Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

NOE fitting to improve MD population data - a case study with *S*-Adenosylmethionine

Jessica Bame, Casper Hoeck, Matthew J. Carrington, Craig P. Butts, Christof M. Jäger,* Anna K. Croft**

S1 Simulation details:

All simulation coordinates, input and important analysis files (in particular the MD interatomic distance analysis) are also available online on the figshare repository (DOI https://doi.org/10.6084/m9.figshare.c.3842356). Additional analysis scripts (perl) and cpptraj input files are also presented there.

Atomic coordinates, atomic charges and Amber force field atom types for SAM molecules:

HF parametrisation, net charge +1:

1	N1	5.8160	-1.7950	0.9590	N3	1 SAM	-0.750981
2	C1	5.3750	-0.4450	0.4850	CT	1 SAM	0.162598
3	C2	4.3220	0.1060	1.4730	C	1 SAM	0.727601
4	O1	4.2270	-0.4940	2.5450	O2	1 SAM	-0.694426
5	O2	3.7230	1.1070	1.0780	O2	1 SAM	-0.694426
6	C3	4.9860	-0.4330	-0.9930	CT	1 SAM	-0.180320
7	C4	3.8150	-1.2990	-1.4680	CT	1 SAM	-0.111002
8	S1	2.1450	-0.8390	-0.8740	SP	1 SAM	0.298959
9	C5	1.1710	-2.1340	-1.6770	CT	1 SAM	-0.369872
10	C6	1.7590	0.6450	-1.8610	CT	1 SAM	-0.456582
11	C7	0.5490	1.4170	-1.3510	CT	1 SAM	0.395147
12	O3	-0.5430	0.5330	-1.1900	OS	1 SAM	-0.516763
13	C8	0.6950	2.1780	-0.0160	CT	1 SAM	0.259632
14	O4	0.5030	3.5390	-0.2920	OH	1 SAM	-0.680308
15	C9	-0.4440	1.6040	0.8400	CT	1 SAM	0.094263
16	O5	-1.0000	2.5840	1.6580	OH	1 SAM	-0.670709
17	C10	-1.4050	1.0770	-0.2240	CT	1 SAM	0.220550
18	N2	-2.3460	0.0840	0.2050	N*	1 SAM	-0.097438
19	C11	-2.0980	-1.0950	0.8700	CK	1 SAM	0.160221
20	N3	-3.1450	-1.8070	1.0680	NB	1 SAM	-0.526332
21	C12	-4.1740	-1.0760	0.5000	CB	1 SAM	-0.103906
22	C13	-5.5510	-1.3000	0.3780	CA	1 SAM	0.800080
23	N4	-6.1610	-2.3810	0.8910	N2	1 SAM	-0.919968
24	N5	-6.2850	-0.3820	-0.2490	NC	1 SAM	-0.787572
25	C14	-5.6840	0.7010	-0.7260	CQ	1 SAM	0.617823
26	N6	-4.4100	1.0200	-0.6700	NC	1 SAM	-0.785587
27	C15	-3.6920	0.0920	-0.0430	CB	1 SAM	0.509904
28	H1	-1.1100	-1.3670	1.1750	H5	1 SAM	0.125616
29	H2	-6.3280	1.4040	-1.2220	H5	1 SAM	0.074061
30	H3	5.1060	-2.5040	0.8570	H	1 SAM	0.408815
31	H4	6.0030	-1.7440	1.9500	H	1 SAM	0.408815
32	H5	6.6480	-2.1210	0.4900	H	1 SAM	0.408815
33	H6	6.2480	0.1840	0.5850	HP	1 SAM	0.106726
34	H7	5.8350	-0.7620	-1.5810	HC	1 SAM	0.119085
35	H8	4.8230	0.6020	-1.2560	HC	1 SAM	0.119085
36	H9	3.9200	-2.3360	-1.1790	H1	1 SAM	0.110268
37	H10	3.7750	-1.2690	-2.5470	H1	1 SAM	0.110268
38	H11	1.3800	-3.0580	-1.1590	H1	1 SAM	0.187908
39	H12	1.4400	-2.2070	-2.7200	H1	1 SAM	0.187908
40	H13	0.1320	-1.8700	-1.5680	H1	1 SAM	0.187908

```
41 H14        1.6050    0.3200   -2.8800 H1        1 SAM      0.220351
42 H15        2.6350    1.2770   -1.8220 H1        1 SAM      0.220351
43 H16        0.3150    2.1450   -2.1190 H1        1 SAM      0.075273
44 H17        1.6550    2.0210    0.4500 H1        1 SAM      0.061440
45 H18        0.3800    4.0070    0.5260 HO        1 SAM      0.458490
46 H19       -0.0620    0.7810    1.4290 H1        1 SAM      0.057657
47 H20       -1.3160    2.2090    2.4700 HO        1 SAM      0.468950
48 H21       -1.9820    1.8910   -0.6400 H2        1 SAM      0.145537
49 H22       -5.6150   -3.1570    1.1910 H         1 SAM      0.418044
50 H23       -7.1030   -2.5660    0.6220 H         1 SAM      0.418044
```

B3LYP PCM parameterisation, net charge +1:

```
 1 N1         5.8160   -1.7950    0.9590 N3        1 SAM     -0.601004
 2 C1         5.3750   -0.4450    0.4850 CT        1 SAM      0.101680
 3 C2         4.3220    0.1060    1.4730 C         1 SAM      0.678724
 4 O1         4.2270   -0.4940    2.5450 O2        1 SAM     -0.701919
 5 O2         3.7230    1.1070    1.0780 O2        1 SAM     -0.701919
 6 C3         4.9860   -0.4330   -0.9930 CT        1 SAM     -0.131748
 7 C4         3.8150   -1.2990   -1.4680 CT        1 SAM     -0.165702
 8 S1         2.1450   -0.8390   -0.8740 SP        1 SAM      0.304571
 9 C5         1.1710   -2.1340   -1.6770 CS        1 SAM     -0.401482
10 C6         1.7590    0.6450   -1.8610 CT        1 SAM     -0.437318
11 C7         0.5490    1.4170   -1.3510 CT        1 SAM      0.432890
12 O3        -0.5430    0.5330   -1.1900 OS        1 SAM     -0.488541
13 C8         0.6950    2.1780   -0.0160 CT        1 SAM      0.141171
14 O4         0.5030    3.5390   -0.2920 OH        1 SAM     -0.636579
15 C9        -0.4440    1.6040    0.8400 CT        1 SAM      0.023806
16 O5        -1.0000    2.5840    1.6580 OH        1 SAM     -0.615195
17 C10       -1.4050    1.0770   -0.2240 CT        1 SAM      0.258320
18 N2        -2.3460    0.0840    0.2050 N*        1 SAM     -0.140035
19 C11       -2.0980   -1.0950    0.8700 CK        1 SAM      0.194256
20 N3        -3.1450   -1.8070    1.0680 NB        1 SAM     -0.576177
21 C12       -4.1740   -1.0760    0.5000 CB        1 SAM     -0.029569
22 C13       -5.5510   -1.3000    0.3780 CA        1 SAM      0.656657
23 N4        -6.1610   -2.3810    0.8910 N2        1 SAM     -0.757974
24 N5        -6.2850   -0.3820   -0.2490 NC        1 SAM     -0.762493
25 C14       -5.6840    0.7010   -0.7260 CQ        1 SAM      0.561196
26 N6        -4.4100    1.0200   -0.6700 NC        1 SAM     -0.777049
27 C15       -3.6920    0.0920   -0.0430 CB        1 SAM      0.546488
28 H1        -1.1100   -1.3670    1.1750 H5        1 SAM      0.121211
29 H2        -6.3280    1.4040   -1.2220 H5        1 SAM      0.057540
30 H3         5.1060   -2.5040    0.8570 H         1 SAM      0.383791
31 H4         6.0030   -1.7440    1.9500 H         1 SAM      0.383791
32 H5         6.6480   -2.1210    0.4900 H         1 SAM      0.383791
33 H6         6.2480    0.1840    0.5850 HP        1 SAM      0.120759
34 H7         5.8350   -0.7620   -1.5810 HC        1 SAM      0.112726
35 H8         4.8230    0.6020   -1.2560 HC        1 SAM      0.112726
36 H9         3.9200   -2.3360   -1.1790 H1        1 SAM      0.143697
37 H10        3.7750   -1.2690   -2.5470 H1        1 SAM      0.143697
38 H11        1.3800   -3.0580   -1.1590 H1        1 SAM      0.199040
39 H12        1.4400   -2.2070   -2.7200 H1        1 SAM      0.199040
40 H13        0.1320   -1.8700   -1.5680 H1        1 SAM      0.199040
41 H14        1.6050    0.3200   -2.8800 H1        1 SAM      0.208899
42 H15        2.6350    1.2770   -1.8220 H1        1 SAM      0.208899
43 H16        0.3150    2.1450   -2.1190 H1        1 SAM      0.070558
44 H17        1.6550    2.0210    0.4500 H1        1 SAM      0.104387
```

```
45 H18         0.3800    4.0070    0.5260 HO       1 SAM      0.438221
46 H19        -0.0620    0.7810    1.4290 H1       1 SAM      0.093864
47 H20        -1.3160    2.2090    2.4700 HO       1 SAM      0.460604
48 H21        -1.9820    1.8910   -0.6400 H2       1 SAM      0.114488
49 H22        -5.6150   -3.1570    1.1910 H        1 SAM      0.382089
50 H23        -7.1030   -2.5660    0.6220 H        1 SAM      0.382089
```

PCM parameterisation, R-R isomer, net charge +1:

```
 1 N1           5.79812   -1.78965    0.86352 N3       1 SAM     -0.601004
 2 C1           5.27477   -0.47898    0.45368 CT       1 SAM      0.101680
 3 C2           4.08564   -0.06510    1.32881 C        1 SAM      0.678724
 4 O1           4.38073   -0.20568    2.68691 O2       1 SAM     -0.701919
 5 O2           3.75507    1.25827    1.07447 O2       1 SAM     -0.701919
 6 C3           4.95638   -0.43053   -1.07683 CT       1 SAM     -0.131748
 7 C4           3.90575   -1.45145   -1.61113 CT       1 SAM     -0.165702
 8 S1           2.34120   -0.89421   -2.46693 SP       1 SAM      0.304571
 9 C5           1.16755   -2.08815   -1.70758 CS       1 SAM     -0.401482
10 C6           1.96899    0.82187   -1.87282 CT       1 SAM     -0.437318
11 C7           0.61607    1.41460   -1.36777 CT       1 SAM      0.432890
12 O3          -0.44619    0.49551   -1.25030 OS       1 SAM     -0.488541
13 C8           0.71411    2.22916   -0.05466 CT       1 SAM      0.141171
14 O4           0.41673    3.58225   -0.32344 OH       1 SAM     -0.636579
15 C9          -0.37030    1.59264    0.78364 CT       1 SAM      0.023806
16 O5          -0.97048    2.48767    1.68635 OH       1 SAM     -0.615195
17 C10         -1.29649    1.08770   -0.29862 CT       1 SAM      0.258320
18 N2          -2.24600    0.10045    0.20244 N*       1 SAM     -0.140035
19 C11         -1.98008   -1.11943    0.89927 CK       1 SAM      0.194256
20 N3          -3.16733   -1.86763    1.17150 NB       1 SAM     -0.576177
21 C12         -4.21031   -1.08237    0.59477 CB       1 SAM     -0.029569
22 C13         -5.64744   -1.33511    0.47241 CA       1 SAM      0.656657
23 N4          -6.26414   -2.47656    1.07197 N2       1 SAM     -0.757974
24 N5          -6.43757   -0.35270   -0.22720 NC       1 SAM     -0.762493
25 C14         -5.84807    0.83118   -0.79028 CQ       1 SAM      0.561196
26 N6          -4.44115    1.07694   -0.66947 NC       1 SAM     -0.777049
27 C15         -3.64924    0.11297    0.02433 CB       1 SAM      0.546488
28 H1          -0.98909   -1.45808    1.18206 H5       1 SAM      0.121211
29 H2          -6.47528    1.54675   -1.31278 H5       1 SAM      0.057540
30 H3           5.05844   -2.52594    0.82269 H        1 SAM      0.383791
31 H4           6.15785   -1.74407    1.84386 H        1 SAM      0.383791
32 H5           6.59509   -2.06799    0.25219 H        1 SAM      0.383791
33 H6           6.08793    0.26805    0.63296 HP       1 SAM      0.120759
34 H7           5.92918   -0.68915   -1.55486 HC       1 SAM      0.112726
35 H8           4.75733    0.59094   -1.43381 HC       1 SAM      0.112726
36 H9           3.78487   -2.33422   -0.96135 H1       1 SAM      0.143697
37 H10          4.40763   -1.90695   -2.49238 H1       1 SAM      0.143697
38 H11          1.55779   -2.58817   -0.80253 H1       1 SAM      0.199040
39 H12          1.03344   -2.89882   -2.45431 H1       1 SAM      0.199040
40 H13          0.16198   -1.68654   -1.53744 H1       1 SAM      0.199040
41 H14          1.85877    1.12114   -2.94219 H1       1 SAM      0.208899
42 H15          2.82951    1.47411   -1.63643 H1       1 SAM      0.208899
43 H16          0.27468    2.14852   -2.13940 H1       1 SAM      0.070558
44 H17          1.68355    2.16615    0.46884 H1       1 SAM      0.104387
45 H18          0.66418    4.10027    0.48556 HO       1 SAM      0.438221
46 H19          0.07939    0.71599    1.30505 H1       1 SAM      0.093864
47 H20         -1.41192    1.93634    2.38483 HO       1 SAM      0.460604
48 H21         -1.84787    1.92889   -0.78179 H2       1 SAM      0.114488
```

```
     49 H22          -5.61153   -3.29542    1.04334 H        1 SAM      0.382089
     50 H23          -7.14022   -2.71453    0.55816 H        1 SAM      0.382089
```

Force field modification for Amber to adjust SAM parameters by Saez and Voehringer:[1]

```
frcmod file for Adomet(SAM) with parameters of Markham et al. Biochemistry 2002,
Saez et al 2015, Yildirim et al 2010
MASS
SP 32.7798       2.900
CS 12.010        0.878

BOND
CT-SP  345.28   1.78     same as CT-SP,Markham 2002
SP-CS  345.28   1.78     same as CT-SP,Markham 2002
CS-H1  340.00   1.090    same as CT-H1

ANGLE
CT-CT-SP    96.058    96.058    Markham 2002
CT-SP-CS   289.219   100.546    Markham 2002
CT-SP-CT   289.219   100.546    Markham 2002
SP-CT-H1    33.045   104.439    Markham 2002
SP-CS-H1    33.045   104.439    Markham 2002
H1-CS-H1    35.000   109.500    Markham 2002

DIHE
CT-CT-SP-CT   2    0.400       315.000          -3.000      Saez 2015
CT-CT-SP-CT   2    0.200        50.000          -2.000      Saez 2015
CT-CT-SP-CT   2    1.000       350.000           1.000      Saez 2015
CT-CT-SP-CS   2    0.300       100.000          -3.000      Saez 2015
CT-CT-SP-CS   2    0.100       325.000          -2.000      Saez 2015
CT-CT-SP-CS   2    0.900       300.000           1.000      Saez 2015
CT-SP-CS-H1   1    0.000         0.000           0.000      ATTN, need revision
CT-SP-CT-H1   1    0.000         0.000           0.000      ATTN, need revision
CS-SP-CT-H1   1    0.000         0.000           0.000      ATTN, need revision
OS-CT-N*-CK   2    0.152425      0.000          -4.000      Yildirim 2010
OS-CT-N*-CK   2   -1.699430      0.000          -3.000      Yildirim 2010
OS-CT-N*-CK   2    0.504875      0.000          -2.000      Yildirim 2010
OS-CT-N*-CK   2    1.355570      0.000          -1.000      Yildirim 2010
CT-CT-N*-CK   2    0.228818      0.000          -4.000      Yildirim 2010
CT-CT-N*-CK   2    1.267980      0.000          -3.000      Yildirim 2010
CT-CT-N*-CK   2   -0.278197      0.000          -2.000      Yildirim 2010
CT-CT-N*-CK   2    1.603540      0.000           1.000      Yildirim 2010

IMPROPER
CT-O2-C -O2        10.5          180.0         2.0
H5-N*-CK-NB         1.1          180.0         2.0
CA-CB-CB-NB         1.1          180.0         2.0
CA-H -N2-H          1.0          180.0         2.0
H5-NC-CQ-NC         1.1          180.0         2.0
CB-N*-CB-NC         1.1          180.0         2.0

NONBON
  SP          2.3890  0.0053
  CS          1.9080  0.1094
```

S2 Simulation analysis:

Table S1: Average NOE relevant interatomic distances (for comparison with r^{-6} averaged distances from Table 1 in the main publication) from 500 ns simulation (PCM point charge parameterisation).

PCM-500ns		1	2	3	4	5	6
8-1'	H28toH48	2.538	2.532	3.731	3.728	3.7	3.76
8-2'	H28toH46	4.411	3.943	3.381	3.875	2.351	2.311
8-3'	H28toH44	5.971	5.406	4.885	3.336	4.518	2.873
8-5'	H28toH42	5.94	6.173	3.695	3.25	4.496	4.436
1'-2'	H48toH46	3.02	2.845	3.025	2.834	3.024	2.817
1'-3'	H48toH43	3.173	3.019	2.895	3.035	2.971	3.129
2'-3'	H46toH44	2.345	2.333	2.326	2.358	2.333	2.368
2'-5'	H46toH42	2.786	4.482	3.208	4.6	3.047	4.522
3'-4'	H44toH43	2.797	3.018	2.877	3.024	2.85	3.015
3'-5'	H44toH42	2.821	2.589	2.761	2.687	2.796	2.631
3'-5''	H44toH41	2.514	2.915	2.568	2.889	2.562	2.873
4'-5'	H43toH42	3.057	3.025	3.049	3.024	3.05	3.042
4'-5''	H43toH41	2.54	2.542	2.563	2.59	2.556	2.562
5'-5''	H42toH41	1.816	1.816	1.815	1.816	1.817	1.817
5'gamma	H42toH36	3.494	3.477	3.532	3.467	3.461	3.448

Table S2: Analysis of pucker population of 500 ns simulation (PCM point charge parameterisation). Analysis according to nomenclature outlined by Altona and Sundaralingam.[2]

	C2'exo	**C3'endo**	C4'exo	O4'endo	C1'exo	**C2'endo**	C3'exo
average pucker angle [°]	-10.0	22.3	31.9	89.1	128.8	160.9	-172.1
occupation [%]							
all	3.3	38.3	31.9	4.3	8.1	12.5	1.5
c1	0.0	0.5	1.0	3.4	18.1	69.1	7.9
c2	2.7	46.8	46.0	2.9	0.8	0.5	0.1
c3	0.0	0.8	2.7	15.2	50.5	28.2	2.5
c4	3.2	52.1	38.4	3.8	1.8	0.6	0.0
c5	0.1	1.0	2.7	8.6	43.4	37.8	5.8
c6	15.0	47.2	28.3	4.3	2.2	0.9	0.1

Table S3. Differences of cluster averaged interatomic distances and mean absolute distance difference (MADD) between 300ns and 500ns simulation of SAM using PCM parameterisation. All values in Ångstrom.

	c1	c2	c3	c4	c5	c6
8-1'	0.000	0.000	-0.005	-0.001	-0.001	0.000
8-2'	0.002	-0.002	0.019	0.003	0.001	0.003
8-3'	0.003	-0.002	0.023	0.006	0.001	0.010
8-5'	-0.006	0.006	0.000	0.013	-0.001	0.004
1'-2'	0.001	-0.001	0.001	0.000	0.000	0.002
1'-3'	0.004	0.003	0.004	0.002	0.004	-0.002
2'-3'	0.001	0.000	0.002	0.000	0.000	0.000
2'-5'	-0.014	-0.008	-0.012	0.011	-0.002	-0.003
3'-4'	-0.002	0.000	-0.002	0.000	-0.001	0.000
3'-5'	0.004	-0.014	0.002	0.020	0.004	0.008
3'-5''	-0.005	0.015	-0.004	-0.010	0.000	-0.007
4'-5'	0.000	0.003	0.001	-0.010	-0.001	-0.003
4'-5''	0.001	-0.009	0.000	0.010	0.002	0.004
5'-5''	0.000	0.000	0.000	0.000	0.000	0.000
5'gamma	-0.009	-0.005	0.010	0.003	0.000	-0.007
average MADD	**0.003**	**0.005**	**0.006**	**0.006**	**0.001**	**0.004**
						0.004

Table S4. Differences of cluster averaged interatomic distances and mean absolute distance difference (MADD) between 500ns simulation of SAM using PCM and HF parameterisation. All values in Ångstrom.

	c1	c2	c3	c4	c5	c6
8-1'	-0.0020	-0.0030	0.0000	-0.0030	-0.0190	-0.0070
8-2'	-0.0090	-0.0060	-0.0060	-0.0210	0.0100	-0.0190
8-3'	-0.0090	-0.0120	0.0050	0.0150	0.0030	0.1010
8-5'	0.0130	0.0730	0.0360	0.0510	0.1530	0.0990
1'-2'	0.0020	0.0110	0.0050	0.0130	0.0020	0.0200
1'-3'	-0.0130	-0.0560	-0.0360	-0.0360	-0.0550	-0.0660
2'-3'	0.0010	0.0000	-0.0020	-0.0020	-0.0030	-0.0010
2'-5'	-0.0340	-0.0060	-0.0830	-0.0280	0.0380	-0.0360
3'-4'	0.0040	0.0030	0.0050	0.0000	0.0140	0.0010
3'-5'	-0.0540	-0.0500	-0.1120	-0.0440	-0.0460	-0.0490
3'-5''	-0.0040	0.0520	0.0050	-0.0050	-0.0070	0.0110

4'-5'	0.0020	-0.0010	0.0290	0.0030	-0.0030	0.0030
4'-5''	-0.0110	-0.0180	-0.0360	-0.0020	0.0000	-0.0080
5'-5''	-0.0020	-0.0010	-0.0020	-0.0010	0.0000	-0.0010
5'gamma	0.1970	0.2190	0.3050	0.2630	0.2100	0.2460
average MADD	**0.024**	**0.034**	**0.044**	**0.032**	**0.038**	**0.045**
						0.036

Table S3. Differences of cluster averaged interatomic distances and mean absolute distance difference (MADD) between 500ns simulation of SAM using PCM parameterisation for both S-isomeric structures. All values in Ångstrom.

	c1	**c2**	**c3**	**c4**	**c5**	**c6**
8-1'	-0.001	0.001	0.005	0.003	-0.016	-0.004
8-2'	-0.002	-0.029	-0.023	-0.010	0.003	-0.013
8-3'	-0.001	-0.014	0.000	-0.001	0.001	-0.030
8-5'	-0.005	0.003	0.007	-0.088	0.011	-0.133
1'-2'	0.000	-0.004	0.003	0.002	0.000	-0.005
1'-3'	-0.008	0.025	-0.020	-0.005	0.005	0.016
2'-3'	0.001	0.000	-0.001	0.000	0.001	0.004
2'-5'	-0.027	-0.147	-0.044	-0.090	-0.040	-0.084
3'-4'	0.002	-0.002	0.003	-0.002	0.000	0.001
3'-5'	-0.027	-0.214	-0.052	-0.129	-0.030	-0.150
3'-5''	0.018	0.132	0.026	0.067	0.014	0.071
4'-5'	0.002	0.072	0.012	0.052	0.009	0.063
4'-5''	-0.003	-0.102	-0.012	-0.051	-0.007	-0.063
5'-5''	0.000	0.000	-0.001	0.001	0.001	0.001
5'gamma	0.089	0.056	0.097	0.051	0.014	0.039
average MADD	**0.012**	**0.053**	**0.020**	**0.037**	**0.010**	**0.045**
						0.030

S3 NOE-distance determinations

S3.1 Calculation of NOE-derived distances from experimental NOE data

The experimental NOE-derived interproton distances (shown in Table 1, "NOE-derived distances, This Study" column) of SAM were determined by using equation 1 [3] and the 1D DPFSE NOESY spectra acquired in this report.

$$r_n^{exp} = r_{ref}[\frac{\eta_{NOE}}{\eta_{ref}}]^{-\frac{1}{6}} \quad (1)$$

Where r is interproton distance and η is NOE signal intensity derived from the NOESY spectrum. In order to compare NOE intensities across all 1D NOESY spectra and to correct for different rates of external relaxation for each proton in the molecule, the PANIC (Peak Amplitude Normalisation for Improved Cross-relaxation [4-6]) technique was employed by setting the integral of irradiated peak in each 1D NOESY spectrum to a single value (arbitrarily 1000 intensity units for this analysis) and taking each NOE signal intensity, η, from the resulting integrals for the corresponding peak in the spectrum.

Butts *et al.* [3],[7-8] have demonstrated that starting from a single known distance, r_{ref}, between two protons in the same molecule, the distances between other NOE-related protons in the rest of the molecule, r_{NOE}, can be accurately determined by comparing the PANIC-corrected relative intensity of the peak of interest (η_{NOE}) and the reference NOE (η_{ref}). For SAM the initial reference distance (r_{ref}) was chosen to be the distance between methine protons H1' and H2', however to maximise accuracy of the fit to the experimental data (and to allow the reference distance to have some reasonable deviation from ideality) the reference distance is then adjusted to give the minimum standard deviation between calculated and experimental NOE-distances across all of the measured data (resulting in a small, few %, error in the reference distance).

S3.2 Calculation of NOE-derived Interproton distances from MD data

[Note that these "NOE distances" should not be confused with real linearly time-averaged distances arising from MD simulations. Instead, the "NOE distances" are r^{-6} averaged and thus allow the direct comparison of the MD simulation results with the experimental NOE-distance data described in Section 3.1.]

1. Calculate mean NOE distances $r(NOE)_n^c$ *(1…**n**)* for each cluster *(1…**c**)*, (shown in Table 1 in columns labelled "Cluster 1" to "Cluster 6") consisting of ***l*** structures each with population *MDfract(l)* from MD simulation using Equation 2:

$$r(NOE)_n^c = \left(\sum_1^l \left(r_n^{l\,-6} * MDfract(l) \right)_n^l \right)^{-\left(\frac{1}{6}\right)} \tag{2}$$

2. The overall effective NOE-distance predicted *based on the MD populations* ($\langle r_n \rangle^{MD/NOE}$, shown in Table 1 in the main manuscript, right-hand most column) is calculated by weighting the averaged NOE distances $r(NOE)_n^c$ (again *via* r^{-6} averaging) with the population of each cluster derived directly from the MD simulations ($MDfract(c)$) using Equation 3:

$$\langle r_n \rangle^{MD/NOE} = \left(\sum_1^c \left(r(NOE)_n^c \right)^{-6} * MDfract(c) \right)^{-1/6} \tag{3}$$

3. The overall population averaged NOE distances *based on re-weighting the populations to give the best fit to the NOE experimental data* ($\langle r_n \rangle^{NOE}$, Table 1 in the main manuscript, second-to-last column) are found in a two-step process:

Firstly, new populations (*NOEfract(c)*) for clusters 1-6 are calculated that best fit the experimental NOE data. This is achieved by back-calculating relative ΔG values of each cluster (from $MDfract(c)$) and refining these free energies (and thus providing new populations, *NOEfract(c)*) in a GRG non-linear least squares optimisation of the standard deviation between experimental and calculated NOE-distances (using the Solver element in Microsoft Excel 2016). In order to ensure physical and chemical sense in the refinement, a constraint ($\Sigma\,(\Delta\Delta G)^2$ <400) was enforced and this prevents unreasonably large changes in ΔG (<20kJ/mol) away from the original MD-calculated populations/energies.

Secondly, the population-averaged NOE distances ($\langle r_n \rangle^{NOE}$) are then recalculated using the new cluster populations (*NOEfract(c)*) using Equation 4:

$$\langle r_n \rangle^{NOE} = \left(\sum_1^c \left(r(NOE)_n^c \right)^{-6} * NOEfract(c) \right)^{-1/6} \tag{4}$$

S3.3 Calculation of Mean Absolute Deviations (MADs) between NOE-derived distances from experimental NOE and MD data:

The mean absolute deviations remaining after NOE fitting MAD^{NOE} and directly from the MD data MAD^{MD} and experiment can then be calculated with Equation 5:

$$MAD = \left(\sum_{1}^{n}\left(\frac{\left|\langle r_n\rangle - r_n^{exp}\right|}{r_n^{exp} * 100}\right)\right) * n^{-1} \quad (5)$$

References

1. D. A. Saez and E. Vöhringer-Martinez, *J. Comput.-Aid. Mol. Design*, 2015, **29**, 951-961.
2. C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, 1973, **95**, 2333-2344.
3. C. R. Jones, C. P. Butts and J. N. Harvey, *Beilstein J. Org. Chem.*, 2011, **7**, 145.
4. H. Hu and K. Krishnamurthy, *J. Magn. Reson.,* 2006, **182**, 173–177.
5. S. Macura, B. T. Farmer and L. R. Brown, *J. Magn. Reson* 1986*,* **70**, 493–499.
6. E. Esposito and A. Pastore, J. *Magn. Reson.* 1988, **76**, 331–336.
7. C. P. Butts, C. R. Jones, E. C. Towers, J. L. Flynn, L. Appleby and N. J. Barron, *Org. Biomol. Chem*., 2011, **9**, 177-184.
8. C. P. Butts, C. R. Jones and J. N. Harvey, *Chem. Commun*, 2011, **47**, 1193-1195.