Supporting Information: “Development and Validation of a DFT Based Force Field for a Hydrated Homoalanine Poly-peptide”

Ying Yuan,¹ Zhonghua Ma,¹ and Feng Wang¹,*

¹Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, Arkansas, 72701, USA
1. fitting of dispersion parameters

Dispersion parameters were fitted to Grimme’s D3 dispersion forces using three types of fragments as described in the text. The three types of fragments are summarized in Figure S1. When creating these fragments, the covalent bonds were cut. These fragments are thus radicals and won’t be appropriate for a symmetry adapted perturbation theory (SAPT) based fitting. However, the use of radicals won’t lead to problems with empirical D3 based fitting.

![Figure S1. Fragments used for D3 calculations. (a) N terminus, (b) Ala, (c) C terminus.](image)

As described in the text, the SRD dispersion is used to eliminate the divergent behavior at short range. The van der Waals (vdW) radii used for SRD is shown in Table S1. Figure S1 provides a comparison of the SRD dispersion with the popular Tang-Tonnies dispersion. The SRD offers stronger damping in the short-range and also lead to smaller root mean square difference (RMSD) of the fit.

Table S1. The vdW radius for each atom type of Ala.

| Type | Rvdw (Bohr) | hybridization |
|------|-------------|---------------|
| C1   | 3.22        | sp$^3$        |
| C2   | 3.22        | sp$^3$        |
| C3   | 3.34        | sp$^2$        |
| C4   | 3.22        | sp$^3$        |
| C5   | 3.22        | sp$^3$        |
| C6   | 3.34        | sp$^2$        |
| N1   | 3.18        | sp$^3$        |
| N2   | 3.18        | sp$^2$        |
| O1   | 3.09        | sp$^2$        |
| O2   | 3.09        | sp$^2$        |
Figure S2. A comparison of dispersion curves modeled with the SRD and Tang-Tonnies damping. The undamped form is also shown for comparison. The undamped dispersion used for this figure is -1847 kcal/mol at 1 Å. The sum of van der Waals radii of the two atoms are 3.41 Å. The dispersion energy at 3.41 Å was damped by 1% for Tang-Tonnies and 4% for SRD damping compared to the undamped form.

2. force field parameters

In summary, there are a total of 455 parameters for the AFM2020 zwitterionic alanine model. 74 of these are intermolecular parameters, 381 of these are intramolecular parameters. Out of the intramolecular parameters, 87 of these are torsional parameters with each torsional term having at least three parameters.

2.1 all force field parameters for the hydrated Ala$_7$.

| Type | Partial-charge(e) |
|------|-------------------|
| H1   | 0.15681           |
| H2   | 0.40227           |
| H3   | 0.55263           |
| H4   | 0.14652           |
| O1   | -0.52345          |
| O2   | -0.69732          |
| C1   | -0.12120          |
| C2   | -0.47844          |
| C3   | 0.44795           |
| C4   | 0.00385           |
| C5   | 0.26460           |
| C6   | 0.67818           |
| N1   | -1.22751          |
| N2   | -0.47713          |
nonbonded parameters

ENERGY EXPRESSION:

\[ \text{EXP}=A \exp(-B \cdot r) \]

\[ \text{SRD}=C_6/(r^6+R_0^6) \]

\[ \text{POW}=A/r^{12} \]

| Ala-Ala repulsion and dispersion parameters |
|-------------------------------------------|
| atom1 | atom2 | Repulsion A (kcal/mol) | B (Å⁻¹) | Dispersion C⁶ (kcal/(mol Å⁶)) | R⁰ (Å) |
|-------|-------|------------------------|--------|-------------------------------|-------|
| H₄    | H₄    | EXP                    | 385.136| 3.381                         |       |
| H₄    | H₁    | EXP                    | 8176.257| 4.483                         |       |
| H₄    | O₂    | EXP                    | 8777.945| 3.500                         |       |
| H₄    | O₁    | EXP                    | 918.488 | 2.998                         |       |
| H₄    | C₁    | EXP                    | 8857.687| 3.500                         |       |
| H₄    | C₂    | EXP                    | 3100.462| 3.400                         |       |
| H₄    | C₃    | EXP                    | 2961.385| 3.500                         |       |
| H₄    | C₄    | EXP                    | 926.765 | 3.500                         |       |
| H₄    | N₂    | EXP                    | 12841.598| 3.500                         |       |
| H₁    | H₁    | EXP                    | 761.310 | 3.738                         |       |
| H₁    | O₂    | EXP                    | 5276.153| 4.134                         |       |
| H₁    | O₁    | EXP                    | 7249.907| 4.015                         |       |
| H₁    | C₆    | EXP                    | 5355.539| 3.500                         |       |

| C₂    | O₂    | EXP                    | 6757.217| 3.500                         |       |
| C₂    | O₁    | EXP                    | 1766.806| 3.400                         |       |
| C₂    | C₃    | EXP                    | 3768.702| 3.300                         |       |
| C₂    | N₂    | EXP                    | 10473.568| 3.500                         |       |
| C₂    | N₁    | EXP                    | 11770.635| 3.500                         |       |
| C₃    | O₂    | EXP                    | 6893.693| 3.419                         |       |
| C₃    | O₁    | EXP                    | 399.259 | 2.416                         |       |
| C₃    | C₂    | EXP                    | 4314.027| 3.300                         |       |
| C₃    | C₃    | EXP                    | 2251.230| 3.500                         |       |
| C₃    | C₄    | EXP                    | 8283.029| 3.500                         |       |
| C₃    | C₅    | EXP                    | 1492.781| 3.500                         |       |
| C₃    | N₂    | EXP                    | 11770.635| 3.500                         |       |
| C₃    | N₁    | EXP                    | 10473.568| 3.500                         |       |
| C₃    | N₂    | EXP                    | 369.527 | 3.085                         |       |
| C₄    | O₂    | EXP                    | 197.956 | 3.536                         |       |
| C₄    | O₁    | EXP                    | 9281.277| 4.100                         |       |
| C₄    | C₂    | EXP                    | 5355.539| 3.500                         |       |
| C₄    | C₃    | EXP                    | 12035.240| 3.500                         |       |
| C₄    | O₁    | EXP                    | 1051.128| 3.600                         |       |
| C₄    | C₆    | EXP                    | 27773.501| 3.600                         |       |
| C₄    | C₁    | EXP                    | 31471.540| 3.700                         |       |
| C₄    | C₂    | EXP                    | 186737.415| 3.600                         |       |
| C₄    | C₃    | EXP                    | 20668.664| 3.500                         |       |
| C₄    | C₅    | EXP                    | 51155.421| 3.600                         |       |
| C₄    | N₁    | EXP                    | 197705.286| 3.600                         |       |
| C₅    | O₂    | EXP                    | 411853.138| 3.600                         |       |
| C₅    | O₁    | EXP                    | 29105.048| 3.600                         |       |
| C₅    | C₂    | EXP                    | 1051.128| 3.600                         |       |
| C₅    | C₃    | EXP                    | 27773.501| 3.600                         |       |
| C₅    | C₄    | EXP                    | 31471.540| 3.700                         |       |
| C₅    | C₅    | EXP                    | 186737.415| 3.600                         |       |
| C₅    | N₂    | EXP                    | 20668.664| 3.500                         |       |
| C₅    | N₁    | EXP                    | 51155.421| 3.600                         |       |
| C₅    | N₁    | EXP                    | 20668.664| 3.500                         |       |
| C₅    | N₂    | EXP                    | 411853.138| 3.600                         |       |
| C₅    | N₃    | EXP                    | 29105.048| 3.600                         |       |
| C₅    | N₄    | EXP                    | 1051.128| 3.600                         |       |
| C₅    | N₅    | EXP                    | 27773.501| 3.600                         |       |
| C₅    | N₆    | EXP                    | 31471.540| 3.700                         |       |
| C₅    | N₇    | EXP                    | 186737.415| 3.600                         |       |
| C₅    | N₈    | EXP                    | 20668.664| 3.500                         |       |
| C₅    | N₉    | EXP                    | 51155.421| 3.600                         |       |
| C₅    | N₁₀   | EXP                    | 20668.664| 3.500                         |       |
### Ala-water repulsion and dispersion parameters

| atom1 | atom2 | Repulsion A(kcal/mol Å²) | Dispersion C6(kcal/(mol Å⁶)) | R0(Å) |
|--------|-------|---------------------------|-------------------------------|-------|
| HW O2  | EXP   | 1083.973                  | 3.314                         |       |
| HW O1  | EXP   | 1719.978                  | 3.792                         |       |
| HW C1  | EXP   | 142404.189                | 5.137                         |       |
| HW C2  | EXP   | 2651.640                  | 2.980                         |       |
| HW N2  | EXP   | 588.854                   | 2.536                         |       |
| HW N1  | EXP   | 15304.476                 | 3.322                         |       |
| OW H3  | EXP   | 1168.238                  | 3.503                         |       |
| OW H2  | EXP   | 376.060                   | 2.912                         |       |
| OW H1  | EXP   | 5015.246                  | 3.674                         |       |
| OW H4  | EXP   | 2758.441                  | 3.239                         |       |
| OW N1  | EXP   | 425637.178                | 4.323                         | -1038.686 | 1.984 |
| OW N2  | EXP   | 825262.448                | 4.493                         | -487.852  | 1.984 |
| OW C5  | EXP   | 3622219.890               | 5.363                         | -609.539  | 1.997 |
| OW C4  | EXP   | 2022922.681               | 5.081                         | -867.488  | 1.997 |
| OW C3  | EXP   | 84291.773                 | 3.795                         | -438.930  | 2.035 |
| OW C2  | EXP   | 62333.711                 | 3.494                         | -1124.531 | 1.997 |
| OW C1  | EXP   | 427188.493                | 4.894                         | -913.293  | 1.997 |
| OW C6  | EXP   | 296188.378                | 4.009                         | -586.588  | 2.035 |
| OW O1  | EXP   | 185032.327                | 4.199                         | -333.497  | 1.956 |
| OW O2  | EXP   | 548862.558                | 4.678                         | -306.997  | 1.956 |

### bonded parameters:

Gromacs unit, can be used for Gromacs topology file directly.

| Bonded parameters: |
|--------------------|
| U=k/2*(r-r0)^2 |
| r0(nm) |
| k([kJ/(mol nm^2)]) |

| C6 O2  | 0.1271 | 476083.900072 |
| C2 H1  | 0.1108 | 293563.217912 |
| C4 H4  | 0.1103 | 285821.935088 |
| C3 N2  | 0.1351 | 355163.2332   |
| C4 N2  | 0.1446 | 251766.043248 |
| C1 N2  | 0.1443 | 286362.197816 |
| C3 O1  | 0.1248 | 568275.181656 |
| C5 H4  | 0.1898 | 292236.187504 |
| H3 N1  | 0.1839 | 386536.245464 |
| C1 C2  | 0.1589 | 198699.78088  |
| C3 C5  | 0.1444 | 160546.887368 |
| C3 C4  | 0.148  | 157980.338088 |
| H2 N2  | 0.1027 | 369963.7148   |
| C2 C4  | 0.1519 | 195899.369432 |
| C2 C5  | 0.1536 | 281958.7512   |
| C6 C1  | 0.1482 | 145147.805072 |
| C5 N1  | 0.1471 | 187137.818208 |
| C1 H4  | 0.1184 | 287718.38748  |
### Angles

\[ U = k/2*(\theta - \theta_0)^2 \]

| Atom 1 | Atom 2 | Atom 3 | U   | k (kJ/(mol rad^2)) |
|--------|--------|--------|-----|-------------------|
| C6     | C1     | N2     | 99.43 | 388.466744       |
| C4     | C2     | H1     | 104.798 | 355.267624       |
| C4     | C3     | O1     | 147.292 | 442.428344       |
| C2     | C4     | C3     | 107.624 | 338.242928       |
| C5     | C3     | O1     | 145.813 | 437.294944       |
| C2     | C5     | N1     | 115.126 | 685.807808       |
| H3     | N1     | H3     | 110.5  | 302.013672       |
| C4     | N2     | H2     | 109.661 | 250.395664       |
| C2     | C4     | H4     | 108.113 | 283.84256        |
| H4     | C5     | N1     | 103.271 | 399.53016        |
| C5     | N1     | H3     | 107.624 | 338.242928       |
| O2     | C6     | O2     | 152.789 | 704.50192        |
| C3     | C5     | N1     | 106.5  | 456.22336        |
| C4     | C3     | N2     | 109.661 | 250.395664       |
| C6     | C1     | H4     | 105.49 | 317.327112       |
| H4     | C1     | N2     | 104.427 | 399.078288       |
| C5     | C3     | N2     | 142.14 | 318.113704       |
| C2     | C5     | H2     | 109.121 | 320.963008       |
| C3     | C5     | H2     | 108.371 | 328.95026        |
| C1     | C2     | H1     | 104.113 | 283.84256        |

### Dihedral

\[ U = A*(1+\cos(n*\phi - \delta)) \]

| Atom 1 | Atom 2 | Atom 3 | Atom 4 | U   | A (kJ/mol) | n |
|--------|--------|--------|--------|-----|------------|---|
| C2     | C1     | N2     | C3     | 0   | 1.787072   | 3 |
| C2     | C1     | N2     | C3     | 0   | 1.979032   | 2 |
| C2     | C1     | N2     | C3     | 0   | 5.317864   | 1 |
| C3     | C4     | N2     | C3     | 0   | 1.313776   | 3 |
| C3     | C4     | N2     | C3     | 0   | 1.937912   | 2 |
| C4     | C3     | N2     | C3     | 0   | 1.280084   | 1 |
| N2     | C3     | C5     | N1     | 0   | 0.225936   | 3 |
| N2     | C3     | C5     | N1     | 180 | 3.631712   | 2 |
| N2     | C3     | C5     | N1     | 0   | 5.669320   | 1 |
| N2     | C3     | C4     | N2     | 180 | 0.774800   | 3 |
| N2     | C3     | C4     | N2     | 180 | 3.978984   | 2 |
| N2     | C3     | C4     | N2     | 0   | 3.438880   | 1 |
| C4     | C3     | N2     | C1     | 180 | 17.593720  | 2 |
| C5     | C3     | N2     | C4     | 180 | 16.003800  | 2 |
| C4     | C3     | N2     | C4     | 180 | 20.137592  | 2 |
| C4     | C3     | N2     | H2     | 180 | 7.928680   | 2 |
| C5     | C3     | N2     | H2     | 180 | 8.146248   | 2 |
| O1     | C3     | N2     | H2     | 180 | 18.786160  | 2 |
| O1     | C3     | N2     | C4     | 180 | 15.874592  | 2 |
| O1     | C3     | N2     | C4     | 180 | 15.568664  | 2 |
| O1     | C3     | C4     | H4     | -11.018 | 1.121312  | 3 |
| N2     | C3     | C4     | C2     | 126.331 | 0.192464  | 3 |
| C2     | C4     | N2     | C3     | 44.428 | 2.694496   | 3 |
| N2     | C3     | C5     | C2     | 13.07  | 0.719648   | 3 |
| C2     | C1     | N2     | C3     | 47.316 | 2.606632   | 3 |
| H1     | C2     | C4     | C3     | 24.673 | 0.552288   | 3 |
| H1     | C2     | C5     | C3     | 23.642 | 0.443584   | 3 |
2.2 Ace-(Ala)$_n$-NMe parameters

Gas phase simulations were performed with Ace-(Ala)$_n$-NMe to avoid the unphysical attractions between charged termini. The atom types and parameters for Ace and NMe are borrowed from the original AFM2020 except that the charges on CH$_3$ were slightly adjusted to make Ace and NMe neutral. The adjusted charges and the atom types are summarized in Figure S2.

![Figure S3. Ace-Ala-NMe atom types and the adjusted charges on CH$_3$ groups.](image)

2.3 GROMACS input files for simulations

The gromacs input files for hydrated zwitterionic Ala$_7$ and for Ace-(Ala)$_7$-NMe in vacuum can be found at https://wanglab.uark.edu/models/Alanine.htm

3. Computation of J-coupling constants

The J-coupling constants were calculated with the Karplus equation$^2$

\[ J(\phi) = A \cos^2(\phi + \theta) + B \cos(\phi + \theta) + C \]  

(S1)

where $\phi$ is the backbone dihedral angle. The coefficients $A$, $B$, $C$ and offset angle $\theta$ are shown in Table
S2 along with the estimated errors $\sigma_i$ used to calculate $\chi^2$.

The following equation was used to calculate $^3J(H_N,C_\alpha)^3$

$$^3J_{H_N,C_\alpha} (\phi, \psi_{i-1}) = -0.23 \cos \phi - 0.20 \cos \psi_{i-1} + 0.07 \sin \phi + 0.08 \sin \psi_{i-1}$$

$$+ 0.07 \cos \phi \cos \psi_{i-1} + 0.12 \cos \phi \sin \psi_{i-1}$$

$$- 0.08 \sin \phi \cos \psi_{i-1} - 0.14 \sin \phi \sin \psi_{i-1} + 0.54$$  \hspace{1cm} (S2)

Table S2. Parameters for Karplus equations used in this work.$^{3-6}$

| J-coupling       | $\phi$ | $A$ (Hz) | $B$ (Hz) | $C$ (Hz) | $\theta$ (°) | $\sigma$ (Hz) |
|------------------|--------|----------|----------|----------|--------------|--------------|
| $^3J(H_N,H_\alpha)$ | $\phi_i$ | 7.09     | -1.42    | 1.55     | -60          | 0.70         |
| $^3J(H_N,C')$    | $\phi_i$ | 4.29     | -1.01    | 0.00     | 180          | 0.45         |
| $^3J(H_\alpha,C')$ | $\phi_i$ | 3.72     | -2.18    | 1.28     | 120          | 0.29         |
| $^3J(H_N,C_\beta)$ | $\phi_i$ | 3.06     | -0.74    | 0.13     | 60           | 0.30         |
| $^1J(N,C_\alpha)$ | $\psi_i$ | 1.70     | -0.98    | 9.51     | 0            | 0.59         |
| $^2J(N,C_\alpha)$ | $\psi_{i-1}$ | -0.66   | -1.52    | 7.85     | 0            | 0.50         |
| $^3J(H_N,C_\alpha)$ | $\phi_{i,\psi_{i-1}}$ |          |          |          | Eq. S2       | 0.10         |
4. J-coupling constants of each residue of Zwitterionic Ala$_3$ and Ala$_5$

Detailed analysis of the J-coupling constants for zwitterionic Ala$_3$ and Ala$_5$ at 300 K from simulations with the AFM2020 force field. Experimental numbers$^7$ are also shown for convenience.

Table S3. J-coupling constants for Ala$_3$ and Ala$_5$ in water at 300 K calculated from simulations with and experiments.

| J-coupling constants/Hz | Ala3     | Ala5     |
|------------------------|---------|---------|
|                        | AFM2020 | Exp     | AFM2020 | Exp     |
| residue                | type    |         |         |
| A2                     | $^3J$(HN,H$_\alpha$) ($\phi_2$) | 5.87    | 5.68    | 5.84    | 5.59    |
|                        | $^3J$(HN,C') ($\phi_2$)       | 1.04    | 1.13    | 1.05    | 1.13    |
|                        | $^3J$(HN,C$\beta$) ($\psi_2$) | 2.02    | 2.39    | 2.03    | 2.30    |
|                        | $^1J$(N,C$_\alpha$) ($\psi_2$) | 11.26   | 11.34   | 11.41   | 11.36   |
|                        | $^3J$(HN,H$_\alpha$) ($\phi_3$) |         | 6.02    | 5.74    |
|                        | $^3J$(HN,C') ($\phi_3$)       | 1.03    | N/A     |
|                        | $^3J$(HN,C$\alpha$) ($\phi_3$) | 1.65    | 1.86    |
| A3                     | $^3J$(HN,C$\beta$) ($\phi_3$) | 1.97    | 2.24    |
|                        | $^1J$(N,C$_\alpha$) ($\phi_3$) | 11.34   | 11.26   |
|                        | $^2J$(N,C$_\alpha$) ($\psi_2$) | 8.38    | 8.55    |
|                        | $^3J$(HN,C$_\alpha$)($\phi_3$, $\psi_2$) | 0.45    | 0.68    |
|                        | $^3J$(HN,H$_\alpha$) ($\phi_4$) | 5.94    | 5.98    |
|                        | $^3J$(HN,C$\beta$) ($\phi_4$) | 1.07    | 1.15    |
|                        | $^3J$(HN,C') ($\phi_4$)       | 1.64    | 1.89    |
| A4                     | $^3J$(HN,C$\beta$) ($\phi_4$) | 1.97    | 2.14    |
|                        | $^1J$(N,C$_\alpha$) ($\psi_4$) | 10.82   | 11.25   |
|                        | $^2J$(N,C$_\alpha$) ($\psi_3$) | 8.38    | 8.40    |
|                        | $^3J$(HN,C$_\alpha$)($\phi_4$, $\psi_3$) | 0.46    | 0.69    |

N/A: data are not available.
5. convergence of simulations

In this work, all the peptides in water were simulated for 1 µs at 300 K. To check the convergence, running average of $\chi^2$ as a function of simulation time for two 500 ns trajectories are shown for Ala$_7$. (Figure S5) The results show that the properties of interest converge around 50 ns. Also shown is the relative free energy based on two separate 500 ns trajectories (Figure S6) for Ala$_7$. For the more important regions of the conformation space, the free energy converges well within 500 ns.

Figure S5. Running average of the $\chi^2$ for two independent trajectories calculated from simulations of Ala$_7$ with AFM2020.

Figure S6. Relative free energies of Ala$_7$ calculated from two 500 ns simulation trajectories with AFM2020.
6. reference conformation and gradient data

A ref file is provided that contains conformation and gradient data for adaptive force matching. The file can be used by the CRYOFF code available from https://wanglab.uark.edu/CRYOFF.

This file can be used to refit the AFM2020. This is useful for further refinement of the potential with additional conformations in the training set.

Due to its large size, the file can be found at https://wanglab.uark.edu/models/Alanine.htm
References

(1) Anatole von Lilienfeld, O.; Tkatchenko, A. Two- and Three-Body Interatomic Dispersion Energy Contributions to Binding in Molecules and Solids. J. Chem. Phys. 2010, 132, 234109.

(2) Karplus, M. Vicinal Proton Coupling in Nuclear Magnetic Resonance. J. Am. Chem. Soc. 1963, 85, 2870-2871.

(3) Hennig, M.; Bermel, W.; Schwalbe, H.; Griesinger, C. Determination of ψ Torsion Angle Restraints from 3J(Cα,Cα) and 3J(Cα,HN) Coupling Constants in Proteins. J. Am. Chem. Soc. 2000, 122, 6268-6277.

(4) Hu, J.-S.; Bax, A. Determination of φ and χ1 Angles in Proteins from 13C−13C Three-Bond J Couplings Measured by Three-Dimensional Heteronuclear NMR. How Planar Is the Peptide Bond? J. Am. Chem. Soc. 1997, 119, 6360-6368.

(5) Wirmer, J.; Schwalbe, H. Angular Dependence of 3J(Ni,Cα) and 3J(Ni,Cα(i−1)) Coupling Constants Measured in J-modulated HSQCs. J. Biomol. NMR 2002, 23, 47-55.

(6) Ding, K.; Gronenborn, A. M. Protein Backbone 1H−13Cα and 15N−13Cα Residual Dipolar and J Couplings: New Constraints for NMR Structure Determination. J. Am. Chem. Soc. 2004, 126, 6232-6233.

(7) Graf, J.; Nguyen, P. H.; Stock, G.; Schwalbe, H. Structure and Dynamics of the Homologous Series of Alanine Peptides: A Joint Molecular Dynamics/NMR Study. J. Am. Chem. Soc. 2007, 129, 1179-1189.