

Supporting Information

Degree of *N*-Methylation of Nucleosides and Metabolites Controls Binding Affinity to Pristine Silica Surfaces

Mouzhe Xie^{1,2*} and Rafael Bruschweiler^{1,3,4*}

¹ Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, USA

² Current address: Pritzker School of Molecular Engineering, the University of Chicago, Chicago, Illinois, 60637, USA

³ Campus Chemical Instrument Center, The Ohio State University, Columbus, Ohio 43210, USA

⁴ Department of Biological Chemistry and Pharmacology, The Ohio State University, Columbus, Ohio 43210, USA

*To whom correspondence should be addressed:

Dr. Mouzhe Xie, E-mail: xiemouzhe@uchicago.edu

Dr. Rafael Bruschweiler, E-mail: bruschweiler.1@osu.edu

NMR sample preparation

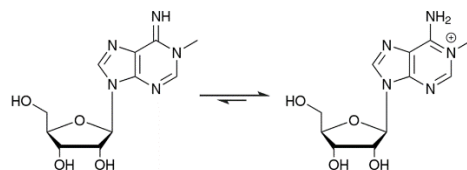
Glycine (GLY, 99.5% purity), sarcosine (SAR, 98%), *N,N*-dimethylglycine (DMG, 99%), and *N,N,N*-trimethylglycine (TMG, 98%), which were obtained from J.T. Baker (GLY) and Sigma-Aldrich (SAR, DMG, TMG), were individually dissolved in 20 mM sodium phosphate buffer (pH 7.0 ± 0.1 , also contained 10% D₂O as lock signal) to achieve a final concentration of 2.0 ± 0.1 mM. For samples with SNPs present, Bindzil 2040 colloidal silica nanoparticles was dialyzed for > 16 h into the same buffer and mixed directly with molecule solutions at a final 0.84 μ M concentration. These SNPs carry a negative surface charge (ζ -potential -23.0 ± 7.4 mV) and have a relatively narrow size distribution (19.5 ± 5.3 nm in diameter by transmission electron microscopy and 25.3 ± 8.9 nm by dynamic light scattering). Their particle concentration was estimated based on mass concentration (gravimetric analysis after evaporation) assuming 20 nm diameter, spherical morphology, monodispersity, and 2.196 g/cm³ amorphous silica density. Detailed characterization of these SNPs was previously reported.^[1]

Likewise, ribonucleosides and their derivatives were prepared in 60 mM sodium phosphate buffer (pH 7.0 ± 0.1 , 10% D₂O) at a final concentration of 12 ± 1 mM (except guanosine at 0.6 times of its saturation concentration at room temperature due to limited solubility estimated to be ~ 1.5 mM) either with or without the presence of 3.36 μ M SNPs. As an exception, a lower SNP concentration of 0.661 μ M was used for 1-methyladenosine as it experiences significantly stronger relaxation enhancement, and the ΔR_2 value was normalized with respect to 3.36 μ M SNPs for the comparison with other nucleosides.

The following list contains all compounds used in this study along with their CAS number, manufacturers, catalog numbers, and specified minimal purities:

Adenosine (CAS 58-61-7, Sigma-Aldrich Cat. A9251, 99%), guanosine (CAS 118-00-3, Sigma-Aldrich Cat. G6264, 98%), cytidine (CAS 65-46-3, Sigma-Aldrich Cat. C122106, 99%), uridine (CAS 58-96-8, Sigma-Aldrich Cat. U3003, 99%), inosine (CAS 58-63-9, Sigma-Aldrich Cat. I4125, 99%), 1-methyladenosine (CAS 15763-06-1, Cayman Chemical Cat. 16937, 98%), *N*⁶-methyladenosine (CAS 1867-73-8, Ark Pharm Cat. AK153126, 96%), *N*⁶,*N*⁶-dimethyladenosine (CAS 2620-62-4, Sigma-Aldrich Cat. COM448650878, 97%), 7-methylguanosine (CAS 20244-86-4, Cayman Chemical Cat. 15988, 98%), 3-methylcytidine (CAS 21028-20-6, Cayman Chemical Cat. 21064, 95%), *N*⁴-methylcytidine (CAS 10578-79-7, Ark Pharm Cat. AK120861, 98%), *N*³-methyluridine (CAS 2140-69-4, CarboSynth Cat. NM06185, 98%), and 5-methyluridine (CAS 1463-10-1, Ark Pharm Cat. AK88466, 97%), *N*⁶-methyladenosine-5'-monophosphate sodium salt (CAS 81921-35-9, Fisher Scientific Cat. 50-491-880, 98.5%), *N*⁶-methyl-2'-deoxyadenosine (CAS 2002-35-9, Sigma-Aldrich Cat. M2389, 98%), and 2'-deoxyadenosine (CAS 16373-93-6, Sigma-Aldrich, Cat. D7400, 99%).

For each compound, two 500 μL samples (with or without SNPs) were prepared and transferred to 5 mm NMR tubes for NMR measurements. In most cases, samples were stable for days and certainly for the duration of the NMR measurements, which typically lasted less than an hour, depending on the range of chemical shift dispersion of resonances. We noticed a shift of 1-methyladenosine from a neutral to a cationic isoform



induced by the presence of SNPs over the course of several hours. The two isoforms are in slow exchange as manifested by two distinct sets of NMR resonances. R_2 relaxation rates were only quantified for the predominant cationic isoform after equilibrium was reached.

^{13}C NMR spin relaxation

NMR data were collected at 298 K on a Bruker AVANCE HD spectrometer equipped with a cryogenic probe operating at 850 MHz ^1H resonance frequency (19.97 T). The effective R_2 relaxation rates for the molecules, either in the absence or presence of SNPs, were measured in a pseudo-2D fashion, where series of ^1H - ^{13}C HSQC-type spectra (without t_1 evolution along the ^{13}C dimension) were recorded at 8 relaxation delay intervals, which included 2 replicates. A proton-decoupled Carr-Purcell-Meiboom-Gill (PD-CPMG) pulse program (see Figure S1) was adapted from Yuwen & Skrynnikov^[2] to measure ^{13}C relaxation, with CPMG frequency $\nu_{\text{CPMG}} = 109$ Hz where the ^{13}C -radio-frequency carrier was set on-resonance for each ^{13}C resonance to be measured. The rf field strength was set to 3.125 kHz for DIPSI-2 decoupling. The delay intervals were individually adjusted to account for different relaxation rates of the various molecules. Generally, the longest delays were chosen so that the peak intensities decreased approximately 2.7-fold compared to the peak intensity at zero delay. For SNP-free samples, the longest relaxation delay was set to 737 ms. To extract R_2 -relaxation rates, the peak intensities quantified from series of 1D spectra for each resolved resonance were determined as a function of the relaxation time interval and fitted to a single exponential with zero offset. All NMR spin relaxation data were processed with NMRPipe. The 1D ^1H NMR data were visualized with TopSpin. Experimental uncertainties were estimated by error propagation from (i) intensity fitting against discrete spectral traces, (ii) replicates measured at the same relaxation intervals, and (iii) exponential R_2 fitting. More details can be found elsewhere.^[3,4]

Prediction of R_2^b of a tightly surface-anchored ^{13}C spin

The R_2 relaxation rate of a ^{13}C spin that is rigidly anchored on the SNP surface is dominated by magnetic dipole-dipole and chemical shielding anisotropy (CSA) relaxation. It follows Redfield theory according to (see, e.g., Gu et al.^[5] and references therein)

$$R_2 = \frac{1}{2} d_{\text{XH}} [4J(0) + 3J(\omega_{\text{X}}) + J(\omega_{\text{H}} - \omega_{\text{X}}) + 6J(\omega_{\text{H}}) + 6J(\omega_{\text{H}} + \omega_{\text{X}})] \\ + \frac{1}{6} c_{\text{XH}} \omega_{\text{X}}^2 [4J(0) + 3J(\omega_{\text{X}})] \quad (\text{S1})$$

where $d_{\text{CH}} = (1/20)(\mu_0/4\pi)^2(h/2\pi)^2[\gamma_{\text{C}}\gamma_{\text{H}}(r_{\text{CH}})^{-3}]^2$, and $c_{\text{CH}} = (1/15) \Delta\sigma^2$. μ_0 is the permeability of vacuum, h is Planck's constant, γ_{H} and γ_{C} are the gyromagnetic ratios of the corresponding nuclei, and the C–H bond length was set to $r_{\text{CH}} = 1.10 \text{ \AA}$. Note that in the case of a $^{13}\text{CH}_2$ methylene group, the dipole-dipole relaxation contribution to R_2 in Eq. (S1) needs to be multiplied by a factor 2. The chemical shielding anisotropy (CSA) $\Delta\sigma$ is around 25 ppm for a methylene ^{13}C spin, such as for glycine and its derivatives of Figure 1 (see, e.g., Table 1 of Havlin et al.^[6]). Although the CSA can vary between different sites, ^{13}C spin relaxation is almost entirely dominated by dipole-dipole interaction(s): at 850 MHz ^1H frequency, the CSA contribution to methylene carbons, such as those depicted in Figure 1, is of the order of a few percent of the total ^{13}C - R_2 . Therefore, the effect of the CSA term on the bound populations p_b extracted from ΔR_2 can be neglected in good approximation. $J(\omega)$ is the spectral density function of a ^{13}C - ^1H group rigidly attached to an isotropically tumbling nanoparticle with correlation time τ_c

$$J(\omega) = \frac{2\tau_c}{1 + \omega^2\tau_c^2} \quad (\text{S2})$$

sampled at angular frequency ω (in rad/s). τ_c can be estimated from the Stokes-Einstein-Debye relationship

$$\tau_c = \frac{V\eta}{k_{\text{B}}T} \quad (\text{S3})$$

where V is the hydrodynamic volume of a single nanoparticle, η is the shear viscosity of the solvent at absolute temperature T , and k_{B} is the Boltzmann constant. Using the parameters $r_{\text{SNP}} = 10 \text{ nm}$, $T = 298 \text{ K}$, and $\eta = 0.890 \text{ mPa}\cdot\text{s}$, for a ^{13}C - ^1H group we estimate $R_2^b \approx 3.8 \times 10^3 \text{ s}^{-1}$.

Conversion of R_2 relaxation rates to nanoparticle-bound populations

Changes in transverse R_2 relaxation rates upon the addition of nanoparticles are directly related to the bound population as follows^[3]

$$\Delta R_2 = p_b \cdot R_2^b \quad (\text{S4})$$

where p_b is the population of a molecule bound to the nanoparticle. The ratio between the ΔR_2 values at corresponding sites of molecules X and Y can then be directly related to the ratio of their bound populations

$$\frac{\Delta R_{2,X}}{\Delta R_{2,Y}} = \frac{p_{b,X}}{p_{b,Y}} \quad (\text{S5})$$

This relationship forms the basis for the conversion of NMR R_2 relaxation rates to nanoparticle binding constants and free energy changes in Eqs. (3) – (5) of the main text.

Table S1. pK_a values of GLY, SAR, DMG, and TMG. Under pH 7.0 experimental condition, these compounds are > 99.5% in their zwitterionic forms.^a

Compound	pK_a (amine H)	pK_a (–COOH)
Glycine (GLY)	9.60	2.34
Sarcosine (SAR)	10.2	2.12
<i>N,N</i> -Dimethylglycine (DMG)	9.94	2.15
<i>N,N,N</i> -Trimethylglycine (TMG)	not applicable	1.83

^a pK_a data taken from J. A. Dean *Lange's Handbook of Chemistry*

Table S2. ΔR_2 changes of transverse ^{13}C spin relaxation together with free energy changes $\Delta\Delta G$ upon *N*-methylation of metabolites and nucleosides studied here in their binding affinity to anionic silica nanoparticles.

Molecules <i>X/Y</i>	Mean ^{13}C - ΔR_2 (s^{-1}) of <i>X</i>	Mean ^{13}C - ΔR_2 (s^{-1}) of <i>Y</i>	$\Delta\Delta G^a$ (kJ/mol)	Note
SAR/GLY	0.76	0.16	-3.87	
DMG/SAR	2.47	0.76	-2.93	
TMG/DMG	6.28	2.47	-2.28	
$\text{m}^6\text{A/A}$	5.55	1.38	-3.45	
$\text{m}^6_2\text{A/m}^6\text{A}$	13.82	5.55	-2.26	
$\text{m}^6\text{dA/dA}$	18.32	4.44	-3.51	
$\text{m}^4\text{C/C}$	4.44	1.29	-3.06	
$\text{m}^3\text{U/U}$	0.30	0.09	-2.98	
Average \pm std.			-3.04 \pm 0.57	
$\text{m}^5\text{U/U}$	0.53	0.09	-4.39	C-methylation
$\text{m}^7\text{G/G}$	11.71	0.30	-9.08	+1 charge
$\text{m}^3\text{C/C}$	15.68	1.29	-6.19	+1 charge
$\text{m}^1\text{A/A}$	41.46	1.38	-8.43	+1 charge

^a $\Delta\Delta G$ values were obtained according to Eq. (5) as discussed in the main text.

Table S3. Mean $^{13}\text{C}-\Delta R_2$ (s^{-1}) values averaged over all ^{13}C atoms or ribose and nucleobase subgroups of adenosine without and with N^6 -methylation.

Adenosine	Total mean $^{13}\text{C}-\Delta R_2$ (s^{-1})	Mean ribose $^{13}\text{C}-\Delta R_2$ (s^{-1})	Mean base $^{13}\text{C}-\Delta R_2$ (s^{-1})
A	1.38	1.00	2.14
m^6A	5.55	3.78	9.11
m^6_2A	13.83	8.74	24.00

Table S4. Entire experimental ^{13}C - R_2 data of nucleosides. For a few sites quantitative relaxation data could not be obtained due to resonance overlap (OL) or relaxation that is too fast (TF).

Adenosine (A)

Carbon site	R_2 (s^{-1}) without SNPs	R_2 (s^{-1}) with SNPs	ΔR_2 (s^{-1})
P2	1.31	3.52	2.21
P8	1.64	3.71	2.07
S1	1.47	2.62	1.15
S3	1.68	2.61	0.93
S4	1.43	2.32	0.89
S5	2.87	3.92	1.05

Guanosine (G)

Carbon site	R_2 (s^{-1}) without SNPs	R_2 (s^{-1}) with SNPs	ΔR_2 (s^{-1})
P8	1.40	1.86	0.46
S1	1.50	1.57	0.08
S3	1.47	2.62	1.15
S4	1.68	2.61	0.93
S5	1.43	2.32	0.89

Uridine (U)

Carbon site	R_2 (s^{-1}) without SNPs	R_2 (s^{-1}) with SNPs	ΔR_2 (s^{-1})
P5	OL	OL	--
P6	1.57	1.92	0.35
S1	OL	OL	--
S2	1.42	1.39	-0.03
S3	1.24	1.32	0.07
S4	1.25	1.46	0.22
S5	2.27	2.13	-0.14

Cytidine (C)

Carbon site	R_2 (s^{-1}) without SNPs	R_2 (s^{-1}) with SNPs	ΔR_2 (s^{-1})
P5	1.29	3.51	2.22
P6	1.64	4.18	2.54
S1	1.40	2.38	0.98
S2	1.58	2.21	0.64
S3	1.35	2.07	0.72
S4	1.10	2.09	1.00
S5	2.35	3.28	0.93

Inosine (I)

Carbon site	R_2 (s⁻¹) without SNPs	R_2 (s⁻¹) with SNPs	ΔR_2 (s⁻¹)
P2	2.01	2.77	0.77
P8	1.84	2.31	0.47
S1	1.42	1.73	0.31
S3	1.40	1.51	0.11
S4	1.17	1.48	0.31
S5	2.58	2.74	0.16

***N*⁶-Methyladenosine (m⁶A)**

Carbon site	R_2 (s⁻¹) without SNPs	R_2 (s⁻¹) with SNPs	ΔR_2 (s⁻¹)
P2	12.16	21.77	9.61
P8	2.12	10.73	8.61
S1	2.84	7.29	4.45
S3	2.16	5.39	3.23
S4	2.10	5.51	3.41
S5	3.36	7.38	4.02
Met	172	TF	--

***N*⁶,*N*⁶-Dimethyladenosine (m⁶₂A)**

Carbon site	R_2 (s⁻¹) without SNPs	R_2 (s⁻¹) with SNPs	ΔR_2 (s⁻¹)
P2	2.14	26.75	24.61
P8	2.35	25.74	23.39
S1	2.02	12.48	10.46
S3	1.94	8.84	6.90
S4	1.95	10.37	8.42
S5	3.14	12.32	9.18
Met	TF	TF	--

***N*¹-Methyladenosine (m¹A), cationic isoform**

Carbon site	R_2 (s⁻¹) without SNPs	R_2 (s⁻¹) with SNPs	ΔR_2 (s⁻¹) *Normalized
P2	2.25	19.76	70.07
P8	7.33	21.28	55.79
S1	1.53	9.59	32.22
S3	1.05	9.65	34.37
S4	1.19	8.80	30.44
S5	1.94	8.40	25.84
Met	TF	TF	--

***N*⁷-Methylguanosine (m⁷G)**

Carbon site	<i>R</i>₂ (s⁻¹) without SNPs	<i>R</i>₂ (s⁻¹) with SNPs	Δ<i>R</i>₂ (s⁻¹)
P8	3.87	22.93	19.07
S1	1.33	12.69	11.36
S2	1.40	8.84	7.44
S3	1.33	10.76	9.44
S4	1.31	11.72	10.41
S5	2.29	16.92	14.63
Met	1.07	10.69	9.62

5-Methyluridine (m⁵U)

Carbon site	<i>R</i>₂ (s⁻¹) without SNPs	<i>R</i>₂ (s⁻¹) with SNPs	Δ<i>R</i>₂ (s⁻¹)
P6	1.62	2.89	1.27
S1	1.36	2.13	0.77
S2	1.52	1.76	0.24
S3	1.33	1.64	0.31
S4	1.19	1.70	0.51
S5	2.39	2.58	0.18
Met	0.52	0.93	0.41

***N*³-Methyluridine (m³U)**

Carbon site	<i>R</i>₂ (s⁻¹) without SNPs	<i>R</i>₂ (s⁻¹) with SNPs	Δ<i>R</i>₂ (s⁻¹)
P5	1.02	1.59	0.57
P6	1.62	2.33	0.71
S1	1.26	1.64	0.38
S2	1.43	1.46	0.03
S3	1.26	1.37	0.11
S4	1.06	1.32	0.26
S5	2.20	2.37	0.17
Met	0.78	0.98	0.21

***N*⁴-Methylcytidine (m⁴C)**

Carbon site	<i>R</i>₂ (s⁻¹) without SNPs	<i>R</i>₂ (s⁻¹) with SNPs	Δ<i>R</i>₂ (s⁻¹)
P5	3.46	9.88	6.42
P6	4.57	12.12	7.54
S1	2.62	6.59	3.97
S2	2.00	4.86	2.87
S3	1.63	4.60	2.97
S4	1.28	5.38	4.09
S5	2.63	6.35	3.72
Met	3.20	7.14	3.94

***N*³-Methylcytidine (m³C)**

Carbon site	<i>R</i>₂ (s⁻¹) without SNPs	<i>R</i>₂ (s⁻¹) with SNPs	ΔR_2 (s⁻¹)
P5	1.96	28.39	26.43
P6	2.72	32.01	29.29
S1	1.47	13.29	11.82
S2	1.41	10.61	9.20
S3	1.22	11.11	9.89
S4	1.22	12.97	11.75
S5	2.10	15.24	13.14
Met	0.87	14.77	13.90

2'-Deoxyadenosine (dA)

Carbon site	<i>R</i>₂ (s⁻¹) without SNPs	<i>R</i>₂ (s⁻¹) with SNPs	ΔR_2 (s⁻¹)
P2	1.24	8.23	6.98
P8	1.72	7.96	6.25
S1	1.76	5.63	3.87
S2	1.46	6.94	5.48
S3	1.33	3.90	2.57
S4	1.38	4.69	3.30
S5	2.67	5.31	2.63

***N*⁶-Methyl-2'deoxyadenosine (m⁶dA)**

Carbon site	<i>R</i>₂ (s⁻¹) without SNPs	<i>R</i>₂ (s⁻¹) with SNPs	ΔR_2 (s⁻¹)
P2	10.91	41.25	30.34
P8	2.02	30.25	28.23
S1	3.57	18.91	15.34
S2	2.44	20.67	18.23
S3	1.61	15.42	13.81
S4	1.65	13.83	12.18
S5	2.95	13.04	10.09
Met	TF	TF	--

***N*⁶-Methyladenosine monophosphate (m⁶AMP)**

Carbon site	<i>R</i>₂ (s⁻¹) without SNPs	<i>R</i>₂ (s⁻¹) with SNPs	ΔR_2 (s⁻¹)
P2	11.88	13.31	1.43
P8	3.35	5.83	2.49
S1	2.22	3.55	1.33
S3	2.44	3.22	0.77
S4	2.17	3.48	1.30
S5	2.73	4.24	1.51

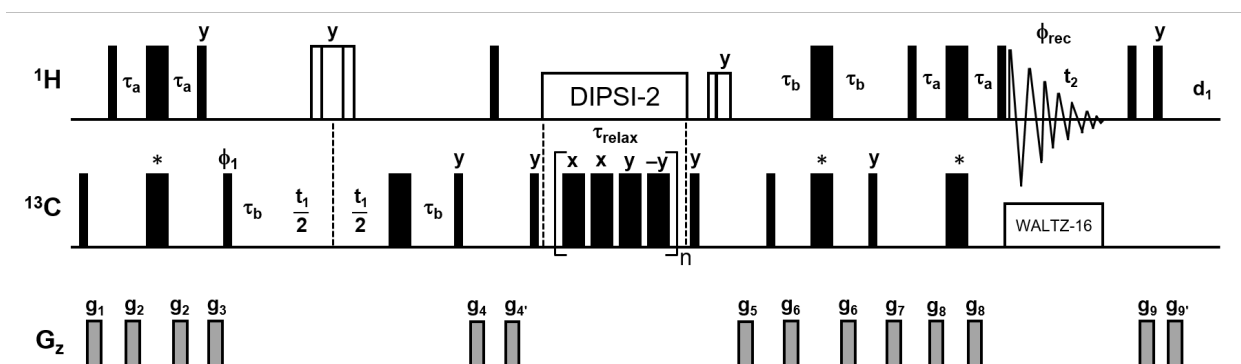


Figure S1. NMR pulse sequence used for ^{13}C - R_2 relaxation experiments used in this work. The sequence was adapted from the ^{15}N - R_2 relaxation experiment of Ref.^[2] Delays τ_a and τ_b were chosen to select for the desired spin system: CH ($\tau_a = 1.72$ ms, $\tau_b = 1.85$ ms), CH_2 ($\tau_a = 1.85$ ms, $\tau_b = 925$ μs), CH_3 ($\tau_a = 2$ ms, $\tau_b = 784$ μs), and aromatic CH ($\tau_a = \tau_b = 1.35$ ms). A minimal phase cycle is required with $\phi_1 = (x, -x)$ and $\phi_{\text{rec}} = (x, -x)$, while quadrature detection is achieved by States-TPPI phase cycling of ϕ_1 . Axial peaks are shifted to the edges of the spectrum by inverting ϕ_1 and ϕ_{rec} for each t_1 complex point. The open ^1H pulse shape during t_1 is a composite 90_x - 180_y - 90_x pulse. The open ^1H pulse shape after the DIPSI-2 element are 2 ms and 3 ms purge pulses applied at the DIPSI-2 power level. ^{13}C pulses with an asterisk (*) are adjustable band-selective pulses with a REBURP shape. In this work, a pulse width of 300 μs was used. Gradient strengths in G/cm (lengths in ms) are $g_1 = 21.0$ (0.5), $g_2 = 11.2$ (0.3), $g_3 = -32.9$ (1.0), $g_4 = -5.8$ (1.0), $g_{4'} = -6.8$ (1.0), $g_5 = 39.4$ (0.5), $g_6 = 17.7$ (0.3), $g_7 = 26.3$ (0.8), $g_8 = -7.3$ (0.3), $g_9 = 5.3$ (0.5), $g_{9'} = 10.5$ (0.5).

Additional references

- [1] B. Zhang, M. Xie, L. Bruschweiler-Li, K. Bingol, R. Bruschweiler, Use of Charged Nanoparticles in NMR-Based Metabolomics for Spectral Simplification and Improved Metabolite Identification. *Anal. Chem.* **2015**, *87*, 7211-7217.
- [2] T. Yuwen, N. R. Skrynnikov, Proton-Decoupled CPMG: a Better Experiment for Measuring ^{15}N R_2 Relaxation in Disordered Proteins. *J. Magn. Reson.* **2014**, *241*, 155-169.
- [3] M. Xie, A. L. Hansen, J. Yuan, R. Bruschweiler, Residue-Specific Interactions of an Intrinsically Disordered Protein with Silica Nanoparticles and Their Quantitative Prediction. *J. Phys. Chem. C* **2016**, *120*, 24463-24468.
- [4] M. Xie, Probing and Modeling Biomolecule-Nanoparticle Interactions by Solution Nuclear Magnetic Resonance Spectroscopy. Ph.D. Dissertation, The Ohio State University, **2018**.
- [5] Y. Gu, D.-W. Li, R. Bruschweiler, NMR Order Parameter Determination from Long Molecular Dynamics Trajectories for Objective Comparison with Experiment. *J. Chem. Theory Comput.* **2014**, *10*, 2599-2607.
- [6] R. H. Havlin, H. Le, D. D. Laws, A.C. deDios, E. Oldfield, An *ab Initio* Quantum Chemical Investigation of Carbon-13 NMR Shielding Tensors in Glycine, Alanine, Valine, Isoleucine, Serine, and Threonine: Comparisons between Helical and Sheet Tensors, and the Effects of χ_1 on Shielding. *J. Am. Chem. Soc.* **1997**, *119*, 11951-11958.