Supporting Information

Section 1: Molecular Dynamics

Molecular dynamics (MD) simulations were performed using GROMACS 2021.2 on the β-alanine-water-IPA system.[1] Simulation boxes were constructed to represent two different weight fractions of β-alanine: 0.15 and 0.5 at different IPA weight fractions, 0.56, 0.37, 0.18, 0, and 0.2, 0.05, 0, respectively. The SPC/E water model was employed together with the OPLS-AA force field, both available within the GROMACS database. OPLS-AA (Optimized Potential for Liquid Simulations All Atom) is an all-atom force field, reliable for amino acids, that accounts for both inter- and intra-molecular interactions, with functional forms described by van der Waals, electrostatic, and mixing rules, as summarized below.

$$U_{LJ}(r_{ij}) = 4\,\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

$$U_{Col}(r_{ij}) = \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}}$$

Here, the minimum value of potential energy between atoms i and j is given by ϵ_{ij}, σ is the distance at which the potential energy is zero, r is the distance between atoms, ε_0 is the electric constant, and q is the charge on atoms. Lorentz-Berthelot mixing rules were used to calculate ϵ_{ij} and σ_{ij} from independent values.

The β-alanine topology was generated using gmx pdb2gmx with the OPLS-AA force field and SPC/E water model, with hydrogen atoms regenerated. [2] Isopropanol (IPA) parameters were generated using ACPYPE and manually fine-tuned to OPLS-AA atom types. A cubic simulation box of 6 nm was constructed by first placing the β-alanine molecules with gmx insert-molecules, then inserting the IPA molecules with 100 placement attempts per molecule. The box edge length is 6 times the 1 nm real-space cutoff, ensuring rapid convergence while allowing proper temperature fluctuations. The system was solvated using gmx solvate with the SPC/E 216-molecule equilibrated water box; the topology file was updated to include IPA parameters and reflect the final molecular counts. The numbers of each component were determined by scaling experimental solubility data of β-alanine in water-IPA mixtures by the box volume to achieve target concentrations, as detailed in **Table S1.**

Classical force fields predict solubilities that deviate from experiment, since rigorous evaluation of solubility requires a full thermodynamic cycle involving crystal lattice dissociation, solvent cavity formation, and solute insertion. In this cycle, the solvation free energy is a critical contribution. For alanine, the experimental hydration free energy is reported as +1.94 kcal mol^{-1}, while OPLS-AA with SPC/E water model predicts +2.27 kcal mol^{-1}, in reasonable agreement but slightly less favorable than experiment. [3-5] For IPA, the experimental

hydration free energy is –4.76 kcal mol^{-1}, whereas calculated values give –4.15 kcal mol^{-1}, again showing modest deviations but correct qualitative trends.[6] These benchmarks highlight that while OPLS captures relative solvation trends, deviations from experimental free energies can propagate into predicted desolvation barriers and growth kinetics.

After box assembly, energy minimization was performed using the steepest descent algorithm (step size 0.01 nm) to remove unfavorable contacts until the maximum force fell below 100 kJ/mol/nm. Following energy minimization, the system underwent NVT and NPT equilibration for 1 ns each, followed by a 5 ns production run, for a total simulation of 7 ns. The equations of motion were integrated using the leapfrog algorithm with a 2 fs timestep. All bonds involving hydrogen atoms were constrained using the LINCS algorithm. [7, 8] The system temperature was maintained at 300K using the V-rescale thermostat with a coupling time constant of τ_T = 0.1 ps. The β-alanine and the mixed solvent were coupled to separate temperature baths to ensure proper thermalization. Long-range electrostatics were treated with the Particle Mesh Ewald method (1 nm real-space cutoff, Verlet scheme) with a Van der Waals cutoff of 1.0 nm. Periodic boundary conditions were applied in all three dimensions.[9] Pressure coupling in the NPT ensemble employed the Parrinello–Rahman barostat to maintain an isotropic pressure of 1.0 bar, applying a coupling time constant of τ_P = 2.0 ps and the isothermal compressibility of $5.75 \times 10^{-5} bar^{-1}$. The exact molecular counts for each simulated composition are detailed in Table S1. These trajectories served as the basis for the analyses of RDF, SDF, coordination number, binding energy, and activation energy presented in this work.

Section 2: MD simulation output analysis

Radial Distribution Functions (RDFs)

Radial distribution functions (RDFs) were calculated using the center of mass of β-alanine as the reference. The 'gmx rdf' command, coupled with index files (.ndx) specifying water and IPA molecules, was used to compute RDFs. The first minima in the water RDFs across all simulations occurred at ~0.65 nm, which was taken as the solvation shell cutoff distance for subsequent analyses. Per-molecule RDFs were computed over all frames of the 5 ns production trajectory at 1 ps intervals and then averaged across all β-alanine molecules in each system. To examine temporal evolution, the production trajectory was partitioned into sequential time blocks and RDFs were recomputed for each block (**Figure S7**). A consistent approach was used for the BAL – BAL in **Figure S2**; these solute-solute RDFs were calculated separately for **fully solvated** (FS) and **partially desolvated** (PS) states.

Spatial Distribution Functions (SDF)

Spatial distribution functions (SDFs) were generated to visualize the three-dimensional distribution of water and IPA molecules around β-alanine using the TRAVIS package.[10, 11] Distributions were averaged over all simulation frames with a binning of 100 points per axis and normalized with respect to the bulk density of

each solvent species. The resulting relative density maps were visualized in VMD, where isovalue adjustment was used to highlight specific density levels for water and IPA.

Solvation State Classification

At each simulation frame, the pairwise distance between all β-alanine molecules, the force acting on each β-alanine molecule, and the number of water and IPA molecules within the solvation shell were computed. Based on this information, molecules were classified as **fully solvated** or **partially desolvated** as shown in **Figure 1** of the main text. Molecules showing an increasing cumulative moving average (CMA) of the force norm with respect to time, a pairwise separation shorter than twice the solvation shell thickness, and the minimum number of water molecules in the shell were identified as **partially desolvated**. The CMA, or running average, at each time step was defined as the arithmetic mean of the force magnitudes from the first frame to the current frame. Unlike simple moving averages that use a fixed window, the CMA incorporates every new data point, making it ideal for tracking long-term trends and overall molecular behavior rather than short-term fluctuations. Molecules showing a constant or decreasing running average of the force norm, a pairwise separation larger than twice the solvation shell thickness, and the maximum number of water molecules in the shell were identified as **fully solvated**. These classifications were then used to analyze activation barriers.

Binding Energy Calculations

The binding energy of β-alanine with water and IPA molecules within the solvation shell was calculated using the MM-PBSA method, which applies a linear Poisson–Boltzmann solver to decompose energy contributions.[12] Six equally spaced frames (every 1000 ps from 0 to 5000 ps) were sampled, and the per-molecule contributions of water and IPA to binding energy were extracted separately. These values were then plotted against solvent composition, as shown in Figure 5c of the main text and in Figure S3 of the SI.

Double-Well Potential Calculations

Double-well potential (DWP) calculations were performed following the protocol described in our earlier work.[13] The construction follows the approach of Shustorovich, originally developed for adsorption, in which the dominant energy contribution arises from electron exchange between the adsorbent and the adsorbate and is identified by superimposing the potential energy profiles (potential wells) of the adsorbate near and away from the surface.[14] In crystallization, the dominant energy contribution instead comes from the energy required to remove the solvation shell from a solute molecule; the same superposition of potential wells, here of the fully solvated and partially desolvated states, therefore yields the desolvation activation barrier. β-alanine molecules were classified as either fully solvated or partially desolvated based on the pairwise distance, and potential energy profiles were computed separately for each state and superimposed (**Figure S4**).

Because the two solvation states contain different numbers of water and IPA molecules, the potential energy surfaces were normalized to account for solvent occupancy. To accurately account for solvent occupancy, the partially desolvated profile was shifted upward by the binding energy required to remove the excess solvent (water) and antisolvent (IPA) molecules, calculated utilizing the MM-PBSA method. The activation energy was then defined as the difference between the minimum of the fully solvated profile and the intersection point with the partially desolvated profile.

To extend this analysis, the molecular configuration at the exact intercept point was rotated to all possible orientations, and the lowest-energy configuration was identified. The rotational landscape (Figure S5) confirms that lattice integration is orientation-dependent, with a well-defined energy minimum identifying the lattice-compatible configuration. The lowest was used to construct the second peak in the energy profile.

The crystallization coordinate is plotted on the x-axis for figures 6a and 6b, showing this multi-step pathway. A coordinate of 0 denotes the fully solvated state, where the distance between BAL molecules is large and intermolecular interactions are negligible. A coordinate of 1 refers to the fully desolvated state, corresponding to the exact distance and alignment between molecules in the crystal lattice. The intermediate states, governed by the desolvation and orientational barriers, were positioned at the intersections of the superimposed energy profiles.

Hydrogen Network Analysis

The hydrogen bond network within the BAL solvation shell was analyzed dynamically across the trajectories using a graph-theoretical approach.[15] To properly account for solvent exchange and bulk diffusion over time, the instantaneous solvation shell was redefined at every trajectory frame to include all water molecules residing within 0.65 nm of the target BAL residue. Within this dynamic network, each resident water molecule was treated as a node, and each intermolecular water–water hydrogen bond was treated as an edge. Hydrogen bonds were continuously identified, applying a donor–acceptor distance cutoff of 0.35 nm and a donor–hydrogen–acceptor angle cutoff of 150°. The mean node degree, ⟨k⟩ was computed for each frame utilizing the handshaking lemma.[16] The resulting network statistics across all seven compositions are reported in **Figure S6f**.

Section 3: Probability of crystallization

The transition of β-alanine molecules from a fully solvated state (A) to a crystalline state (C) through dense states (D_1, D_2) can be written as:

$$A \overset{k_1}{\rightarrow} D_1 \quad (1)$$

$$D_1 \overset{k_D}{\rightarrow} D_2 \quad (2)$$

$$D_1 \overset{k_2}{\rightarrow} C \tag{3}$$

The system initially at a fully solvated state (A) strips the solvent molecules to form the partially desolvated dense state (D_1). The dense state refers to an amorphous cluster of partially desolvated solute molecules that have aggregated but lack long-range crystalline order, where BAL molecules are in proximity due to reduced solvent coverage, but their relative arrangement remains disordered. However, after this point, because the system is at its lowest energy in the dense state, it either remains in a second dense state (D_2) or forms a crystal lattice (C). The crystal state refers to an ordered cluster in which the solute molecules have paid the energetic cost of reorienting from the disordered cluster to adopt the periodic arrangement characteristic of the crystal lattice. Rate constants k_1, k_D, and k_2 are calculated using transition-state theory based on the energy values from Figure 6. Finally, the rate equations are written as follows:

$$\frac{d[A]}{dt} = -\overline{k_1}[A] \tag{4}$$

$$\frac{d[D_1]}{dt} = \overline{k_1}[A] - \overline{k_D}[D_1] - \overline{k_2}[D_1] \tag{5}$$

$$\frac{d[D_2]}{dt} = \overline{k_D}[D_1] \tag{6}$$

$$\frac{d[C]}{dt} = \overline{k_2}[D_1] \tag{7}$$

Rate constants are made dimensionless by dividing them with the least value of all rate constants, while time is made dimensionless by multiplying it. Square brackets represent the probability of finding the molecule in the respective state.

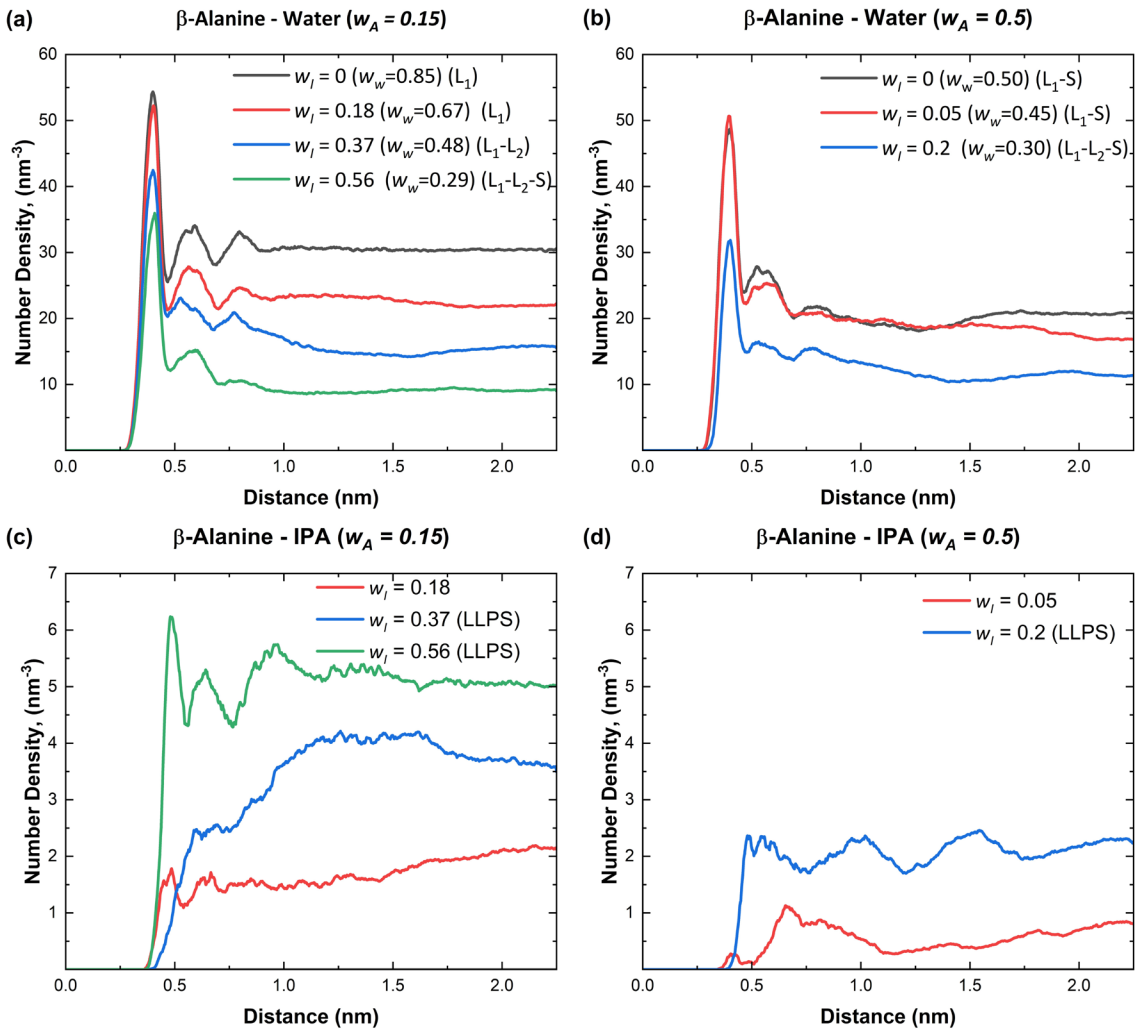

Figure S1: Number density radial distribution of water and IPA around β-alanine at two different weight fractions $\boldsymbol{w_A}$=0.15 and $\boldsymbol{w_A}$=0.5. (a) RDF of water around β-alanine at $\boldsymbol{w_A}$=0.15 (b) RDF of water around β-alanine at $\boldsymbol{w_A}$=0.5 (c) RDF of IPA around β-alanine at $\boldsymbol{w_A}$=0.15 (d) RDF of IPA around β-alanine at $\boldsymbol{w_A}$=0.5. The legend denotes the weight fraction of the antisolvent IPA ($\boldsymbol{w_I}$) and water ($\boldsymbol{w_w}$).

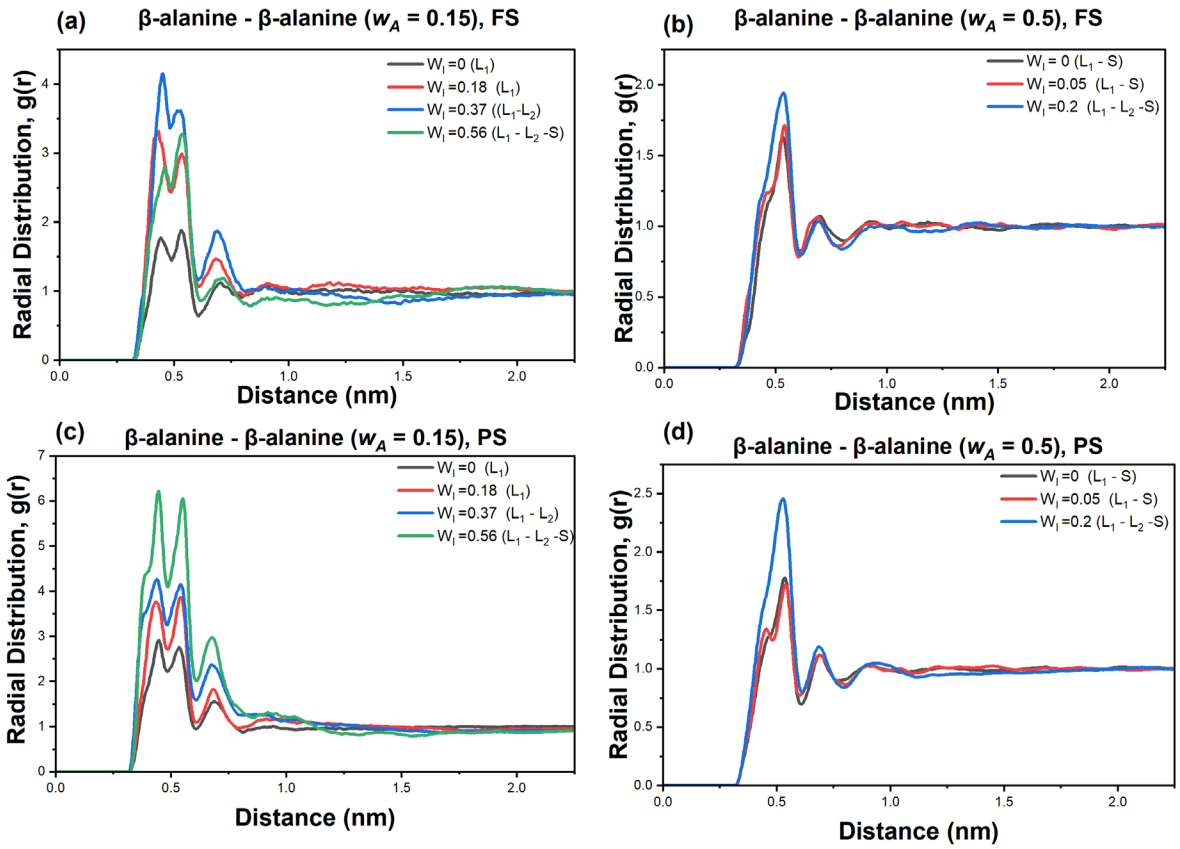

Figure S2 : Radial distribution functions of β-alanine around β-alanine (BAL-BAL) for fully solvated (FS) states at (a) w_A = 0.15 and (b) w_A = 0.5, and for partially desolvated (PS) states at (c) w_A = 0.15 and (d) w_A = 0.5 across different IPA weight fractions. At both solute concentrations, PS states exhibit enhanced BAL-BAL association compared to FS states, with the strongest clustering observed in the LLPS regime.

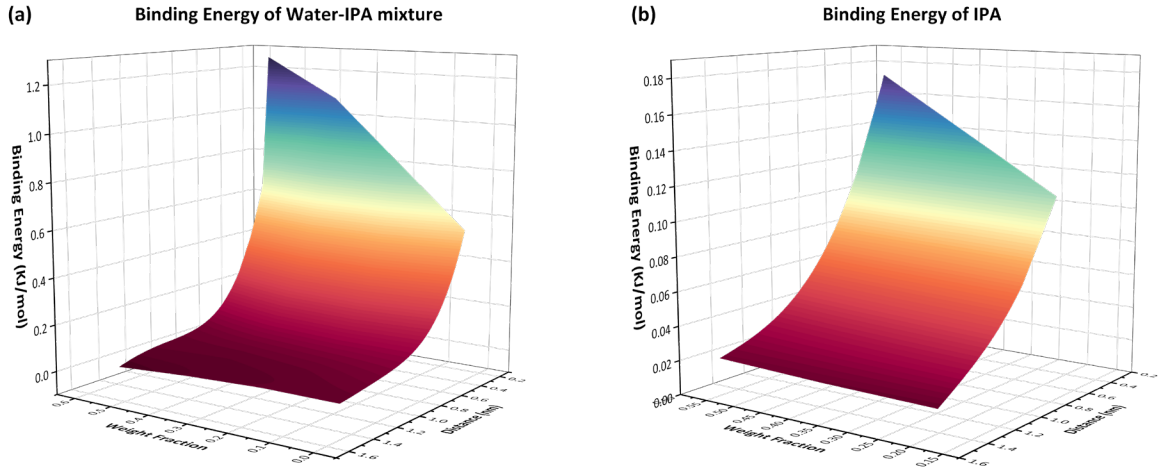

Figure S3: (a) Binding energy of water-IPA mixture to the solute molecule (b) Binding energy of IPA molecules to the solute. Binding energies are calculated using the MM-PBSA method across solvent compositions.

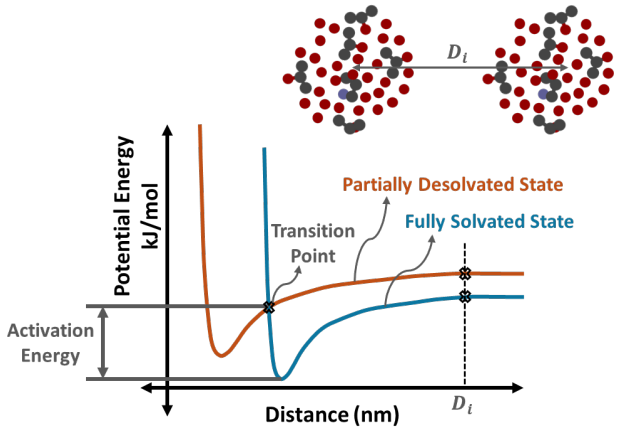

Figure S4: Double well potential approach. At any distance $\boldsymbol{D_i}$ the system of molecules is separated as shown and potential energy is calculated at each such distances. Superimposing the two energy profiles gives us the activation energy required to shift from a fully solvated to a transition point.

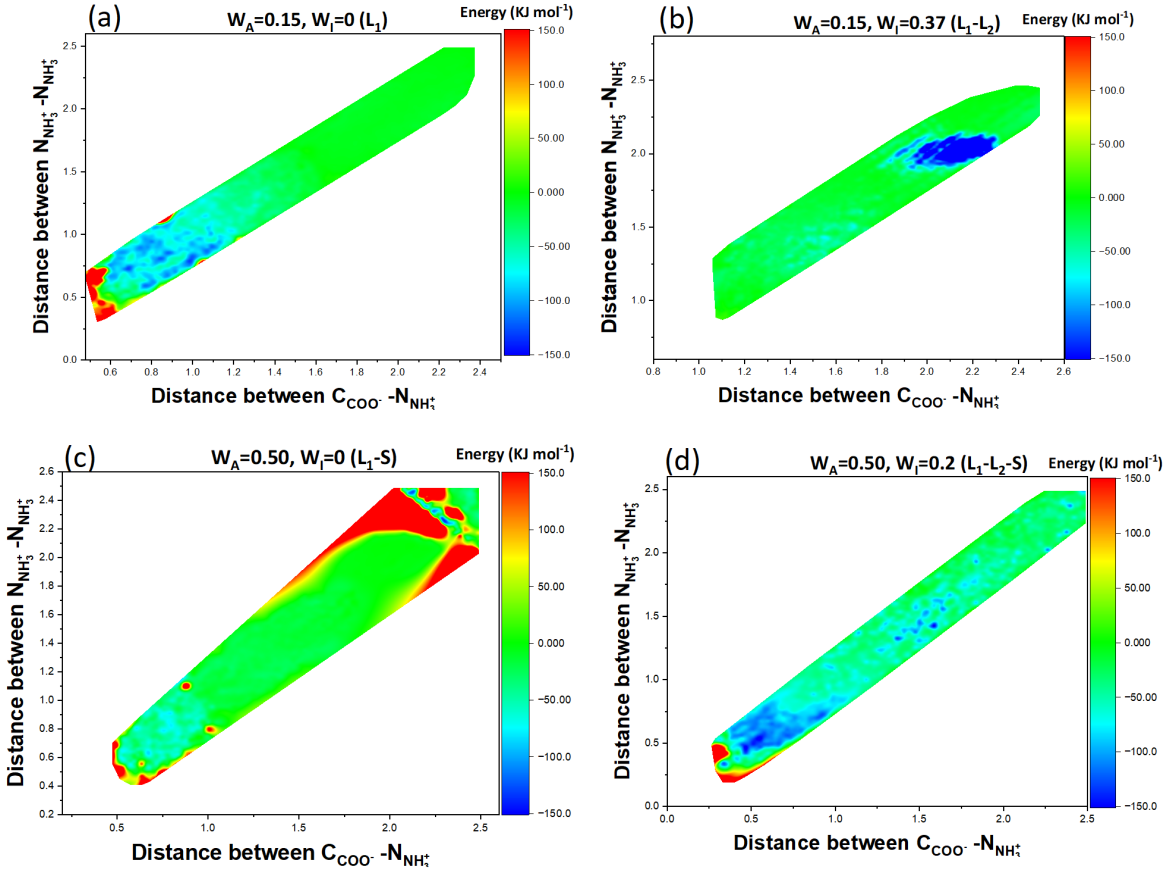

Figure S5: Rotational energy landscapes of β-alanine (BAL) in the first solvation shell as a function of intermolecular distance, for four solvent compositions: (a) $\boldsymbol{w_A}$ = 0.15, $\boldsymbol{w_I}$ = 0; (b) $\boldsymbol{w_A}$= 0.15,$\boldsymbol{w_I}$ = 0.37; (c) $\boldsymbol{w_A}$ = 0.50, $\boldsymbol{w_I}$= 0; (d) $\boldsymbol{w_A}$ = 0.50, $\boldsymbol{w_I}$= 0.2. The x-axis is the distance between the carboxyl carbon (C_{COO^-}) of one BAL molecule and the

ammonium nitrogen (N_{NH3}^+) of the neighboring BAL; the y-axis is the $N_{NH3}^+ - N_{NH3}^+$ distance between the two BAL zwitterions. The color scale is the total pair interaction energy in kJ mol^{-1}.

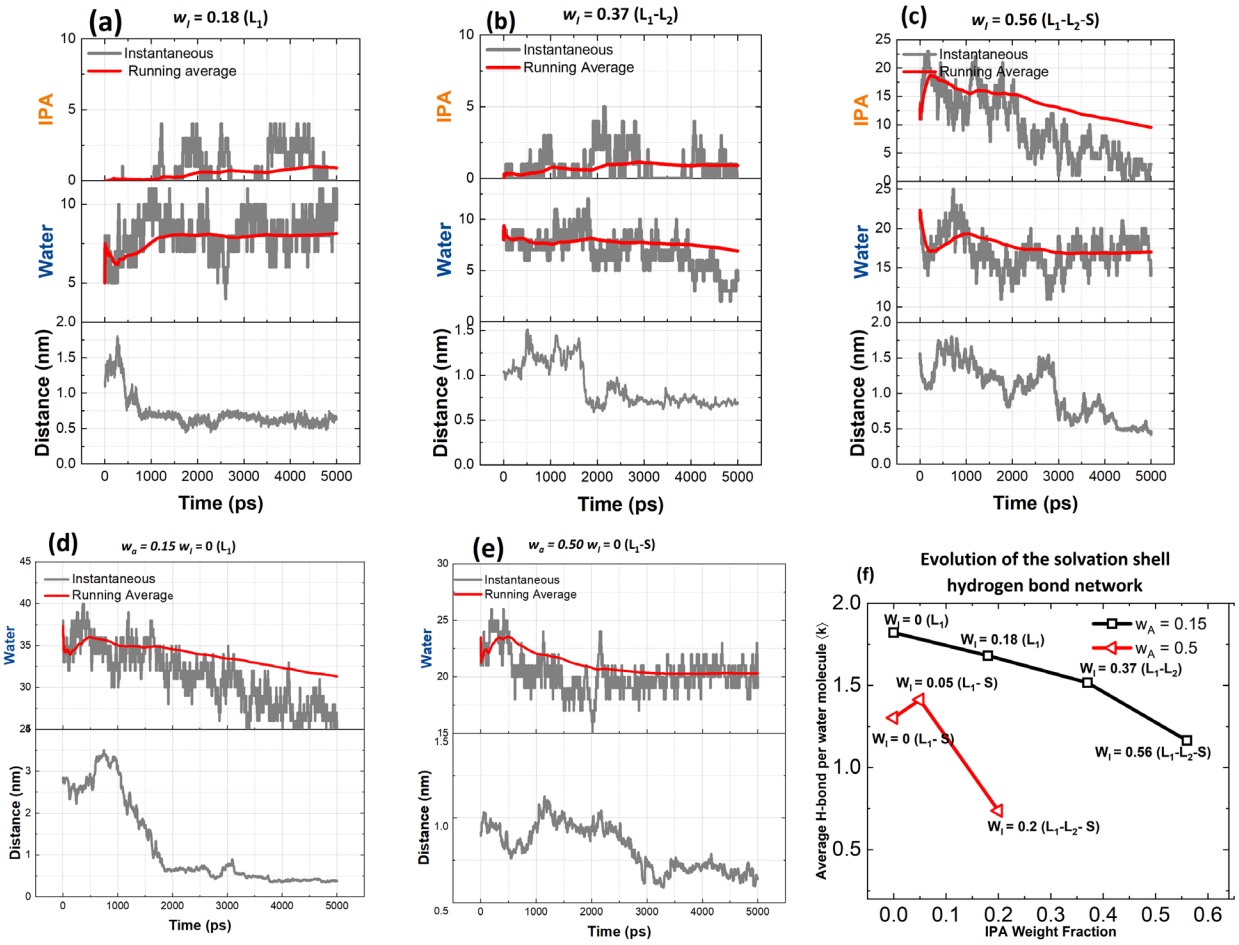

Figure S6: Time evolution of solvent coordination numbers and solute–solute distance for β-alanine across varying IPA weight fractions. The top row shows systems at a solute weight fraction of w_A = 0.15 for (a) w_I = 0.18 (L_1), (b) w_I = 0.37 (L_1-L_2), and (c) w_I = 0.56 (L_1-L_2-S). The bottom row shows pure water conditions for (d) w_A = 0.15, w_I = 0 (L_1) and (e) w_A = 0.50, w_I = 0 (L_1-S). Panels (a)–(e) display instantaneous (gray lines) and running-average (red lines) coordination numbers for IPA (top sub-panels), water (middle sub-panels), and the pairwise distance between β-alanine molecules (bottom sub-panels). Panel (f) shows the evolution of the mean hydrogen-bond degree ⟨k⟩ per water molecule within the primary solvation shell as a function of IPA weight fraction for both solute concentrations.

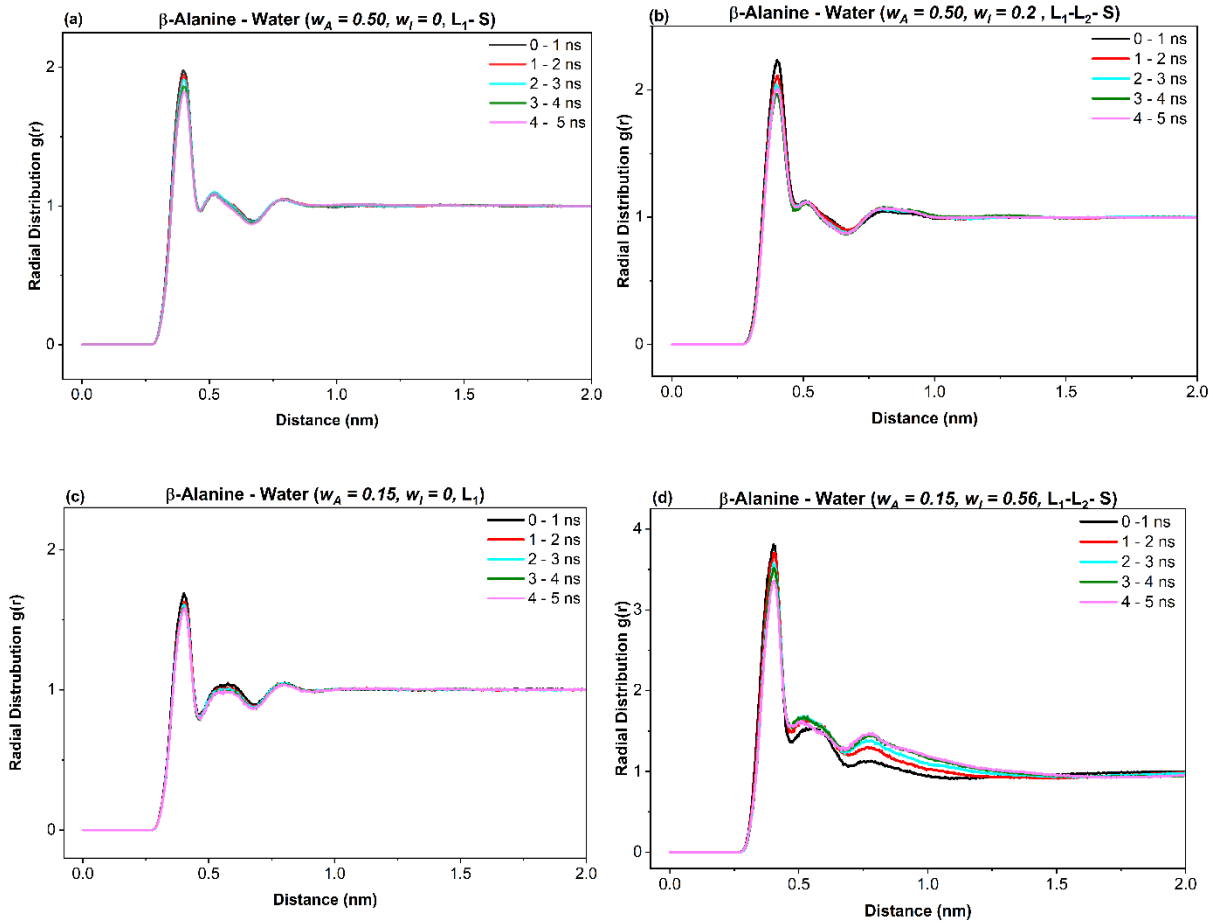

Figure S7: Time-block analysis of β-alanine–water radial distribution functions for four compositions spanning the non-LLPS, weak-LLPS, and strong-LLPS regimes: (a) w_A = 0.50, w_I = 0 (non-LLPS, L_1–S); (b) w_A = 0.50, w_I = 0.2 (weak LLPS, L_1–L_2–S); (c) w_A = 0.15, w_I = 0 (non-LLPS, L_1); and (d) w_A = 0.15, w_I = 0.56 (strong LLPS, L_1–L_2–S)

Table S1: Simulated system compositions, detailing the target weight fractions and the corresponding number of β-alanine, water, and isopropanol (IPA) molecules in each simulation box.

Target β-alanine (wt fraction)	Target IPA (wt fraction)	β-alanine Molecules	Water Molecules	IPA Molecules
0.15	0.56	113	1055	638
0.15	0.376	122	1771	430
0.15	0.18	122	2552	224
0.15	0	124	3475	0
0.50	0.20	432	1299	254
0.50	0.05	429	1914	64
0.50	0	431	2146	0

Table S2: Average coordination numbers of water and IPA molecules within the first solvation shell of β-alanine. Coordination numbers were obtained by integrating the number density RDFs to the first shell minimum. Values are reported as mean ± standard deviation across all BAL molecules in each system.

β-alanine (wt fraction)	IPA (wt fraction)	Coordination number (Water)	Coordination number (IPA)
0.15	0	30.9 ± 2.2	0
0.15	0.18	24.8 ± 2.1	1.1 ± 0.4
0.15	0.376	21.2 ± 2.2	1.7 ± 0.6
0.15	0.56	15.7 ± 2.0	3.2 ± 0.9
0.50	0	22.6 ± 2.6	0
0.50	0.05	20.4 ± 2.8	0.5 ± 0.5
0.50	0.2	13.1 ± 2.3	1.7 ± 0.7

Table S3: Average per-molecule MM-PBSA binding energy of water to β-alanine across all simulated compositions. Values are reported as mean ± standard deviation across all water molecules. Binding energies were computed using 6 equally spaced frames (every 1000 ps) of the production trajectory.

β-alanine (wt fraction)	IPA (wt fraction)	Number of Water Molecules	Binding Energy (kJ/mol)
0.15	0	3475	-0.015 ± 0.126
0.15	0.18	2552	-0.010 ± 0.137
0.15	0.37	1771	-1.905 ± 1.421
0.15	0.56	1055	-0.040 ± 0.210
0.50	0	2146	0.069 ± 0.150
0.50	0.05	1914	-0.011 ± 0.132
0.50	0.20	1299	0.006 ± 0.110

References

(1) Abraham, M. J. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19.

(2) Burley, S. K.; Bhikadiya, C.; Bi, C.; Bittrich, S.; Chen, L.; Crichlow, G. V.; Christie, C. H.; Dalenberg, K.; Di Costanzo, L.; Duarte, J. M.; et al. RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res* **2021**, *49* (D1), D437-d451. DOI: 10.1093/nar/gkaa1038 From NLM.

(3) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *The Journal of Physical Chemistry* **1987**, *91* (24), 6269-6271. DOI: 10.1021/j100308a038.

(4) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society* **1996**, *118* (45), 11225-11236. DOI: 10.1021/ja9621760.

(5) Shirts, M. R.; Pande, V. S. Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. *J Chem Phys* **2005**, *122* (13).

(6) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *Journal of chemical theory and computation* **2010**, *6* (5), 1509-1519.

(7) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **1997**, *18* (12), 1463-1472. DOI: https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

(8) Van Gunsteren, W. F.; Berendsen, H. J. C. A Leap-frog Algorithm for Stochastic Dynamics. *Molecular Simulation* **1988**, *1* (3), 173-185. DOI: 10.1080/08927028808080941.

(9) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98* (12), 10089-10092. DOI: 10.1063/1.464397 (acccessed 4/20/2026).

(10) Brehm, M.; Kirchner, B. TRAVIS-a free analyzer and visualizer for Monte Carlo and molecular dynamics trajectories. ACS Publications: 2011.

(11) Brehm, M.; Thomas, M.; Gehrke, S.; Kirchner, B. TRAVIS—A free analyzer for trajectories from molecular simulation. *The Journal of Chemical Physics* **2020**, *152* (16). DOI: 10.1063/5.0005078 (acccessed 10/3/2025).

(12) Kumari, R.; Kumar, R.; Consortium, O. S. D. D.; Lynn, A. g_mmpbsa- A GROMACS tool for high-throughput MM-PBSA calculations. *Journal of chemical information and modeling* **2014**, *54* (7), 1951-1962.

(13) Dighe, A. V.; Podupu, P. K. R.; Coliaie, P.; Singh, M. R. Three-Step Mechanism of Antisolvent Crystallization. *Crystal Growth & Design* **2022**, *22* (5), 3119-3127. DOI: 10.1021/acs.cgd.2c00014.

(14) Shustorovich, E. Energetics of metal-surface reactions: Back-of-the-envelope theoretical modeling. *Journal of Molecular Catalysis* **1989**, *54* (3), 301-311.

(15) Bakó, I.; Bencsura, Á.; Hermannson, K.; Bálint, S.; Grósz, T.; Chihaia, V.; Oláh, J. Hydrogen bond network topology in liquid water and methanol: a graph theory approach. *Physical Chemistry Chemical Physics* **2013**, *15* (36), 15163-15171, 10.1039/C3CP52271G. DOI: 10.1039/C3CP52271G.

(16) Newman, M. *Networks: An Introduction*; Oxford University Press, 2010. DOI: 10.1093/acprof:oso/9780199206650.001.0001.