate which method performed better for the bin. Numbers indicate the number of site pairs that showed improvement over the compared method.

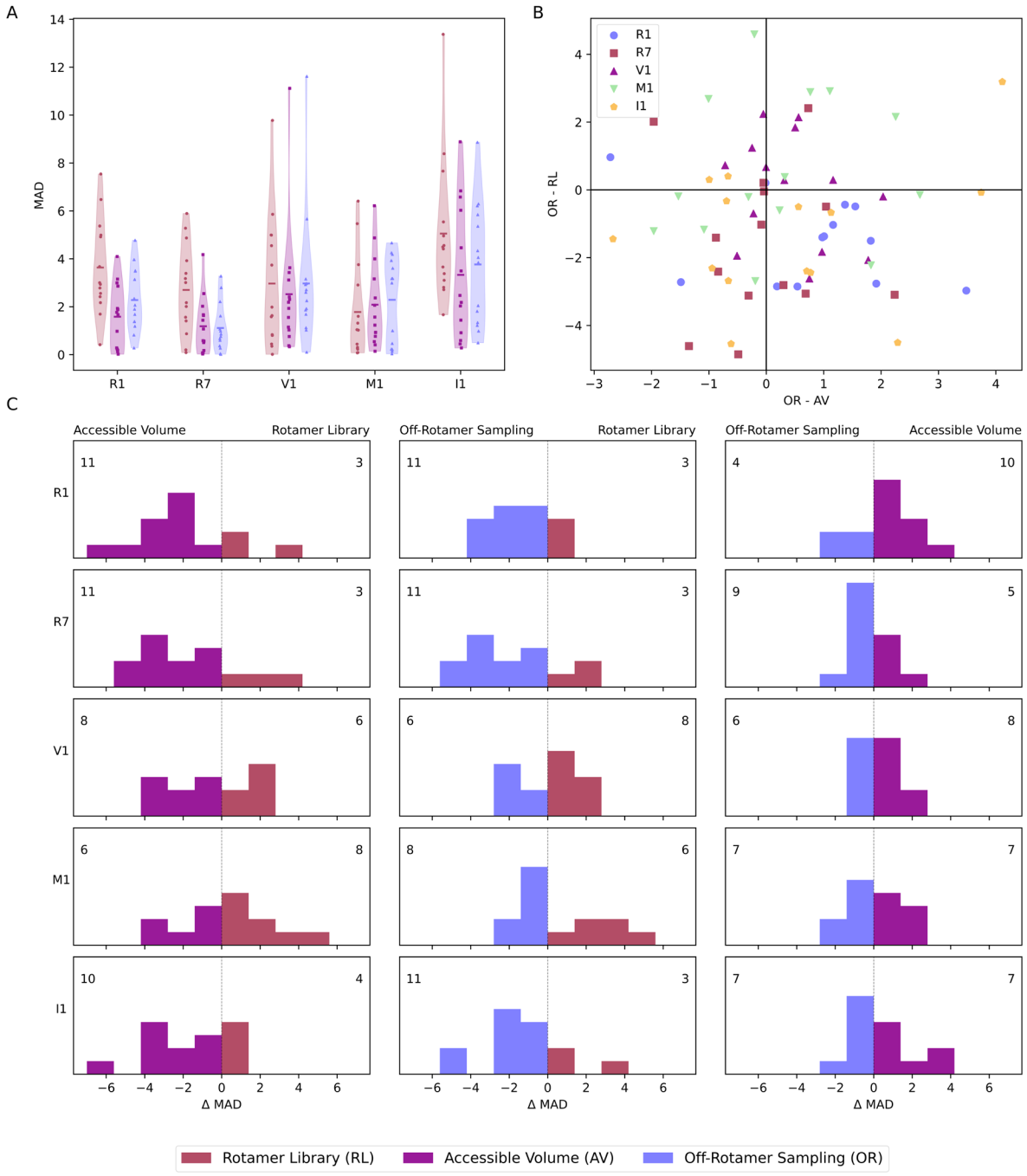


Figure S14. Extended comparison of prediction mean absolute deviation (MAD). A) Violin plots comparing all methods and all spin labels. Dots indicate individual MADs, and horizontal lines indicate the mean MAD of the group. Shaded bands are a density estimate of all data points in the group, indicating the characteristics of the distribution of the data B) Scatterplot shows difference between off-rotamer sampling and other methods. C) Histograms of the change in MAD between modeling methods for each site pair. Colors indicate which method performed better for the bin. Numbers indicate the number of site pairs that showed improvement over the compared method.

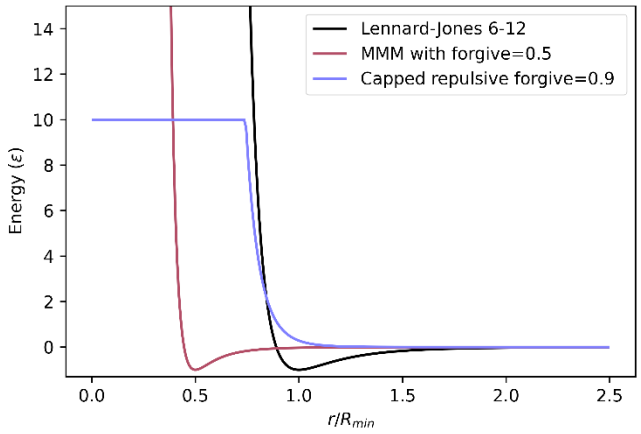


Figure S15. Comparison of clash evaluation potentials

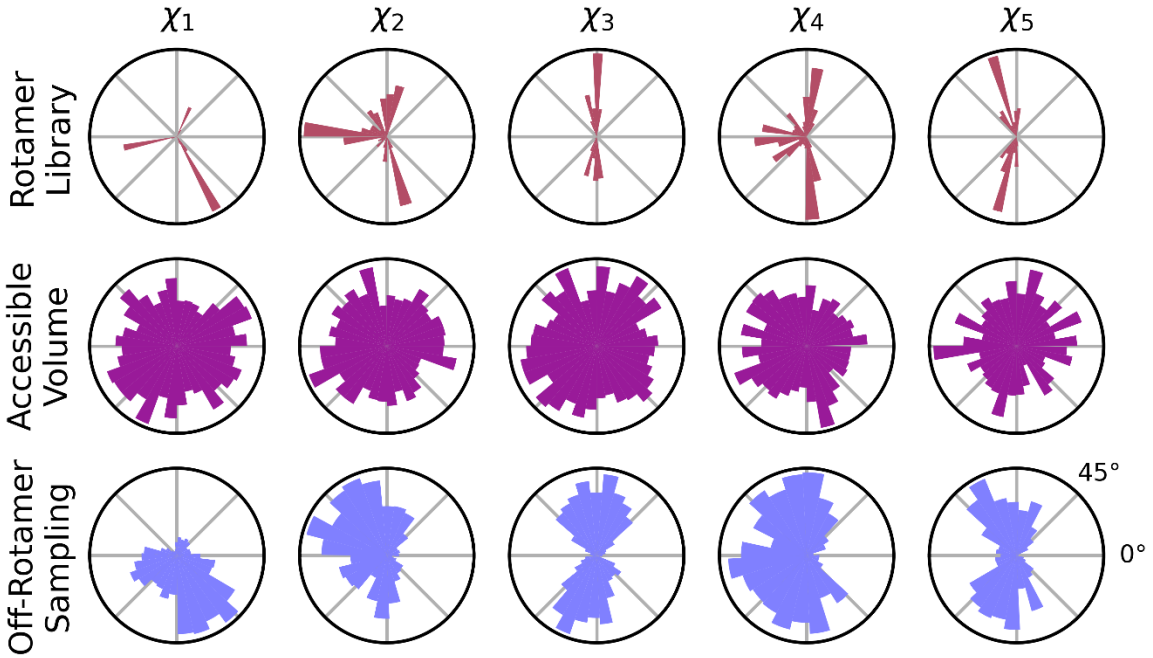


Figure S16. Polar histograms of spin label model dihedral angles for R1. Rows group the spin label modeling method and columns group the dihedral angels. Dihedral definitions for R1 are as follows (χ_1: N-CA-CB-SG, χ_2: CA-CB-SG-SD, χ_3:CB-SG-SD-CE, χ_4: SG-SD-CE-C3, χ_5: SD-CE-C3-C4) using the MMM atom names.

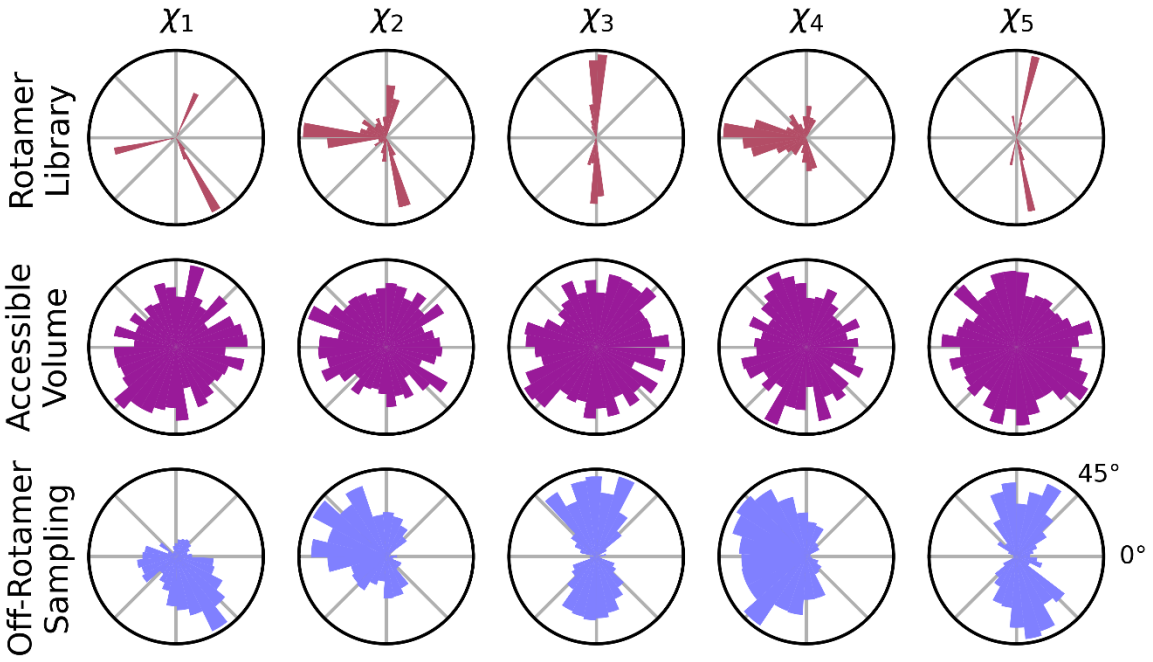


Figure S17. Polar histograms of spin label model dihedral angles for R7. Rows group the spin label modeling method and columns group the dihedral angels. Dihedral definitions for R7 are as follows (χ_1: N-CA-CB-SG, χ_2: CA-CB-SG-S2, χ_3:CB-SG-S2-C4, χ_4: SG-S2-C4-C5, χ_5: S2-C4-C5-C6) using the MMM atom names.

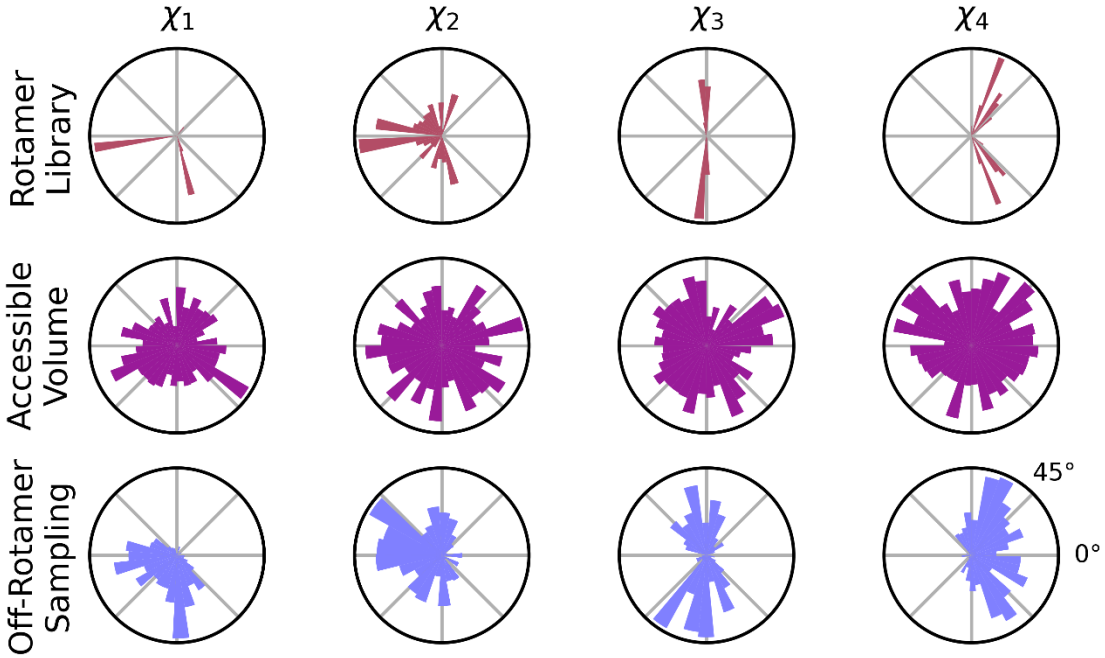


Figure S18. Polar histograms of spin label model dihedral angles for V1. Rows group the spin label modeling method and columns group the dihedral angels. Dihedral definitions for V1 are as follows (χ_1: N-CA-CB-SG, χ_2: CA-CB-SG-S2, χ_3:CB-SG-S2-C4, χ_4: SG-S2-C4-N2) using the MMM atom names.

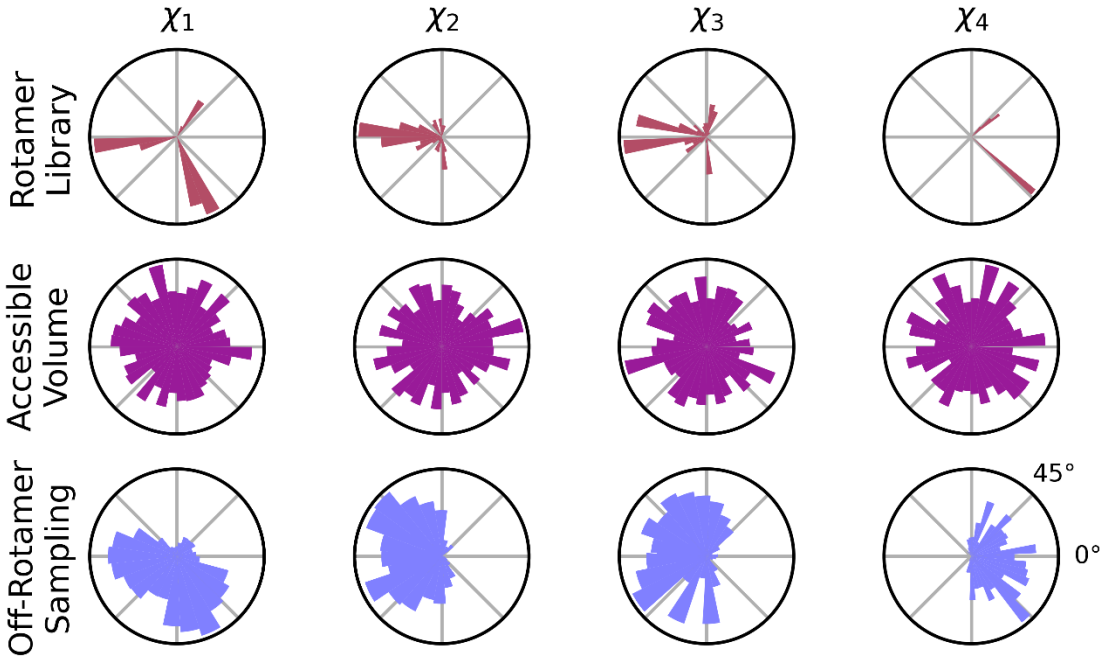


Figure S19. Polar histograms of spin label model dihedral angles for M1. Rows group the spin label modeling method and columns group the dihedral angels. Dihedral definitions for M1 are as follows (χ_1: N-CA-CB-SG, χ_2: CA-CB-SG-C4, χ_3:CB-SG-C4-C5, χ_4: C5-N2-C8-C9) using the MMM atom names.

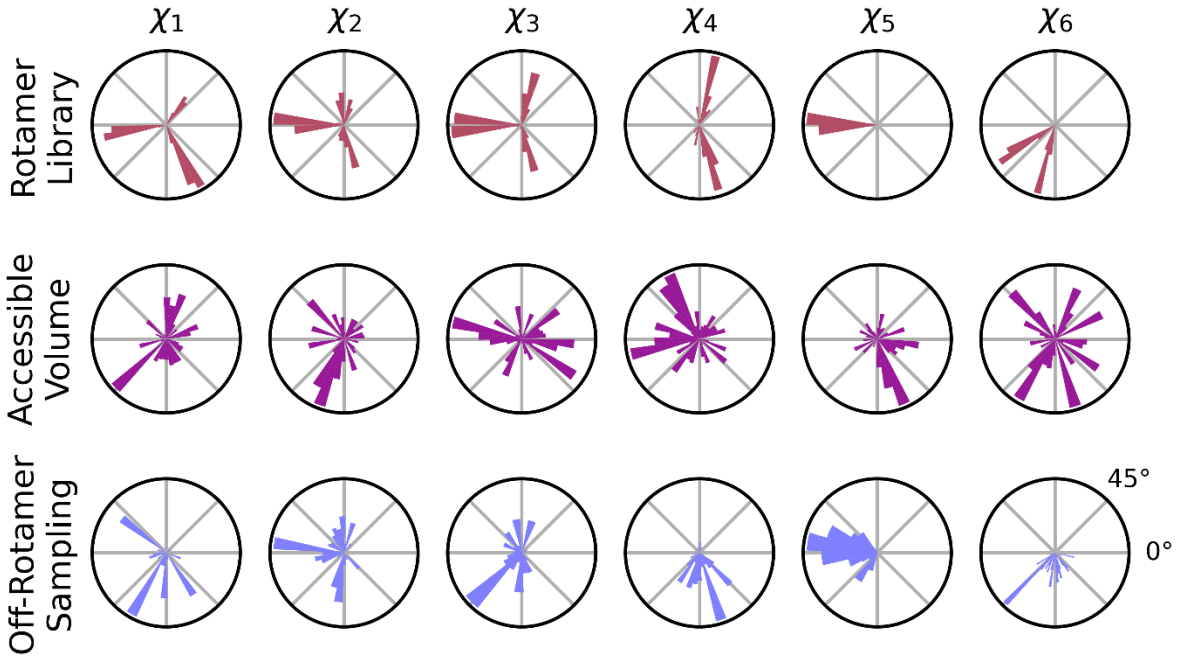


Figure S20. Polar histograms of spin label model dihedral angles for I1. Rows group the spin label modeling method and columns group the dihedral angels. Dihedral definitions for I1 are as follows (χ_1: N-CA-CB-SG, χ_2: CA-CB-SG-CD, χ_3:CB-SG-CD-CE, χ_4: SG-CD-CE-NZ, χ_5: CD-CE-NZ-C3, χ_6: CE-NZ-C3-C2) using the MMM atom names.

Description of metrics

For all metrics described below, *p(r)* and *q(r)* are the probability distribution functions being compared, *P(r)* and *Q(r)* are their respective cumulative distribution functions and ε is the machine epsilon. All metrics used for comparing distributions were computed as follows

Overlap

$$Overlap = \sum_{r} \min(\mathrm{p}(r_i), q(r_i))$$

Earth Mover's Distance (Wasserstein distance)

$$EMD = \sum_{r} |P(r_i) - Q(r_i)|$$

Kullback–Leibler divergence

$$KLD = \sum_{r} \max(p(r_i), \varepsilon) \log\left(\frac{\max(p(r_i), \varepsilon)}{\max(q(r_i), \varepsilon)}\right)$$

Mean Absolute Deviation

$$MAD = \left| \sum_{r} r_i p(r_i) - \sum_{r} r_i q(r_i) \right|$$

Table S1. List of DEER parameters.

[a] Fitted modulation depth of the experimental DEER trace.

[b] Semiquantitative labeling efficiency on a with * being the lowest efficiency and ***** being the highest.

[c] Signal to noise ratio.

[d] four pulsed DEER pump pulse time increment.

Sample	Label	State	Mod Depth[a]	Labeling Efficiency[b]	SNR[c]	Δt (ns)[d]	t_0 offset (ns)	τ_2 (ns)[e]	Scans
MBP L311C T345C I1 Apo	I1	Apo	0.20	**	33.3	22	110	7000	13
MBP L311C T345C R1 Holo	R1	Holo	0.50	*****	39.5	22	88	7000	39
MBP L311C T345C M1 Apo	M1	Apo	0.46	*****	99.4	22	83.6	7000	17
MBP L311C T345C R7 Holo	R7	Holo	0.43	****	38.1	22	90.2	7000	36
MBP L311C T345C M1 Holo	M1	Holo	0.46	*****	111.1	28	100.8	8000	30
MBP L311C T345C I1 Holo	I1	Holo	0.20	**	33.3	22	99	7000	26
MBP L311C T345C R7 Apo	R7	Apo	0.35	****	105.7	22	74.8	7000	41
MBP L311C T345C V1 Apo	V1	Apo	0.18	**	51.3	22	44	7000	40
MBP L311C T345C V1 Holo	V1	Holo	0.19	**	41.7	22	85.8	7000	125
MBP L311C T345C R1 Apo	R1	Apo	0.49	*****	98.4	22	105.6	7000	33
MBP L20C S238C R1 Holo	R1	Holo	0.50	*****	195.1	16	88	4000	93
MBP L20C S238C M1 Apo	M1	Apo	0.19	**	49.3	8	88.8	7000	135
MBP L20C S238C V1 Holo	V1	Holo	0.16	**	20.6	8	96	7000	106
MBP L20C S238C I1 Holo	I1	Holo	0.45	*****	45.8	14	74.2	4000	45
MBP L20C S238C R7 Apo	R7	Apo	0.35	****	35.2	8	111.2	7000	34
MBP L20C S238C I1 Apo	I1	Apo	0.53	*****	34.4	12	81.6	3500	62
MBP L20C S238C V1 Apo	V1	Apo	0.10	*	14.7	8	112	7000	258
MBP L20C S238C M1 Holo	M1	Holo	0.19	**	69.8	8	107.2	7000	144
MBP L20C S238C R7 Holo	R7	Holo	0.38	****	33.5	8	84.8	7000	34
MBP L20C S238C R1 Apo	R1	Apo	0.52	*****	47.7	8	97.6	7000	20
MBP L20C T345C R1 Holo	R1	Holo	0.51	*****	31.7	10	90	7000	77
MBP L20C T345C R1 Apo	R1	Apo	0.46	*****	50.5	10	95	7000	85
MBP L20C T345C V1 Holo	V1	Holo	0.24	**	21.2	10	95	7000	69
MBP L20C T345C R7 Apo	R7	Apo	0.46	*****	29.3	10	101	7000	27
MBP L20C T345C M1 Apo	M1	Apo	0.17	**	18.4	10	107	7000	22
MBP L20C T345C R7 Holo	R7	Holo	0.43	****	22.5	10	103	7000	28
MBP L20C T345C V1 Apo	V1	Apo	0.28	***	48.9	10	115	7000	81
MBP L20C T345C I1 Apo	I1	Apo	0.48	*****	75.5	16	32	5000	44
MBP L20C T345C I1 Holo	I1	Holo	0.47	*****	67.9	16	80	5000	38
MBP L20C T345C M1 Holo	M1	Holo	0.19	**	30.5	10	83	7000	44
MBP S238C L311C I1 Holo	I1	Holo	0.12	*	54.2	8	110.4	7000	97
MBP S238C L311C R7 Apo	R7	Apo	0.45	*****	36.5	8	95.2	7000	11
MBP S238C L311C R7 Holo	R7	Holo	0.43	****	32.3	8	84	7000	27
MBP S238C L311C M1 Apo	M1	Apo	0.50	*****	54.4	10	101	7000	6
MBP S238C L311C I1 Apo	I1	Apo	0.14	*	26.3	8	96.8	7000	18

MBP S238C L311C V1 Holo	V1	Holo	0.15	**	70.4	8	90.4	7000	160
MBP S238C L311C V1 Apo	V1	Apo	0.13	*	35.3	8	90.4	7000	77
MBP S238C L311C M1 Holo	M1	Holo	0.50	*****	54.6	8	110.4	7000	6
MBP S238C L311C R1 Holo	R1	Holo	0.53	*****	53.5	8	92.8	7000	11
MBP S238C L311C R1 Apo	R1	Apo	0.50	*****	72.0	8	99.2	7000	16
MBP L20C S211C I1 Apo	I1	Apo	0.49	*****	62.3	12	80.4	4000	20
MBP L20C S211C V1 Apo	V1	Apo	0.27	***	70.0	8	76	4000	27
MBP L20C S211C R7 Apo	R7	Apo	0.28	***	54.8	8	69.6	3000	21
MBP L20C S211C I1 Holo	I1	Holo	0.47	*****	118.0	12	74.4	4000	52
MBP L20C S211C M1 Holo	M1	Holo	0.11	*	23.0	8	96	7000	41
MBP L20C S211C V1 Holo	V1	Holo	0.27	***	49.7	4	77.6	2300	22
MBP L20C S211C R1 Apo	R1	Apo	0.43	****	63.0	8	80.8	3000	11
MBP L20C S211C R1 Holo	R1	Holo	0.44	****	145.6	8	80	3000	139
MBP L20C S211C R7 Holo	R7	Holo	0.27	***	75.9	8	75.2	4000	50
MBP L20C S211C M1 Apo	M1	Apo	0.10	*	22.2	8	96.8	7000	28
MBP S211C T345C M1 Apo	M1	Apo	0.44	****	138.8	8	102.4	7000	13
MBP S211C T345C M1 Holo	M1	Holo	0.43	****	166.9	8	85.6	7000	31
MBP S211C T345C I1 Apo	I1	Apo	0.39	****	93.5	8	119.2	7000	22
MBP S211C T345C R1 Holo	R1	Holo	0.51	*****	136.5	8	94.4	7000	97
MBP S211C T345C V1 Holo	V1	Holo	0.22	**	45.5	8	111.2	7000	242
MBP S211C T345C R1 Apo	R1	Apo	0.48	*****	119.7	8	100	7000	29
MBP S211C T345C I1 Holo	I1	Holo	0.49	*****	182.5	8	95.2	7000	78
MBP S211C T345C V1 Apo	V1	Apo	0.22	**	94.9	8	99.2	7000	131
MBP S211C T345C R7 Holo	R7	Holo	0.53	*****	167.0	8	88.8	7000	114
MBP S211C T345C R7 Apo	R7	Apo	0.43	****	94.1	8	77.6	7000	29
MBP S238C L275C M1 Apo	M1	Apo	0.48	*****	59.2	26	166.4	8000	14
MBP S238C L275C R1 Holo	R1	Holo	0.51	*****	139.7	22	85.8	7000	93
MBP S238C L275C V1 Apo	V1	Apo	0.28	***	60.2	16	70.4	5000	304
MBP S238C L275C R1 Apo	R1	Apo	0.50	*****	40.1	26	197.6	8000	10
MBP S238C L275C M1 Holo	M1	Holo	0.45	****	74.6	20	116	7000	9
MBP S238C L275C V1 Holo	V1	Holo	0.11	*	12.5	20	152	7000	121
MBP S238C L275C I1 Holo	I1	Holo	0.40	****	93.0	20	60	7000	18
MBP S238C L275C R7 Apo	R7	Apo	0.50	*****	39.8	26	218.4	8000	23
MBP S238C L275C I1 Apo	I1	Apo	0.40	****	53.9	26	182	8000	21
MBP S238C L275C R7 Holo	R7	Holo	0.48	*****	88.4	26	91	7000	50
MBP S238C T345C R1 Holo	R1	Holo	0.47	*****	144.3	22	107.8	7000	42
MBP S238C T345C V1 Holo	V1	Holo	0.26	***	76.0	22	94.6	7000	113
MBP S238C T345C M1 Apo	M1	Apo	0.44	****	407.6	22	103.4	7000	153
MBP S238C T345C I1 Holo	I1	Holo	0.48	*****	357.3	22	105.6	7000	127
MBP S238C T345C V1 Apo	V1	Apo	0.22	**	158.7	22	110	7000	533
MBP S238C T345C R1 Apo	R1	Apo	0.48	*****	245.8	22	103.4	7000	175
MBP S238C T345C I1 Apo	I1	Apo	0.45	****	88.7	22	112.2	7000	10
MBP S238C T345C M1 Holo	M1	Holo	0.49	*****	90.6	22	92.4	7000	12

MBP S238C T345C R7 Apo	R7	Apo	0.45	****	199.6	16	97.6	7000	87
MBP S238C T345C R7 Holo	R7	Holo	0.46	*****	126.0	16	107.2	7000	84
MBP D41C S238C R1 Holo	R1	Holo	0.43	****	275.0	8	99.2	7000	96
MBP D41C S238C R1 Apo	R1	Apo	0.43	****	333.8	8	100.8	7000	68
MBP D41C S238C I1 Holo	I1	Holo	0.48	*****	216.2	8	98.4	7000	33
MBP D41C S238C V1 Apo	V1	Apo	0.10	*	25.2	12	72	4000	18
MBP D41C S238C R7 Holo	R7	Holo	0.47	*****	196.5	8	83.2	7000	140
MBP D41C S238C M1 Apo	M1	Apo	0.45	*****	537.1	8	94.4	7000	21
MBP D41C S238C V1 Holo	V1	Holo	0.10	*	20.0	12	78	4000	36
MBP D41C S238C R7 Apo	R7	Apo	0.44	****	177.8	8	88.8	7000	47
MBP D41C S238C I1 Apo	I1	Apo	0.49	*****	284.3	16	99.2	7000	40
MBP D41C S238C M1 Holo	M1	Holo	0.41	****	234.5	8	96.8	7000	10