POLITECNICO DI TORINO
Repository ISTITUZIONALE

Understanding glucagon aggregation: in silico insights and experimental validation

(Article begins on next page)

24 April 2026

Understanding Glucagon Aggregation: *In-silico* Insights and Experimental Validation

Roberto Pisano[1,*,‡], *Andrea Arsiccio*[1,†,‡], *Valerie Collins*[2,‡], *Patrick King*[2], *Marco Macis*[3],

Walter Cabri[3,4], *Antonio Ricci*[3,*]

[1] Department of Applied Science and Technology, Politecnico di Torino, 24 corso Duca degli Abruzzi, IT-10129, Torino, Italy

[2] Redshift Bioanalytics, 80 Central Street, Boxborough, MA-01719, USA

[3] Fresenius Kabi iPSUM, 108 via Roma, IT-20081, Cassina de Pecchi, Italy

[4] Center for Chemical Catalysis Department of Chemistry "Giacomo Ciamician", University of Bologna, 40126 Bologna, Italy

[†]Present address: Coriolis Pharma Research, 18B Fraunhoferstraße, DE-82152, Martinsried, Germany

[‡]These authors contributed equally to the manuscript editing.

[*]Corresponding authors: roberto.pisano@polito.it, antonio.ricci@fresenius-kabi.com

1

KEYWORDS. glucagon aggregation, in-silico simulations, experimental validation, peptide

stability, 2-HPβCD, lactose

ABSTRACT

Peptide aggregation poses a significant challenge in biopharmaceutical development and neurodegenerative diseases. This study combines computational simulations and experimental validation to uncover the underlying mechanisms and countermeasures for the aggregation of glucagon, a peptide with a high tendency to aggregate. *In silico* simulations demonstrate that lactose and 2-hydroxypropyl-β-cyclodextrin (2-HPβCD) influence glucagon aggregation differently: lactose stabilizes glucagon by increasing α-helical content, while 2-HPβCD disrupts protein-protein interactions. According to the simulations, 2-HPβCD is particularly effective at preserving the monomeric form of glucagon. Experimental validation with Microfluidic Modulation Spectroscopy (MMS) confirms these findings, showing that glucagon in the presence of 2-HPβCD remains structurally stable, supporting the anti-aggregation effect of this excipient. This research provides essential insights into glucagon aggregation obtained through a new powerful tool for the monitoring of the critical property of peptide aggregation, suggesting new strategies for addressing this challenge in therapeutic peptide development.

1.   INTRODUCTION

Peptide aggregation is a fundamental concern in the field of biochemistry. The propensity of peptides to aggregate, forming amyloid fibrils and misfolded structures, can significantly impact their stability and biological function.[1-2] This phenomenon has risen significant attention due to its profound implications in various fields, including the development of biopharmaceuticals and the understanding of neurodegenerative diseases. To provide an example, the aggregation of beta-amyloid peptides has been associated with Alzheimer's disease. These beta-amyloid peptides, particularly β-amyloid 42 (Aβ42), tend to aggregate and form insoluble plaques within the brain.[3] These plaque formations are believed to contribute to the deterioration of neurons and the characteristic cognitive decline observed in Alzheimer's disease.[4] The accumulation of these aggregated beta-amyloid peptides remains a distinctive hallmark of Alzheimer's disease pathology.

Additionally, the proclivity of peptides to aggregate has been associated with challenges in the stability and functionality of therapeutic peptides, such as insulin[5] and glucagon.[6] As a result, elucidating the mechanisms of peptide aggregation and developing strategies to inhibit this process have become critical objectives in peptide research.

Peptide aggregation characterization still lacks optimization. Commonly used techniques like Size Exclusion Chromatography (SEC) and ultracentrifugation (UC) have limitations; SEC employs a stationary phase and denaturant solvents like acetonitrile, while UC is not sensitive enough to track small peptide aggregates. A recent study highlighted Asymmetric Field Flow Fractionation (Af4) as a potent tool for monitoring the aggregation pattern of peptides under native condition.[7] However, while Af4 measures the aggregates' size, it does not reveal details about their

secondary/tertiary structures. Consequently, having an experimental tool to monitor the structural aspects of aggregates in their native state would complement understanding peptide aggregation.

In the case of glucagon, this aggregation tendency has raised substantial challenges in its development as a therapeutic agent for treating severe hypoglycemia.[6-8] Therefore, understanding the mechanisms of peptide aggregation and exploring strategies to mitigate these effects is a crucial endeavor. In this article, we investigated the complex issue of glucagon aggregation, utilizing both experimental and computational methods, to shed light on potential solutions for this challenging problem.

Glucagon is a vital peptide hormone consisting of 29 amino acids, and it plays a critical role in regulating glucose levels. Its primary function involves promoting glycogenolysis and gluconeogenesis by binding to the glucagon receptor[9] and, thus, is used for the treatment of severe hypoglycemia associated with diabetes.[10-12]

Glucagon exhibits constraints in terms of both its physical and chemical stability, and it has a significant inclination toward aggregation. Mishandling of glucagon has been demonstrated to lead to the creation of amyloidogenic fibrils,[13-14] and this process has been linked to the development of β-sheet structures.[15] The mechanisms underlying the aggregation of glucagon have been extensively investigated. Researchers have explored its behavior in various conditions, including the presence of metals[16], varying pH levels, and different filtration methods[17], as well as investigations into new formulations for drug products.[18] Furthermore, numerous scientific studies have examined the morphological changes of glucagon, from its individual monomeric form to oligomeric and fibrillation states.[19,20]

Numerous initiatives have been undertaken to mitigate the risk of aggregation in glucagon. These efforts encompass optimizing its formulation,[18-21] modifying its chemical structure,[22] and creating assays for the detection of aggregates.[23-24]

The primary aim of this research is to investigate the likelihood of glucagon aggregation under different formulation conditions. In addition to proposing alternative methods to prevent aggregation, this study seeks to use computational tools to better understand how the molecule behaves at the molecular level. Furthermore, an innovative Microfluidic Modulation Spectroscopy technique is used to monitor how glucagon clumps together in real-time under various conditions, with a focus on its secondary structure. Ultimately, our aim is to link predictions from molecular dynamics simulations with experimental observations.

As concerns the formulation, we investigated the use of lactose or 2-hydroxypropyl-β-cyclodextrin (2-HPβCD) as excipients. 2-HPβCD is composed of 7 α-glucopyranose monomers bonded together to create a truncated cone shape. The oxygen atoms in the glucopyranose monomers are connected to hydroxypropyl derivations, which enhance water solubility and provide amphiphilic properties.[25-27]

Lactose and HPβCD were selected due to their contrasting behavior. Lactose is repelled from the protein surface, thereby acting as a stabilizer, in line with the preferential exclusion theory.[28-29] In contrast, HPβCD exhibits a preference for interacting with the peptide backbone, effectively acting as a mild denaturant.[30] Furthermore, various authors showed that cyclodextrins serve as protective agents against protein aggregation.[31,32]

To analyze the aggregation of glucagon, a combination of experimental and simulation approaches was employed. The first-generation MMS instrument, created by RedShift BioAnalytics and relying on infrared (IR) spectroscopy, was used for the experimental part. Circular dichroism (CD) is another useful tool for the measurement of protein secondary structures and is more sensitive to α-helical proteins as it utilizes the differential absorption of left and right-handed polarised light to make measurements. It is much more difficult to deconvolute spectra when samples contain other types of secondary structures in combination, as they often overlay in similar regions of the spectrum and can be positive or negative in magnitude. This is particularly true for two of the most useful types of secondary structure to measure - intramolecular beta sheet, and intermolecular beta sheet. Intermolecular beta sheet is especially important to monitor as it is very commonly associated with structural changes that accompany structural rearrangement leading to aggregation, whereas intramolecular beta sheet is often part of the natively-folded protein. A good example of the utility of quantifying both types of beta-sheets is for stability or formulation experiments in the analysis of monoclonal antibodies (mAbs), whose complex structures contain both inter and intra molecular beta-sheets.

Additionally, due to the low wavelength of light used for CD measurements, scattering effects are prominent and make measurements difficult in all but very dilute protein concentrations, without aggregates, and in simple, low-concentration buffers. While employing larger path-length cells can help address the first issue, using this method exacerbates the second problem even further. Infrared (IR) spectroscopy of protein secondary structure does not suffer from any of these problems. Signals for all secondary structures are well-defined and of the same magnitude using

transmission measurements, meaning deconvolution is straightforward, and secondary structures are quantifiable. Using IR, the wavelengths used for absorbance measurements in the amide I band, the region containing the most information for protein secondary structure, are more than an order of magnitude longer than for CD (in the approximate range 5800-6200nm), meaning scattering effects are very minimal, even with larger complexes or aggregates present. Buffer scattering is also minimal, meaning measurements can be made *in situ*, at the concentration and in the buffer of interest directly. The major issue with IR spectroscopy of protein secondary structure in the past has been that the signal to noise is relatively low compared to some other types of spectroscopy, requiring a relatively high protein concentration for repeatable measurements to be made. By using a powerful and precise quantum cascade laser in combination with a Microfluidic rapid referencing method, Microfluidic Modulation Spectroscopy (MMS) solves this problem, and so brings the benefits of IR analysis to a much lower and therefore more useful range of protein concentrations (0.1 to >200 mg/ml *in situ*) for the first time.

This experimental analysis was further enhanced by integrating an advanced implicit solvent molecular dynamics technique, specifically designed for examining interactions between proteins and excipients.[33]

2. MATERIALS AND METHODS

2.1. Implicit Solvent Simulations

2.1.1. Theoretical Background

The free energy of a solvated protein G^{tot} can be described as the sum of the following contributions,

$$G^{tot} = E^{vac} + G^{el} + G^{np} + G^{tr}(c) \tag{1}$$

where E^{vac} is the energy of the protein in vacuum, which results from both internal contributions (bond and angle stretching, dihedral angles interactions) and van der Waals energy terms. G^{np} is the contribution from non-polar solvation in pure water, i.e., the free energy of hydration for a molecule from which all charges have been removed. G^{el} is the electrostatic part, calculated as the free energy for turning on the partial charges in the solution. These first three terms describe a protein in pure water. The term $G^{tr}(c)$ is used to take into account the presence of co-solutes at a concentration c, and is based on the free energy of transfer approach MIST (Model with Implicit Solvation Thermodynamics), that was recently developed.[33]

The generalized Born equation[34] was employed to compute G^{el}, using the OBC(II) model[35] to estimate the Born radii. The dielectric constant was set to 78.4, which is the value for water at 298 K according to the equations in Bradley and Pitzer.[36]

The non-polar contribution G^{np} was expressed as,

$$G^{np} = \gamma \, SASA \tag{2}$$

where γ is a surface tension, while SASA is the total solvent accessibility of the protein. In this study, a specific value of γ, which is 5 cal mol^{-1} Å^{-2}, was employed.[37]

The transfer free energy term $G^{tr}(c)$, that considers the effect of co-solutes, is described as,[33]

$$G^{tr}(c) = \sum_{k=1}^{n_r} g_{k,sc}^{tr}(c)\alpha_{k,sc} + g_{bb}^{tr}(c)\sum_{k=1}^{n_r} \alpha_{k,bb} \tag{3}$$

where the summations run over the n_r residues of the protein, and $g_{k,sc}^{tr}(c)$ and $g_{bb}^{tr}(c)$ are the amino acids side chains (subscript *sc*) and backbone (subscript *bb*) contributions, respectively.

For $G^{tr}(c)$, each contribution is weighed by the fractional solvent accessible surface area SASA_k of residue k,

$$\alpha_k = \frac{\text{SASA}_k}{\text{SASA}_{k,Gly-k-Gly}} \tag{4}$$

where $\text{SASA}_{k,Gly-k-Gly}$ is the solvent accessibility of residue k in the tripeptide $Gly-k-Gly$. The fractional solvent accessible surface area α_k was calculated using the fast algorithm proposed by Hasel et al.[38] The script for the computation of the fractional solvent accessible surface area and of the energetic term $G^{tr}(c)$ is currently part of the 2.8 version of Plumed as a separate module called SASA.

2.1.2. Simulation protocol

The Amber 99SB-ILDN force field[39] was used to describe the behavior of glucagon (PDB ID: 1GCN,[40] Figure 1a) under specific conditions. Simulations were carried out at a pH of 2.7 with a net charge of +4. The AMBER 20 simulation suite,[41] in combination with Plumed 2.4.7,[42] was employed for this purpose.

The simulations involved observing how glucagon aggregated in three different environments: pure water, water containing 10.7% (w/w) lactose, and water with 10.7% (w/w) 2-HPβCD. The water or water-cosolute solutions were described implicitly, as detailed in the Theoretical Background section.

To model the behavior of glucagon and its interaction with its surroundings, we used the generalized Born/surface area model of AMBER 20 for simulating G^{el} and G^{np} in Eq. (1). Furthermore, the free energy of transfer term $G^{tr}(c)$ was introduced as an external bias using Plumed. To optimize computational efficiency and reduce the computational cost associated with the $G^{tr}(c)$ biasing force, a multiple time step approach was used[43] and, specifically, this term was added every two integration steps.
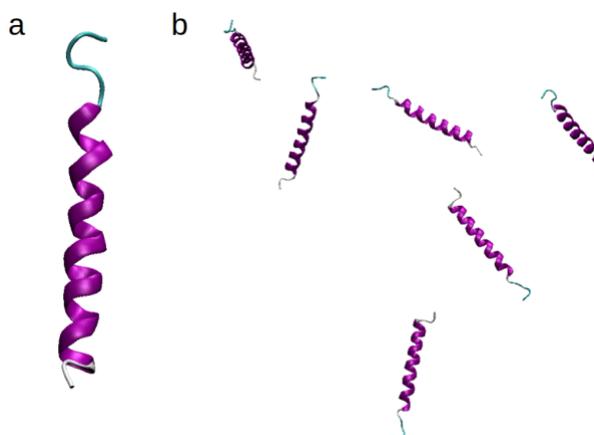

Figure 1. The starting configurations of (a) the glucagon molecule (PDB ID: 1GCN[34]) and (b) the simulated system comprising six glucagon monomers for the aggregation study. These visual representations were created using Visual Molecular Dynamics (VMD).[44]

In order to study the aggregation process, 6 glucagon molecules were randomly placed within a spherical cage of radius $r_{\text{sphere}} = 14$ nm, as depicted in Figure 1b. This configuration corresponded to a concentration of 3 g/L. A harmonic potential U_{sphere} was applied to prevent the peptides from drifting away from the borders of the spherical container.

$$\alpha_k = \begin{cases} k_{sphere}\left(|\boldsymbol{r}_i - \boldsymbol{r}_{center}| - r_{sphere}\right)^2 & if \ |\boldsymbol{r}_i - \boldsymbol{r}_{center}| > r_{sphere} \\ 0 & otherwise \end{cases} \tag{5}$$

where $\boldsymbol{r}_i$ is the position of the center of mass of peptide i, while $\boldsymbol{r}_{\text{center}}$ is the position of the center of the sphere. k_{sphere} was set to 10 kcal/(mol nm^2).

During the production runs, the simulations were conducted for a total of 300 ns. The initial 100 ns were designed as an equilibration period and were not considered for subsequent analyses. Only the last 200 ns of the simulation data were used for further analysis.

Langevin dynamics was employed to regulate the temperature at 298 K, with a collision frequency of 1.0 ps^{-1}. The SHAKE algorithm[45] was applied to restrict all bonds linking to hydrogen atoms, and a time step of 2.0 fs was used. The system's configuration was saved every 2 ps, and the center of mass translation and rotation were removed every 500 steps (1 ps). There were no cut-off values applied for handling Coulombic and Lennard-Jones interactions.

The $g_{k,sc}^{tr}(c)$ and $g_{bb}^{tr}(c)$ parameters, which will be utilized in Eq. (3), for the two distinct co-solutes under examination in this study, namely, lactose and HPβCD, were derived from all-atom explicit solvent simulations performed employing the ADD force field.[46] Specifically, the values for HPβCD were obtained from our previous computational investigation,[30] and are listed in Table S2. The free energy of transfer terms for lactose were instead extracted as described in the Supporting Information section.

2.2. Analysis of the Trajectories

2.2.1 α-helix content

The α-helix (α) content, as employed in the subsequent sections of this study, is defined as the
count of six-residue segments within the peptide exhibiting an α-helical configuration,[47]

$$\alpha = \sum_k g\left[r_{dist}\left(\{R_i\}_{i\in\Omega_\mu}, \{R^0\}\right)\right] \qquad (6)$$

The summation runs over all possible segments contributing to the α-helix, while $\{R_i\}_{i\in\Omega_\mu}$
represents the atomic coordinates of a specific set Ω_μ of six-residue fragments of the protein, and
$g(r_{dist})$ is the following switching function,

$$g(r_{dist}) = \frac{1 - \left(\frac{r_{dist}}{r_0}\right)^8}{1 - \left(\frac{r_{dist}}{r_0}\right)^{12}} \qquad (7)$$

A cutoff distance of $r_0 = 0.08$ nm was used, and r_{dist} represents the root mean square deviation
from a reference α-helix structure denoted as $\{R^0\}$.

2.2.2 Hydrogen bond analysis

The number of hydrogen bonds within each system was also measured. For the detection of a
hydrogen bond, a geometric criterion was applied, which required that the distance between the
donor and acceptor was less than 0.30 nm, and that the angle formed between the acceptor,
hydrogen, and donor atoms was greater than 135°.

2.3. Experimental validation

2.3.1 Sample preparation

Glucagon samples (provided as lyophilized powder by Fresenius Kabi) were prepared by
dissolving enough glucagon to make approximately 2 mL of 1 mg/mL in each of the following
solutions at room temperature and another set was prepared at 4°C: water adjusted to pH 2.7 with

HCl, 10.7% (w/v) lactose (Sigma 61339), and 10.7% (w/v) 2-hydroxypropyl-β-cyclodextrin

(HPβCD Sigma 389145). Samples were immediately run on an MMS system and then incubated

for 24 hours at room temperature and at 4°C and run again. The list of samples analyzed is given

in

Table 1.

Table 1. Full list of samples analyzed. Each sample was then incubated at either room temperature

(RT) or 4°C overnight and run again after 24 hours.

Sample, #	Description
1	1 mg/mL glucagon dissolved in pH 2.7 water at RT
2	1 mg/mL glucagon dissolved in pH 2.7 water at 4°C
3	1 mg/mL glucagon dissolved in 10.7% lactose at RT
4	1 mg/mL glucagon dissolved in 10.7% lactose at 4°C
5	1 mg/mL glucagon dissolved in 10.7% HPβCD at RT
6	1 mg/mL glucagon dissolved in 10.7% HPβCD at 4°C

2.3.2   MMS methods

Samples were run on a first-generation microfluidic modulation spectroscopy (AQS³pro MMS

system, RedShiftBio, Boxborough, MA, USA) system with sweep scan functionality immediately

after preparation, and again after the 24-hour incubation. MMS is an IR tool used to measure

protein secondary structure. With the use of a quantum cascade laser as the light source and constantly modulating between the sample and the references buffer, MMS can accurately measure aqueous samples at very low concentrations, even in complex or strongly-absorbing backgrounds.

All samples were loaded in sample-reference buffer pairs in a 24-well plate. A backing pressure of 5 psi was used to move the samples into the flow cell (pathlength of 22.3 µm and temperature controlled to 25°C) and 1 Hz rate was used to take readings and modulate between sample and reference buffer. The laser scanned across the amide I band from 1588 to 1712 cm^{-1} and collected and averaged 8 spectra per replicate. Three replicates were collected for each sample and the data presented are the averages of all replicates.

Lysozyme was used as the model protein for comparing and baseline matching. Savitzky-Golay smoothing with a 19-wavenumber window was applied to the second derivative plot. The second derivative plot was inverted and the baseline subtracted to yield the similarity plot, which is used for Gaussian curve fitting. Eleven Gaussian curves were used to quantitate the fractional contribution of secondary structures, as shown in the bar chart in Figure 5**Error! Reference source not found.**D.

3.   RESULTS

3.1 *In-silico* analysis

The average number of interpeptide hydrogen bonds between glucagon molecules and the average α-helical content were evaluated during the last 200 ns of the implicit solvent simulations. The results of this analysis are shown in Figure 2 where errors were estimated by block averaging.

Briefly, the equilibrated trajectories were divided into four blocks, and the standard deviation was computed over the average values of the properties in each of the blocks.

We observed that both lactose and HPβCD increased the retention of α-helical content in glucagon and decreased the formation of hydrogen bonds between peptides when compared to glucagon in pure water. These results are noteworthy because lactose and HPβCD typically exhibit opposing effects on protein conformations.

Lactose acted as a stabilizer, yielding a positive $g_{bb}^{tr}(c)$, while HPβCD had a slight denaturing effect, resulting in a negative $g_{bb}^{tr}(c)$. However, their overall impact on the aggregation propensity of glucagon was found to be similar.

Notably, the α-helical content of glucagon remained unchanged in the presence of cyclodextrin, contrary to expectations. Our interpretation of this finding is that HPβCD may act to prevent glucagon aggregation. Typically, aggregation leads to a loss of secondary structure through protein-protein interactions. However, in the presence of HPβCD, these interactions may be reduced, thereby preserving the α-helical content. This hypothesis is supported by the observation of favorable interactions between HPβCD and the protein backbone, as indicated by negative free energy of transfer values. Such interaction patterns align with those of classical denaturants, as discussed in relevant literature[48,49], where denaturants typically reduce protein-protein interactions by interacting with the protein surface. Thus, these results suggest that HPβCD may exhibit denaturant-like properties, potentially preventing secondary structure loss even at low concentrations.

These observations are corroborated by the average distribution of aggregate sizes recorded in the simulations, as shown in Figure 3. In the case of pure water, only two glucagon molecules were, on average, found in their monomeric state, while the remaining peptides tended to form higher-order clusters, especially dimers and trimers.

When either lactose or HPβCD was introduced, the number of glucagon molecules existing in the monomeric state increased to around three, with cyclodextrin exhibiting a more pronounced effect in preventing aggregation than lactose. Additionally, some 4-mers and 5-mers were sparsely populated in the presence of lactose, albeit with low occurrence.
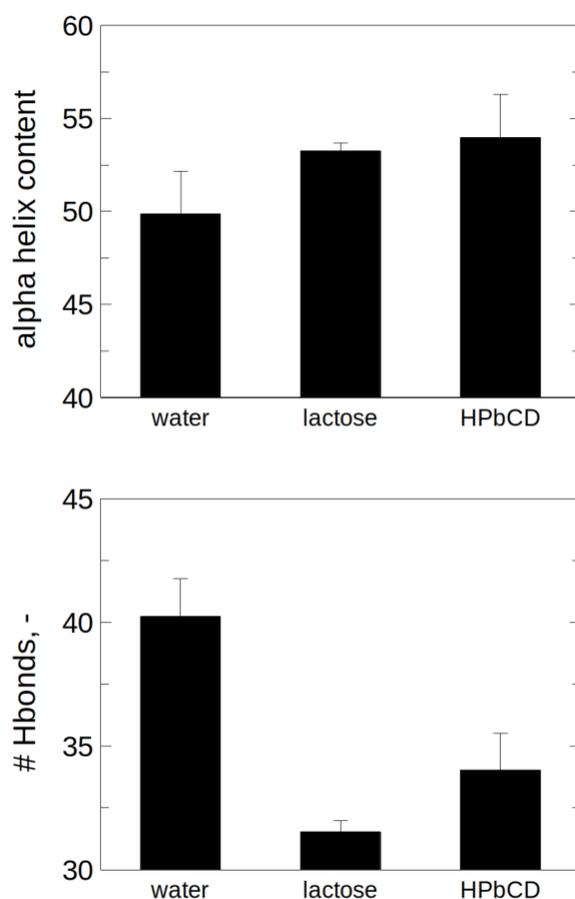
Figure 2. Average α-helical content (top graph) and number of interpeptide hydrogen bonds

(bottom graph) for the three systems investigated. The error bars show one standard deviation, as
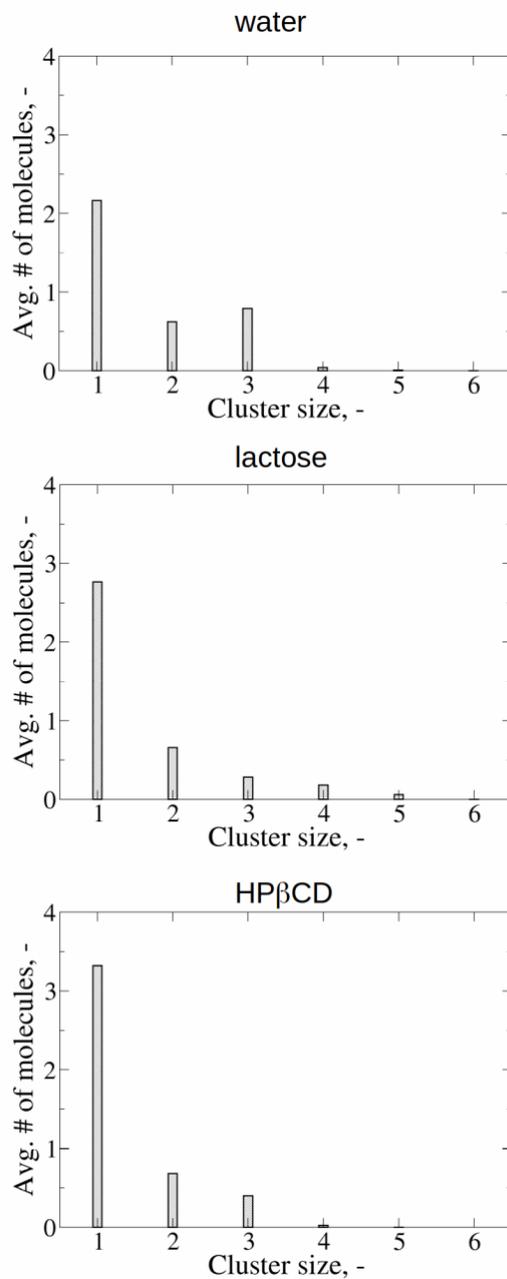
determined through the block averaging method.

Figure 3. The average number of glucagon molecules within clusters in relation to cluster size, as observed across all the solution conditions under investigation.

Lactose, acting as a stabilizer, tends to be preferentially excluded from the protein surface. As a result, lactose thermodynamically favors states that expose a reduced surface area, such as the native protein fold (in comparison to the unfolded state) or multimeric aggregate species (in comparison to the monomeric, isolated form). This preference might elucidate why 4-mers and 5-mers were more frequently observed in the presence of lactose than in a pure water environment. Although lactose slightly promotes the formation of 4-mers and 5-mers compared to water, Figure 2 demonstrates that this oligomer formation does not result in a decrease in α-helical content, nor does it increase the number of intermolecular hydrogen bonds between glucagon molecules. This result suggests that lactose, being preferentially excluded from the protein surface, shifts the equilibrium towards protein species that minimize surface area exposure. The native state is more compact than the unfolded one and is, therefore, stabilized.

Similarly, native oligomers expose less surface than monomers and are thus also stabilized by preferentially excluded excipients like lactose. These observations imply that the oligomers might be linked by relatively weak forces, making them capable of reversibly converting to the monomeric state. The formation of reversible oligomers is commonly observed with excipients that function through the preferential exclusion mechanism, such as lactose. Additionally, Figure 3 indicates that the average number of glucagon molecules in the monomeric state is higher in the presence of lactose than in water, further suggesting that lactose hinders the irreversible aggregation of glucagon. If the unfolding process is the determining step for aggregation, preferentially excluded excipients like lactose are beneficial in counteracting the aggregation cascade.

Figure 4 shows snapshots of the most likely aggregated conformations as extracted from the implicit solvent simulations. These initial oligomeric forms of glucagon did not exhibit a significant β-sheet content. Intermolecular β-sheets likely form later on during the aggregation process, in timescales that could not be reached by our molecular dynamics simulations. Instead, the first oligomers that we observed were primarily characterized by a reduced α-helical content compared to the native glucagon, and turn and coil structures served as connectors between different peptides within the clusters.

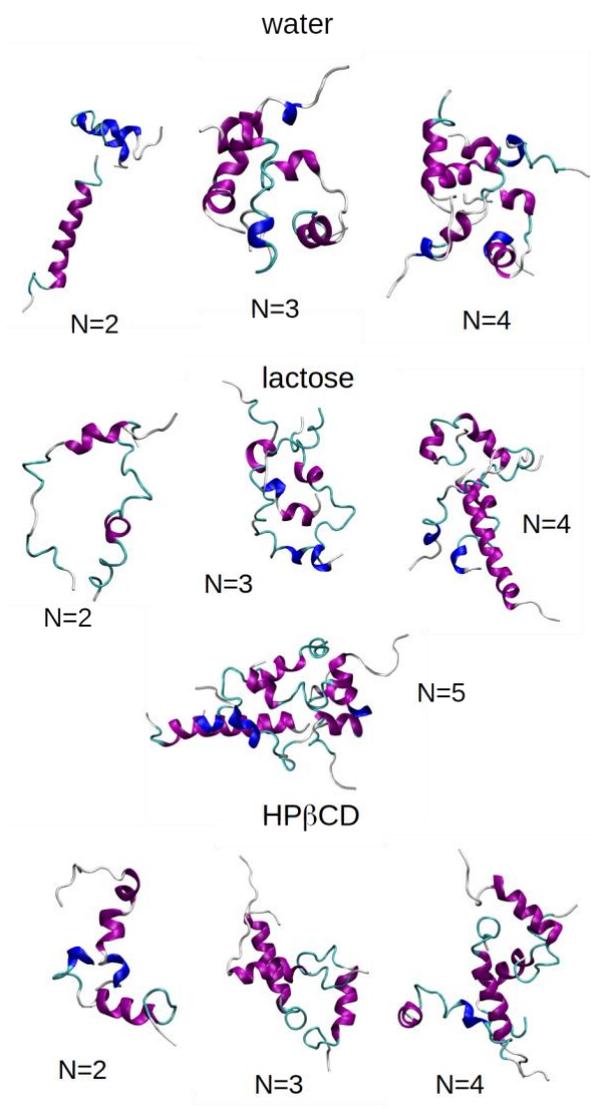Figure 4. The most likely conformations of each cluster, encompassing all the examined solution conditions. In this representation, α-helices are depicted in purple, 3_{10}-helices in blue, turns in cyan, and coils in white.

Drawing from these simulation results, it can be inferred that both lactose and, to an even greater extent, HPβCD foster the monomeric state of glucagon when compared to water. Given that lactose

and HPβCD exhibit contrasting behaviors, with lactose acting as a stabilizer and HPβCD as a denaturant, it is clear that their mechanisms of action differ.

Based on these findings, we hypothesize that denaturation and aggregation are closely interconnected phenomena in the case of glucagon. This hypothesis is substantiated by the snapshots of aggregated oligomers in Figure 4, where the helix content is at least partially lost upon aggregation. Consequently, we can postulate that lactose inhibits aggregation by promoting the folded conformation of glucagon. This excipient reduces the population of partially unfolded conformations, which are the species driving the aggregation process. As a result, this leads to a diminished tendency for aggregation and the stabilization of the monomeric form.

The mechanism is different for HPβCD. As previously established, and in accordance with prior findings, this cyclodextrin exhibits surfactant-like properties. Consequently, it impedes aggregation through its favorable interaction with the protein surface, that obstructs protein-protein interactions. This interference with protein-protein contacts diminishes the aggregation-induced loss of α-helical content, ultimately preserving the native secondary structure, as can be seen in Figure 2.

3.2 Experimental validation

Experimental validation was carried out through MMS to corroborate the insights gained from the previous molecular simulations. As determined by MMS, the native structure of glucagon when freshly dissolved in water, lactose, and HPβCD was mostly turn structures and α-helix, which was expected from the MD simulations. After the 24-hour incubation at RT, the glucagon samples in

water and lactose showed an increase in β-sheet content, indicating instability and a propensity for
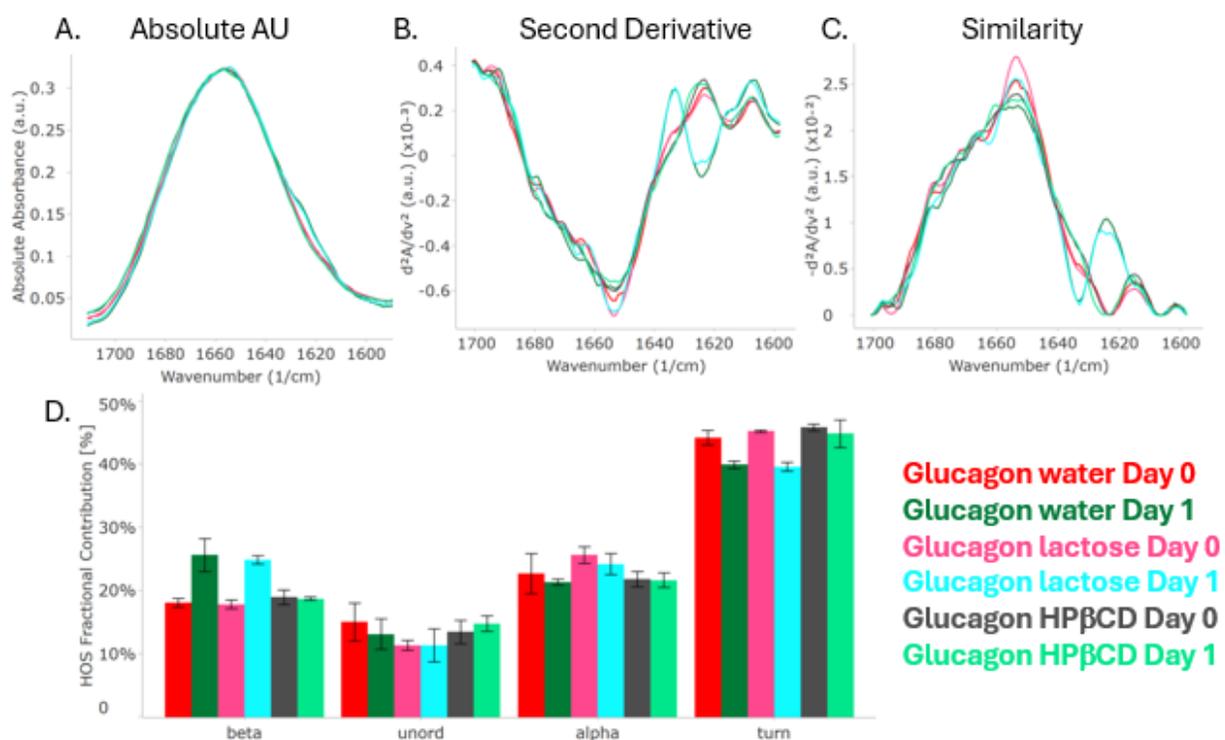
aggregation (Figure 5).

Figure 5. All samples incubated at room temperature: A) The absolute absorbance spectra for each

sample shows a similar structure for the samples, where the main peak is located near 1656 cm^{-1}.

B) Second derivative analysis reveals smaller disparities between the samples, specifically around

1620 cm^{-1} where the samples in water and lactose after 24-hour incubation show significant

structural change. C) the similarity plot is generated from the inverted and baselined-subtracted

second derivative spectra to facilitate Gaussian curve-fitting. D) The fractional contribution of

secondary structures for glucagon under these 6 conditions shows that the samples in water and

lactose, after 24-hour incubation, had an increase in β-sheet content typically associated with aggregation, and that the sample in HPβCD was resistant to the aggregation.

Glucagon samples incubated at 4°C in water and lactose (Figure 6) also showed an increase in β-sheet content, but the extent was reduced, demonstrating that the colder temperature slowed down the aggregation process. In addition, the increase in β-sheets content in lactose was less pronounced than in water during the 24-h incubation at 4°C, suggesting that lactose tends to slow down the aggregation process. Interestingly, in the presence of HPβCD, glucagon did not undergo structural changes during the 24-hour incubation at either RT or 4°C, showing that HPβCD is very effective at inhibiting aggregation. The ranking of glucagon stability within the tested formulations as determined via MMS (HPβCD > lactose > water) matches, therefore, the molecular dynamics simulations outputs.
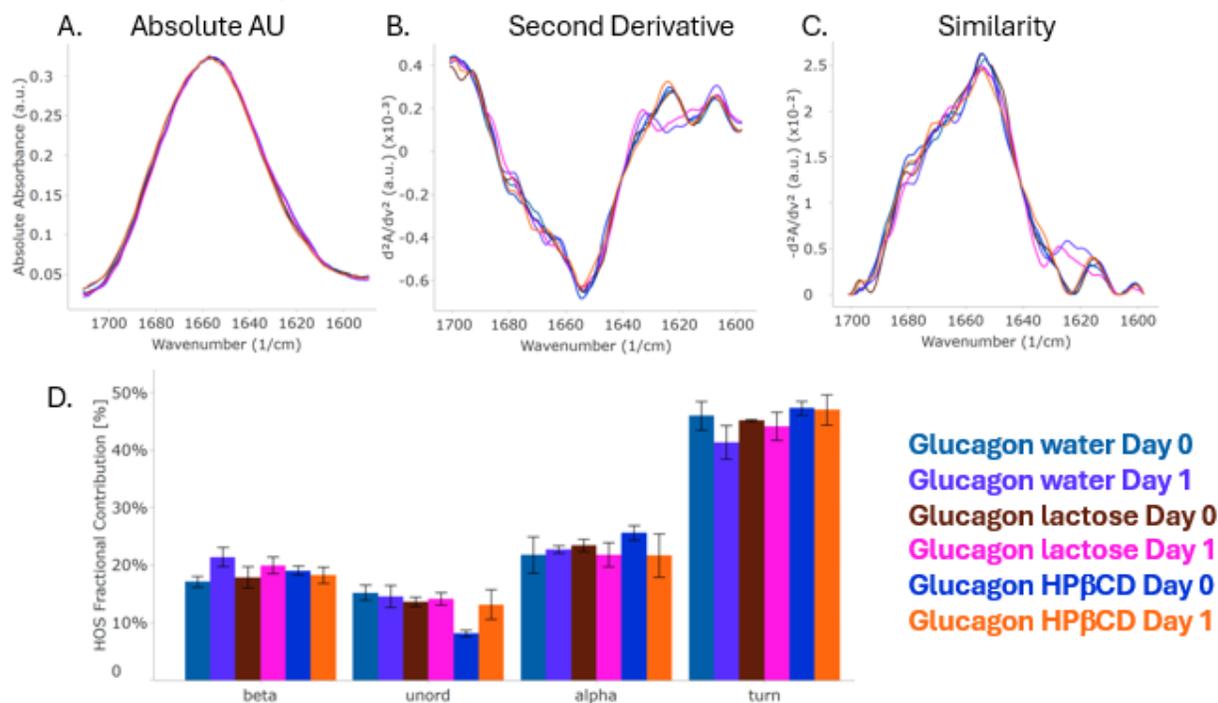
Figure 6. All samples incubated at 4°C: A) Absolute absorbance spectra have been baselined and normalized for concentration. B) Second derivative plots highlight regions of small change, specifically the region near 1620 cm^{-1} indicating structural changes in the samples dissolved in water and lactose after the 24-hour incubation. C) The similarity plot is the inverted and baselined second derivative that is used to Gaussian curve fit to get the higher order structure (HOS) percentages. D) HOS bar chart shows that glucagon after 24 hours in water increases the β-sheet content, consistent with aggregation. In the presence of 10.7% lactose, that increase in β-sheet content is less, and in the presence of HPβCD, the β-sheet content does not increase after 24 hours.

As part of the MMS software, the "displacement factor" is calculated, which refers to the amount of buffer that is displaced by the protein, governed by the volume of the protein and that is separate

26

from the structural measurements. Table 2 shows the calculated values for each sample. Glucagon dissolved in water displaced the least buffer, followed by lactose, and lastly, glucagon dissolved in HPβCD displaced the most buffer. This observation supports the hypothesis that HPβCD inhibits aggregation by binding to the external surface of glucagon, which restricts protein-protein interactions responsible for aggregation.

Table 2. The calculated displacement factors for each sample run on MMS. Glucagon in water displaces the least water and glucagon in HPβCD displaces the most water.

Sample	Displacement Factor
Water RT day 0	0.29
Water RT day 1	0.16
Water 4°C day 0	0.29
Water 4°C day 1	0.25
Lactose RT day 0	0.88
Lactose RT day 1	0.49
Lactose 4°C day 0	1.01
Lactose 4°C day 1	0.84
HPβCD RT day 0	1.24
HPβCD RT day 1	1.10
HPβCD 4°C day 0	1.18
HPβCD 4°C day 1	1.34

4. CONCLUSIONS

Through a synergy of *in-silico* simulations and experimental validation, we unraveled the intricate mechanisms of glucagon aggregation and identified strategies to mitigate its adverse effects.

The *in-silico* simulations revealed that lactose and 2-HPβCD, two excipients with opposing behaviors, play crucial roles in modulating glucagon's aggregation propensity. Lactose, as a stabilizer, fosters the monomeric state of glucagon by increasing its α-helical content, while 2-HPβCD, functioning as a surfactant, impedes aggregation by inhibiting protein-protein interactions. These findings not only enhance our understanding of the underlying mechanisms but also offer actionable strategies for therapeutic peptide development.

To validate the *in-silico* results, we employed Microfluidic Modulation Spectroscopy (MMS), demonstrating that glucagon in the presence of 2-HPβCD remains structurally stable, affirming the anti-aggregation potential of this excipient. Additionally, the displacement factor analysis further supported our hypothesis by showing that 2-HPβCD inhibits aggregation through its interaction with the external surface of glucagon, hindering protein-protein interactions. Lactose is less effective than 2-HPβCD at stabilizing glucagon, and a moderate effect of this excipient is observed in the MMS results only during incubation at 4°C.

These results suggest that linear saccharides may stabilize glucagon similarly to lactose, though the extent of stabilization is likely dependent on the specific molecule considered. Cyclodextrin, by contrast, appears to prevent glucagon aggregation through its surfactant-like properties. Consequently, we anticipate that surfactants may exhibit a similar mechanism of action, with the degree of stabilization varying based on the specific surfactant used.

   This research provides essential insights into the multifaceted world of glucagon aggregation and presents a blueprint for addressing this formidable challenge in therapeutic peptide development. By better understanding the dynamics of peptide aggregation, we move one step closer to the development of safe and effective therapeutic peptides, offering hope in the quest to combat both severe hypoglycemia and neurodegenerative diseases. The limitations of the implicit solvation method used in this study are acknowledged, particularly the restricted availability of transfer free energy data, which limits the number of formulations that can be simulated. Additionally, this method is not easily extendable to coarse-grained simulations. Future research should prioritize creating a more extensive dataset of transfer free energies for relevant excipients and aim to extend the proposed approach to coarse-grained simulations. Addressing these limitations will enhance the robustness and applicability of the findings, providing a more comprehensive understanding of excipient behavior in preventing glucagon aggregation.

SUPPORTING INFORMATION

Explicit solvent simulations for lactose, values of transfer free energies employed for the implicit solvent runs.

ACKNOWLEDGMENT

REFERENCES

1. Sipe, J.D. and Cohen, A.S. Review: history of the amyloid fibril. *J. Struct. Biol.* **2000**, *130*, 2-3, 88–98.

2. Chiti, F. and Dobson, C.M. Protein misfolding, functional amyloid, and human disease. *Ann. Rev. Biochem.* **2006**, *75*, 333–366.

3. Selkoe, D.J. and Hardy, J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol. Med.* **2016**, *8*, 6, 595–608.

4. Long, J.M. and Holtzman, D.M. Alzheimer disease: an update on pathobiology and treatment strategies. *Cell* **2019**, *179*, 2, 312–339.

5. Westermark, P., Andersson, A., Westermark, G.T. Islet amyloid polypeptide, islet amyloid, and diabetes mellitus. *Physiol. Rev.* **1986**, *86*, 3, 695–707.

6. Boyer, S.K. and Laine, E.S. Glucagon as a potential biopharmaceutical for treating hypoglycemia. *Diabetes* **1984**, *33*, 8, 757–765.

7. Marassi, V.; Macis, M.; Giordani, S.; Ferrazzano, L.; Tolomelli, A.; Roda, B.; Zattoni, A.; Ricci, A.; Reschiglian, P.; Cabri, W. Application of Af4-multidetection to Liraglutide in its formulation: preserving and representing native aggregation. *Molecules* **2022**, *27*, 5485.

8. Bromer, W.W.; Sinn, L.G.; Chakrabarti, S. Therapeutic use of glucagon. *Drug Develop. Res.* **1983**, *3*, 5, 473–485.

9. Unson, C. G. Molecular determinants of glucagon receptor signaling. *Peptide Sci.* **2002**, *66*, 218–235.

10. Hall-Boyer, K.; Zaloga, G. P.; Chernow, B. Glucagon: hormone or therapeutic agent? *Crit. Care Med.* **1984**, *12*, 584–589.

11. Bromer, W.W. Chemical characteristics of glucagon. In *Glucagon I*, 1st ed.; Lefèbvre, P.J., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1983; pp 1–22.

12. FDA. Glucagon for injection: highlights of prescribing information. U.S. Food and Drug Administration, 2022.

    https://www.accessdata.fda.gov/drugsatfda_docs/label/2022/201849s013lbl.pdf

13. Onoue, S.; Ohshima, K.; Debari, K.; Koh, K.; Shioda, S.; Iwasa, S.; Kashimoto, K.; Yajima, T. Mishandling of the therapeutic peptide glucagon generates cytotoxic amyloidogenic fibrils. *Pharm. Res.* **2004**, *21*, 1274–1283.

14. Pedersen, J. S. The nature of amyloid-like glucagon fibrils. *J. Diabetes Sci. Technol.* **2010**, *4*, 1357–1367.

15. De Jong, K. L.; Incledon, B.; Yip, C. M.; DeFelippis, M. R. Amyloid fibrils of glucagon characterized by high-resolution atomic force microscopy. *Biophys. J.* **2006**, *91*, 1905–1914.

16. Zhou, X.; Tan, J.; Zheng, L.; Pillai, S.; Li, B.; Xu, P.; Zhang, B.; Zhang, Y. The opposite effects of Cu(II) and Fe(III) on the assembly of glucagon amyloid fibrils. *RSC Advances* **2012**, *2*, 5418–5423.

17. Hoppe, C.C.; Nguyen, L.T.; Kirsch, L.E.; Wiencek, J.M. Characterization of seed nuclei in glucagon aggregation using light scattering methods and field-flow fractionation, *J. Biol. Eng.* **2008**, *2*: 1–11.

18. Bakhtiani, P. A.; Caputo, N.; Castle, J. R.; Youssef, J. E.; Carroll, J. M.; David, L. L.; Charles T. Roberts, J.; Ward, W. K. A novel, stable, aqueous glucagon formulation using ferulic acid as an excipient. *J. Diabetes Sci. Technol.* **2015**, *9*, 17–23.

19. Košmrlj, A.; Cordsen, P.; Kyrsting, A.; Otzen, D.E.; Oddershede, L.B., Jensen, M.H. A monomer-trimer model supports intermittent glucagon fibril growth. *Sci. Rep.* **2015**, *5*, Article no. 9005.

20. Onoue, S.; Iwasa, S.; Kojima, T.; Katoh, F.; Debari, K.; Koh, K.; Matsuda, Y.; Yajima, T. Structural transition of glucagon in the concentrated solution observed by electrophoretic and spectroscopic techniques. *J. Chromatogr. A* **2006**, *1109*: 167–173.

21. Pohl, R.; Li, M.; Krasner, A.; Souza, E. D. Development of stable liquid glucagon formulations for use in artificial pancreas. *J. Diabetes Sci. Technol.* **2015**, *9*, 8–16.

22. Chabenne, J. R.; DiMarchi, M. A.; Gelfanov, V. M.; DiMarchi, R. D. Optimization of the native glucagon sequence for medicinal purposes. *J. Diabetes Sci. Technol.* **2010**, *4*, 1322–1331.

23. Oliveira, C. L. P.; Behrens, M. A.; Pedersen, J. S.; Erlacher, K.; Otzen, D.; Peder- sen, J. S. A SAXS study of glucagon fibrillation. *J. Mol. Biol.* **2009**, *387*, 147–161.

24. Stimple, S. D.; Kalyoncu, S.; Desai, A. A.; Mogensen, J. E.; Spang, L. T.; Asgreen, D. J.; Staby, A.; Tessier, P. M. Sensitive detection of glucagon aggregation using amyloid fibril-specific antibodies. *Biotechnol. Bioeng.* **2019**, *116*, 1868–1877.

25. Serno, T.; Carpenter, J. F.; Randolph, T. W.; Winter, G. Inhibition of agitation- induced aggregation of an IgG-antibody by hydroxypropyl-β-cyclodextrin. *J. Pharm. Sci.* **2010**, *99*, 1193–1206.

26. Tavornipas, S.; Tajiry, S.; Hirayama, F.; Arima, H.; Uekama, K. Effects of Hydrophilic cyclodextrins on aggregation of recombinant human growth hormone. *Pharm. Res.* **2004**, *21*, 2369–2376.

27. Serno, T.; Härtl, E.; Besheer, A.; Miller, R.; Winter, G. The role of polysorbate 80 and HPβCD at the air-water interface of IgG solutions. *Pharm. Res.* **2013**, *30*, 117–130.

28. Timasheff, S. Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9721–9726.

29. Timasheff, S. The control of protein stability and association by weak interactions with water: How do solvents affect these processes? *Ann. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 67–97.

30. Arsiccio, A.; Rospiccio, M.; Shea, J.-E.; Pisano, R. Force field parameterization for the description of the interactions between hydroxypropyl-β-cyclodextrin and proteins. *J. Phys. Chem. B* **2021**, *125*, 7397–7405.

31. Oliveri, V. and Vecchio, G. Cyclodextrins as protective agents of protein aggregation: an overview. *Chem. Asian J.* **2016**, *11*(5), 592–602.

32. Niccoli, M.; Oliva, R.; Castronuovo, G. Cyclodextrin–protein interaction as inhibiting factor against aggregation. *J Therm Anal Calorim* **2017**, *127*, 1491–1499.

33. Arsiccio, A.; Ganguly, P.; Shea, J.-E. A transfer free energy based implicit solvent model for proteins simulations in solvent mixtures: urea-induced denaturation as a case study. *J. Phys. Chem. B* **2022**, *126*, 24, 4472–4482.

34. Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate born radii. *J. Phys. Chem. A* **1997**, *101*, 3005–3014.

35. Onufriev, A.; Bashford, D.; Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Structure, Function, and Bioinformatics* **2004**, *55*, 383–394.

36. Bradley, D. J.; Pitzer, K. S. Thermodynamics of electrolytes. 12. Dielectric properties of water and Debye-Hueckel parameters to 350 °C and 1 kbar. *J. Phys. Chem.* **1979**, *83*, 1599–1603.

37. Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.

38. Hasel, W.; Hendrickson, T. F.; Still, W. C. A rapid approximation to the solvent accessible surface areas of atoms. *Tetrahedron Comput. Methodol.* **1988**, *1*, 103–116.

39. Lindor-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber 99SB protein force field. *Proteins: Structure, Function, and Bioinformatics* **2010**, *78*, 1950–1958.

40. Sasaki, K.; Dockerill, S.; Adamiak, D. A.; Tickle, I. J.; Blundell, T. X-ray analysis of glucagon and its relationship to receptor binding. *Nature* **1975**, *257*, 751–757.

41. Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; De- Bolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a package of computer pro- grams for applying molecular mechanics, normal mode analysis, molecular dynam- ics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* **1995**, *91*, 1–41.

42. Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604–613.

43. Ferrarotti, M. J.; Bottaro, S.; Pérez-Villa, A.; Bussi, G. Accurate multiple time step in biased molecular simulations. *J. Chem. Theory Comput.* **2015**, *11*, 139–146.

44. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

45. 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.

46. Arsiccio, A.; Ganguly, P.; La Cortiglia, L.; Shea, J.-E.; Pisano, R. ADD force field for

   sugars and polyols: predicting the additivity of protein-osmolyte interaction. *J. Phys. Chem.*

   B **2020**, *124*, 7779–7790.

47. Pietrucci, F.; Laio, A. A collective variable for the efficient exploration of protein beta-sheet

   structures: application to SH3 and GB1. *J. Chem. Theory .Comput.* **2009**, *5*, 2197–2201.

48. Auton, M.; Holthauzen, L.M.F.; Bolen, D.W. Anatomy of energetic changes accompanying

   urea-induced protein denaturation. *PNAS* **2007**, *104*. 15317–15322.

49. Rospiccio, M.; Arsiccio, A.; Winter, G.; Pisano, R. The role of cyclodextrins against

   interface-induced denaturation in pharmaceutical formulations: a molecular dynamics

   approach. *Molecular Pharmaceutics* **2021**, *18*, 2322–2333.

TOC graph: