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

Supporting information for: Mode Specific THz Spectra of Solvated Amino Acids using the AMOEBA Polarizable Force Field

Alexander Esser[1],[†] Saurabh Belsare[1],[‡] Dominik Marx,[*],[†] and
Teresa Head-Gordon[*],[¶]

Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany,
The UCB-UCSF Graduate Program in Bioengineering, University of California Berkeley,
CA 94720, United States, and Kenneth S. Pitzer Center for Theoretical Chemistry,
Department of Chemistry, Department of Bioengineering, Department of Chemical and
Biomolecular Engineering, University of California Berkeley, CA 94720, United States

E-mail: dominik.marx@rub.de; thg@berkeley.edu

---

[1] *These two authors contributed equally*

[*] To whom correspondence should be addressed

[†] Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany

[‡] The UCB-UCSF Graduate Program in Bioengineering, University of California Berkeley, CA 94720, United States

[¶] Kenneth S. Pitzer Center for Theoretical Chemistry, Department of Chemistry, Department of Bioengineering, Department of Chemical and Biomolecular Engineering, University of California Berkeley, CA 94720, United States

Structures

Figure S1 shows the reference structures used for the mode decomposition where each atom is labeled.

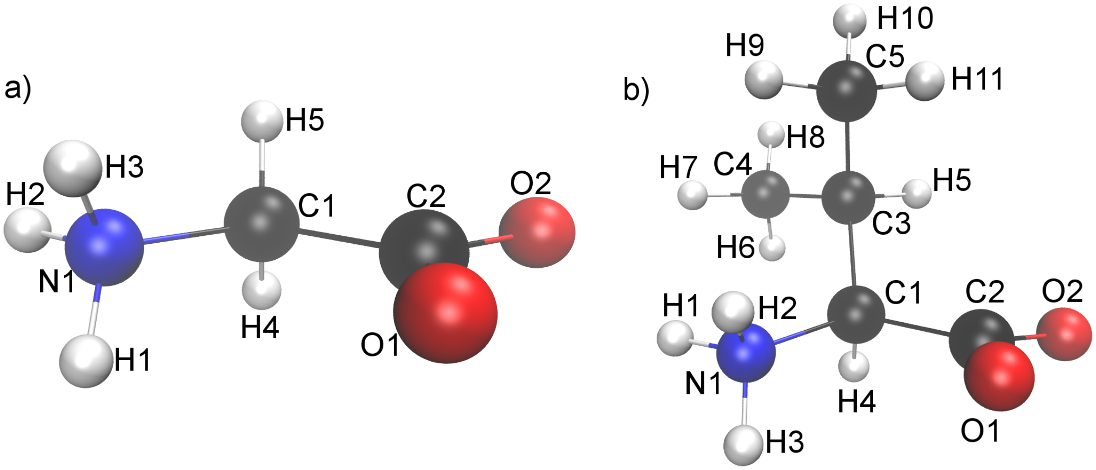


Figure S1: Reference structures of (a) glycine-only and (b) valine-only atom labels.

Radial and Angular Distribution Functions

The radial distribution functions (RDFs) of valine in water show the radial arrangement of water molecules around the protonated amino and deprotonated carboxyl groups. Akin to the results shown for glycine in water in the main text the original AMOEBA parametrization leads to an outward shift compared to AIMD also in case of valine which is largely corrected when using the modified AMOEBA model that is shown here; note that the corresponding AIMD RDFs of glycine are overall in reasonable accord with experimental data as demonstrated in the main text and thus are expected to serve as benchmarks also for valine.

Angular distribution functions for glycine and valine are shown in Fig. S3 and Fig. S4, respectively. The angles as obtained from the original AMOEBA model nicely match for both glycine and valine those of the corresponding angles obtained from AIMD so that no optimization was carried out in the modified model.

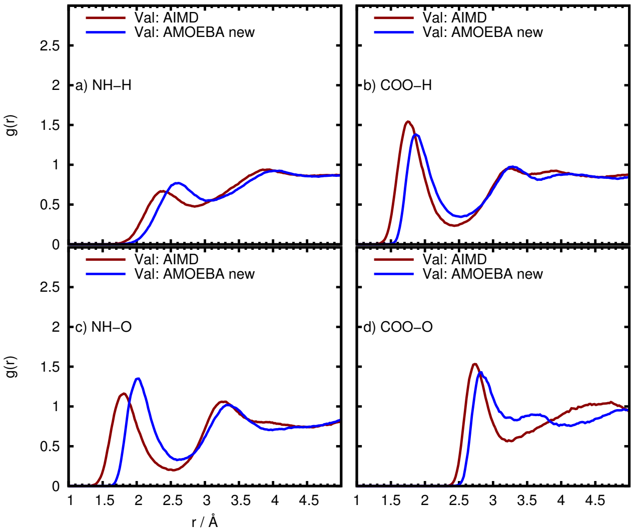


Figure S2: Radial distribution functions between the protonated amino and deprotonated carboxyl groups of valine in water with respect to the water molecules.

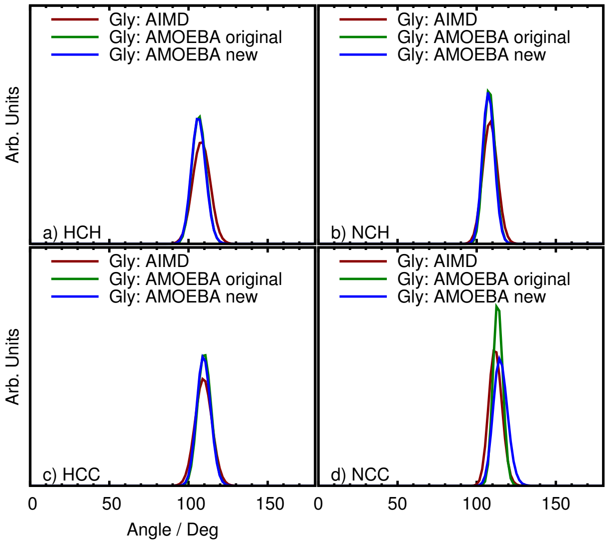


Figure S3: Angular distribution functions for glycine. (a) H4-C1-H5, (b) N1-C1-H4/H5, (c) H4/H5-C1-C2, (d) N1-C1-C2.

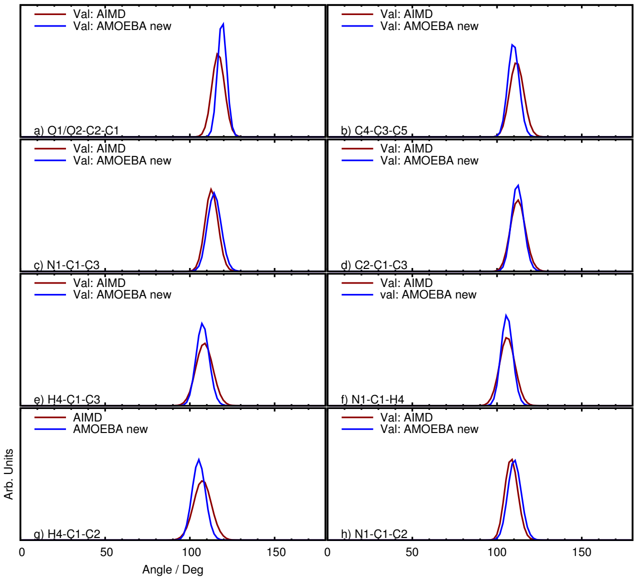


Figure S4: Angular distribution functions for Valine. (a) O1/O2-C2-C1, (b) C4-C3-C5, (c) N1-C1-C3, (d) C2-C1-C3, (e) H4-C1-C3, (f) N1-C1-H4, (g) H4-C1-C2, (h) N1-C1-C2.

The dihedral distribution functions are shown in Fig. S5 for glycine and Fig. S6 for valine. Overall the dihedral angles show similar tendencies to the ones from AIMD but are generally too stiff in AMOEBA.

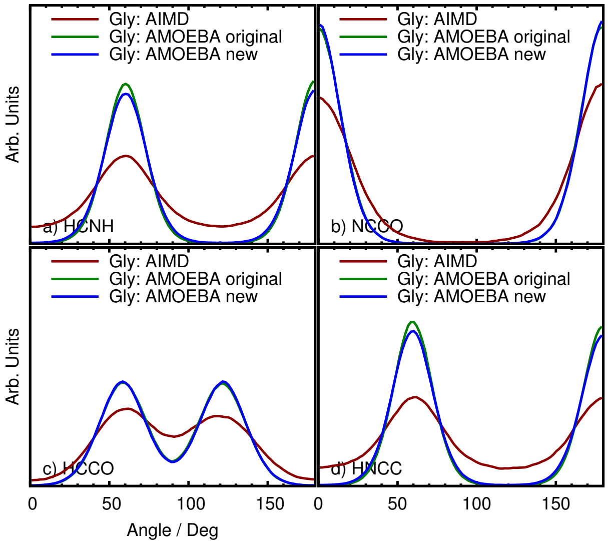


Figure S5: Dihedral distribution functions for glycine. (a) HCNH, (b) NCCO, (c) HCCO, (d) HNCC.

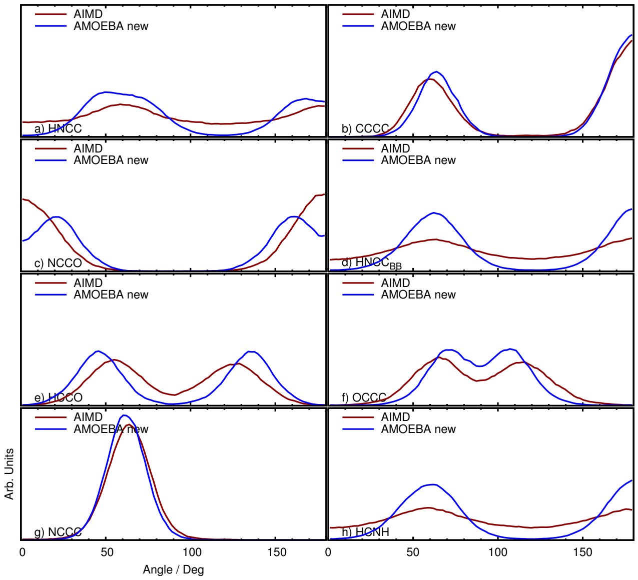


Figure S6: Dihedral distribution functions for valine. (a) amide H and water H, (b) carboxyl O and water H, (c) amide H and water O, (d) carboxyl O and water O.

Comparison of Mode Displacement Patterns derived from Decomposing VDOS versus IR Spectra

Decomposing vibrational spectra in terms of modes is well–known to be not unique except in the fully (i.e. both mechanical and electrical) harmonic limit. In order to explicitly assess our particular approach, we compute and decompose the vibrational (or phonon) density of states (VDOS) using the identical approach as used to dissect the IR/THz spectra, but apply it to the $\sqrt{\text{mass}}$–weighted velocity cross-correlation matrix[S1,S2] of all atoms or nuclei (instead of using the charge current, i.e. dipole velocity cross-correlation matrix, involving not only the nuclei but also the electrons as represented by the centers of their Wannier functions, see main text). Thus, the VDOS exclusively considers the correlated motion of the atoms (in case of force field simulations or nuclei in AIMD) in three-dimensional space. It is noted in passing that the VDOS is experimentally accessible to inelastic neutron scattering but lacks, of course, the lineshape of IR/THz spectra, which results from dipolar fluctuations that depend on the electrons in the first place and thus on polarization and charge transfer contributions. The corresponding frequency-dependent weighting of IR/THz spectra is particularly important in the low-frequency regime where, for instance, the prominent water network mode at roughly 200 cm^{-1} is simply absent in any VDOS–based spectrum[S1,S2].

We applied our full decomposition procedure to severall representative classes of modes, namely to (i) essentially intramolecular motion (using the “NCCO open/close mode”, see Fig. S7), (ii) largely intermolecular hydrogen bond stretching and hydrogen bond bending motion (using the HB-stretch I and HB-bend II modes, respectively, in Figs. S8 and S9), as well as to (iii) intimately coupled inter/intramolecular motion (using the “C-C twist + HB stretch” mode, see Fig. S10) in order to extract VDOS–based modes in direct comparison to IR–based modes both obtained using SSC(+) analysis. Moreover, we compute the respective AMOEBA as well as AIMD modes which allows for comparison also at the level of the interaction model. The detailed comparisons compiled in Figs. S7–S10 make clear that the

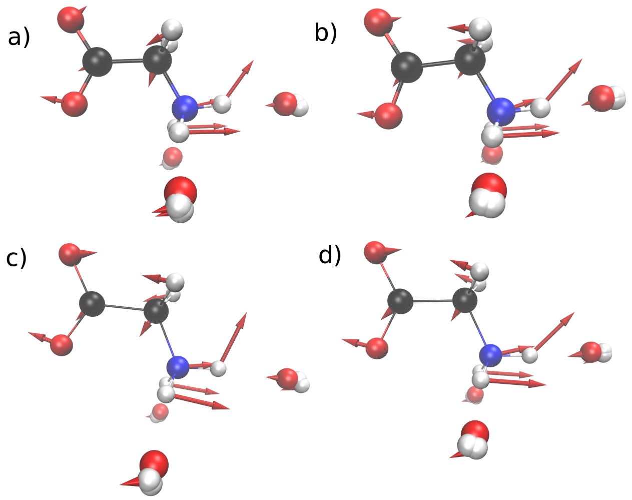


Figure S7: Glycine VDOS mode displacement patterns of the NCCO open/close mode for a) AMOEBA and b) AIMD. The corresponding IR modes are shown for c) AMOEBA and d) AIMD.

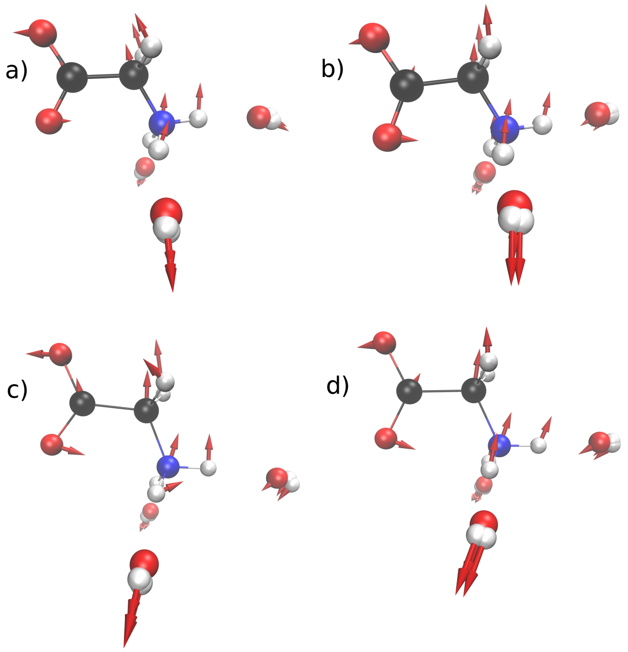


Figure S8: Glycine VDOS mode displacement patterns of the HB-stretch I mode for a) AMOEBA and b) AIMD. The corresponding IR modes are shown for c) AMOEBA and d) AIMD.

atomic displacement patterns in real–space as obtained from the IR/THz decomposition compare well to those obtained from the VDOS for all classes of modes, in particular also including those modes where intermolecular solute-solvent couplings are greatly contributing to the motion or even dominate. Moreover, the excellent match between the AMOEBA and AIMD mode displacement patterns support the approach devised in the main text in order to compute the effective molecular dipoles in the former case, which provide an approximate lineshape weighting of IR/THz spectra within the AMOEBA approach according to Eqs. (18) and (19).

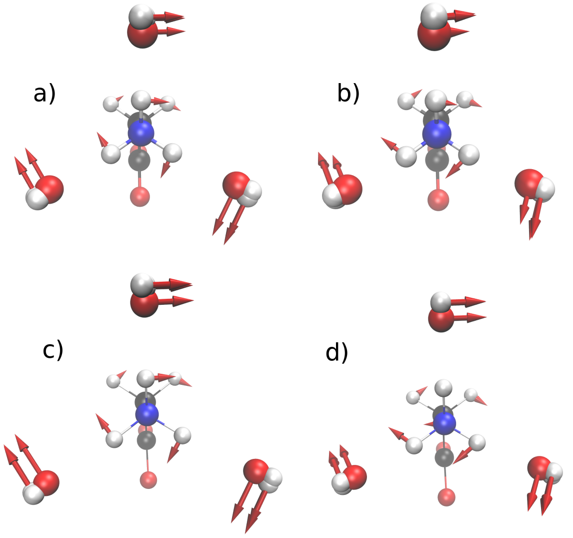


Figure S9: Glycine VDOS mode displacement patterns of the HB-bend II mode for a) AMOEBA and b) AIMD. The corresponding IR modes are shown for c) AMOEBA and d) AIMD.

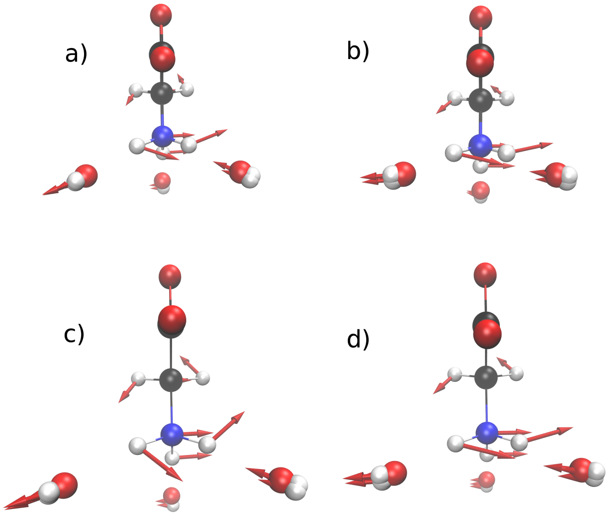


Figure S10: Glycine VDOS mode displacement patterns of the C-C twist + HB stretch mode for a) AMOEBA and b) AIMD. The corresponding IR modes are shown for c) AMOEBA and d) AIMD.

Modes of Glycine According to SSC(+) Analysis

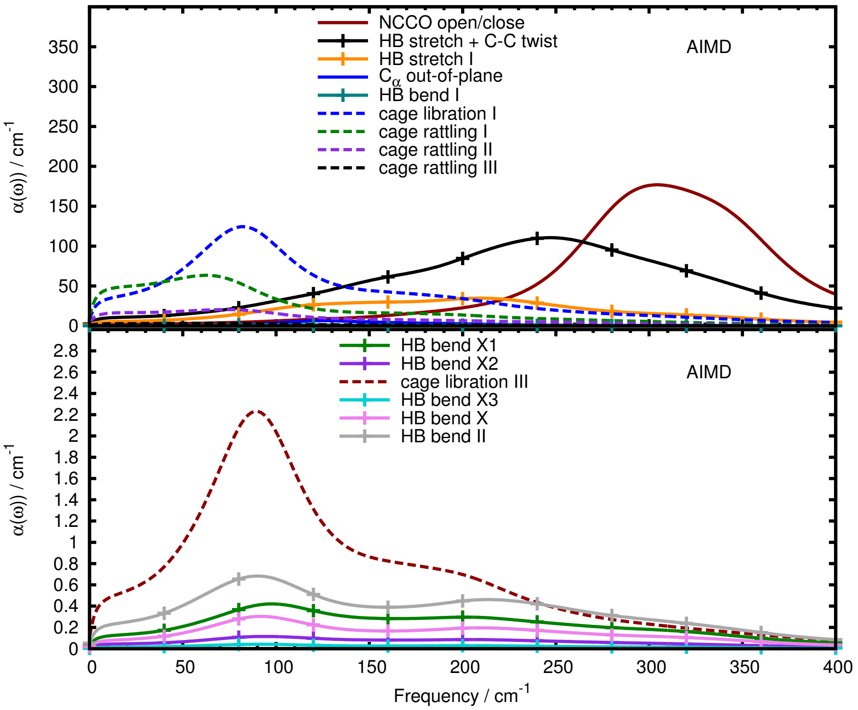


Figure S11: THz mode intensities of glycine within the supermolecular solvation complex at the amino group (SSC(+)) computed via AIMD solvated by 30 water molecules. The top panel shows high THz intensities, the bottom shows low intensities.

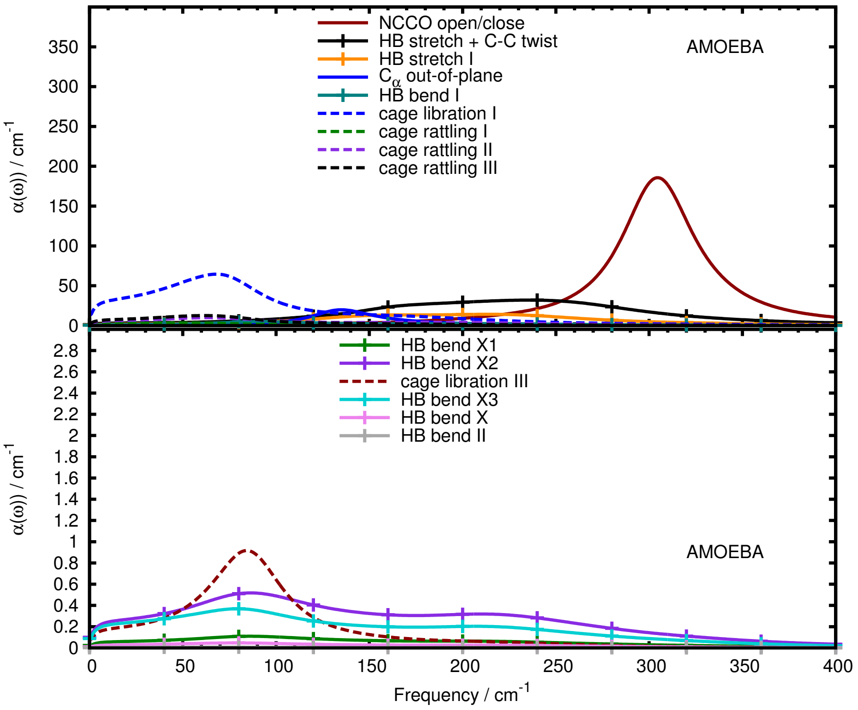


Figure S12: THz mode intensities of glycine within the supermolecular solvation complex at the amino group (SSC(+)) computed via AMOEBA solvated by 30 water molecules. The top panel shows high THz intensities, the bottom shows low intensities.

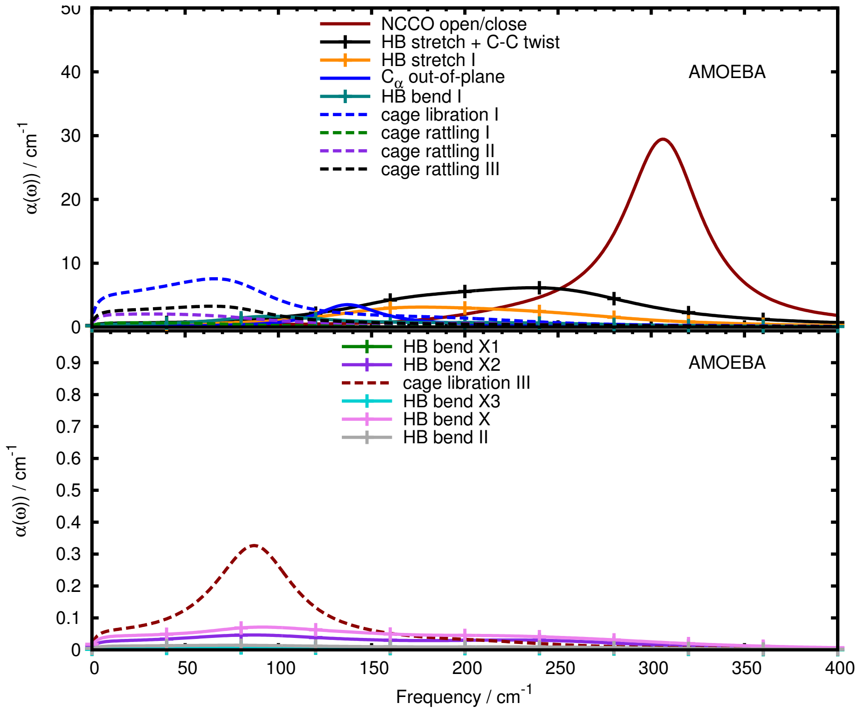


Figure S13: THz mode intensities of glycine within the supermolecular solvation complex at the amino group (SSC(+)) computed via AMOEBA solvated by 256 water molecules. The top panel shows high THz intensities, the bottom shows low intensities.

Modes of Valine According to SSC(+) Analysis

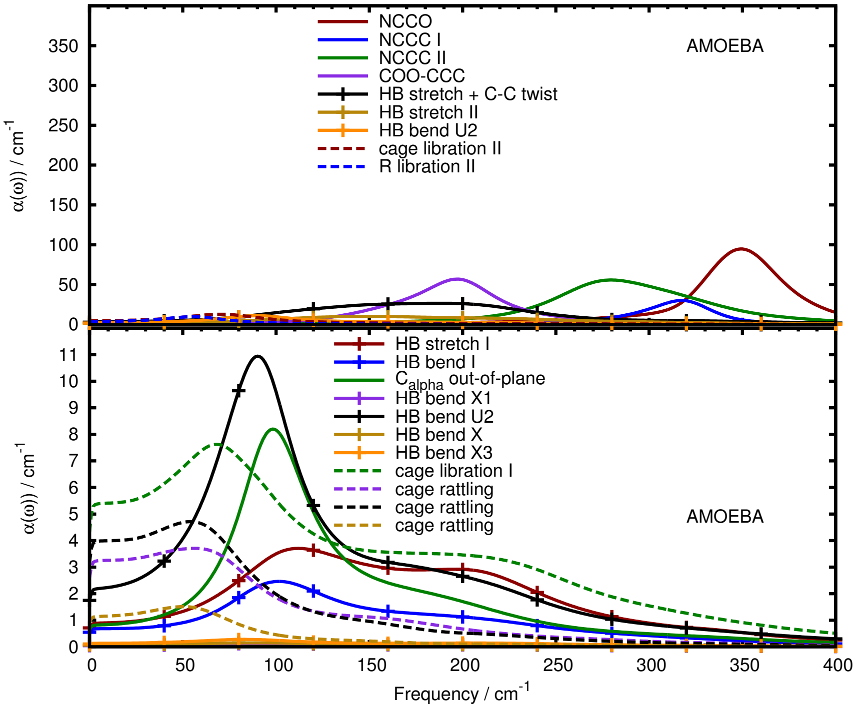


Figure S14: THz mode intensities of valine within the supermolecular solvation complex at the amino group (SSC(+)) computed via AIMD solvated by 60 water molecules. The top panel shows high THz intensities, the bottom shows low intensities.

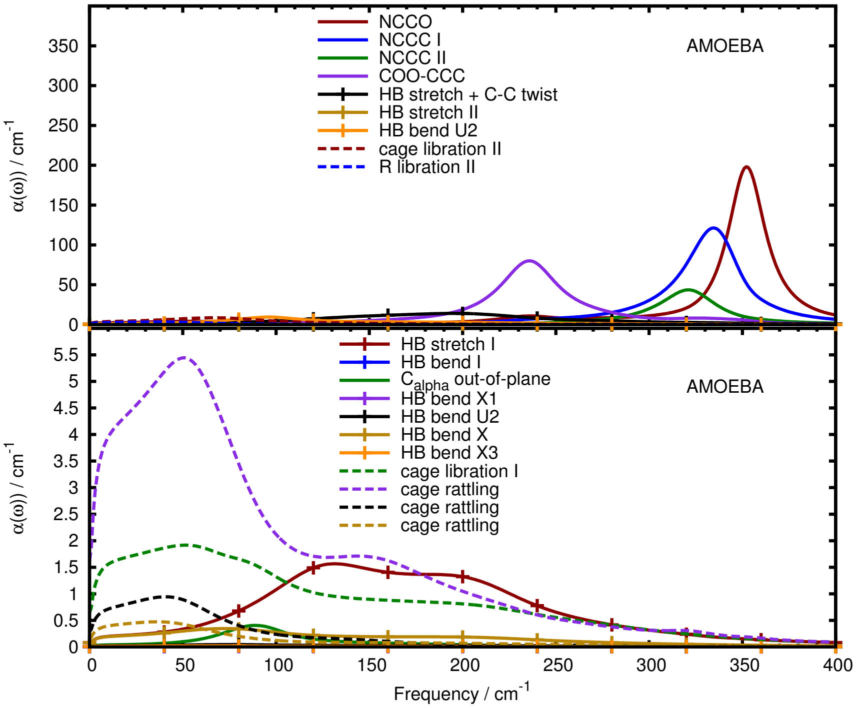


Figure S15: THz mode intensities of valine within the supermolecular solvation complex at the amino group (SSC(+)) computed via AMOEBA solvated by 60 water molecules. The top panel shows high THz intensities, the bottom shows low intensities.

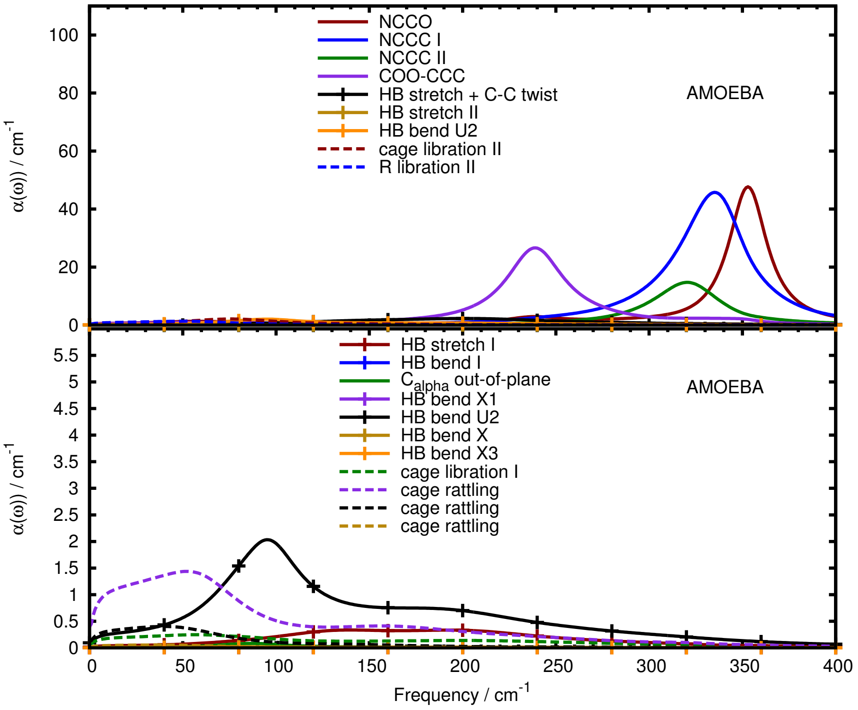
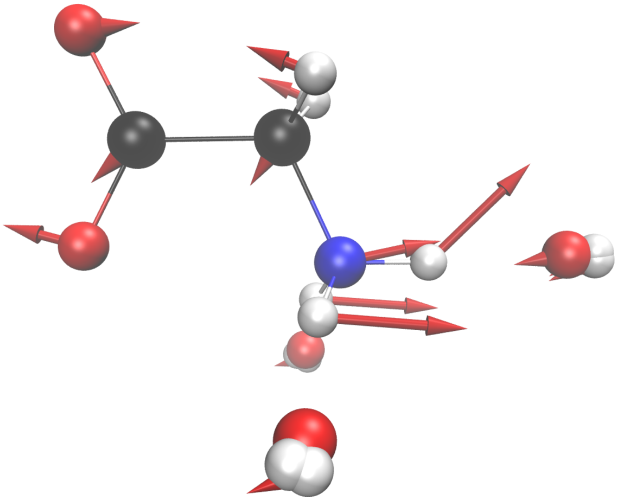
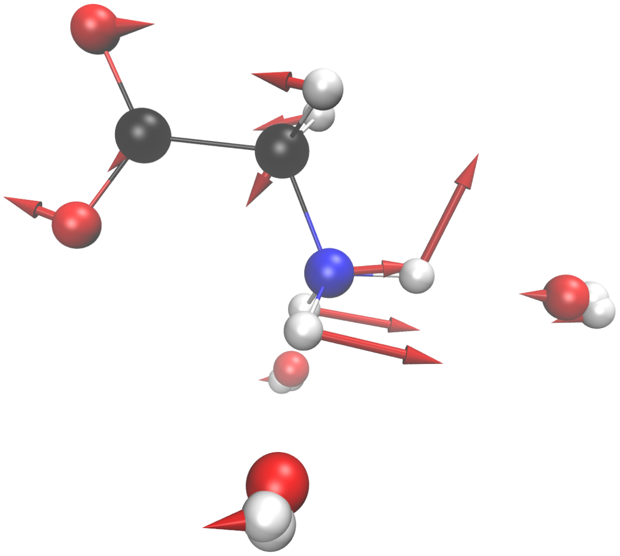


Figure S16: THz mode intensities of valine within the supermolecular solvation complex at the amino group (SSC(+)) computed via AMOEBA solvated by 256 water molecules. The top panel shows high THz intensities, the bottom shows low intensities.

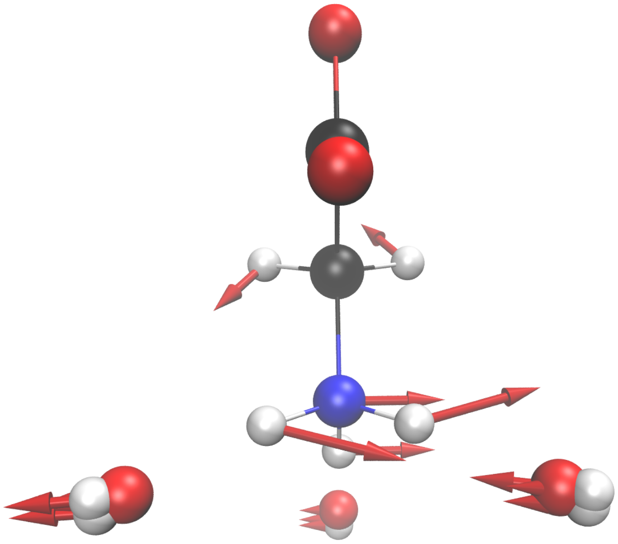
Mode Displacement Vectors of Glycine



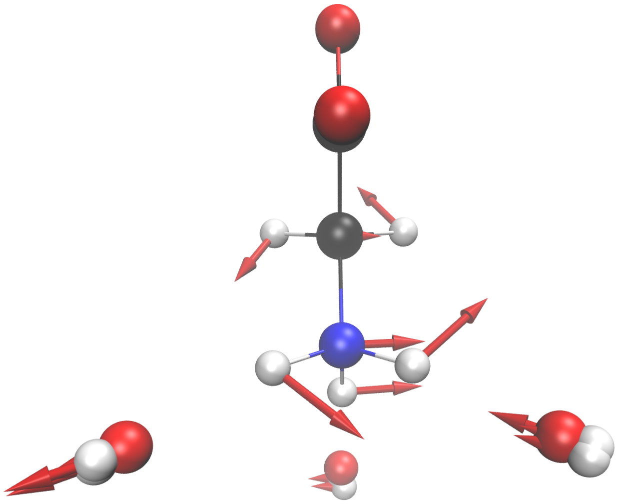
AIMD: NCCO 304 cm^{-1}



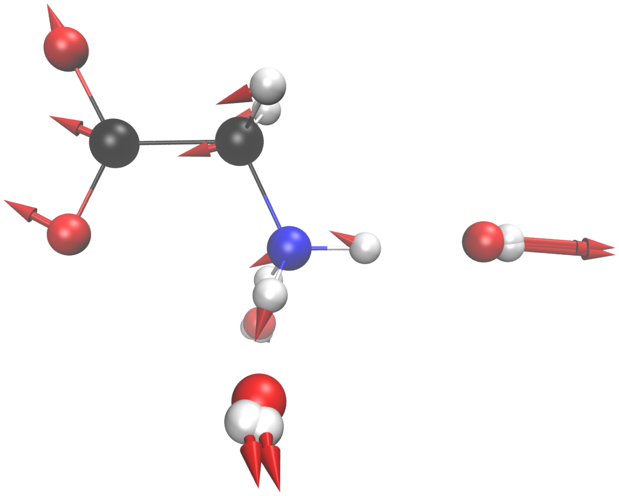
AMOEBA: NCCO 305 cm^{-1}



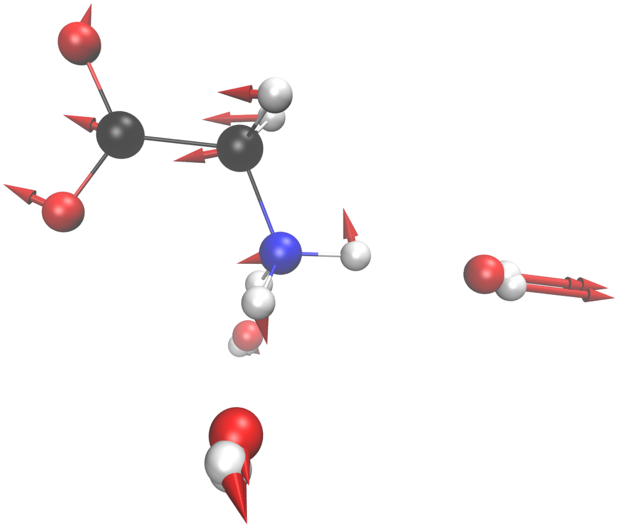
AIMD: HB-stretch-+-CC-Twist 247 cm^{-1}



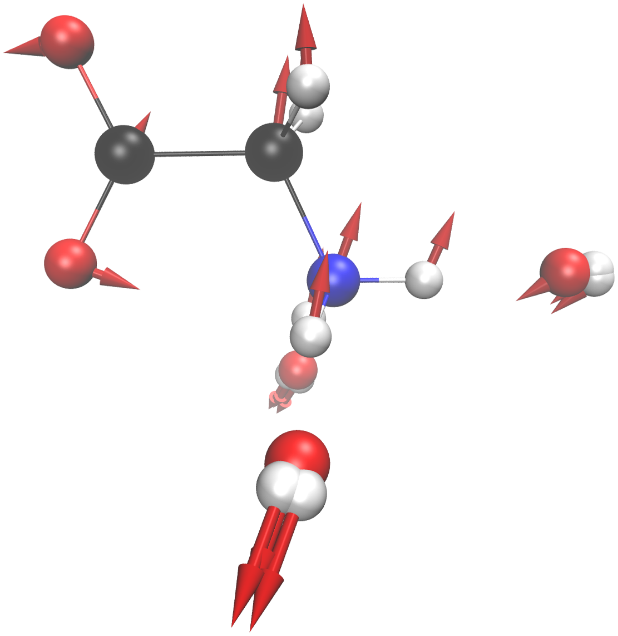
AMOEBA: HB-stretch-+-CCtwist 236 cm^{-1}



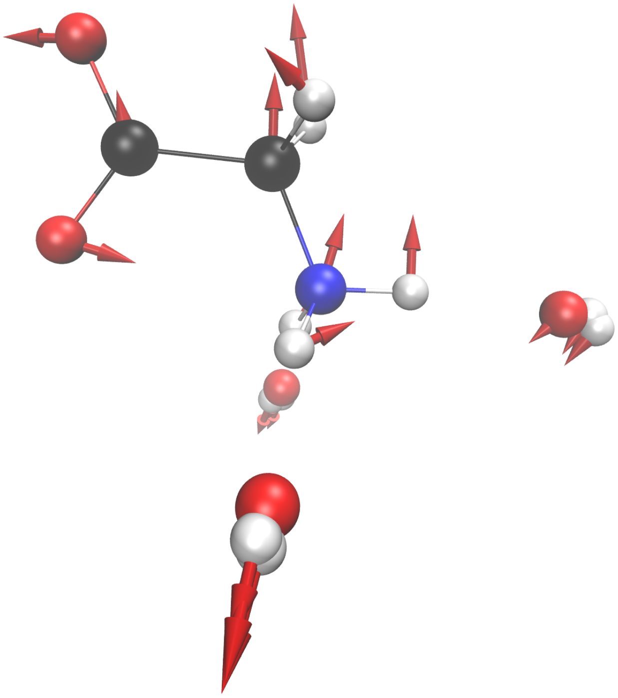
AIMD: HB-stretch-II 218 cm^{-1}



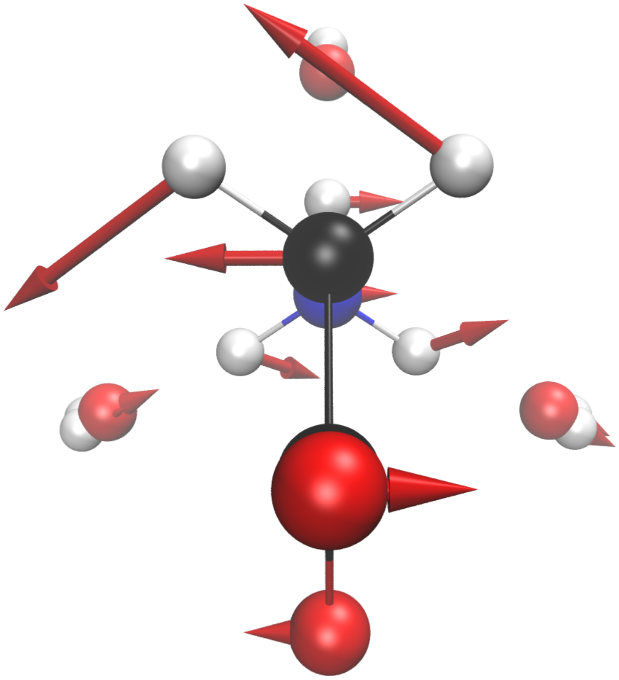
AMOEBA: HB-stretch-II 146 cm^{-1}



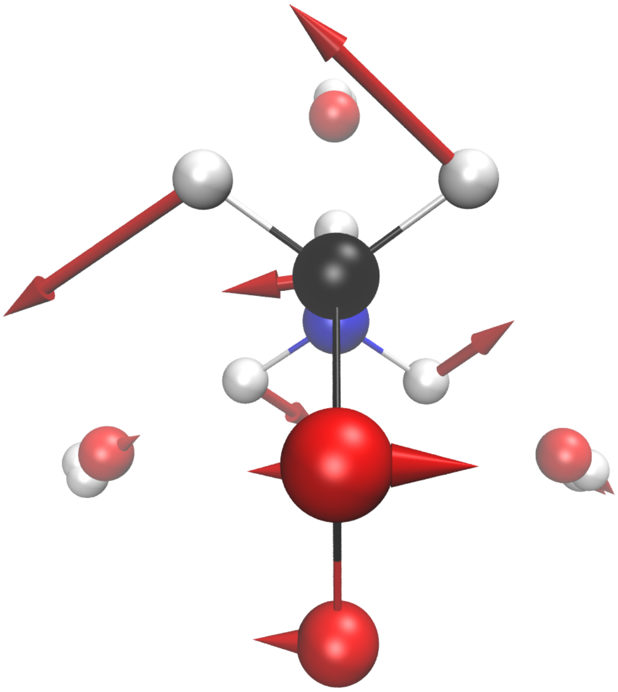
AIMD: HB-stretch-I 210 cm^{-1}



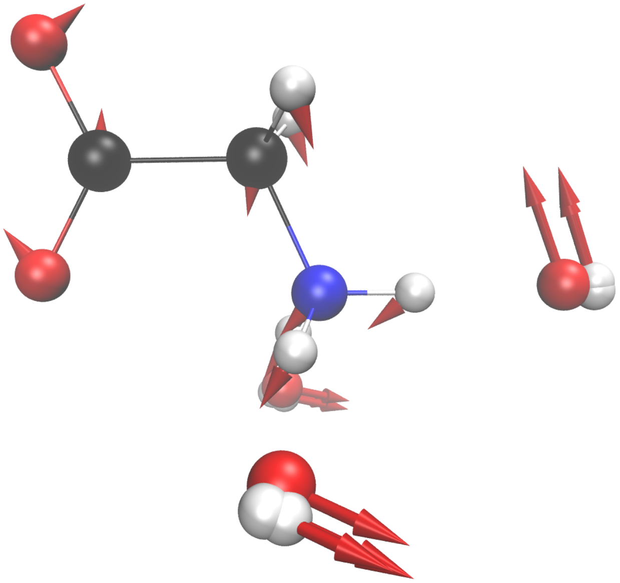
AMOEBA: HB-stretch-I 214 cm^{-1}



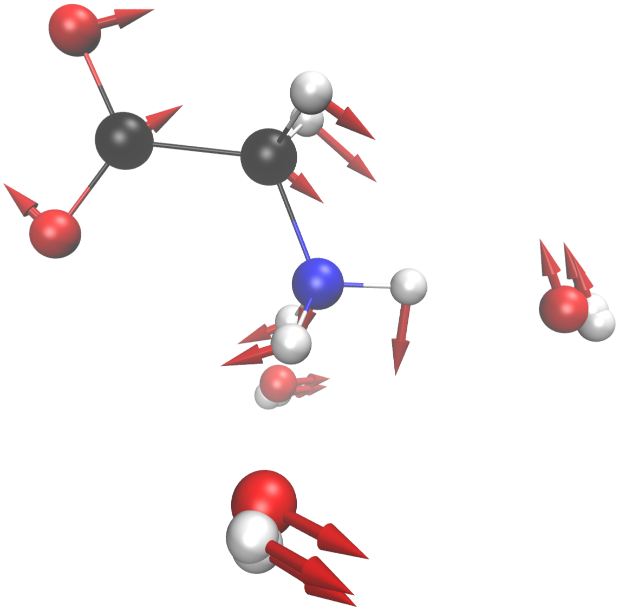
AIMD: C-alpha-oop 125 cm^{-1}



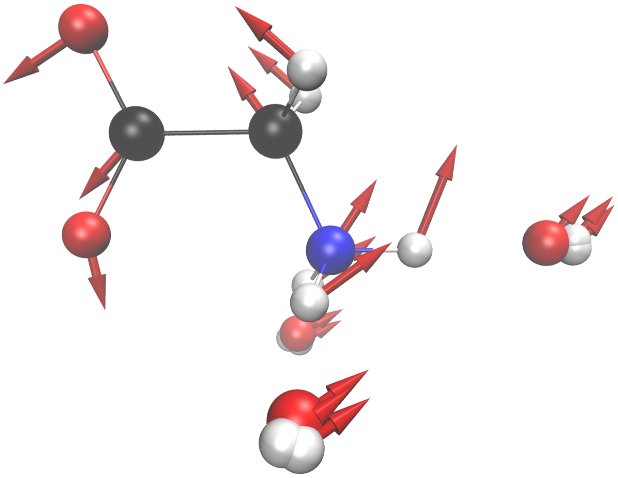
AMOEBA: C-alpha-oop 135 cm^{-1}



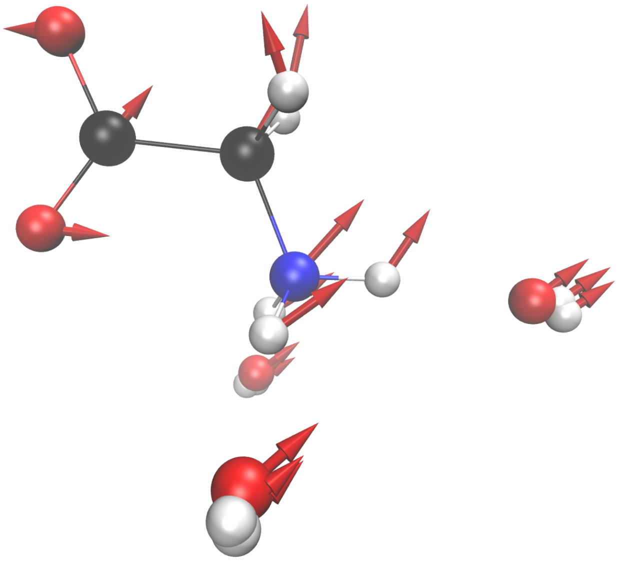
AIMD: HB-bend-I 102 cm^{-1}



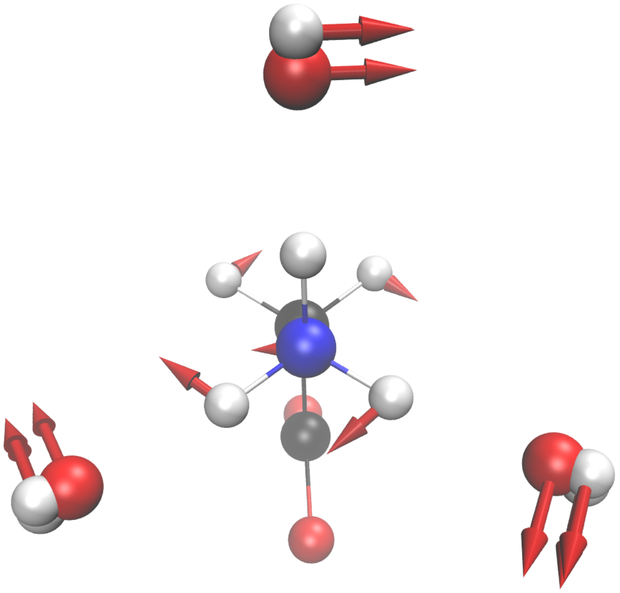
AMOEBA: HB-bend-I 92 cm^{-1}



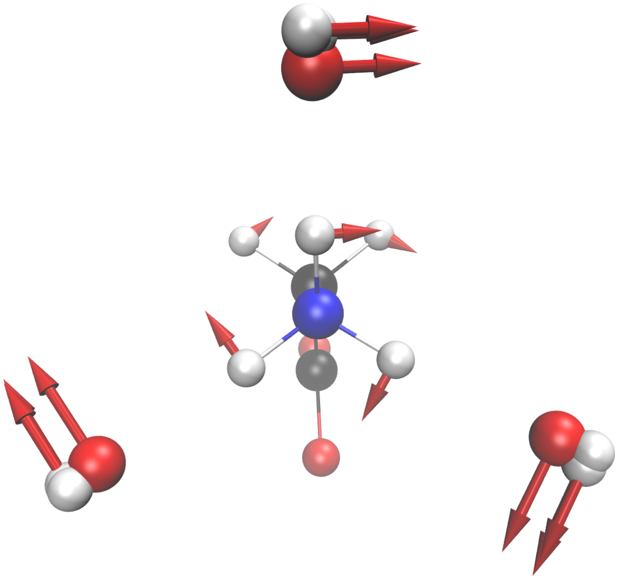
AIMD: cage-libration-II 90 cm^{-1}



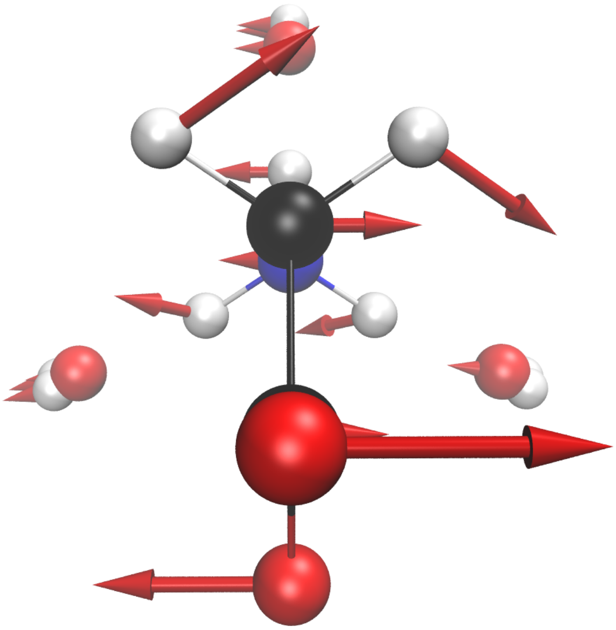
AMOEBA: cage-libration-II 71 cm^{-1}



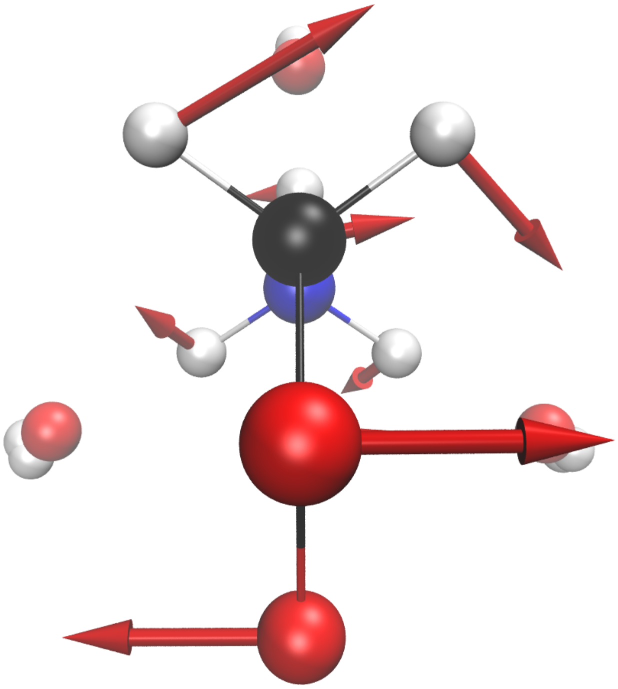
AIMD: HB-bend-II 90 cm^{-1}



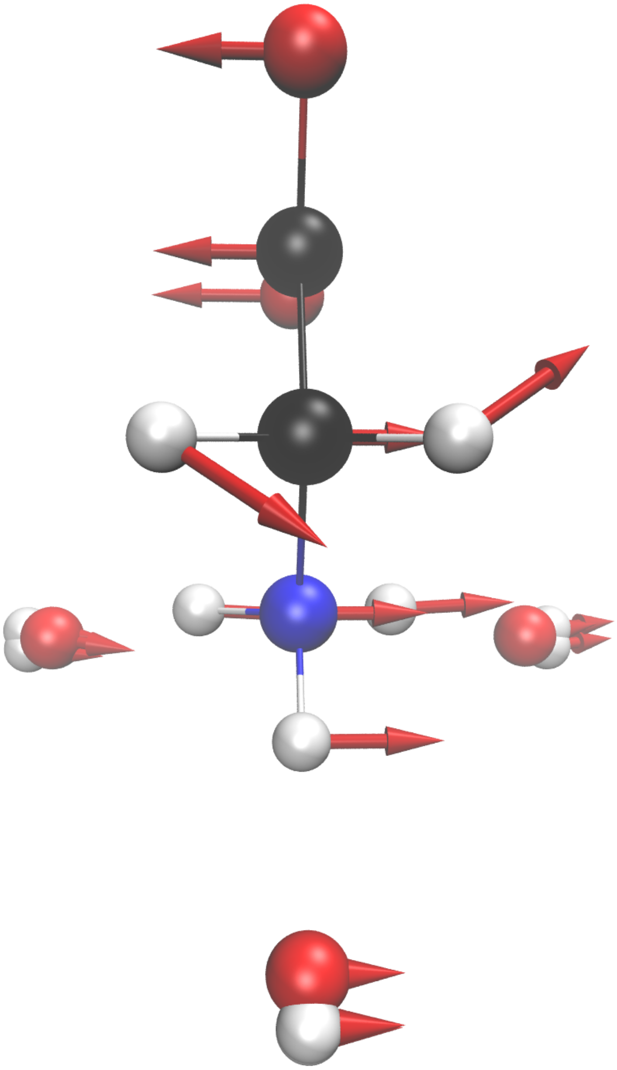
AMOEBA: HB-bend-II 68 cm^{-1}



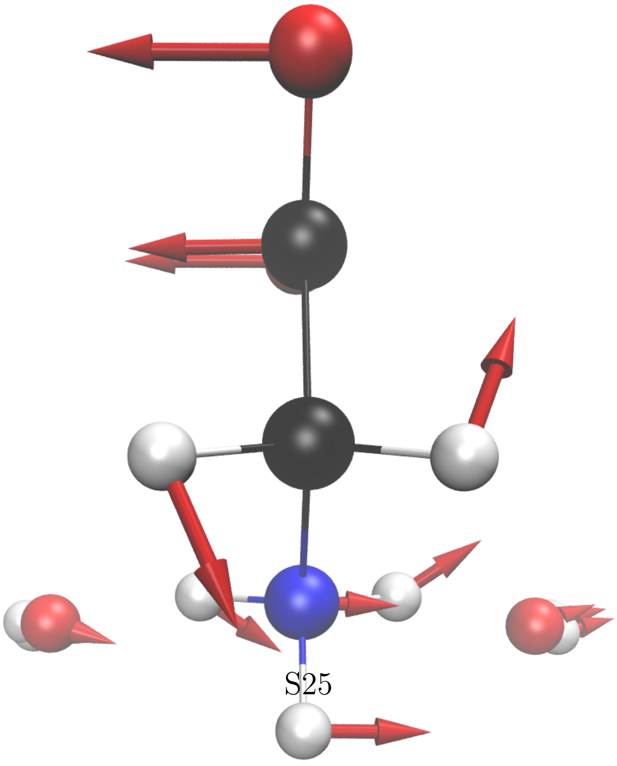
AIMD: cage-libration-III 89 cm^{-1}

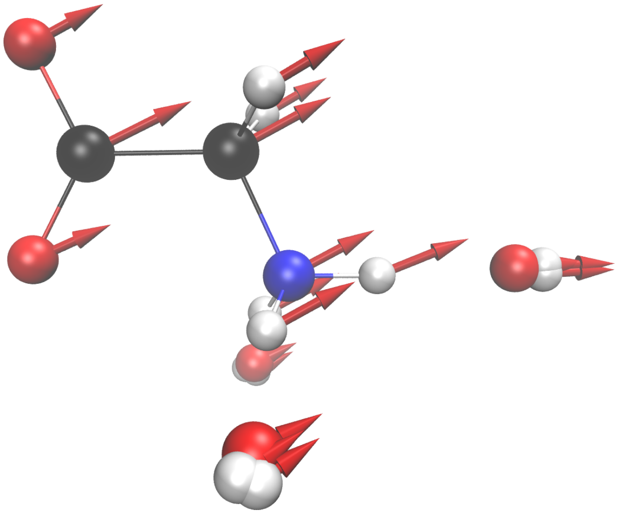


AMOEBA: cage-libration-III 84 cm^{-1}

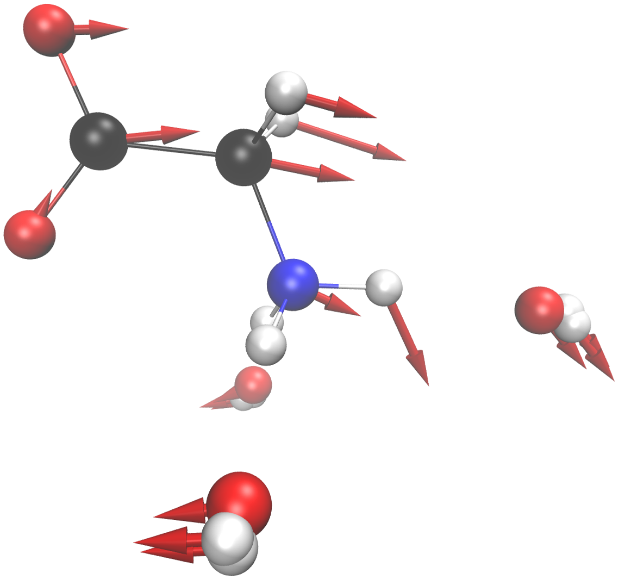


AIMD: cage-libration-I 82 cm^{-1}

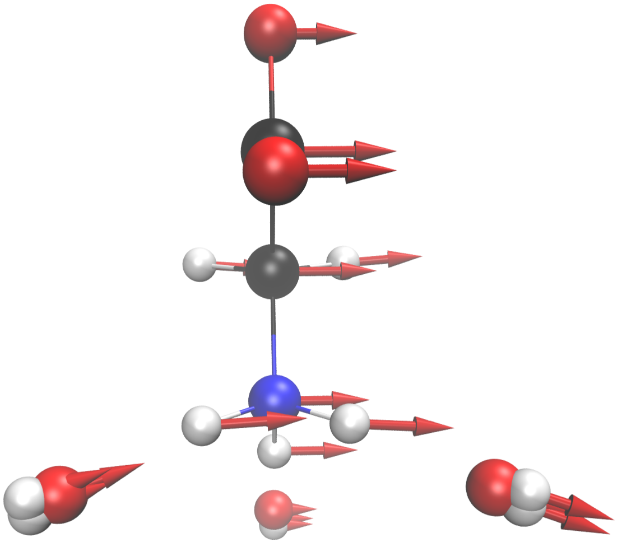




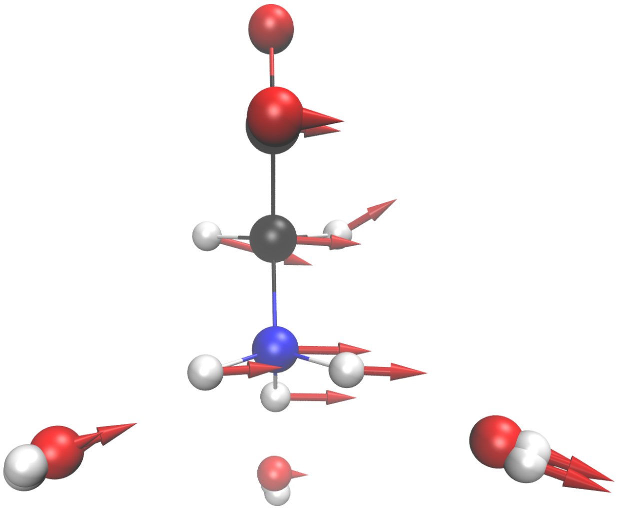
AIMD: cage-rattling-II 73 cm^{-1}



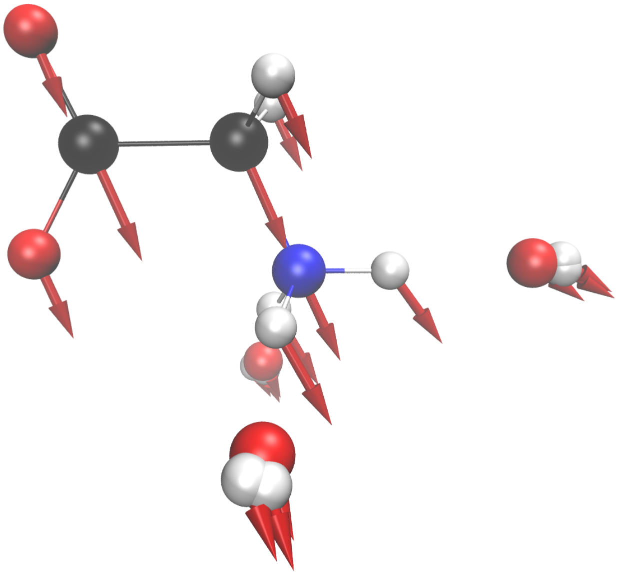
AMOEBA: cage-rattling-II 63 cm^{-1}



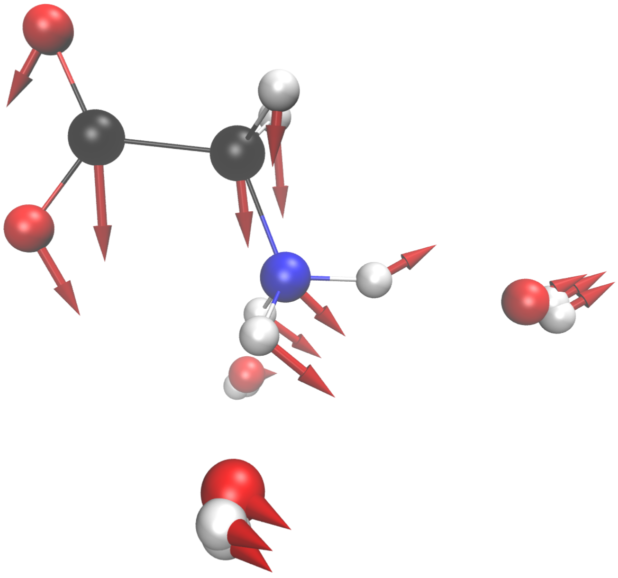
AIMD: cage-rattling-III 64 cm^{-1}



AMOEBA: cage-rattling-III 61 cm^{-1}

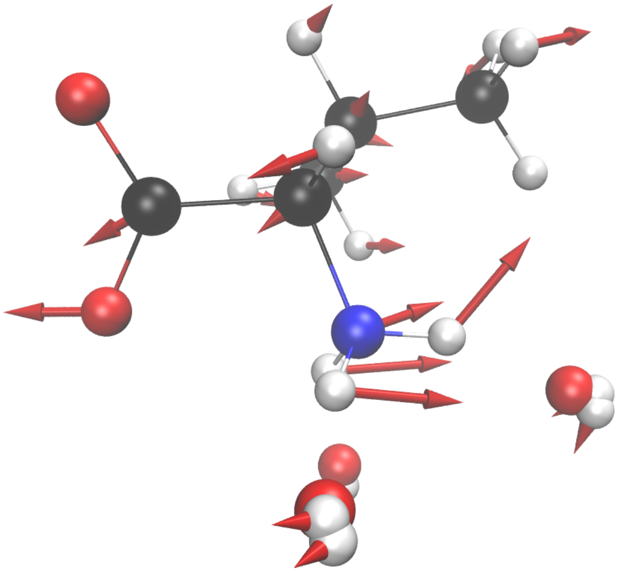


AIMD: cage-rattling-I 62 cm^{-1}

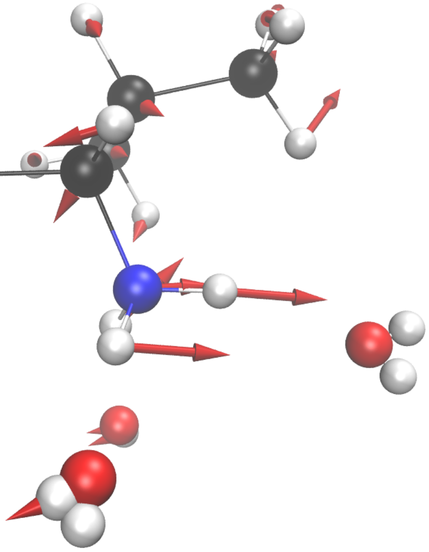


AMOEBA: cage-rattling-I 50 cm^{-1}

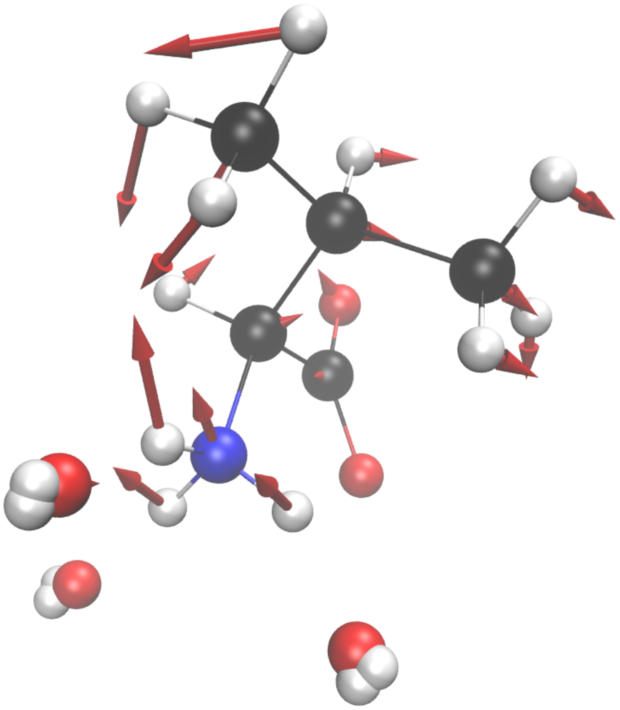
Mode Displacement Vectors of Valine



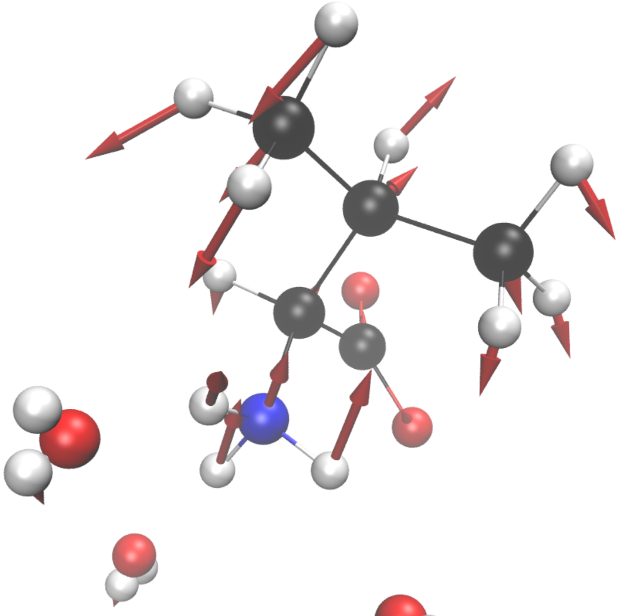
AIMD: NCCO 349



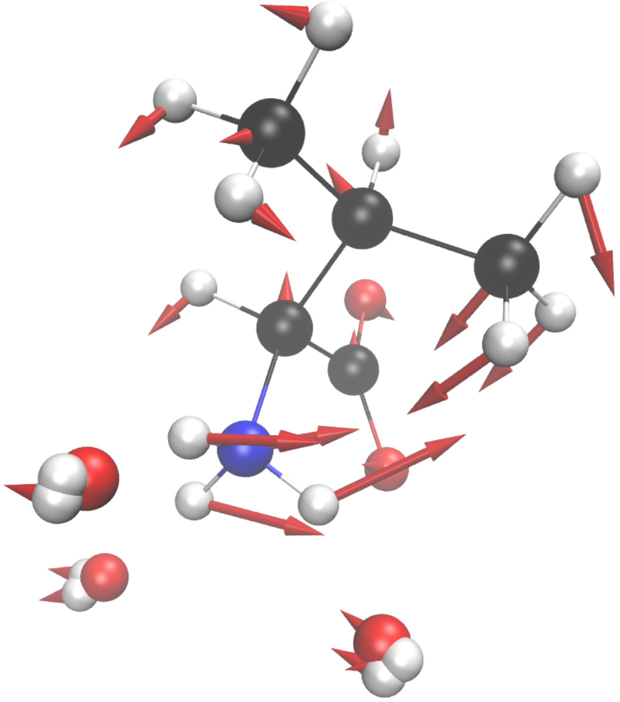
AMOEBA: NCCO 352



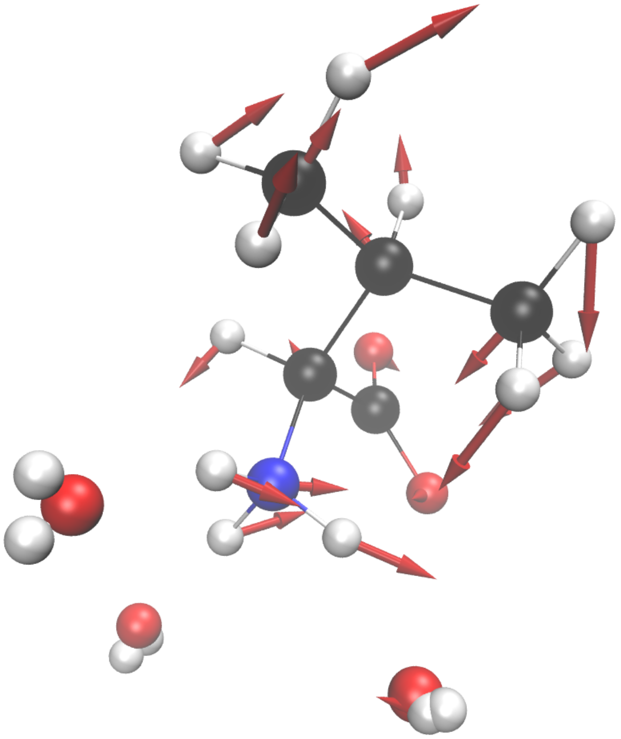
AIMD: NCCC-I 317



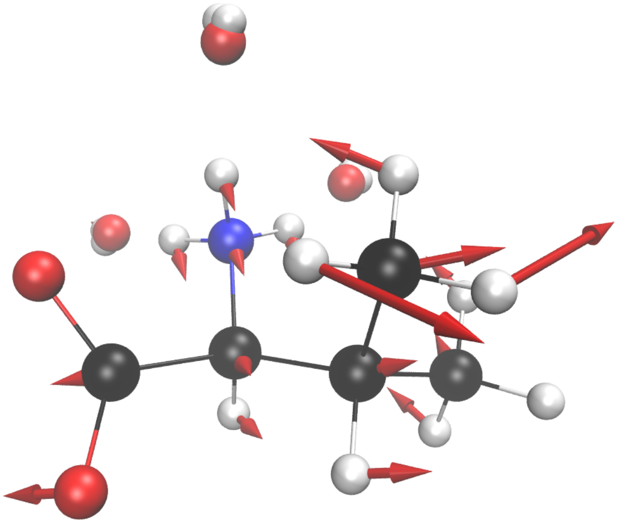
AMOEBA: NCCC-I 335



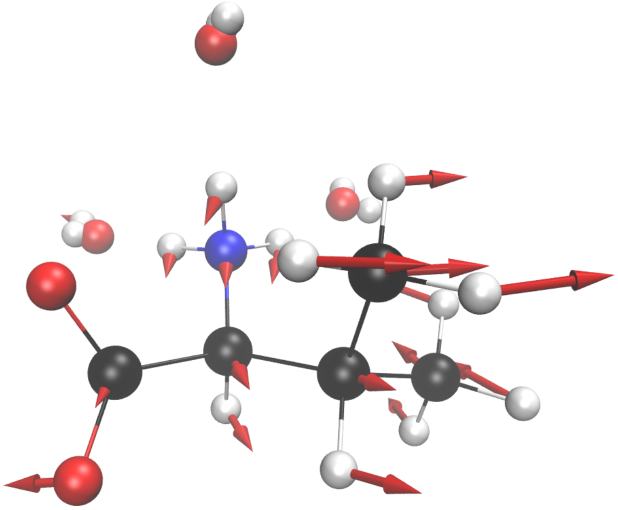
AIMD: NCCC-II 280



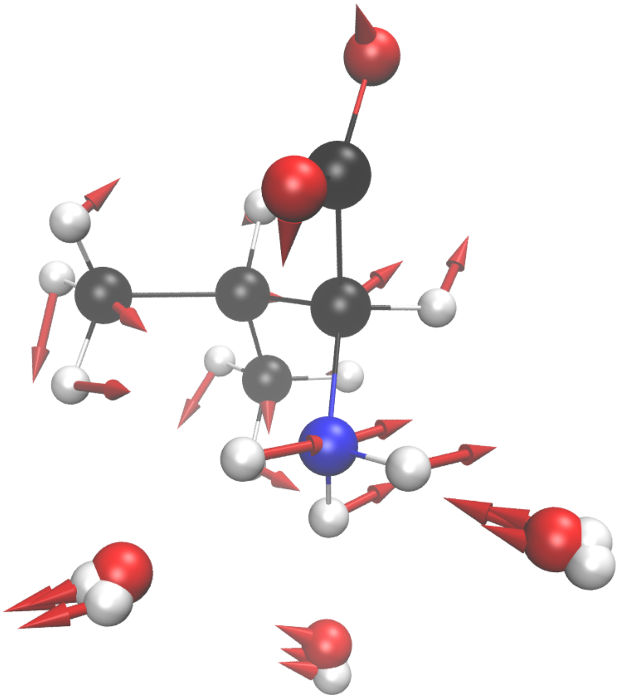
AMOEBA: NCCC-II 321



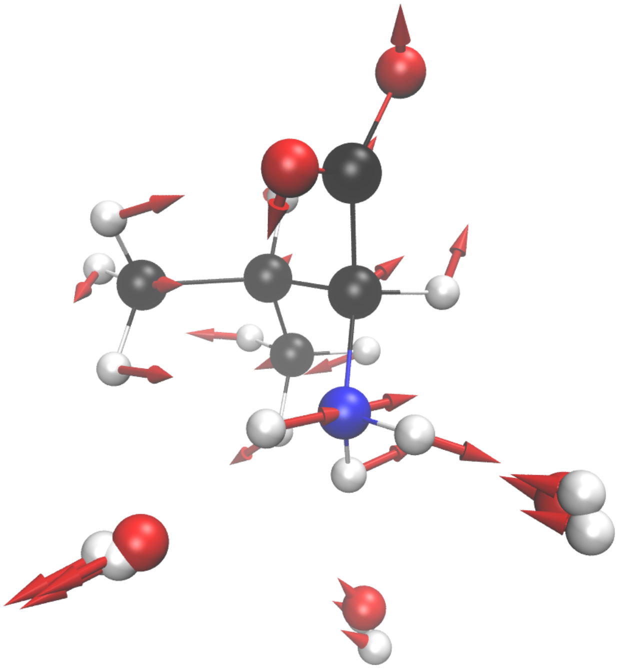
AIMD: COO-CCC 197



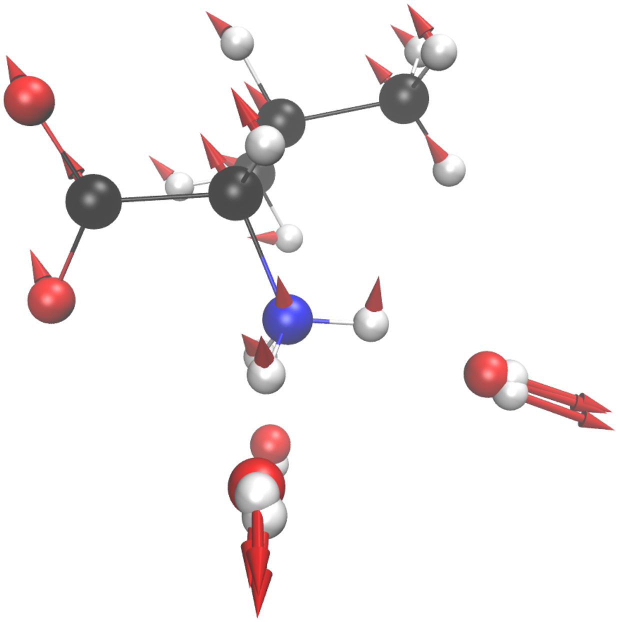
AMOEBA: COO-CCC 236



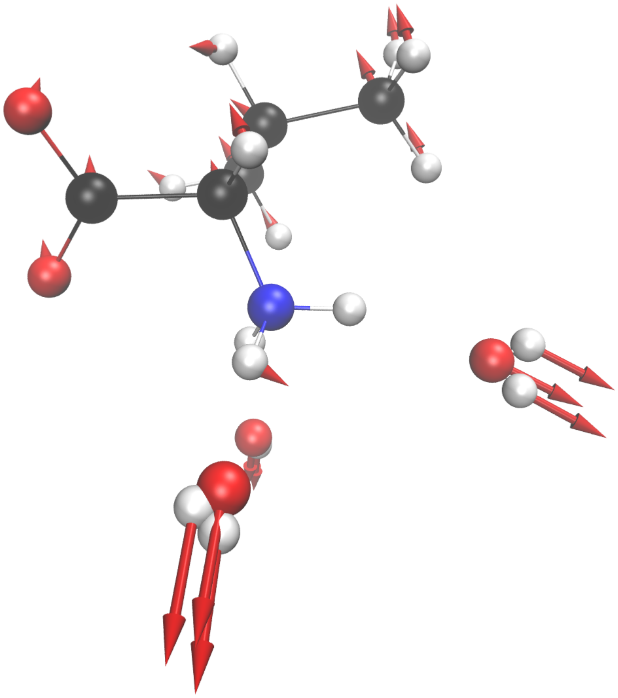
AIMD: CC-Twist-+-HB-Twist 189



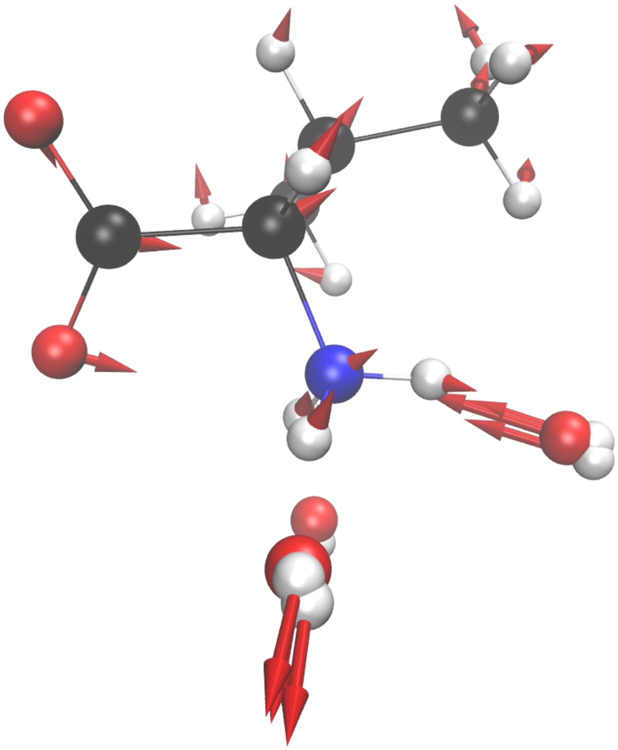
AMOEBA: CC-Twist-+-HB-stretch 196



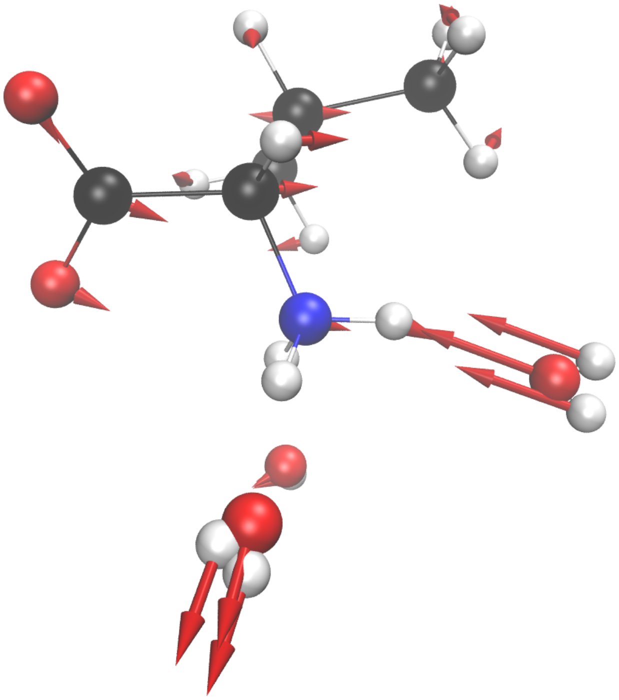
AIMD: HB-stretch-II 148



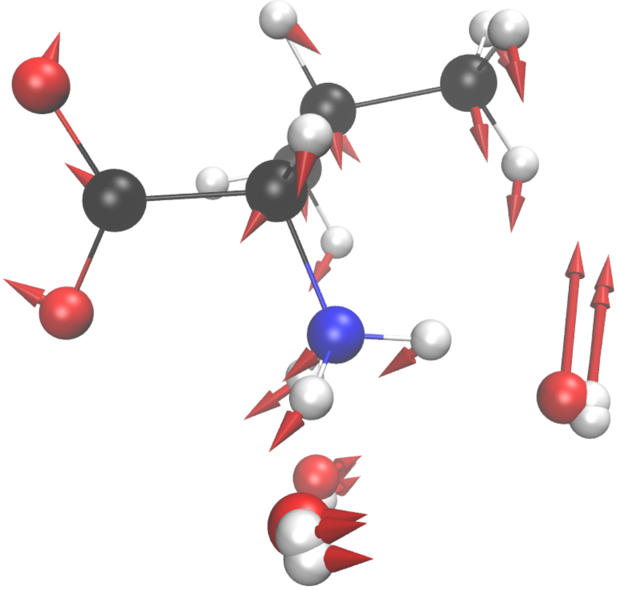
AMOEBA: HB-stretch-II 164



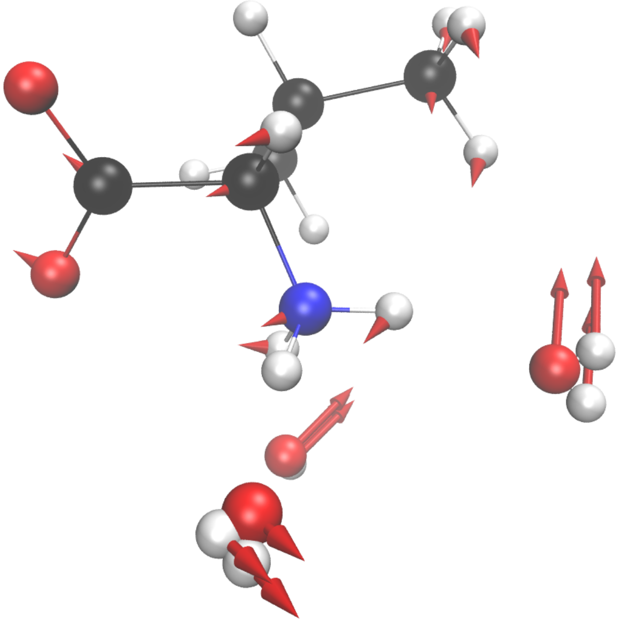
AIMD: HB-stretch-I 112



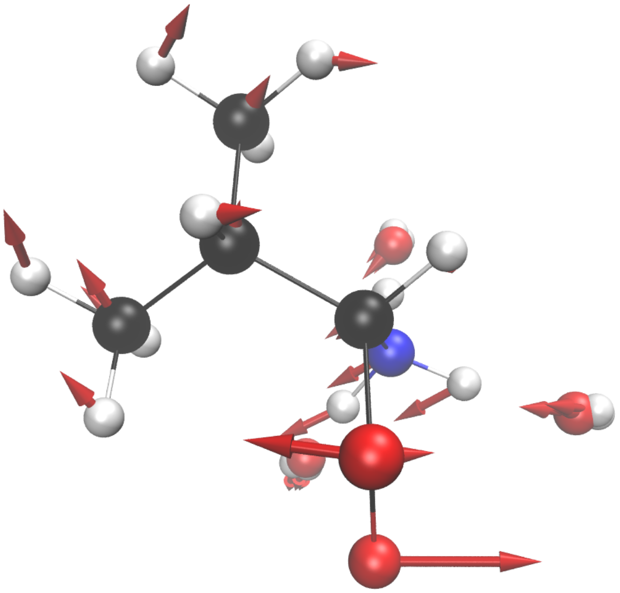
AMOEBA: HB-stretch-I 131



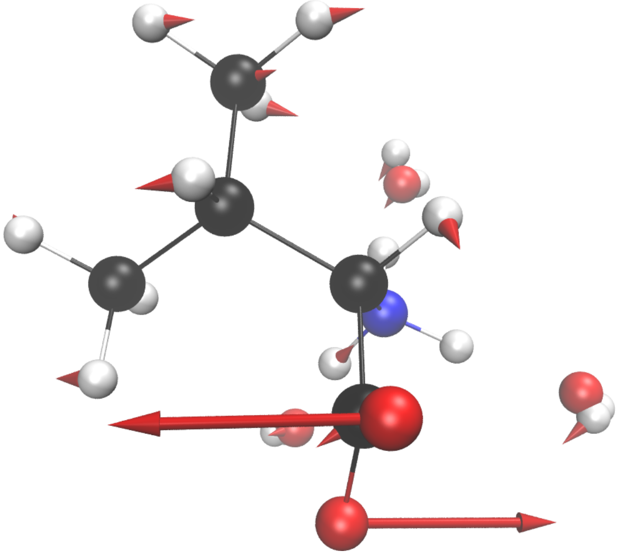
AIMD: HB-bend-I 101



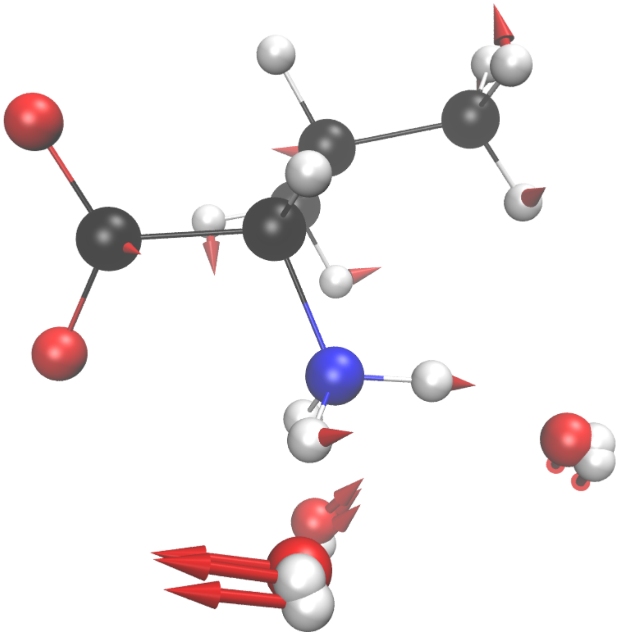
AMOEBA: HB-bend-I 86



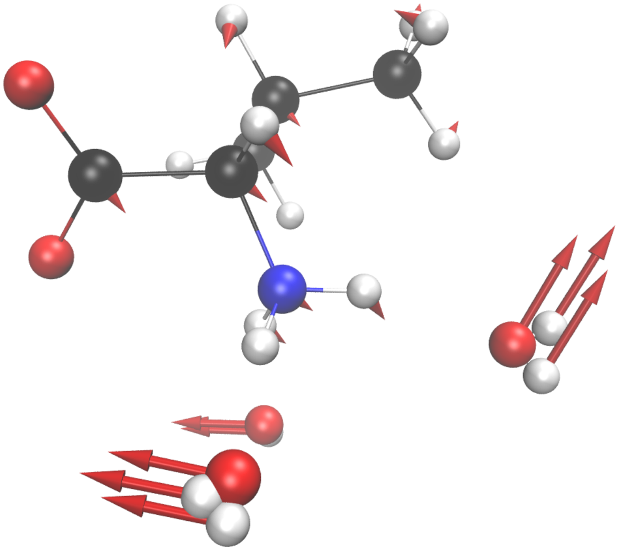
AIMD: Ca-oop-deformed 98



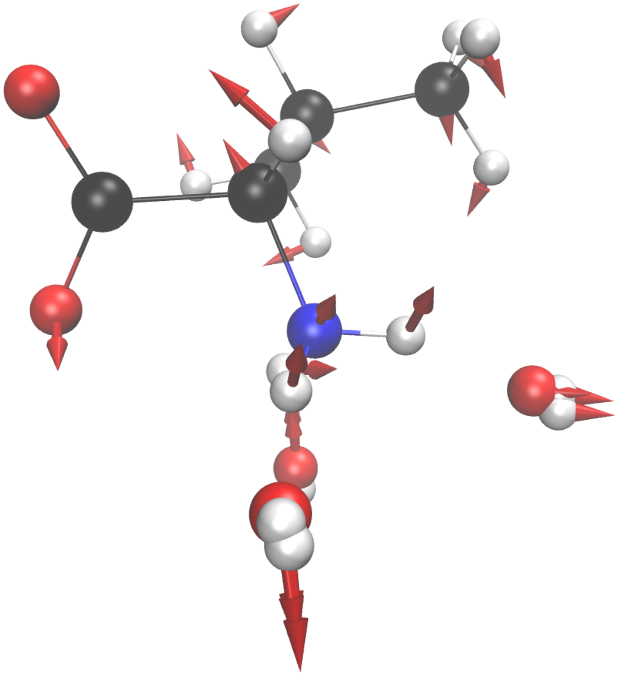
AMOEBA: Caoop 89



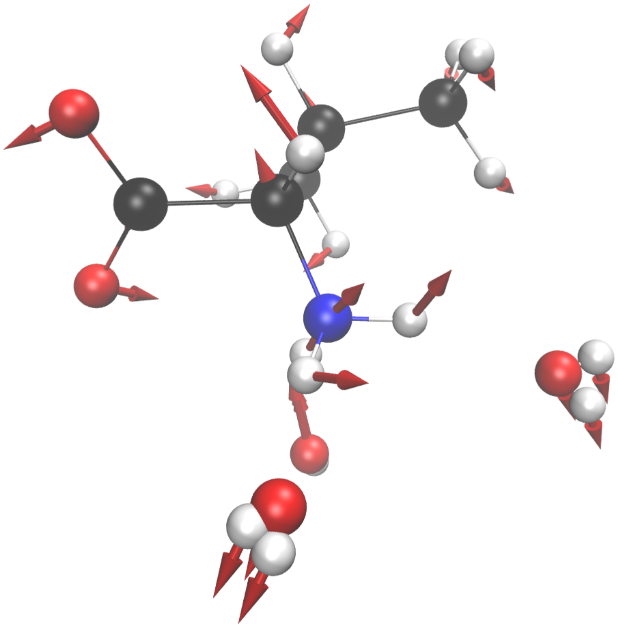
AIMD: HB-bend-X1 94



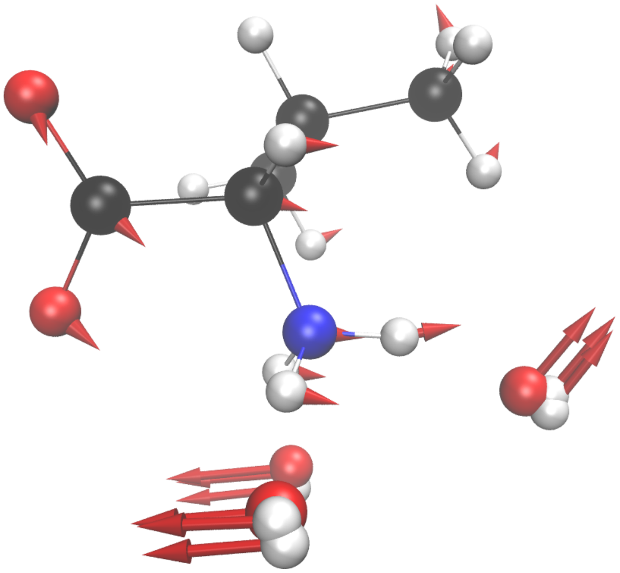
AMOEBA: HB-bend-X1 80



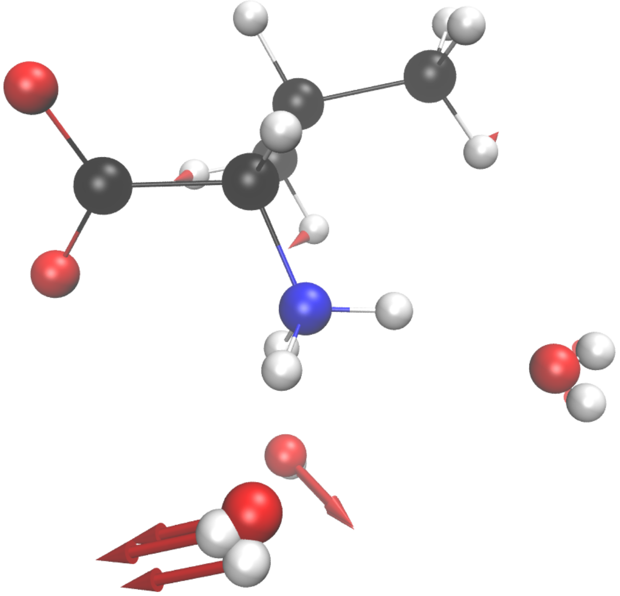
AIMD: HB-bend-U2 90



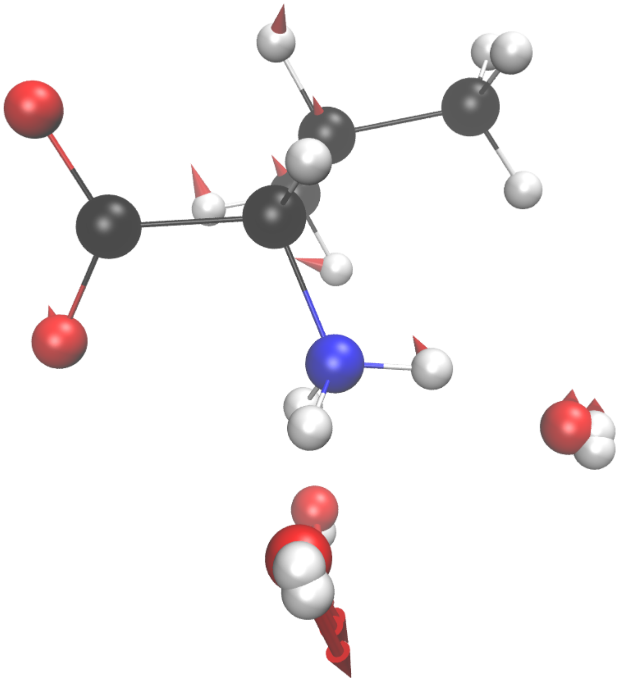
AMOEBA: HB-bend-U2 96



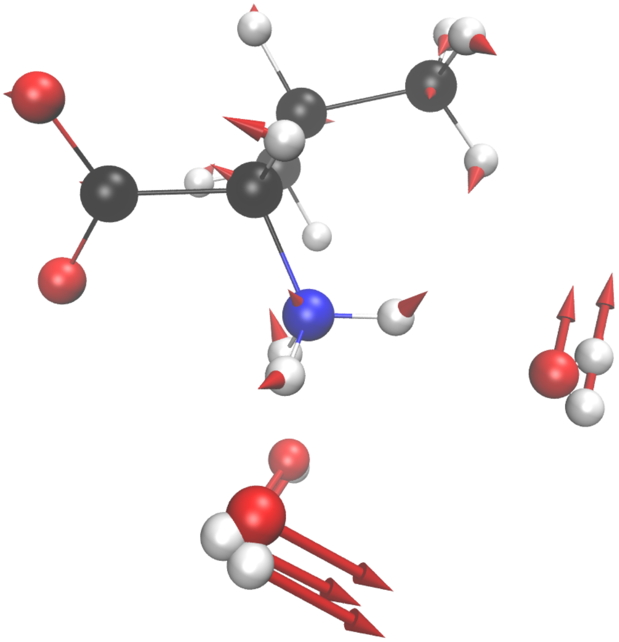
AIMD: HB-bend-U1 88



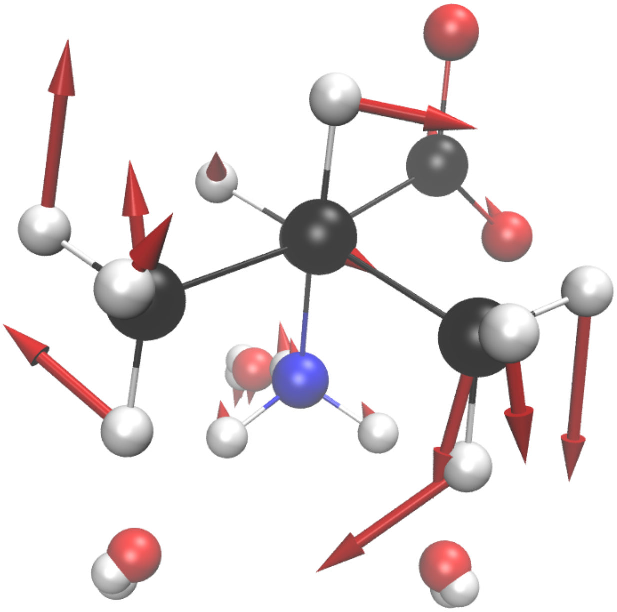
AMOEBA: HB-bend-U1 75



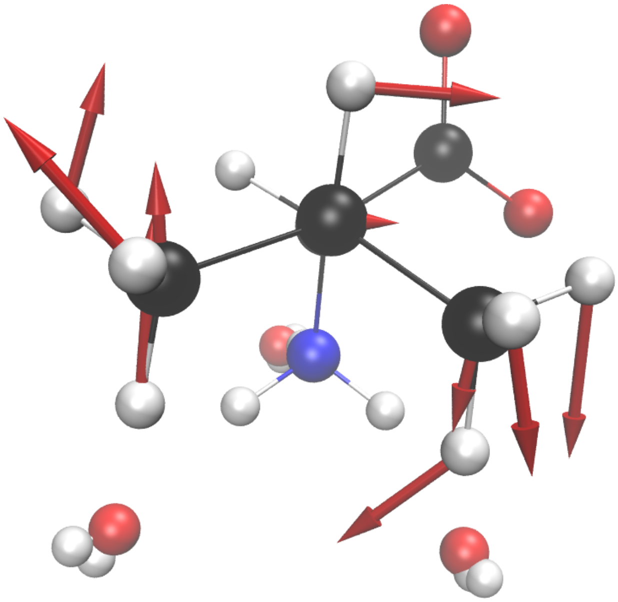
AIMD: HB-bend-X 87



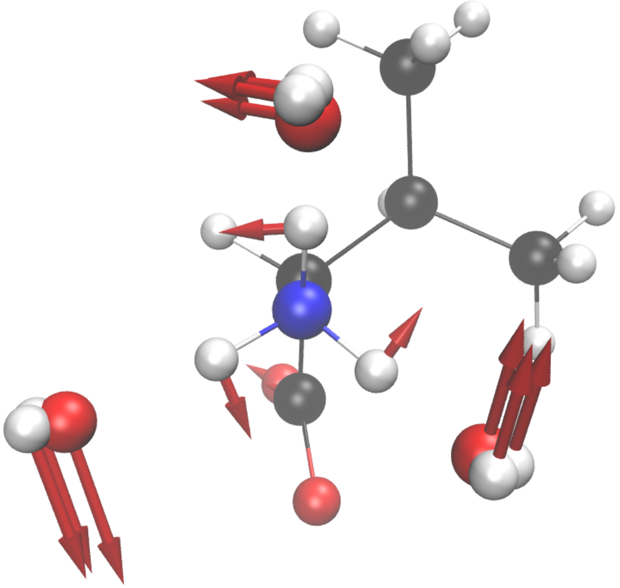
AMOEBA: HB-bend-X 73



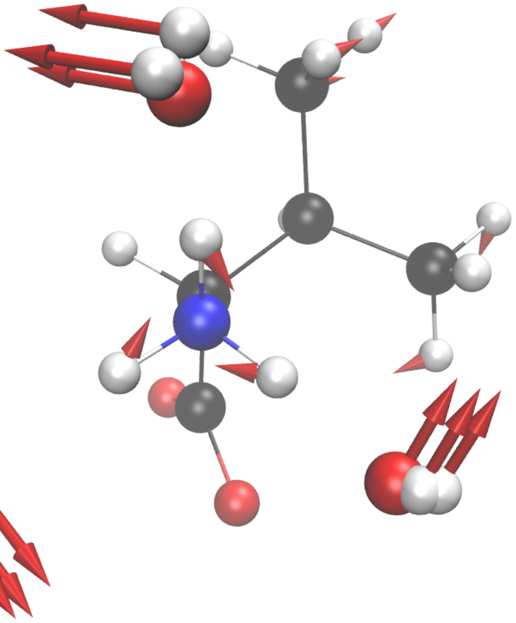
AIMD: R-libration-I 85



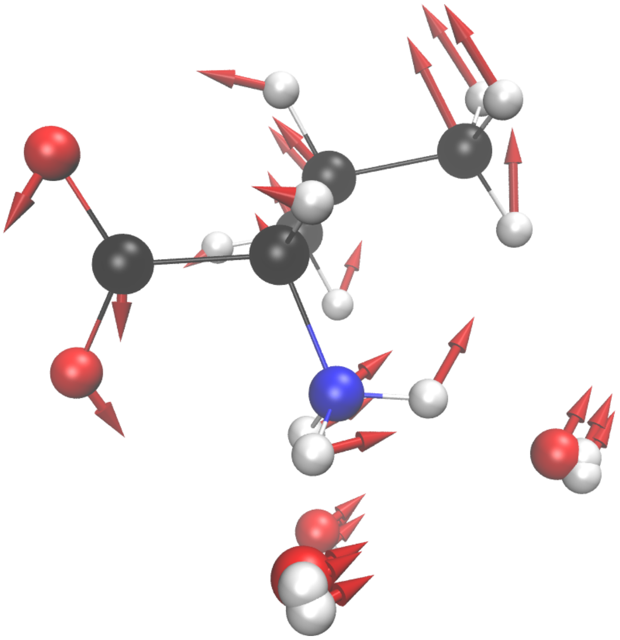
AMOEBA: R-libration-I 75



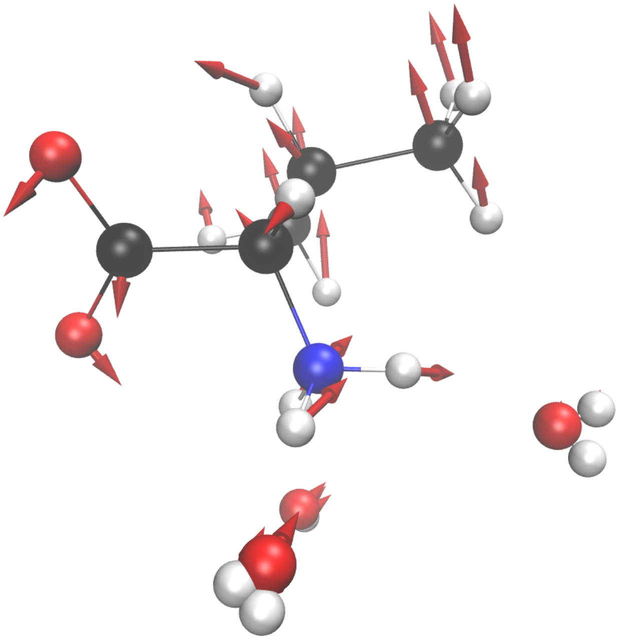
AIMD: HB-bend-X3 83



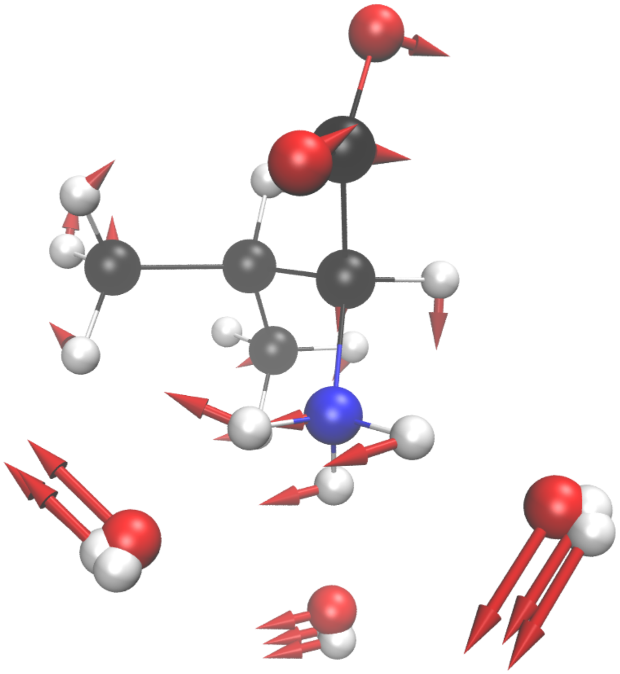
AMOEBA: HB-bend-X3 63



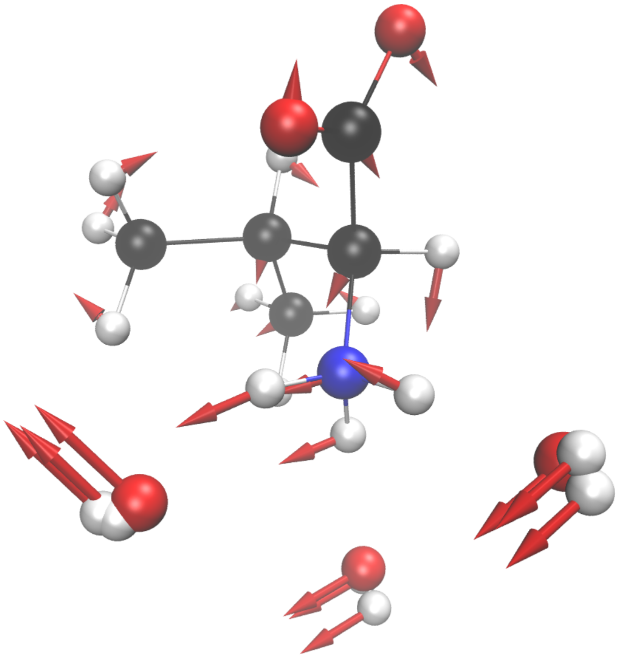
AIMD: cage-libration-II 70



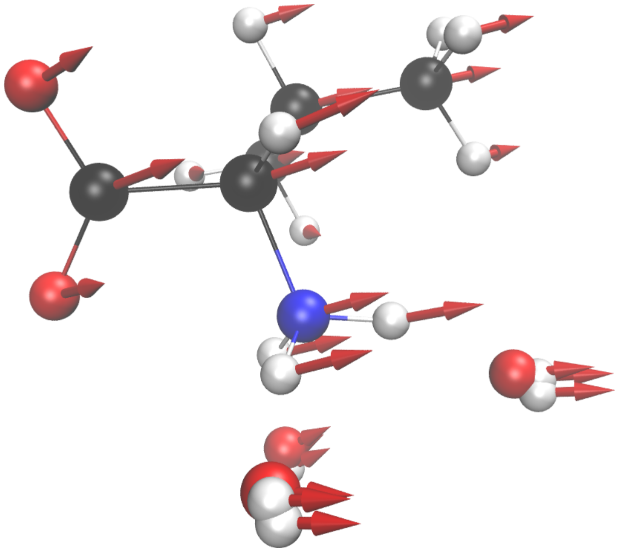
AMOEBA: cage-libration-II 67



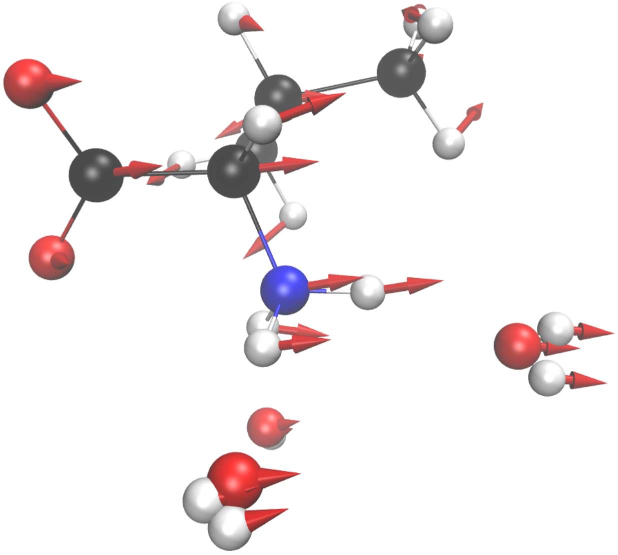
AIMD: cage-libration-I 68



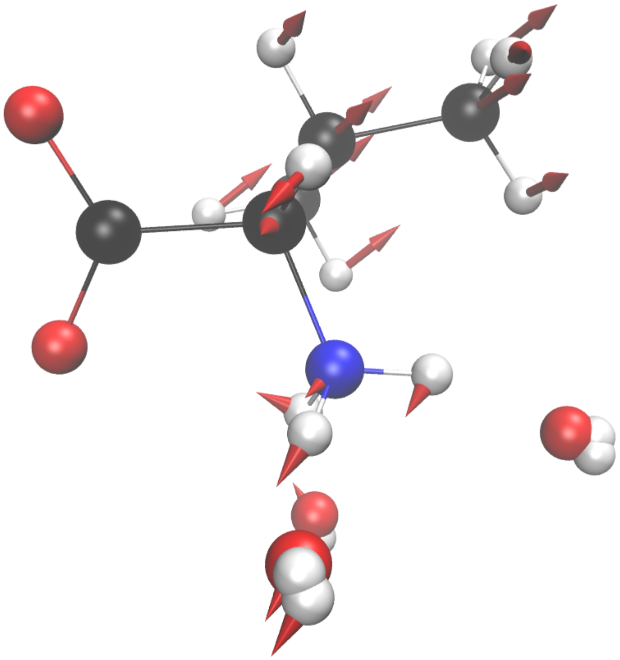
AMOEBA: cage-libration-I 52



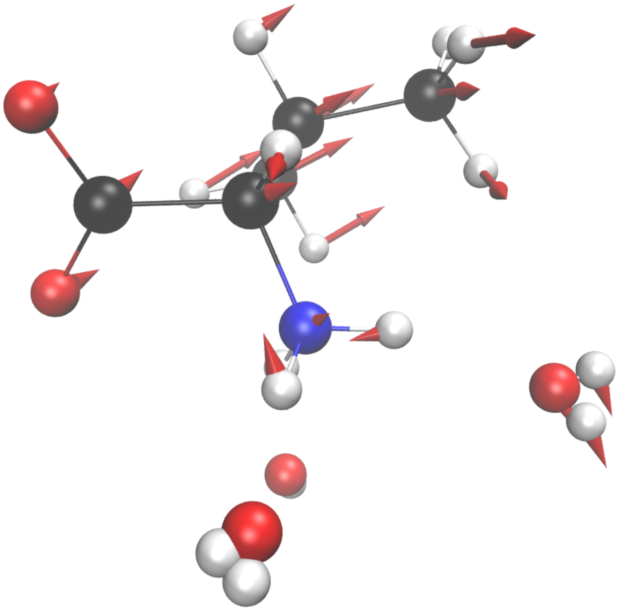
AIMD: cage-rattling-II 56



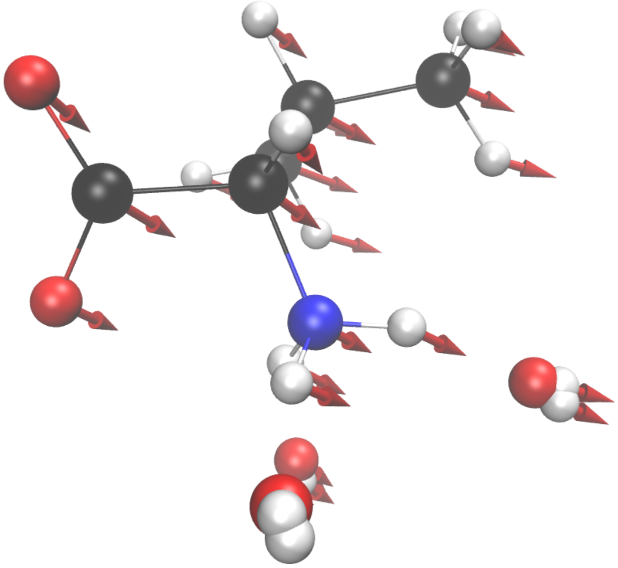
AMOEBA: cage-rattling-def 51



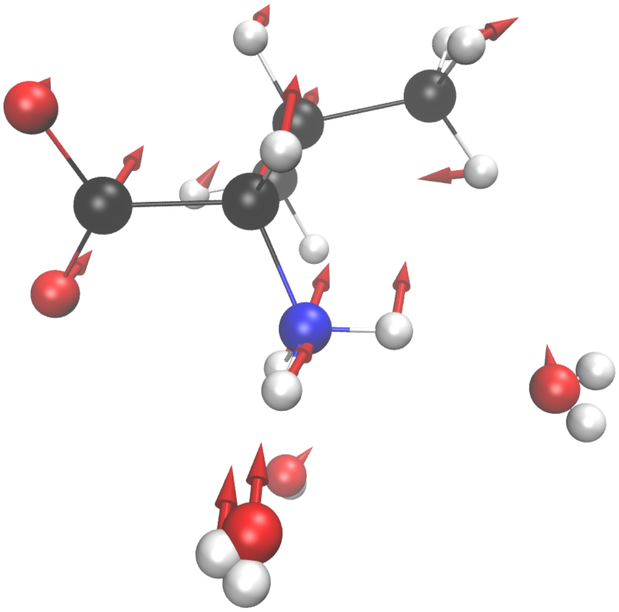
AIMD: R-libration-II 55



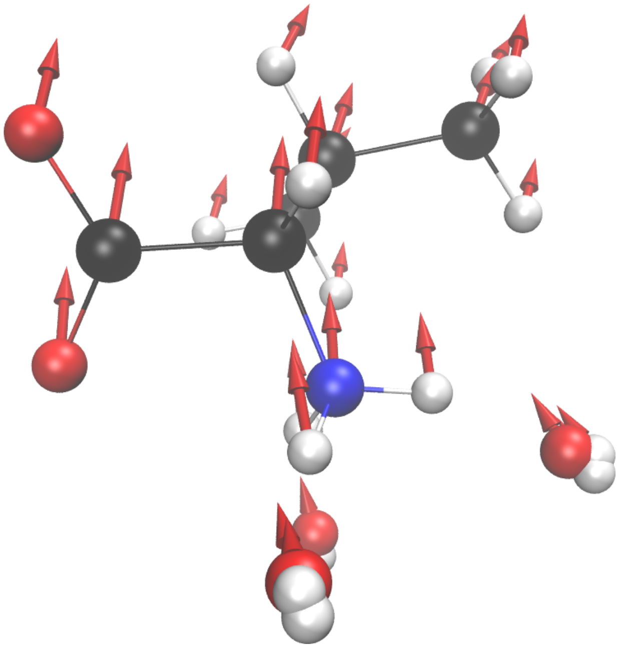
AMOEBA: R-libration-II 38



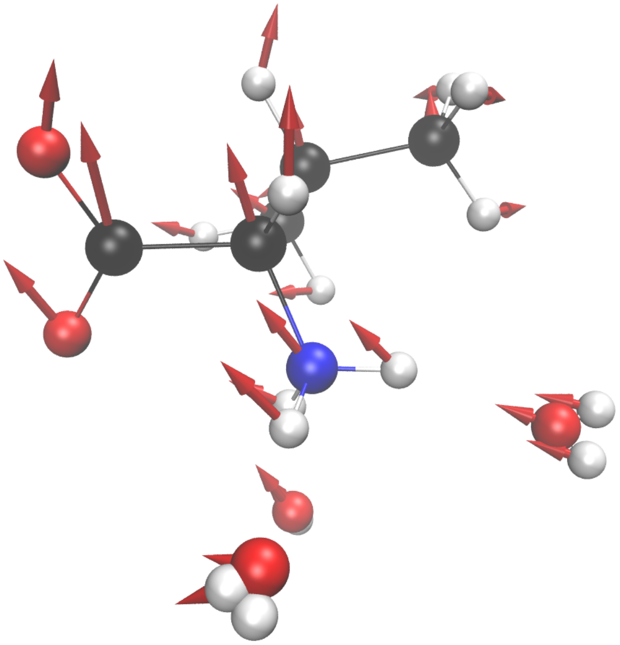
AIMD: cage-rattling-I 54



AMOEBA: cage-rattling-def 41



AIMD: cage-rattling-III 50



AMOEBA: cage-rattling-def 36

Modes of Glycine in the SSC(-)

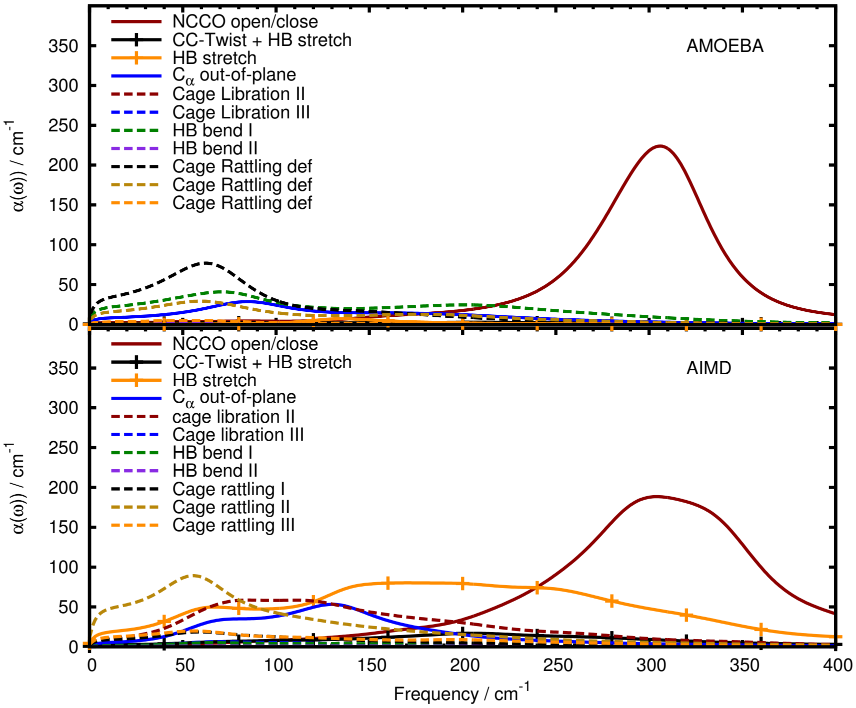


Figure S51: THz mode intensities of glycine within the supermolecular solvation complex at the carboxylate group (SSC(-)) computed via AMOEBA (top) and AIMD (bottom) solvated by 30 water molecules. The top panel shows high THz intensities, the bottom shows low intensities.

Modes of Valine in the SSC(-)

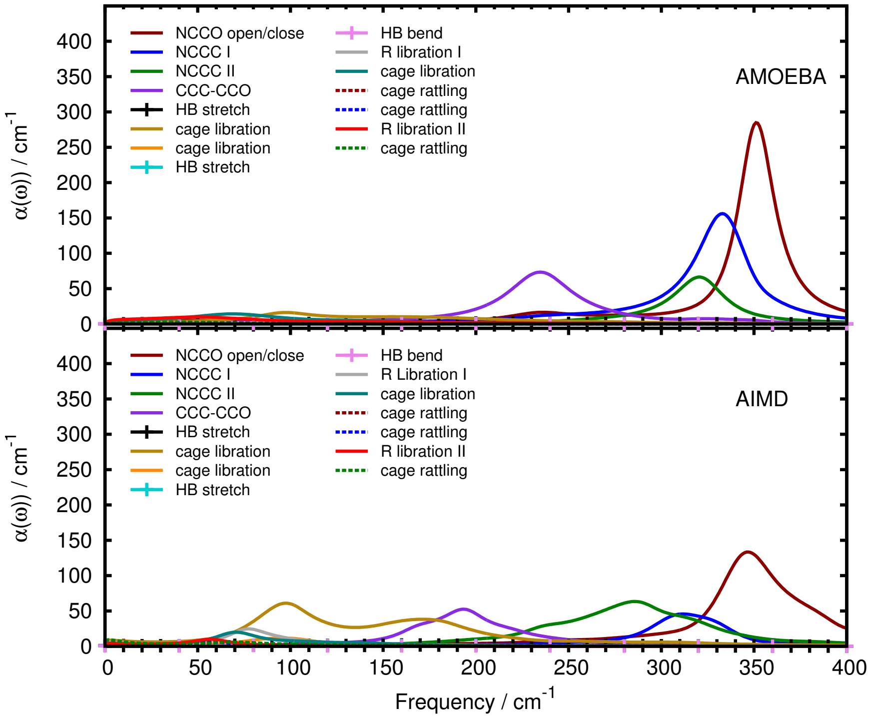


Figure S52: THz mode intensities of valine within the supermolecular solvation complex at the carboxylate group (SSC(-)) computed via AMOEBA (top) and AIMD (bottom) solvated by 60 water molecules. The top panel shows high THz intensities, the bottom shows low intensities.

References

(S1) Heyden, M.; Sun, J.; Funkner, S.; Mathias, G.; Forbert, H.; Havenith, M.; Marx, D. Dissecting the THz spectrum of liquid water from first principles via correlations in time and space. *Proc. Nat. Acad. Sci.* **2010**, *107*, 12068–12073.

(S2) Heyden, M.; Sun, J.; Forbert, H.; Mathias, G.; Havenith, M.; Marx, D. Understanding the Origins of Dipolar Couplings and Correlated Motion in the Vibrational Spectrum of Wat. *J. Phys. Chem. Lett.* **2012**, *3*, 2135–2140.