

Influence of α -Helical Content on the Thermodiffusion of Apomyoglobin

Binny A. Rudani,^{†,‡} Steffen Docter,^{‡,‡} Stephan Schott-Verdugo,[‡] Johan
Buitenhuis,[†] Andreas M. Stadler,^{*,¶,§} Holger Gohlke,^{*,‡,||} and Simone
Wiegand^{*,†,⊥}

[†]*Institute of Biological Information Processing IBI-4: Biomacromolecular Systems and
Processes, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany*

[‡]*Institute of Bio- and Geosciences IBG-4: Bioinformatics, Forschungszentrum Jülich,
52425 Jülich, Germany*

[¶]*Jülich Centre for Neutron Science JCNS-1: Neutron Scattering and Soft Matter,
Forschungszentrum Jülich GmbH, 52425 Jülich, Germany*

[§]*Institute of Physical Chemistry, RWTH Aachen University, Landoltweg 2, 52056 Aachen,
Germany*

^{||}*Institute for Pharmaceutical and Medicinal Chemistry & Bioeconomy Science Center
(BioSC), Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*

[⊥]*Department für Chemie - Institut für Licht und Materialien, Universität zu Köln, 50939
Cologne, Germany*

[#]*B.A.R and S.D. contributed equally*

E-mail: a.stadler@fz-juelich.de; h.gohlke@fz-juelich.de; s.wiegand@fz-juelich.de

Phone: +49-2461-61-6654

October 7, 2025

Abstract

Apo-myoglobin (Apo-Mb) is an extensively studied model system for investigating protein folding due to its distinct stable native, partially folded molten globule (MG), and unfolded states at acidic pH. This study examines the impact of structural conformational changes on the thermodiffusive behavior of Apo-Mb using the infrared thermal diffusion forced Rayleigh scattering (TDFRS) technique. The conformational states were modulated by varying pH and buffer conditions, with their structural changes confirmed via circular dichroism (CD) spectroscopy. The α -helical content decreased with decreasing pH. The thermodiffusion parameter $\Delta S_T(\Delta T)$, a measure of the temperature sensitivity of the Soret coefficient S_T , also showed a decrease, which is typically related to a decreasing hydrophilicity of the solute. Additionally, the buffer composition significantly influenced the thermodiffusive behavior: phosphate buffer promoted Apo-Mb aggregation through electrostatic screening, whereas acetate buffer favored Apo-Mb solubilization. Microsecond-long discrete protonation state constant pH molecular dynamics (CpHMD) simulations support the experimentally observed, pH- and buffer-dependent changes in α -helical content and highlight the differences in protein-buffer interactions for phosphate buffer versus acetate buffer. In conclusion, a strong correlation was observed between the thermodiffusion parameter $\Delta S_T(\Delta T)$ and the α -helical content, with $\Delta S_T(\Delta T)$ increasing alongside hydrophilicity and α -helical content. These findings highlight the role of structural conformation and buffer environment in modulating the thermodiffusive properties of proteins.

keywords: protein conformation, buffer, α -helical content, hydrophilicity, $\log P$, hydrogen bonds, Soret coefficient

Introduction

Thermodiffusion, first observed by Ludwig¹ and later systematically studied by Soret,² describes mass transport in a mixture driven by a temperature gradient.^{3,4} For large macromolecules and colloidal particles, the term thermophoresis is used. The physical effect provides essential information about solute-solvent interactions and is influenced by the size, charge, and structure of the solvation shell surrounding the solute molecules.⁵⁻⁷ This physical phenomenon is used in microscale thermophoresis (MST) to study biomolecular interactions, which is particularly useful for studying subtle surface changes in proteins associated with ligand binding, (un)folding events, and other biomolecular processes.⁷⁻¹⁰

Thermodiffusion in a binary fluid mixture can be described as mass flux $\vec{j}$, which has two contributions. The first is associated with a concentration gradient $-D\vec{\nabla}c$ and the second with a temperature gradient $-D_T\vec{\nabla}T$. The two contribute to thermodiffusion in opposite ways.¹¹

$$\vec{j} = -\rho D\vec{\nabla}c - \rho c(1-c) D_T\vec{\nabla}T \quad (1)$$

where c is the weight fraction, ρ is the mass density, and D_T and D are the thermal and mass diffusion coefficients, respectively. In steady state, $\vec{j} = 0$ and the concentration gradient induced by an applied temperature gradient is proportional to the ratio between D_T and D . This relationship is quantified by the Soret coefficient, $S_T = D_T/D$. The sign of S_T represents the direction of mass flow: a positive sign means that the solute molecules move toward the colder region (thermophobic behavior), while a negative sign indicates that the solute moves toward the warmer region (thermophilic behavior). However, the sign of S_T is unpredictable, especially for aqueous systems, and theoretical models are still under discussion.^{5,12-17}

In particular, in aqueous systems, the Soret coefficient of the solute undergoes a sign shift from negative to positive as the temperature increases.^{5,5,11,18-26} It is common to describe the thermodiffusive behavior of solutes in water by an empirical expression.²⁰ In this study,

the expression is presented in a modified form,

$$S_T(T) = S_T^\infty + A \exp\left(\frac{-T}{T_0}\right) \quad (2)$$

Here, A is the amplitude that measures the temperature sensitivity of S_T . It is equal to $-S_T^\infty \exp(T^*/T_0)$, where S_T^∞ is the S_T value approached at high temperature, T^* is the temperature value where S_T changes its sign, and T_0 describes the bending of the curve.²⁰ A measure of the temperature sensitivity of S_T , the difference at two temperatures $\Delta S_T(\Delta T)$, is negatively correlated with the logarithm of the 1-octanol/water partition coefficient, $\log P$.^{5,23,27} A negative $\log P$ indicates greater hydrophilicity of the solute.^{28,29} Thus, a highly hydrophilic solute exhibits a higher temperature sensitivity than a hydrophobic one.

However, the concept to calculate $\log P$ -values using heuristic or fragment-based approaches only works for relatively small molecules whose parts are always in contact with the solvent (see Supporting Information S5.3). Deviations from the correlation between $\Delta S_T(\Delta T)$ and $\log P$ are expected for more complex proteins. First, experimental determination of $\log P$ is often not possible because proteins can aggregate in pure water and require a buffer for stabilization. Second, since the hydrophilic surface accessible to the solvent depends on the folding state, a fragment-based computation for the protein’s chemical groups is highly likely incorrect. Hence, other properties must be identified to describe the hydrophilicity of the protein as a function of folding state. Systematic studies of several proteins showed that ionic strength and particle charge influence S_T less than temperature effects.^{19,20} Nevertheless, the thermodiffusive behavior of proteins is a complex interplay of surface properties, which are impacted by the ionic strength and pH of the solution.³⁰

To gain deeper insights into the influence of the charge of the solute, several theoretical concepts have been developed to describe the thermodiffusion of singly charged colloidal particles.^{31–38} Both bulk and interfacial effects are considered and assumed to be additive. While the ionic shielding effect can be well described if the radius and the surface charge are

known, the hydration effects have to be determined by fitting the experimental data. The theoretical concepts for colloids can only be applied to a limited extent to proteins because proteins can undergo conformational changes that may vary their hydrophilicity.

Solvent interactions significantly affect protein conformational stability by lowering energy barriers for transitions, damping collective motions, and affecting translational and rotational water entropy. Understanding hydration changes during these transitions remains challenging.³⁹⁻⁴¹ An extensively studied model system for understanding protein folding experimentally and in molecular simulations is Apo-Mb, the heme-free form of myoglobin.⁴²⁻⁴⁴ Apo-Mb can transition between different stable conformational states with distinct structural compactness and secondary structure features.^{40,41,45-48} The globular protein consists of 153 amino acids organized into eight α -helices (A through H) connected by loops. Helix F is in an incompletely folded conformation due to the vacant heme cavity.^{41,42} The protein exhibits three primary structural states that vary with pH. At near-neutral pH (~ 6), Apo-Mb adopts a compact, near native structure; at mildly acidic pH ($\sim 4-4.5$), it shifts to a partially folded *molten globule* (MG) state; at pH ~ 2 , it reaches an acid-unfolded state⁴⁹ in the millisecond timescale.^{49,50} With decreasing pH, the net positive charge of the protein increases.⁵¹⁻⁵⁴ Further details on the calculation of protein charge can be found in Sections S2 and S8 in the Supporting Information. The predominantly α -helical conformation of Apo-Mb is stabilized by strong intramolecular hydrogen bonding.⁵⁵ In an unfolded state, Apo-Mb exposes hydrophobic and hydrophilic residues, increasing water ordering around hydrophobic regions and forming structured solvation shells. These shells, driven by water's hydrogen bonding network, occupy more volume, are less dense, and have a reduced entropy - a *hydrophobic effect*.⁵⁶ While the hydration shells surrounding the unfolded parts of the protein enhance transient water interactions, they are less stable than in a folded state. Hydrophobic exposure reduces electrostatic interactions and effective hydrophilicity, which decreases solvation stability and promotes aggregation.⁵⁷ Anions interact preferentially with the positively charged regions of the protein and effectively shield repulsive forces between

positive charges by binding to them, thus reducing internal repulsion. In particular, anions with higher charge density and stronger affinity for the protein are more effective at inducing transitions between protein folding states than those with lower affinity.^{52,58} The stability of the *intermediate* states of Apo-Mb strongly depends on the net charge of the protein. For Apo-Mb, with a net positive charge, it has been described that the loss of positively charged residues increases stability by reducing internal charge repulsion, while the loss of negatively charged residues decreases stability due to a corresponding imbalance in charge repulsion.⁵⁹

In molecular dynamics (MD) simulations, covalent bonds are described by force field terms, allowing variations in bond lengths and angles, but usually prohibiting bond breaking or forming of new covalent bonds. Therefore, it is common practice to determine the protonation states of titratable amino acids before starting MD simulations and then to maintain the initially set protonation states. However, in this work, effects of pH variations on Apo-Mb at the atomistic scale shall be analyzed with MD simulations. Thus, we performed discrete protonation state constant pH molecular dynamics (CpHMD) simulations in explicit solvent as implemented in the AMBER software suite.^{60–62} In this approach, the MD propagation is periodically interrupted to probe the protonation states of the predefined titratable residues via a stochastic titration method based on the Poisson-Boltzmann equation. Afterward, the protein is held fixed, while the solvent is allowed to reorganize around the newly protonated or deprotonated residues to avoid steric clashes. The MD is then further propagated for all solute and solvent constituents until the next cycle of protonation state adjustments. Applying CpHMD with Apo-Mb at different pH values allows us to investigate pH-induced effects on protein conformation and solvent interactions at picosecond resolution for microsecond-long MD simulations.

This study investigates the thermodiffusive behavior of Apo-Mb at different pH levels using the infrared thermal diffusion forced Rayleigh scattering (TDFRS) technique, highlighting hydration changes associated with the conformational states. Circular dichroism (CD) spectroscopy and CpHMD simulations confirmed changes in the predominantly α -

helical structure of the protein under different conditions. The effect of acetate and phosphate buffers with different anions on Apo-Mb thermodiffusion was measured, suggesting changes in the hydration and stability of Apo-Mb. CpHMD simulations reveal changes in the protein-water hydrogen-bonding capacity and structural changes in the presence of the buffer components. Finally, correlations between α -helical content from CD spectroscopy and Soret coefficients from IR-TDFRS were explored to relate conformational to thermodiffusion properties.

Experimental Section

Sample preparation

Apo-Mb was obtained from horse heart myoglobin (Sigma-Aldrich, St. Louis, MO) using the butanone extraction method to extract the heme group.⁵⁶ The protein in the resulting solution was refolded by dialysis in 20 mM $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ pH 7, followed by dialysis in distilled water. The refolded Apo-Mb was then lyophilized and subsequently freeze-dried. The Apo-Mb powder was then stored at -20°C . For experimental use, a precise concentration of Apo-Mb powder was dissolved in deionized water (Milli-Q system) or buffer solutions. Specifically, 10 mM and 20 mM acetate buffer ($>99.7\%$ acetic acid, Sigma-Aldrich and $>99\%$ sodium acetate, Merck) were used for pH 4, while 20 mM $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ ($>99\%$, Sigma-Aldrich) buffer was used for pH 6. The acetate and phosphate buffers are abbreviated as NaAc and NaP, respectively. The clear supernatant was carefully collected after centrifugation at 29,000 g for 10 minutes to remove larger aggregates. The pH of the solution was adjusted by adding 0.1 M HCl (Sigma Aldrich). The final Apo-Mb concentration was confirmed by UV/vis absorption spectroscopy (Nano-Drop 2000c, Thermo Scientific) with an extinction coefficient ($E_{1\%}$) of 8.25 at 280 nm, determined from the amino acid sequence using the ExPASy web server.⁶³ We characterized the protein's secondary structure using circular dichroism (CD) and determined its thermodiffusion properties with IR-TDFRS.

Additionally, we measured the electrophoretic mobility of Apo-Mb by electrophoresis. The results are summarized in Section S2 of the Supporting Information.

Circular Dichroism (CD)

CD spectra were recorded at 20°C using a MOS-500 spectrophotometer (BioLogic, Science Instruments, France). The ellipticity of 0.9 mg/mL (~ 0.05 mM) Apo-Mb, both in the presence and absence of buffer, was monitored over a wavelength range of 180 to 260 nm at different pH values in a quartz cell that is 0.1 mm thick under constant nitrogen flow. For each sample, three scans were performed and subtracted from the corresponding buffer or water values. The secondary structure content of the protein was estimated using the BeStSel single spectrum analysis software.⁶⁴

Thermal diffusion forced Rayleigh scattering (TDFRS)

The thermal diffusion properties of the protein solution were investigated using the infrared-Thermal diffusion forced Rayleigh scattering (IR-TDFRS) technique.^{65,66} An Apo-Mb solution with a concentration of 7 mg/mL (~ 0.4 mM) was filled into an optical quartz cell (Hellma) with an optical path length of 0.2 mm. Measurements were performed in a temperature range from 15°C to 45°C in steps of 5°C. For pH values of 6, 4, and 2, experiments were performed with and without buffer in solution. In addition, we measured both buffers at a concentration of 0.5 mol/kg. In each experiment, we collected at least 3000 individual measurement signals and calculated their average. We then examined the on and off phases of each signal, resulting in two value sets for S_T - and D .^{65,67} Two fresh samples were measured at least twice. We therefore calculated the mean of at least four data points at each temperature. The standard deviation of the mean is shown as an error bar in the figures.

This transient grating technique uses two infrared laser beams to create a holographic grating within the sample cells. The inherent absorption of water at a wavelength of 980 nm creates a temperature grating that induces particle migration and, thus, generates a

concentration gradient. The result is a refractive index grating probed by the readout laser beam. The normalized heterodyne intensity ($\zeta_{het}(t)$) of the readout beam, which probes the optical contrast of the interference grating, was measured and fitted to the following equation:

$$\zeta_{het}(t) = 1 - \exp\left(-\frac{t}{\tau_{th}}\right) - A_0(\tau - \tau_{th})^{-1} \times \left\{ \tau \left[1 - \exp\left(-\frac{t}{\tau}\right) \right] - \tau_{th} \left[1 - \exp\left(-\frac{t}{\tau_{th}}\right) \right] \right\}. \quad (3)$$

where the steady state amplitude A_0 is given by

$$A_0 = \left(\frac{\partial n}{\partial c} \right)_{p,T} \left(\frac{\partial n}{\partial T} \right)_{p,c}^{-1} S_T c (1 - c), \quad (4)$$

where τ_{th} is the heat diffusion time and τ is the mass equilibrium diffusion time. Note that the diffusion coefficient D and the thermal diffusivity (D_{th}) can be derived from the corresponding equilibrium times using the relation $\tau_{th} = (D_{th}q^2)^{-1}$ and $\tau = (Dq^2)^{-1}$, respectively, where q is the scattering vector. The refractive index gradients as a function of temperature and concentration, denoted as $(\partial n/\partial T)_{c,p}$ and $(\partial n/\partial c)_{T,p}$ respectively, are measured independently (see Supporting Information). In addition, the Soret coefficient can be calculated from the amplitude A (Eq (4)).

Constant pH molecular dynamics (CpHMD) simulations with discrete protonation states in explicit solvent

To obtain representative structural ensembles of Apo-Mb at pH 6, 4, and 2, constant pH molecular dynamics (CpHMD) simulations with explicit solvent were performed.^{61,62} In this method, predefined titratable residues can change their protonation states upon a short Monte-Carlo exchange attempt, which is relevant when considerable conformational changes or unfolding events are associated with a change in pH. The observed changes in protonation

states during the simulations and the estimated pKa values per residue are provided in Table S6 in the Supporting Information. Eleven CpHMD simulation replicas were performed in five explicit solvent conditions with either water solvent with 150 mM sodium chloride, 20 mM sodium phosphate buffer (NaP), or 20 mM acetate buffer (NaAc) using the AMBER24 software suite.⁶⁰ The entire simulation box was neutralized with respect to the positive charge of the protein by adding extra chloride ions. All simulations were based on the X-ray crystal structure of the wild-type horse heart myoglobin (PDB-ID: 2V1K). All water and buffer components and the heme and glycerol molecules from the crystal structure were removed prior to the simulation setup.

For the simulations in water solvent, the protein was placed in a truncated octahedral box of TIP3P water⁶⁸ extending 20 Å around the protein with 150 mM of sodium chloride. The SHAKE algorithm was used to constrain hydrogen atom movements and allows 2 fs simulation time steps.⁶⁹ Energy minimization was performed with 5,000 steps of steepest descent, followed by 5,000 steps of the conjugate gradient method, in three iterations with initial harmonic positional restraints with a force constant of 25 kcal/(mol · Å²) on all protein atoms, followed by a round with restraints of 5 kcal/(mol · Å²), and an unrestrained minimization. The minimized models were initially heated from 0 K to 100 K over 10 ps using restraints of 5 kcal/(mol · Å²) in an NVT ensemble, followed by heating from 100 K to 300 K in NPT conditions with identical restraints over 50 ps, followed by 70 ps at constant 300 K. To adjust the system density at 1 bar, 4.87 ns of unrestrained NPT simulations were performed using a Berendsen barostat⁷⁰, followed by a final round of 10 ns of unrestrained NVT simulations. This totals in 20.13 ns of thermalization for each of the three starting models at pH = 2, 4, and 6, respectively. Simulations containing either 20 mM of NaP or 20 mM of NaAc buffer were prepared using the PACKMOL-Memgen software.⁷¹ to setup cubic simulation boxes, extending 26 Å around the protein to accommodate 20 mM of the respective buffer component, using the TIP3P model⁶⁸ for the water component. Partial charges of the buffer components were determined using RESP fitting with antechamber based on electrostatic

potentials computed for optimized geometries at the Hartree-Fock 6-31G* level of theory with Gaussian09.⁷² Both simulation setups were minimized and thermalized following the protocol for the simulations in water.

The CpHMD simulations for all five equilibrated systems were carried out in 11 replicas for 3 μ s each in NVT conditions. The ff10 force field⁷³ was used for the protein, with the General Amber Force Field 2 (GAFF2)⁷⁴ for the buffer components and Joung-Cheatham-parameters⁷⁵ for sodium and chloride ions. The accumulated simulation time is 165 μ s. Protonation states of all aspartate, glutamate, histidine, lysine, and tyrosine residues were determined every 5 ps, followed by 100 steps of steepest descent minimizations to avoid atom clashes. In all simulations, temperature and pressure were controlled, respectively, using Langevin dynamics⁷⁶ with a friction coefficient of 1 ps⁻¹ at 300 K. No barostat algorithm was used in the production runs, as they were performed in the NVT ensemble.

The CpHMD simulations were analyzed using the cphstats and cpptraj^{60,77} packages with additional in-house Python scripts. Data for simulation analyses were collected in time steps of 200 ps. The analyses used the last 500 ns of the simulations only, for which the average α -helical content became constant for the simulations at pH 6 and 4 (Fig. S7 and S14 in the Supporting Information). This measure was selected, as analysis of the conformational convergence of the simulations via RMS average correlation (RAC) plots of all backbone C α atoms and cluster discovery analysis for both the full 3 μ s trajectories, as well as for the last 500 ns (Fig. S11, S12 and S13 in the Supporting Information) revealed no conformational convergence over the extensive simulation time, due to the slow unfolding process of the protein and the high conformational flexibility of the MG state at pH 4.

Results and discussion

Characterization of the folding states of ApoMb

CD spectroscopy is a widely used technique to study conformational changes of proteins in solution and allows quantification of secondary structures such as α -helices, β -sheets, and random coils.⁴⁷ However, CD spectroscopy has limitations in structural resolution and can be sensitive to environmental factors such as pH, temperature, and ligand presence, which can affect the reliability of its outcome. Compared to high-resolution techniques, such as X-ray crystallography, structural interpretations based on CD spectroscopy can be ambiguous and rely on calibration to standard data. Thus, it lacks the atomic-level detail required for in-depth structural analysis.^{64,78}

The secondary structure content of Apo-Mb under different solution conditions was confirmed by CD spectroscopy. Buffers for the Apo-Mb was selected based on their pKa values, see Supporting Information S5.1). Figure 1 shows the CD spectra of Apo-Mb in the presence and absence of buffer at different pH values. Table 1 summarizes the estimated content of α -helices under these different solution conditions, which agree with the known literature values. Note that the deviations from the literature values reach 20%, and we have found uncertainties of 3-10% in repeated measurements.

In the absence of buffer, CD data show that a significant fraction of α -helical structure is present in Apo-Mb at pH 6 (51%). Apo-Mb folds into a similar topology as the holoprotein at pH 6, except that helix F is incompletely folded.⁴¹ The content of α -helices decreases progressively with decreasing pH, from pH 6 to pH 2. At pH 6 with NaP buffer, Apo-Mb maintains its α -helical content of 50%. However, NaAc buffer at pH 4 leads to significant deviations compared to the unbuffered acidic solution and reduces the α -helical content from 43% in water to 33-34%. This indicates that NaAc buffer promotes partial unfolding of the protein in comparison to an unbuffered acidic solution. At pH 2, where the protein is in the acid-unfolded state, a small fraction of the α -helical content (4%) is retained.

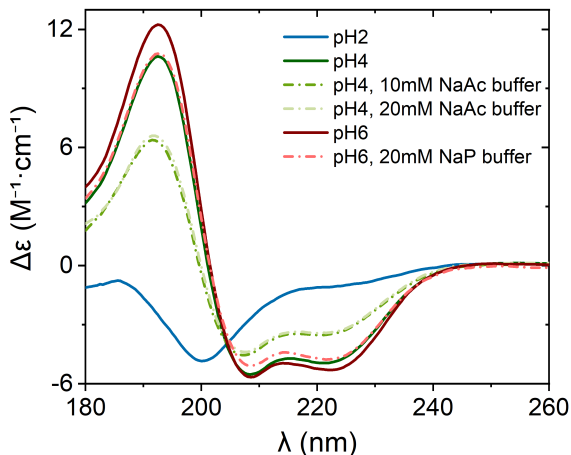


Figure 1: CD spectra of 52.2 μM Apo-Mb observed at 20°C. The solid lines show Apo-Mb in water adjusted to the respective pH values with HCl, while the dashed-dotted lines represent Apo-Mb in buffer. The spectra show that the content of α -helices of Apo-Mb decreases with decreasing pH. The lines show the average of three scans.

To obtain atomic-level information on the structural changes upon pH changes and assess the interactions of the buffer components with the protein surface, CpHMD simulations of Apo-Mb were performed at pH 6 (with and without NaP buffer), 4 (with and without NaAc buffer), and 2, using 11 replicas per solvent condition. As the simulations are started from the holo structure, the initial helical content was higher than the values measured by CD and reached values comparable to those of the unbuffered samples after $\sim 2 \mu\text{s}$ for pH 6 and $2.5 \mu\text{s}$ for pH 4 (see Figure S7 and S14 in the Supporting Information). Over the last 500 ns of simulation time, the average α -helical content found in the simulations overlaps with the range of experimental values determined by CD measurements for the buffered and unbuffered simulations at pH 6 and 4 (Figure 2).

The distribution of the α -helical content from simulations at pH 6 in unbuffered solution is bimodal with an average of $54\% \pm 0.1\%$, with the larger peak agreeing with the experimentally measured range of $51\% \pm 1.6\%$ and the smaller peak representing the higher initial α -helical content of the starting structure. By contrast, simulations with 20 mM NaP buffer at pH 6 yield a unimodal distribution around an average of $48\% \pm 0.1\%$, in agreement with the experimentally measured value of $50\% \pm 0.5\%$. At unbuffered pH 4, the distribution

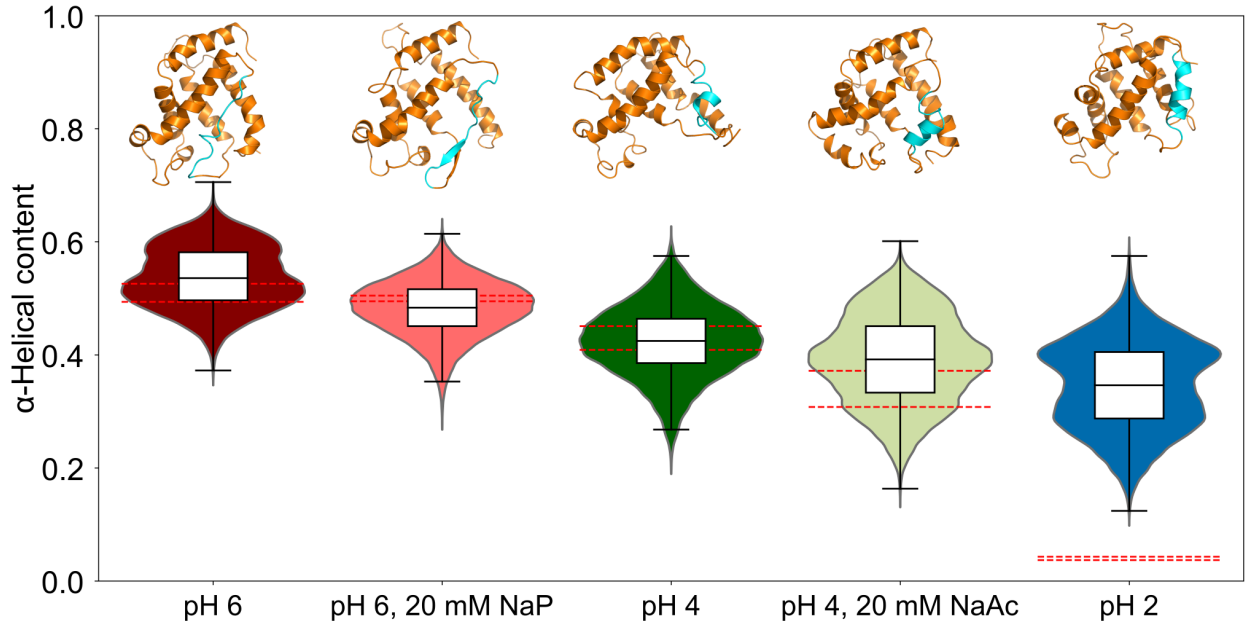


Figure 2: Distribution of the α -helical content of Apo-Mb over the last 500 ns of 3 μ s of CpHMD simulations per replica ($n = 11 \cdot 500 \cdot 5$ datapoints, for 11 replicas, 500 ns and 200 ps of data sampling) at pH 6, 4, or 2 in explicit water, at pH 6 with 20 mM NaP buffer, or at pH 4 with 20 mM NaAc buffer. The same color scheme as in Figure 1 was used. The red dotted lines indicate the experimentally measured α -helical content $\pm$ the standard deviations (Table 1). Representative structures for each condition are shown at the top, with helix F highlighted in cyan to show the model orientations.

is unimodal around an average of $42\% \pm 0.1\%$ and in close agreement with the experimental range of $43\% \pm 2.1\%$. For simulations in NaAc buffer at pH 4, the average α -helical content is $39\% \pm 0.1\%$ but the variations among replicas is high, with 50% of the frames sampling conformations with an α -helical content between 33% and 45%, which encompasses the experimentally measured α -helical content of $34\% \pm 3.2\%$ (Figure 2). When comparing the differences in secondary structure elements, simulations in NaAc buffer reveal a 10-20% reduction in the α -helical content for residues 16-25 in helices A and B, as well as a slight stabilization of helix F for residues 83-91. From these positions, only lysine 77 at the very end of helix E shows a significant decrease (t-test $p < 0.001$) in α -helical content compared to the unbuffered simulation. For the NaP simulations on the other hand, while also moderate deviations were observed, significant changes to the unbuffered simulation were only found in leucine 86 and glutamate 91 in helix F with increased α -helicity (see Figures S8 and S9 in the supporting information). For the unbuffered simulations at pH 2, we again see large variations in the unfolding behaviour between individual replicas, resulting in an average α -helical content of 35%, but find no simulation replica that samples Apo-Mb at the experimental α -helical content of $4\% \pm 0.3\%$. This is also apparent in the representative conformation, depicting the most populated structural cluster over the last 500 ns of the simulations, which shows almost no unfolding. Overall, while a marked decrease in α -helical content was observed at pH 2, also compared to the other simulations, the only partial unfolding of Apo-Mb under this condition is in line with the experimentally determined timescale on the order of (sub-)milliseconds.^{50,79,80} Yet, as suggested for Apo-Mb, a major change was observed in helix F compared to the X-ray crystal structure for Holo-Mb, which partially unfolds in the absence of the heme group in all simulated conditions.⁸¹

Thermodiffusion of the buffer

Since the temperature dependence of S_T provides some information about the hydrophilicity of the solute, we examined the NaP and NaAc buffers at a higher concentration of 0.5 mol/kg.

Table 1: Secondary structure content of different conformational states of Apo-Mb as measured by CD. In the given references, the α -helical content and the state are specified. The experimentally observed α -helical content shows a strong agreement with the different conformational states defined in Ref. 50.

protein	state	α -helical content (%) this work	MD α -helical content (%) this work	α -helical content (%) literature	reference
Apo-Mb at pH2	acid unfolded	4 ± 0.3	(35)*	4-5	54,82
Apo-Mb at pH4	MG	43 ± 2.1	42 ± 0.1	35-43	56,83,84
Apo-Mb at pH4, 10mM NaAc buffer	MG	33 ± 2.2	-	-	85
Apo-Mb at pH4, 20mM NaAc buffer	MG	34 ± 3.2	39 ± 0.1	-	86
Apo-Mb at pH6	folded	51 ± 1.6	54 ± 0.1	49	54
Apo-Mb at pH6, 20mM NaP buffer	folded	50 ± 0.5	48 ± 0.1	55	56

*) structure not fully unfolded within simulation time

This concentration gave a sufficiently strong heterodyne measurement signal. Figure 3 shows D_T , D , and S_T as a function of temperature for both buffers in a temperature range between 15°C and 45°C. While S_T of the NaP buffer shows a strong temperature dependence, S_T of the NaAc buffer is almost temperature-independent. This is consistent with the log P values of -4.7 and -0.28 for NaP and NaAc, respectively,⁸⁷ suggesting that the interaction of the buffer with the protein could have an impact on its thermodiffusion. Only the log P values of the major molecular buffer components in solution were considered: acetic acid for NaAc and monosodium dihydrogen phosphate for NaP (see Supporting Information S5.3). Calculating the log P values shows that NaP is more hydrophilic than NaAc. The temperature sensitivity of S_T of the two buffers follows the same trend as for the non-ionic solutes in Fig. 5 of Ref. 5.

Dielectric measurements also indicate a higher immobilization of water molecules in the case of NaP⁸⁸ compared to NaAc.⁸⁹ The NaP buffer consists of a mole fraction $x \approx 0.9$ of monosodium dihydrogen phosphate and $x \approx 0.1$ of disodium hydrogen phosphate. According to Eiberweiser,⁸⁸ the number of immobilized water molecules is 4 for monosodium dihydrogen phosphate and 11 for disodium hydrogen phosphate at low concentrations. In the case of

sodium acetate ($x \approx 0.2$) and acetic acid ($x \approx 0.8$), the number of immobilized water molecules is 5 and 1, respectively.⁸⁹ This gives approximate values of 5.5 for NaP and 1.8 for NaAc, assuming that the mean value scales with the mole fraction x of the buffer components, further indicating that NaP is more hydrophilic than NaAc.

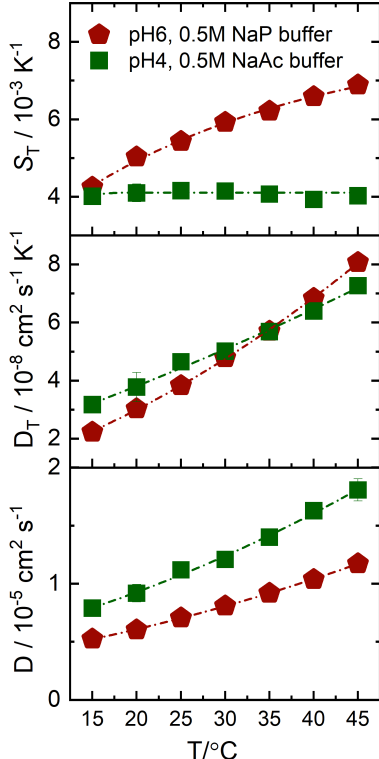


Figure 3: Temperature dependence of D_T , D , and S_T of NaP (pH 6, red pentagon) and NaAc (pH 4, green square) buffers, measured at concentrations of 0.5 mol/kg. The lines are there to guide the eye. Note: For the S_T data plot, the lines are fitted according to Eq.2. The fitting parameters are given in the Supporting Information (Table S4).

Thermodiffusion behavior of Apo-Mb

In the following, we discuss the thermodiffusion data of Apo-Mb in water and in buffered solutions as a function of pH. To provide a basis for understanding the thermodiffusive behavior of the protein, we first examine the IR-TDFRS results of Apo-Mb in buffer-free solutions and later in the presence of buffer. Note that the buffered Apo-Mb solutions were treated as a pseudobinary system and the contribution of the buffer to the measured

signal was neglected. This is justified because we have shown that we do not observe any contribution to the concentration signal at the low buffer concentration of 20 mM (see Supporting Information S5.2 for details).

It is well known that the diffusion properties of proteins strongly depend on their shape, size, and interaction with the environment.⁴⁰ In recent decades, it has been experimentally demonstrated that buffer molecules can selectively adsorb onto charged protein surfaces and, thereby, influence protein-protein interactions.^{90,91} This phenomenon, which has been extensively studied in the context of simple ion adsorption on protein surfaces, is known as the Hofmeister effect.⁹²⁻⁹⁴ Previous research has shown that the precipitation of proteins by different ions is closely related to the hydration properties of the ions.⁹⁵⁻⁹⁷ We then examine the IR-TDFRS results in the presence of buffer compounds to assess how the buffer affects the thermal diffusion properties of the protein. This comparative approach helps to clarify to which extent buffer compounds modulate the thermodiffusive behavior of the protein compared to the unbuffered solutions.

Thermodiffusion of Apo-Mb at different unbuffered HCl-adjusted pH-values

Figure 4 shows the temperature dependence of S_T for Apo-Mb solutions measured at different pH values without buffer. The refractive index increments, D_T and D values used to calculate S_T , are given in Supporting Information section S3 and S4. To minimize particle interactions, the sample concentration was kept at 0.4 mM, well below the lowest estimated overlap concentration of 9 mM, at which protein molecules begin to spatially overlap (see Supporting Information section S1).⁹⁸ The results show that Apo-Mb is thermophilic at pH 2 and pH 6 at lower temperatures, with a negative S_T , but switches to a thermophobic behavior and a positive S_T when the temperature increases above $\sim 20^\circ\text{C}$.²⁷ In contrast, Apo-Mb remains thermophobic at pH 4 over the entire temperature range studied. However, the value of S_T increases with increasing temperature in all solution conditions and can be described by Eq.2. The parameters for the different fits are listed in Table S4 in the

Supporting Information. The parameters S_T^∞ , A , and T_0 decrease with decreasing pH.

Instead of using the parameter A in Eq.2, we used the difference of S_T at a high and a low temperature and define $\Delta S_T(\Delta T) = S_T(40^\circ\text{C}) - S_T(15^\circ\text{C})$ (see Supporting Information section S6). The inset in Figure 4 illustrates the relationship between $\Delta S_T(\Delta T)$ and pH. Apo-Mb appears to follow the same trend as non-ionic solutes, with $\Delta S_T(\Delta T)$ decreasing with decreasing pH.

A similar temperature dependence of S_T with a transition from thermophilic to thermophobic behavior with increasing temperature has also been observed for proteins such as lysozyme, β -lactoglobulin, and streptavidin.^{19,22,38} All these proteins show a positive slope for S_T , with an order of 10^{-2} K^{-1} .^{24,99,100} It is postulated that hydrophobic interactions play an important role in the thermophoresis of proteins.^{8,22} In general, at high temperatures, water interacts preferentially with the charged and polar regions of proteins, promoting thermophobic behavior (positive S_T). However, as temperature decreases, the hydrophobic effects become stronger, causing proteins to exhibit thermophilic behavior by favoring warmer regions.^{19,101} Unfortunately, there is no theory to predict the temperature at which S_T changes sign. Experimental results suggest that the chemical nature of the proteins and the structural properties of water contribute in a complex way.

A $\log P$ value cannot be determined for proteins, neither by calculation nor by experiment. However, for Apo-Mb, the pH value appears to be a reliable indicator of hydrophilicity. Apo-Mb, with an isoelectric point of 7.2, is nearly neutral at pH 6.⁵⁴ At this pH, proteins typically adopt a folded conformation in which most of their hydrophobic regions are buried in the core and most of their hydrophilic or charged regions are exposed to the aqueous environment, so they behave like a hydrophilic molecule in solution.^{39,102,103} As the pH decreases, the protonation of the amino acids increases the net positive charge of the protein, resulting in greater electrostatic repulsion between the protein sidechains. Negatively charged residues at the protein surface get neutralized, generating neutral surface patches, and reducing the interactions with water molecules. This leads to unfolding of the protein and exposure of

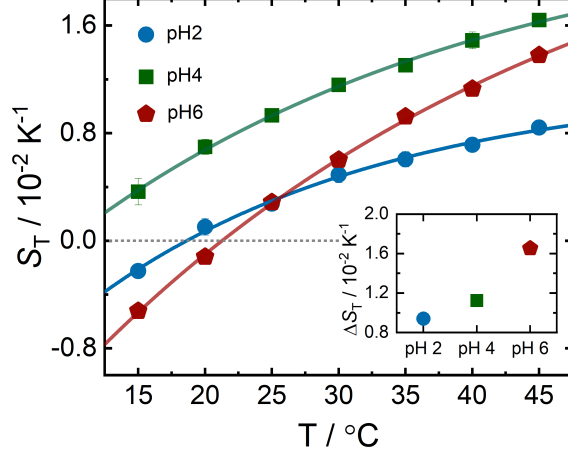


Figure 4: Temperature dependence of S_T of Apo-Mb without buffer at different pH (see table 1). The solid curves represent fits to Eq. 2, which describes the temperature dependence of S_T . The inset shows the temperature sensitivity, $\Delta S_T(\Delta T)$, as a function of pH. Note: $S_T(\Delta T)$ refers to the difference in S_T between high (40°C) and low (15°C) temperatures.

the hydrophobic regions to the solvent.^{56,104} Indeed, in MD simulations of Apo-Mb, the number of hydrogen bonds formed between the protein and water molecules decreases with decreasing pH, whereas the solvent accessible surface area (SASA) increases and the number of intramolecular hydrogen-bonds decreases (Figure 7), in line with a denaturation process taking place due to the pH change. Consequently, the hydrophilicity of the protein decreases when the pH is lowered due to its unfolding. However, at very low pH values (e.g., pH 2), the protein attains a high net positive charge, which can enhance electrostatic hydration, while simultaneously exposing hydrophobic regions due to unfolding. The interplay between these opposing effects, electrostatic hydration and hydrophobic exposure, collectively determines the apparent hydrophilicity of the protein under such acidic conditions. Note that the process of Apo-Mb unfolding is reversible, and the protein can be easily refolded from the state of acidic denaturation.⁴¹

Influence of phosphate buffer at pH 6

Next, we investigated the thermodiffusion behavior of Apo-Mb specifically at pH 6. The temperature dependence of S_T , D_T , and D with and without NaP buffer is shown in Figure

5. The D_T values for Apo-Mb with and without buffer overlap over the entire temperature

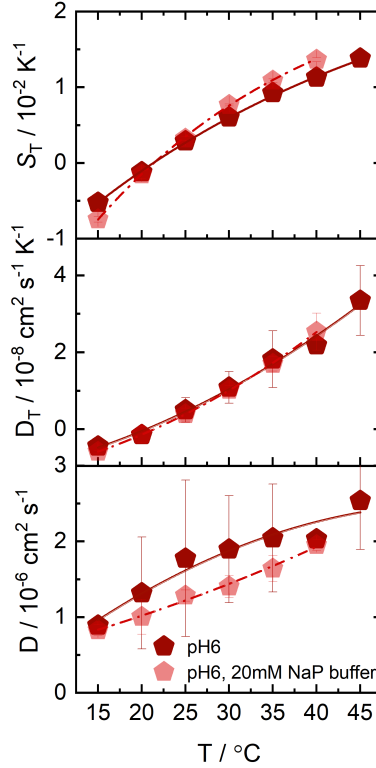


Figure 5: Temperature dependence of S_T , D_T , and D of Apo-Mb with and without buffer, measured at pH 6. The dark red pentagons represent the Apo-Mb solutions without buffer, adjusted with hydrochloric acid, as shown in 4, while the light red pentagons correspond to Apo-Mb in NaP buffer. The curves for S_T are fitted according to Eq. 2 (see Table S4 in Supporting Information), while the lines for D and D_T are a guide for the eye.

range, indicating that the buffer has no significant effect on the thermodiffusion behavior. This is consistent with CD measurements and the MD simulations, where Apo-Mb retains its α -helical and overall secondary structure unchanged under these conditions. In contrast, the D value for Apo-Mb was significantly lower in the presence of NaP buffer than in the unbuffered solutions. This decrease in D is mainly attributed to protein aggregation by the NaP buffer: in the presence of buffer, the light scattering intensity at 45°C increased instantaneously, resulting in increased turbidity, making measurements impossible and indicating protein aggregation. Similar aggregation phenomena have been observed for other proteins in the presence of phosphate buffer.^{58,90,105–107} According to the Hofmeister series, the phosphate ions (e.g., HPO_4^{2-} or PO_4^{3-}) are classified as strong kosmotropes, which are known

to enhance the structuring of water molecules in their vicinity.⁹² This kosmotropic behavior can lead to protein compaction by stabilizing intramolecular interactions and reducing the exposure of hydrophobic residues to the water. Consequently, the solvent-accessible surface of the protein becomes more hydrophilic. However, phosphate ions can also promote protein aggregation due to charge neutralization under specific conditions as shown in our study. For more information on the effect of buffers on the protein, see Supporting Information section S5.4. However, the effects of these buffer ions on protein aggregation can be better understood by identifying the binding site on the protein surface, which has only been discussed on an empirical basis.¹⁰⁷

In CpHMD simulations, aggregation of the protein is not evaluated, as only a single Apo-Mb molecule is simulated. Still, at 20 mM NaP pH 6, the CpHMD simulations reveal that the phosphate ions bind close to and at the vacant heme cavity and at dedicated positions over the protein surface (Figure 6). This could contribute to a difference in the particle diffusion D , as the apparent molecule size increases compared to the unbuffered solution. Interestingly, while the protein has a similar amount of α -helices when buffered with NaP, it displays an increase in the measured SASA when compared with the unbuffered simulations. In addition, as positions on the surface of the protein are occupied by buffer molecules, the number of hydrogen bonds between the protein and water gets reduced compared to the unbuffered solvent, as well as the intramolecular hydrogen bonds of the protein (Figure 7). One has to consider that only direct hydrogen bonds with water are counted as protein-water hydrogen bonds, excluding potential cases where phosphate ions could bridge specific hydrogen bonds between the protein and water. Importantly, the main peak in the hydrogen-bond distribution with NaP shows a slight shift towards more hydrogen bonds compared to the unbuffered simulations, indicating that NaP fosters the formation of protein-water interactions in frames with no direct phosphate-protein interactions. This aligns well with the kosmotropic effect of phosphate ions, fostering protein hydrophilicity as described above.

Influence of acetate buffer at pH 4

The MG states of Apo-Mb at pH 4 have been studied extensively.^{44,49,50} The MG state is in equilibrium between two forms, I_a and I_b , and is positively charged due to the protonation of the amino acid residues.⁴¹ Compared to the state at pH 6, the protein exhibits increased sensitivity to environmental changes due to partial unfolding. CD measurements at pH 4 showed that the NaAc buffer promotes further unfolding of Apo-Mb, resulting in lower α -helical content compared to unbuffered acidic conditions. Because NaAc is near the center of the Hofmeister series (see Fig.S5 in Supporting Information) and is significantly less kosmotropic than sodium phosphate (NaP) buffer, it is expected to interact more strongly with the highly positively charged, partially folded MG state of Apo-Mb. Indeed, the MD simulations with NaAc show strong interactions with the protein surface, particularly when compared with NaP (Figure 6, see below). The thermodiffusion data of Apo-Mb in the presence of NaAc buffer are now discussed. The temperature dependence of S_T of Apo-Mb flattens and the S_T values increase with increasing NaAc concentration (see Figure 8). The thermal diffusion coefficient D_T (see Figure S3 Supporting Information) also increases with increasing NaAc concentration. A similar trend was observed for S_T and D_T of dextran in water with increasing urea concentration.²¹ Despite the differences in these systems, in both cases the addition of a hydrophilic compound leads to a weaker temperature dependence of S_T , which is typically associated with a disruption of hydrogen bonds between the solute and water. We also observed a slight decrease in the diffusion coefficient of Apo-Mb with increasing NaAc concentration (see Figure S3 Supporting Information). The partial unfolding of Apo-Mb enhances its likelihood to aggregate, a process that might further be intensified in NaAc buffer due to charge screening.⁸⁵ Thus, the observed decrease in D likely reflects partial unfolding of the protein and possible contributions from aggregation (see Supporting Information S5.4).

As indicated above, the 20 mM NaAc Apo-Mb pH 4 CpHMD simulations show strong interactions with the buffer components over all protein residues, in stark contrast to the

simulations at NaP pH 6 (Figure 6). Apo-Mb has a considerably higher positive charge at pH 4, which favors further interactions with the negatively charged acetate ions. Although

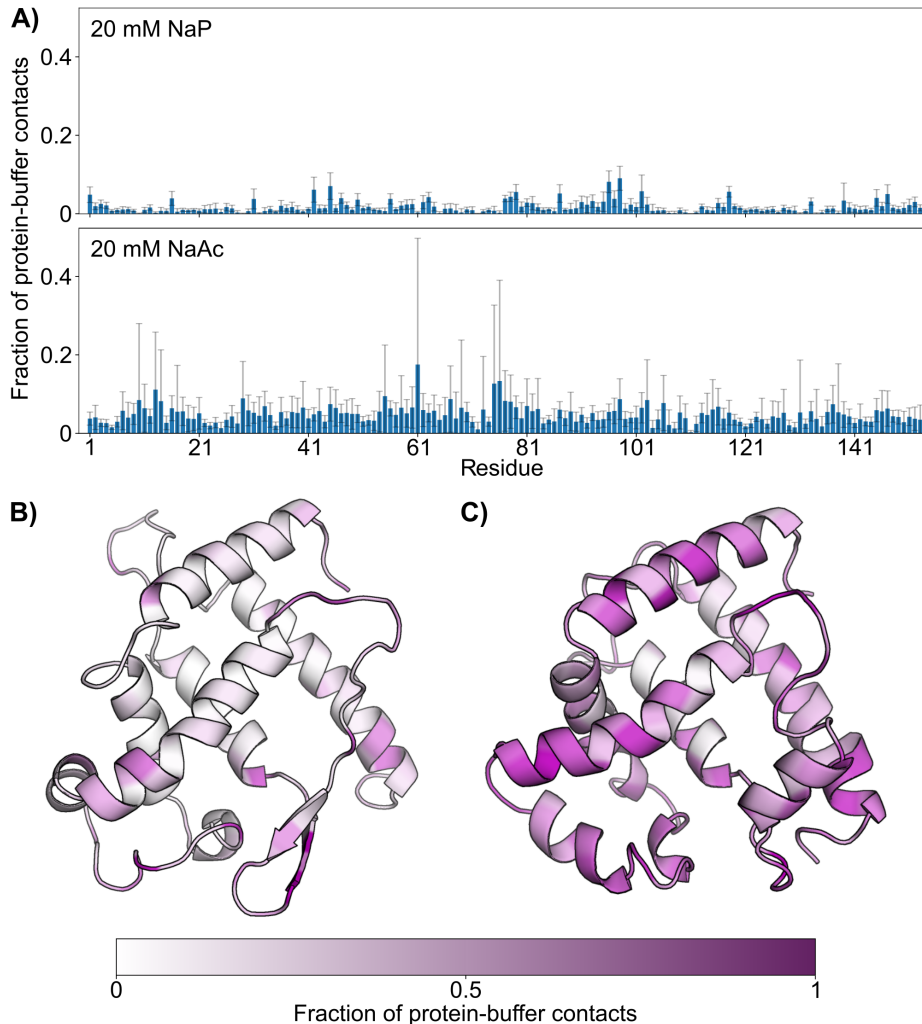


Figure 6: Residue-specific buffer interactions with Apo-Mb determined from CpHMD simulations. A) Per residue average fraction of contacts between Apo-Mb and NaP (top) or NaAc (bottom) buffer molecules. B, C) Values from A shown on representative protein structures obtained from CpHMD simulations with either 20 mM NaP (left) or NaAc buffer (right). Values are averages $\pm$ standard deviation over 11 replicas each.

these multiple interaction positions between NaAc and the protein can, as in the case of NaP, explain changes in D , the additional interaction sites agree with the notion of a molecular shielding for charged residues, which does not seem to be the case for NaP. Similar to NaP, the inclusion of NaAc at pH 4 increased the SASA compared to the unbuffered solution, but

the resulting SASA is higher than that for NaP, which could be a result of the additional reduction in alpha-helical structure shown above. There is a reduction in the number of hydrogen bonds with respect to the unbuffered solution as well, suggesting a decrease in direct interactions with the solvent, with a further reduction in the intramolecular hydrogen bonds than in NaP-buffered or pH 4 unbuffered solution (Figure 7). While at first glance this distribution is similar to that with NaP, it is important to note that the main distribution peak is not shifted towards more hydrogen bonds than in the unbuffered solution, in contrast to the increase observed with NaP. Overall, NaAc increases SASA more than NaP, reduces the number of hydrogen bonds, and simultaneously influences the secondary structure of apo-Mb (Figure 2). The overall NaAc effect with the initial drop in the α - helical content caused by the interactions with the buffer could explain the higher initial S_T and the consequent drop in ΔS_T observed for NaAc in Figure 8. The effects of decreasing pH values and buffer addition on the SASA and hydrogen bonds are also reflected in shifts in the number of water molecules in the 1st and 2nd water shells around the protein, which both increase upon reduction of pH and upon buffer addition (Fig. S15 in the Supporting Information).

The inset in Figure 8 shows that $\Delta S_T(\Delta T)$ decreases with increasing NaAc concentration at pH 4. The decrease in $\Delta S_T(\Delta T)$ is similar for 10 mM and 20 mM NaAc buffers due to their low ionic strength. As evidenced in the MD simulations, NaAc interacts strongly with the protein surface, reducing its effective charge and disrupting intra- and intermolecular hydrogen bonding compared to the unbuffered acidic condition (as evidenced through lower α -helical content). Ultimately, this facilitates the unfolding and exposure of the hydrophobic regions of the protein to the aqueous environment, reducing its overall hydrophilicity. This decrease in hydrophilicity is reflected in decreased $\Delta S_T(\Delta T)$ values and follows the observed trend: the lower the hydrophilicity of the protein, the lower is $\Delta S_T(\Delta T)$.

A similar behavior has also been observed for streptavidin-biotin as ΔS_T of the complex was reduced compared to the isolated streptavidin.²² The streptavidin-biotin complex was less flexible compared to free streptavidin, so that the conformational entropy of the com-

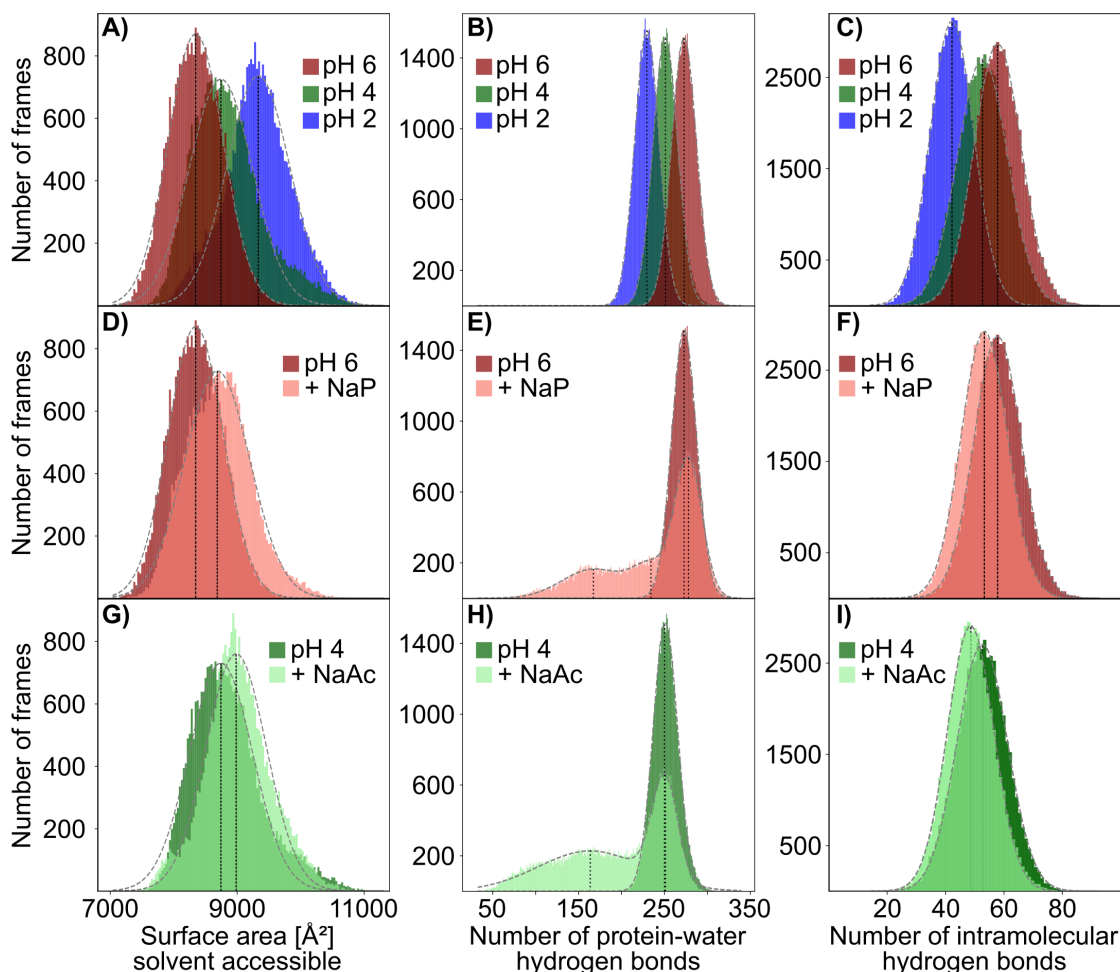


Figure 7: Apo-Mb solvent accessible surface area (SASA) , protein-water hydrogen bonds, and intramolecular hydrogen bonds. Histograms depict the distribution of analyzed simulation frames for unbuffered simulations at pH 6 (red), 4 (green), and 2 (blue), as well as with 20 mM of either NaP buffer at pH 6 (light red) or NaAc buffer (light green). Grey dotted lines indicate either a Gaussian, bimodal, or trimodal fit to the datasets, with black dotted lines highlighting the peak of each fitted curve. At pH 6, 4, and 2 (A, B, C) in unbuffered solution, a decrease in the system pH causes an increase in SASA, and a reduction in the number of hydrogen bonds both with water and within the molecule, indicating a denaturation process for the protein. Upon addition of 20 mM NaP buffer (D,E,F), the protein SASA increases, as fewer intramolecular hydrogen bonds are formed compared to the unbuffered solution. Interestingly, the main peak for protein-water hydrogen bonds shows a slight shift towards more hydrogen bonds when compared to the unbuffered simulations, with two secondary peaks of frames with significantly reduced protein-water hydrogen bonds, not present in the unbuffered solution. NaAc, on the other hand (G, H and I), while also showing an increase in SASA and fewer intramolecular hydrogen bonds, has no significant increment in the main peak for protein-water hydrogen bonds, but shows more frames with low protein-water hydrogen bonds in the secondary peak.

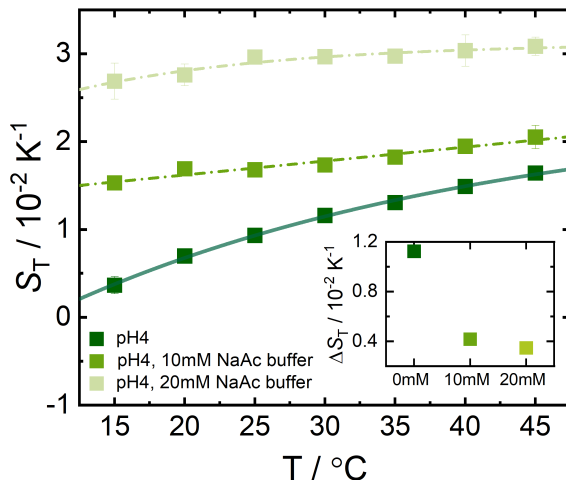


Figure 8: Temperature dependence of S_T of Apo-Mb with and without NaAc buffer at pH 4 (see table 1). The dark green squares represent the ApoMb solution without buffer (pH adjusted with hydrochloric acid), while the middle and light green pentagons correspond to ApoMb in 10 mM and 20 mM NaAc buffer, respectively. The curves represent fits to Eq. 2, which describes the temperature dependence of S_T . The inset shows the temperature sensitivity, $\Delta S_T(\Delta T)$, as a function of increasing NaAc concentration. Note: $S_T(\Delta T)$ refers to the difference in S_T between high (40°C) and low (15°C) temperatures.

plex was substantially reduced, while an increase in the entropy of the hydration layer was observed.²² Liese et al.¹⁰⁸ observed a similar phenomenon in stretched (rigid) versus flexible poly(ethylene glycol) (PEG) chains, where entropic hydration effects nearly compensated for chain conformational entropy. Specifically, water molecules formed fewer hydrogen bonds in the hydration layer of the rigid, stretched PEG than in the flexible PEG coil.

Correlation between Circular Dichroism and IR-TDFRS

The next step was to compare the CD results with the IR-TDFRS results. Considering the change in structure posed by the changes in pH and buffer interactions, we wanted to evaluate if there is a correlation between the calculated α -helical content and the thermophilicity in terms of thermal sensitivity, which is a reliable indicator of the hydrophilicity of solute molecules in water.^{5,11,23,109} To do this, we plotted $\Delta S_T(\Delta T)$ against the α -helical content as shown in Figure 9. We find that $\Delta S_T(\Delta T)$ correlates linearly with the α -helical content of Apo-Mb over a wide range, strongly suggesting that it reflects the folding state and

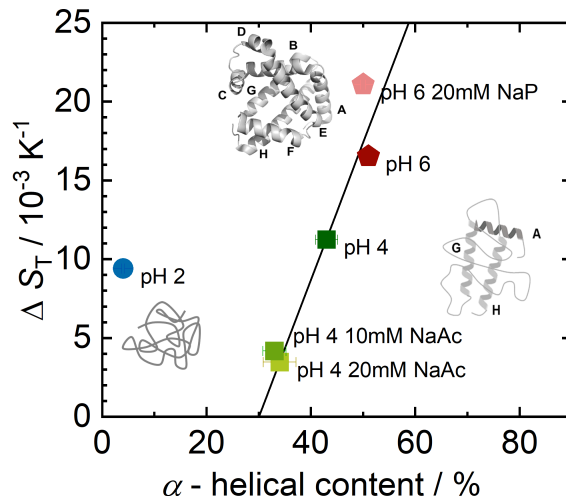


Figure 9: Temperature sensitivity, $\Delta S_T(\Delta T)$, plotted as a function of the α -helical content. The plot shows a strong correlation between the two parameters. Note: $\Delta S_T(\Delta T)$ refers to the difference in S_T between high (40°C) and low (15°C) temperatures.

hydrophilicity of Apo-Mb. This occurs because, in an α -helical conformation, hydrophobic residues are typically buried within the protein core, while hydrophilic residues are exposed to the surrounding aqueous environment. In addition, previous studies show that the α -helical content of a protein is a main determinant of protein diffusion characteristics.⁴⁰ This correlation is also observed in our results. In particular, the temperature sensitivity of the thermodiffusive behavior, expressed by $\Delta S_T(\Delta T)$, decreases in the following order as the hydrophilicity of Apo-Mb decreases (which is achieved by lowering the pH): ΔS_T (pH 6) > ΔS_T (pH 4) > ΔS_T (pH 2).

Our MD simulations suggest that the NaP buffer, compared with the unbuffered solution, has the potential to increase the number of hydrogen bonds of the protein with water, while not affecting the α -helical content of the protein. This results in increased hydrophilicity and a higher value of $\Delta S_T(\Delta T)$.^{48,51} Conversely, NaAc buffer promoted protein solubilization, which was reflected in the reduction of α -helical content and in water hydrogen bonds. At pH 4, the reduced hydrophilicity caused by the increased exposure of the hydrophobic regions led to a significantly lower $\Delta S_T(\Delta T)$ value in the presence of NaAc buffer compared to that in unbuffered acidic solution at the same pH. At pH 2, the protein structure was

largely disrupted, with a significant decrease in the α -helical content and an increase in the net positive charge. This perturbation led to a deviation from the observed trend between $\Delta S_T(\Delta T)$ and the α -helical content.

Conclusion

TDRFS is a highly sensitive technique for probing protein hydration in solution because it is very sensitive to the nature of solute-solvent interactions. This is especially true when hydrophobic, hydrophilic, and charged species are involved in the protein structure, as well as the interaction of the protein with buffer components. In this context, our study investigates the influence of conformational changes on the thermodiffusion behavior of Apo-Mb using the IR-TDRFS technique. By adjusting the pH and changing the buffer conditions, we were able to generate differently folded states of Apo-Mb. Apo-Mb undergoes pH-dependent conformational shifts: It maintains a folded, neutral state at pH 6, transitions to a positively charged, partially folded MG state at pH 4, and becomes strongly positive and acid unfolded at pH 2. These structural changes were confirmed by CD analysis and CpHMD simulations, which revealed a progressive decrease in the α -helical content with decreasing pH (pH 6 > pH 4 > pH 2), and a reduction in the number of hydrogen bonds between the protein and water. We found that the α -helical content is strongly correlated with $\Delta S_T(\Delta T)$, which is a measure of the protein folding state and hydrophilicity.^{5,27,109} Reducing the α -helical content is considered to decrease hydrophilicity (pH6 > pH4 > pH2) and to decrease the temperature sensitivity of S_T (ΔS_T). ($\Delta S_T(\Delta T)$ (pH6) > $\Delta S_T(\Delta T)$ (pH4) > $\Delta S_T(\Delta T)$ (pH2)).

The buffer type also plays a significant role in modulating the structural and diffusion properties of apo-Mb. At pH 6, phosphate buffer retains the α -helical content, but promotes aggregation of the protein due to electrostatic screening, which is reflected in a lower diffusion coefficient. CpHMD simulations showed how NaP tends to interact sparsely at dedicated

spots with the protein, mainly close to the heme cavity, and seems to foster protein-water hydrogen bond formation. By contrast, acetate buffer at pH 4 decreases the α -helical content of the protein below the values obtained in unbuffered conditions, causing a decrease in α -helical content in helices A, E and G. These changes cause a decrease in the temperature sensitivity of S_T and favors the unfolding of the protein in solution. Increasing the concentration of the acetate buffer at pH 4 further decreases the temperature sensitivity of S_T .

Overall, we observed a strong correlation between $\Delta S_T(\Delta T)$ and α -helical content; $\Delta S_T(\Delta T)$ decreases steadily with decreasing hydrophilicity and α -helical content of ApoMb, showing a strong correlation with structural changes evidenced by CpHMD simulations. These results highlight the complex interplay between the structural state of ApoMb, pH, buffer composition, and thermodiffusion behavior and provide valuable insights into protein hydration.

Associated Content

Supporting Information Available

Relation between hydrophilicity and solvent-accessible surface area of proteins including ApoMb, overlap concentration of ApoMb, electrophoresis measurements of ApoMb, refractive index contrast measurements and temperature sensitivity of S_T at different solution conditions, diffusion coefficient and thermal diffusion coefficient at pH 2 and pH 4, additional information on buffers, parameters describing the temperature dependence of S_T . Further information is provided on the simulation step parameters, the estimated pKa values of each titrated sidechain, the average α -helical content per solvent condition, the average net charge of apo-Mb during CpHMD simulations, and the proportion of simulation time that each residue spends as part of an α -helix. Additionally, we evaluated the convergence of

the simulations with RMS average correlation (RAC) plots, the convergence of the average α -helical content per simulation condition, the discovery of conformational clusters over time for both the full 3 μ s simulations and the last 500 ns, and finally the changes in the number of water molecules in the 1st and 2nd hydration shells.

Authors contribution

#Binny A. Rudani and Steffen Docter contributed equally to this work. Binny A. Rudani, experimental investigation, formal analysis, and writing- original draft; Steffen Docter, performed MD simulations and analysis, writing, visualization; Stephan Schott-Verdugo, advised MD simulation setup and analysis, writing, validation; Johan Buitenhuis, Supported circular dichroism and electrophoresis measurements; Andreas M. Stadler, provided the protein, conceptualization, writing; Holger Gohlke, advised MD simulation setup and analysis, writing, funding, resources, conceptualization, administration; Simone Wiegand, advised experiments, conceptualization, writing, and project administration.

Notes

The authors declare no competing financial interest.

Acknowledgement

We thank Wim Briels and Jan Dhont for fruitful discussions. We thank Hartmut Kriegs for technical help with the IR-TDFRS-measurements. We are grateful to Peter Lang for his generous support of our work. BR acknowledges the support of the International Helmholtz Research School of Biophysics and Soft Matter (BioSoft). Calculator plugins were used for the calculation of $\log P$ and pKa values, Marvin 16.5.2.0, 2016, ChemAxon (<http://www.chemaxon.com>). We are grateful for the computational support by the “Zentrum für Informations

und Medientechnologie” at the Heinrich Heine Universität Düsseldorf.

References

- (1) Ludwig, C. Diffusion zwischen ungleich erwärmten Orten gleich zusammengesetzter Lösungen. *Sitz. ber. Math.-Naturw. Clas. Kais. Akad. Wiss.* **1856**, *20*, 539.
- (2) Soret, C. Sur L’etat d’équilibre que prend au point de vue de sa concentration une dissolution saaline primitivement homogène dont deux parties sont portées a des températures différentes. *Arch. Geneve* **1879**, *3*, 48–64.
- (3) de Groot, S. R. *Thermodynamics of irreversible processes*; North Holland: Amsterdam, 1966.
- (4) Kjelstrup, S.; Bedeaux, D.; Johannessen, E.; Gross, J. *Non-equilibrium thermodynamics for engineers*, second edition ed.; World Scientific: Hackensack, NJ, 2017.
- (5) Niether, D.; Wiegand, S. Thermophoresis of biological and biocompatible compounds in aqueous solution. *J. Phys. Condens. Matter* **2019**, *31*, 503003.
- (6) Sarter, M.; Niether, D.; Koenig, B. W.; Lohstroh, W.; Zamponi, M.; Jalarvo, N. H.; Wiegand, S.; Fitter, J.; Stadler, A. M. Strong Adverse Contribution of Conformational Dynamics to Streptavidin-Biotin Binding. *J. Phys. Chem. B* **2020**, *124*, 324–335.
- (7) Fisher, E.; Zhao, Y.; Richardson, R.; Janik, M.; Buell, A. K.; Aigbirhio, F. I.; Tóth, G. Detection and Characterization of Small Molecule Interactions with Fibrillar Protein Aggregates Using Microscale Thermophoresis. *ACS Chem. Neurosci.* **2017**, *8*, 2088–2095.
- (8) Seidel, S. A. I.; Wienken, C. J.; Geissler, S.; Jerabek-Willemsen, M.; Duhr, S.; Reiter, A.; Trauner, D.; Braun, D.; Baaske, P. Label-free microscale thermophoresis dis-

- criminate sites and affinity of protein-ligand binding. *Angew. Chem. Int. Ed.* **2012**, *51*, 10656–10659.
- (9) Alexander, C. G.; Jürgens, M. C.; Shepherd, D. A.; Freund, S. M. V.; Ashcroft, A. E.; Ferguson, N. Thermodynamic origins of protein folding, allostery, and capsid formation in the human hepatitis B virus core protein. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, E2782–E2791.
 - (10) Jerabek-Willemsen, M.; Andre, T.; Wanner, R.; Roth, H. M.; Duhr, S.; Baaske, P.; Breitsprecher, D. MicroScale Thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.* **2014**, *1077*, 101–113.
 - (11) Niether, D.; Kawaguchi, T.; Hovancova, J.; Eguchi, K.; Dhont, J. K. G.; Kita, R.; Wiegand, S. Role of Hydrogen Bonding of Cyclodextrin-Drug Complexes Probed by Thermodiffusion. *Langmuir* **2017**, *33*, 8483–8492.
 - (12) Würger, A. Molecular-Weight Dependent Thermal Diffusion in Dilute Polymer Solutions. *Phys. Rev. Lett.* **2009**, *102*, 078302–1–078302–4.
 - (13) Morozov, K. I.; Köhler, W. Thermophoresis of Polymers: Nondraining vs Draining Coil. *Langmuir* **2014**, *30*, 6571–6576.
 - (14) Zimmermann, N. E. R.; Guevara-Carrion, G.; Vrabec, J.; Hansen, N. Predicting and Rationalizing the Soret Coefficient of Binary Lennard–Jones Mixtures in the Liquid State. *Adv. Theory Simul.* **2022**, 2200311.
 - (15) Hoang, H.; Galliero, G. Predicting thermodiffusion in simple binary fluid mixtures. *The Eur. Phys. J. E* **2022**, *45*, 42.
 - (16) Morozov, K. I. Predicting the Soret Coefficient of Molecular Binary Mixtures. *Ind. Eng. Chem. Res.* **2024**, *63*, 9245–9255.

- (17) Araujo-Rocha, M.; Diaz-Marquez, A.; Stirnemann, G. On the validity of some equilibrium models for thermodiffusion. *Comptes Rendus. Chimie* **2024**, *27*, 1–12.
- (18) Iacopini, S.; Piazza, R. Thermophoresis in protein solutions. *Europhys. Lett.* **2003**, *63*, 247–253.
- (19) Piazza, R.; Iacopini, S.; Triulzia, B. Thermophoresis as a probe of particle-solvent interactions: The case of protein solutions. *Phys. Chem. Chem. Phys.* **2004**, *6*, 1616–1622.
- (20) Iacopini, S.; Rusconi, R.; Piazza, R. The "macromolecular tourist": Universal temperature dependence of thermal diffusion in aqueous colloidal suspensions. *Eur. Phys. J. E* **2006**, *19*, 59–67.
- (21) Kishikawa, Y.; Wiegand, S.; Kita, R. Temperature Dependence of Soret Coefficient in Aqueous and Nonaqueous Solutions of Pullulan. *Biomacromolecules* **2010**, *11*, 740–747.
- (22) Niether, D.; Sarter, M.; Koenig, B. W.; Fitter, J.; Stadler, A. M.; Wiegand, S. Thermophoresis: The Case of Streptavidin and Biotin. *Polymers* **2020**, *12*, 376.
- (23) Mohanakumar, S.; Luettmmer-Strathmann, J.; Wiegand, S. Thermodiffusion of aqueous solutions of various potassium salts. *J. Chem. Phys.* **2021**, *154*, 084506.
- (24) Mohanakumar, S.; Lee, N.; Wiegand, S. Complementary Experimental Methods to Obtain Thermodynamic Parameters of Protein Ligand Systems. *Int. J. Mol. Sci.* **2022**, *23*, 14198.
- (25) Lee, N.; Mohanakumar, S.; Briels, W. J.; Wiegand, S. Non-monotonic Soret coefficients of aqueous LiCl solutions with varying concentrations. *Phys. Chem. Chem. Phys.* **2024**, *26*, 7830–7836.

- (26) Rudani, B. A.; Jakubowski, A.; Kriegs, H.; Wiegand, S. Deciphering the guanidinium cation: Insights into thermal diffusion. *J. Chem. Phys.* **2024**, *160*, 214502.
- (27) Niether, D.; Di Lecce, S.; Bresme, F.; Wiegand, S. Unravelling the hydrophobicity of urea in water using thermodiffusion: implications for protein denaturation. *Phys. Chem. Chem. Phys.* **2018**, *20*, 1012–1020.
- (28) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug. Deliver. Rev.* **2001**, *46*, 3–26.
- (29) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2012**, *64*, 4–17.
- (30) Pal, P.; Chakraborty, S.; Jana, B. Number of Hydrogen Bonds per Unit Solvent Accessible Surface Area: A Descriptor of Functional States of Proteins. *J. Phys. Chem. B* **2022**, *126*, 10822–10833.
- (31) Bringuier, E.; Bourdon, A. Colloid transport in nonuniform temperature. *Phys. Rev. E* **2003**, *67*, 011404.
- (32) Piazza, R. 'Thermal forces': colloids in temperature gradients. *J. Phys.: Condens. Matter* **2004**, *16*, S4195–S4211.
- (33) Dhont, J. K. G.; Wiegand, S.; Duhr, S.; Braun, D. Thermodiffusion of charged colloids: Single-particle diffusion. *Langmuir* **2007**, *23*, 1674–1683.
- (34) Würger, A. Transport in charged colloids driven by thermoelectricity. *Phys. Rev. Lett.* **2008**, *101*, 108302–1–108302–4.
- (35) Ning, H.; Dhont, J. K. G.; Wiegand, S. Thermal-diffusive behavior of a dilute solution of charged colloids. *Langmuir* **2008**, *24*, 2426–2432.

- (36) Dhont, J. K. G.; Briels, W. J. Single-particle thermal diffusion of charged colloids: Double-layer theory in a temperature gradient. *Eur. Phys. J. E* **2008**, *25*, 61–76.
- (37) Wang, Z.; Kriegs, H.; Buitenhuis, J.; Dhont, J. K. G.; Wiegand, S. Thermophoresis of charged colloidal rods. *Soft Matter* **2013**, *9*, 8697–8704.
- (38) Di Pu; Panahi, A.; Natale, G.; Benneker, A. M. A Mode-Coupling Model of Colloid Thermophoresis in Aqueous Systems: Temperature and Size Dependencies of the Soret Coefficient. *Nano Lett.* **2024**, *24*, 2798–2804.
- (39) Qin, Y.; Zhang, L.; Wang, L.; Zhong, D. Observation of the Global Dynamic Collectivity of a Hydration Shell around Apomyoglobin. *J. Phys. Chem. Lett.* **2017**, *8*, 1124–1131.
- (40) Stadler, A. M.; Demmel, F.; Ollivier, J.; Seydel, T. Picosecond to nanosecond dynamics provide a source of conformational entropy for protein folding. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21527–21538.
- (41) Dyson, H. J.; Wright, P. E. How Does Your Protein Fold? Elucidating the Apomyoglobin Folding Pathway. *Acc. Chem. Res.* **2017**, *50*, 105–111.
- (42) Brooks, C. L. Characterization of “native” apomyoglobin by molecular dynamics simulation. *J. Mol. Biol.* **1992**, *227*, 375–380.
- (43) Tirado-Rives, J.; Jorgensen, W. L. Molecular dynamics simulations of the unfolding of apomyoglobin in water. *Biochemistry* **1993**, *32*, 4175–4184.
- (44) Zhang, D.; Lazim, R. Application of conventional molecular dynamics simulation in evaluating the stability of apomyoglobin in urea solution. *Sci. Rep.* **2017**, *7*, 44651.
- (45) Tcherkasskaya, O.; Uversky, V. N. Denatured collapsed states in protein folding: example of apomyoglobin. *Proteins* **2001**, *44*, 244–254.

- (46) Miyashita, Y.; Wazawa, T.; Mogami, G.; Takahashi, S.; Sambongi, Y.; Suzuki, M. Hydration-state change of horse heart cytochrome c corresponding to trifluoroacetic-acid-induced unfolding. *Biophys. J.* **2013**, *104*, 163–172.
- (47) Sales, A. E.; Breydo, L.; Porto, T. S.; Porto, A. L. F.; Uversky, V. N. Hydrophobicity-dependent effects of polymers on different protein conformations. *RSC Adv.* **2016**, *6*, 42971–42983.
- (48) Azami-Movahed, M.; Meratan, A. A.; Ghasemi, A.; Ebrahim-Habibi, A.; Nemat-Gorgani, M. Acetylation of lysine residues in apomyoglobin: Structural changes, amyloid fibrillation, and role of surface charge. *Int. J. Biol. Macromol.* **2018**, *107*, 626–634.
- (49) Mizukami, T.; Xu, M.; Fazlieva, R.; Bychkova, V. E.; Roder, H. Complex Folding Landscape of Apomyoglobin at Acidic pH Revealed by Ultrafast Kinetic Analysis of Core Mutants. *J. Phys. Chem. B* **2018**, *122*, 11228–11239.
- (50) Xu, M.; Beresneva, O.; Rosario, R.; Roder, H. Microsecond Folding Dynamics of Apomyoglobin at Acidic pH. *J. Phys. Chem. B* **2012**, *116*, 7014–7025.
- (51) Wang, F.; Tang, X.-J. Conformational Heterogeneity and Stability of Apomyoglobin Studied by Hydrogen/Deuterium Exchange and Electrospray Ionization Mass Spectrometry. *Biochem.* **1996**, *35*, 4069–4078.
- (52) Goto, Y.; Fink, A. L. Phase diagram for acidic conformational states of apomyoglobin. *J. Mol. Biol.* **1990**, *214*, 803–805.
- (53) Mizukami, T.; Sakuma, Y.; Maki, K. Statistical Mechanical Model for pH-Induced Protein Folding: Application to Apomyoglobin. *J. Phys. Chem. B* **2016**, *120*, 8970–8986.
- (54) Balacescu, L.; Schrader, T. E.; Radulescu, A.; Zolnierczuk, P.; Holderer, O.; Pasini, S.; Fitter, J.; Stadler, A. M. Transition between protein-like and polymer-like dynamic be-

- havior: Internal friction in unfolded apomyoglobin depends on denaturing conditions. *Sci. Rep.* **2020**, *10*, 1570.
- (55) Onufriev, A.; Case, D. A.; Bashford, D. Structural Details, Pathways, and Energetics of Unfolding Apomyoglobin. *J. Mol. Biol.* **2003**, *325*, 555–567.
- (56) Stadler, A. M.; Koza, M. M.; Fitter, J. Determination of Conformational Entropy of Fully and Partially Folded Conformations of Holo- and Apomyoglobin. *J. Phys. Chem. B* **2015**, *119*, 72–82.
- (57) Kuffel, A.; Zielkiewicz, J. Why the Solvation Water around Proteins Is More Dense than Bulk Water. *J. Phys. Chem. B* **2012**, *116*, 12113–12124.
- (58) Brudar, S.; Hribar-Lee, B. Effect of Buffer on Protein Stability in Aqueous Solutions: A Simple Protein Aggregation Model. *J. Phys. Chem. B* **2021**, *125*, 2504–2512.
- (59) Kay, M. S.; Baldwin, R. L. Alternative Models for Describing the Acid Unfolding of the Apomyoglobin Folding Intermediate. *Biochem.* **1998**, *37*, 7859–7868.
- (60) Case, D.A., Aktulga, H.M., Belfon, K., Ben-Shalom, I.Y., Berryman, J.T., Brozell, S.R., Carvahol, F.S., Cerutti, D.S., Cheatham, T.E., III, Cisneros, G.A., Cruzeiro, V.W.D., Darden, T.A., Forouzesh, N., Ghazimirsaeed, M., Giambasu, G., Giese, T., Gilson, M.K., Gohlke, H., Goetz, A.W., Harris, J., Huang, Z., Izadi, S., Izmailov, S.A., Kasavajhala, K., Kaymak, M.C., Kolossváry, I., Kovalenko, A., Kurtzman, T., Lee, T.S., Li, P., Li, Z., Lin, C., Liu, J., Luchko, T., Luo, R., Machado, M., Manathunga, M., Merz, K.M., Miao, Y., Mikhailovskii, O., Monard, G., Nguyen, H., O’Hearn, K.A., Onufriev, A., Pan, F., Pantano, S., Rahnamoun, A., Roe, D.R., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shajan, A., Shen, J., Simmerling, C.L., Skrynnikov, N.R., Smith, J., Swails, J., Walker, R.C., Wang, J., Wang, J., Wu, X., Wu, Y., Xiong, Y., Xue, Y., York, D.M., Zhao, C., Zhu, Q., Kollman, P.A. (2024), Amber 2024, University of California, San Francisco.

- (61) Swails, J. M.; York, D. M.; Roitberg, A. E. Constant pH Replica Exchange Molecular Dynamics in Explicit Solvent Using Discrete Protonation States: Implementation, Testing, and Validation. *J. Chem. Theory Comput.* **2014**, *10*, 1341–1352.
- (62) Mongan, J.; Case, D. A.; McCammon, J. A. Constant pH molecular dynamics in generalized Born implicit solvent. *J. Comput. Chem.* **2004**, *25*, 2038–2048.
- (63) Wilkins, M. R.; Gasteiger, E.; Bairoch, A.; Sanchez, J. C.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F. Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* **1999**, *112*, 531–552.
- (64) Micsonai, A.; Wien, F.; Bulyáki, É.; Kun, J.; Moussong, É.; Lee, Y.-H.; Goto, Y.; Réfrégiers, M.; Kardos, J. BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucl. Acids Res.* **2018**, *46*, W315–W322.
- (65) Wiegand, S.; Ning, H.; Kriegs, H. Thermal diffusion forced Rayleigh scattering setup optimized for aqueous mixtures. *J. Phys. Chem. B* **2007**, *111*, 14169–14174.
- (66) Blanco, P.; Kriegs, H.; Lettinga, M. P.; Holmqvist, P.; Wiegand, S. Thermal Diffusion of a Stiff Rod-Like Mutant Y21M fd-Virus. *Biomacromolecules* **2011**, *12*, 1602–1609.
- (67) Rossmannith, P.; Köhler, W. Polymer polydispersity analysis by thermal diffusion forced Rayleigh scattering. *Macromolecules* **1996**, *29*, 3203–3211.
- (68) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (69) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

- (70) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (71) Schott-Verdugo, S.; Gohlke, H. PACKMOL-Memgen: A Simple-To-Use, Generalized Workflow for Membrane-Protein-Lipid-Bilayer System Building. *J. Chem. Inf. Model.* **2019**, *59*, 2522–2528.
- (72) Gaussian 09, Revision A.02, Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, J. A., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Keith, T., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B., Fox, D. J. , Gaussian, Inc., Wallingford CT, 2016.
- (73) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins Struct. Funct. Bioinf.* **2006**, *65*, 712–725.
- (74) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

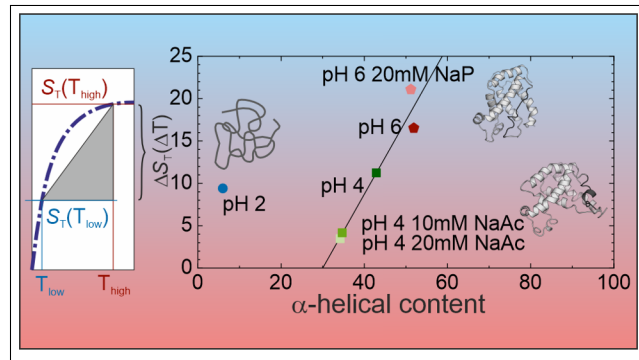
- (75) Joung, I. S.; Cheatham, T. E. I. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.
- (76) Quigley, D.; Probert, M. I. J. Langevin dynamics in constant pressure extended systems. *J. Chem. Phys.* **2004**, *120*, 11432–11441.
- (77) Roe, D. R. and Cheatham, T. E. III PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.
- (78) Yee, A. A.; Savchenko, A.; Ignachenko, A.; Lukin, J.; Xu, X.; Skarina, T.; Evdokimova, E.; Liu, C. S.; Semesi, A.; Guido, V.; Edwards, A. M.; Arrowsmith, C. H. NMR and X-ray Crystallography, Complementary Tools in Structural Proteomics of Small Proteins. *J. Am. Chem. Soc.* **2005**, *127*, 16512–16517.
- (79) Jamin, M.; Yeh, S.; Rousseau, D. L.; Baldwin, R. L. Submillisecond Unfolding Kinetics of Apomyoglobin and its pH 4 Intermediate. *J. Mol. Biol.* **1999**, *292*, 731–740.
- (80) Jamin, M.; Antalík, M.; Loh, S. N.; Bolen, D. W.; Baldwin, R. L. The unfolding enthalpy of the pH 4 molten globule of apomyoglobin measured by isothermal titration calorimetry. *Protein Science* **2000**, *9*, 1340–1346.
- (81) Picotti, P.; Marabotti, A.; Negro, A.; Musi, V.; Spolaore, B.; Zambonin, M.; Fontana, A. Modulation of the structural integrity of helix F in apomyoglobin by single amino acid replacements. *Protein Sci.* **2004**, *13*, 1572–1585.
- (82) Nishii, I.; Kataoka, M.; Goto, Y. Thermodynamic Stability of the Molten Globule States of Apomyoglobin. *J. Mol. Biol.* **1995**, *250*, 223–238.
- (83) Bertagna, A. M.; Barrick, D. Nonspecific hydrophobic interactions stabilize an equi-

- librium intermediate of apomyoglobin at a key position within the AGH region. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12514–12519.
- (84) Jennings, P. A.; Wright P. E. Formation of a Molten Globule Intermediate Early in the Kinetic Folding Pathway of Apomyoglobin. *Science* **1993**, *262*, 892–896.
- (85) Eliezer, D.; Jennings, P. A.; Dyson, H.; Wright, P. E. Populating the equilibrium molten globule state of apomyoglobin under conditions suitable for structural characterization by NMR. *FEBS Lett.* **1997**, *417*, 92–96.
- (86) Eliezer, D.; Yao, J.; Dyson, H. J.; Wright, P. E. Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Mol. Biol.* **1998**, *5*, 148–155.
- (87) Calculator Plugins were used for structure property prediction and calculation, Marvin 24.1.3, 2024, ChemAxon (<http://www.chemaxon.com>).
- (88) Eiberweiser, A. Hydration and Ion Pairing of Aqueous Phosphate Solutions as Observed by Dielectric Spectroscopy. Ph.D. thesis, Universität, Regensburg, 2013.
- (89) Rahman, H. M. A.; Hefter, G.; Buchner, R. Hydration of Formate and Acetate Ions by Dielectric Relaxation Spectroscopy. *J. Phys. Chem. B* **2012**, *116*, 314–323.
- (90) Zbacnik, T. J.; Holcomb, R. E.; Katayama, D. S.; Murphy, B. M.; Payne, R. W.; Coccaro, R. C.; Evans, G. J.; Matsuura, J. E.; Henry, C. S. ; Manning, M. C. Role of Buffers in Protein Formulations. *J. Pharm. Sci.* **2017**, *106*, 713–733.
- (91) Salis, A.; Monduzzi, M. Not only pH. Specific buffer effects in biological systems. *Curr. Opin. Colloid Interface Sci.* **2016**, *23*, 1–9.
- (92) Salis, A.; Ninham, B. W. Models and mechanisms of Hofmeister effects in electrolyte solutions, and colloid and protein systems revisited. *Chem. Soc. Rev.* **2014**, *43*, 7358–7377.

- (93) Lo Nostro, P.; Ninham, B. W. Hofmeister Phenomena: An Update on Ion Specificity in Biology. *Chem. Rev.* **2012**, *112*, 2286–2322.
- (94) Collins, K. D. Why continuum electrostatics theories cannot explain biological structure, polyelectrolytes or ionic strength effects in ion–protein interactions. *Biophys. Chem.* **2012**, *167*, 43–59.
- (95) Muthuselvi, L.; Dhathathreyan, A. Understanding dynamics of myoglobin in heterogeneous aqueous environments using coupled water fractions. *Adv. Colloid. Interfac.* **2009**, *150*, 55–62.
- (96) Kalyuzhnyi, Y. V.; Vlachy, V. Explicit-water theory for the salt-specific effects and Hofmeister series in protein solutions. *J. Chem. Phys.* **2016**, *144*, 215101.
- (97) Janc, T.; Vlachy, V.; Lukšič, M. Calorimetric studies of interactions between low molecular weight salts and bovine serum albumin in water at pH values below and above the isoionic point. *J. Mol. Liq.* **2018**, *270*, 74–80.
- (98) Doi, M.; Edwards, S. F. *The theory of polymer dynamics*, reprint ed.; International series of monographs on physics; Clarendon Press: Oxford, 2013; Vol. 73.
- (99) Machado, L. O.; Reis, D.; Figueiredo Neto, A. M. The Soret coefficient of human low-density lipoprotein in solution: a thermophilic behavior. *Eur. Phys. J. E* **2023**, *46*, 124.
- (100) Wolff, M.; Mittag, J. J.; Herling, T. W.; de Genst, E.; Dobson, C. M.; Knowles, T. P. J.; Braun, D.; Buell, A. K. Quantitative thermophoretic study of disease-related protein aggregates. *Sci. Rep.* **2016**, *6*, 22829.
- (101) Huang, D. M.; Chandler, D. Temperature and length scale dependence of hydrophobic effects and their possible implications for protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8324–8327.

- (102) Narita, M.; Narita, M.; Itsuno, Y.; Itsuno, S. Autonomous Sequences in Myoglobin Emerging from X-ray Structure of Holomyoglobin. *ACS omega* **2019**, *4*, 992–999.
- (103) Vernier, G.; Chenal, A.; Vitrac, H.; Barumandzadhe, R.; Montagner, C.; Forge, V. Interactions of apomyoglobin with membranes: mechanisms and effects on heme uptake. *Protein Sci.* **2007**, *16*, 391–400.
- (104) Weisbuch, S.; Gérard, F.; Padeloup, M.; Cappadoro, J.; Dupont, Y.; Jamin, M. Cooperative Sub-Millisecond Folding Kinetics of Apomyoglobin pH 4 Intermediate. *Biochem.* **2005**, *44*, 7013–7023.
- (105) Katayama, D. S. ; Nayar, R.; Chou, D. K.; Valente, J. J. ; Cooper, J.; Henry, C. S. ; Velde, D. G. V.; Villarete, L.; Liu, C.P.; Manning, M. C. Effect of buffer species on the thermally induced aggregation of interferon-tau. *J. Pharm. Sci.* **2006**, *95*, 1212–1226.
- (106) Kim, S. H.; Yoo, H. J.; Park, E. J.; Na, D. H. Nano Differential Scanning Fluorimetry-Based Thermal Stability Screening and Optimal Buffer Selection for Immunoglobulin G. *Pharmaceuticals* **2022**, *15*, 29.
- (107) Kameoka, D.; Masuzaki, E.; Ueda, T.; Imoto, T. Effect of Buffer Species on the Unfolding and the Aggregation of Humanized IgG. *J. Biochem.* **2007**, *142*, 383–391.
- (108) Liese, S.; Gensler, M.; Krysiak, S.; Schwarzl, R.; Achazi, A.; Paulus, B.; Hugel, T.; Rabe, J. P.; Netz, R. R. Hydration Effects Turn a Highly Stretched Polymer from an Entropic into an Energetic Spring. *ACS Nano* **2017**, *11*, 702–712.
- (109) Niether, D.; Kriegs, H.; Dhont, J. K. G.; Wiegand, S. Peptide model systems: Correlation between thermophilicity and hydrophilicity. *J. Chem. Phys.* **2018**, *149*, 044506.

TOC Graphic



We observe an increase in the thermodiffusion parameter, $\Delta S_T(\Delta T)$, for apomyoglobin with increasing α -helical content, alongside an increase in hydrophilicity.