



## Electrostatics Drive Oligomerization and Aggregation of Human Interferon Alpha-2a

Pohl, Christin; Polimeni, Marco; Indrakumar, Sowmya; Streicher, Werner; Peters, Günther H. J.; Nørgaard, Allan; Lund, Mikael; Harris, Pernille

*Published in:*  
The Journal of Physical Chemistry B

*Link to article, DOI:*  
[10.1021/acs.jpccb.1c07090](https://doi.org/10.1021/acs.jpccb.1c07090)

*Publication date:*  
2021

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Pohl, C., Polimeni, M., Indrakumar, S., Streicher, W., Peters, G. H. J., Nørgaard, A., Lund, M., & Harris, P. (2021). Electrostatics Drive Oligomerization and Aggregation of Human Interferon Alpha-2a. *The Journal of Physical Chemistry B*, 125(50), 13657-13669. <https://doi.org/10.1021/acs.jpccb.1c07090>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Electrostatics Drive Oligomerization and Aggregation of Human Interferon Alpha-2a

Christin Pohl,<sup>\*,#</sup> Marco Polimeni,<sup>#</sup> Sowmya Indrakumar, Werner Streicher, Günther H.J. Peters, Allan Nørgaard, Mikael Lund, and Pernille Harris<sup>\*</sup>



Cite This: *J. Phys. Chem. B* 2021, 125, 13657–13669



Read Online

ACCESS |



Metrics & More

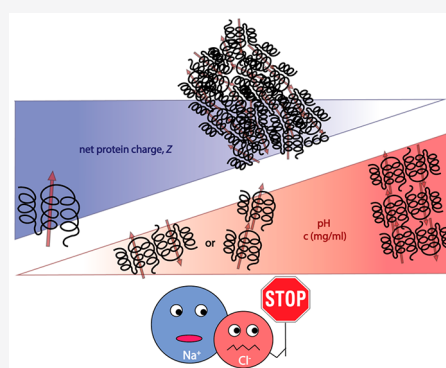


Article Recommendations



Supporting Information

**ABSTRACT:** Aggregation is a common phenomenon in the field of protein therapeutics and can lead to function loss or immunogenic patient responses. Two strategies are currently used to reduce aggregation: (1) finding a suitable formulation, which is labor-intensive and requires large protein quantities, or (2) engineering the protein, which requires extensive knowledge about the protein aggregation pathway. We present a biophysical characterization of the oligomerization and aggregation processes by Interferon alpha-2a (IFN $\alpha$ -2a), a protein drug with antiviral and immunomodulatory properties. This study combines experimental high throughput screening with detailed investigations by small-angle X-ray scattering and analytical ultracentrifugation. Metropolis Monte Carlo simulations are used to gain insight into the underlying intermolecular interactions. IFN $\alpha$ -2a forms soluble oligomers that are controlled by a fast pH and concentration-dependent equilibrium. Close to the isoelectric point of 6, IFN $\alpha$ -2a forms insoluble aggregates which can be prevented by adding salt. We show that monomer *attraction* is driven mainly by molecular anisotropic dipole–dipole interactions that increase with increasing pH. *Repulsion* is due to monopole–monopole interactions and depends on the charge of IFN $\alpha$ -2a. The study highlights how combining multiple methods helps to systematically dissect the molecular mechanisms driving oligomer formation and to design ultimately efficient strategies for preventing detrimental protein aggregation.



## INTRODUCTION

The development of therapeutic protein drugs has increased dramatically in the past decade.<sup>1,2</sup> Owing to their high specificity, they often show fewer side effects compared with small molecule drugs and open many possibilities for the treatment of diseases.<sup>3</sup> However, they are challenging and costly in their development due to their short half-life and low stability compared to small molecules.<sup>4–6</sup> In particular, the formation of oligomers and aggregates remains a challenge as they differ in their characteristics from protein to protein.<sup>7–9</sup> There have been many approaches targeting this problem, but one of the challenges is the diverse behavior of different kinds of protein drugs. In this study, Interferon-alpha-2a (IFN $\alpha$ -2a), which is prone to aggregation<sup>10</sup> is used as a model system to characterize aggregation and oligomer formation. The detailed characterization of oligomers and aggregates formed is essential to understand the mechanisms behind protein aggregation. In the long term, this knowledge can be used to alter proteins specifically to avoid oligomerization and aggregation.<sup>11</sup> IFN $\alpha$ -2a consists of 165 amino acids (19.24 kDa) with 68%  $\alpha$ -helical content. The five  $\alpha$ -helices are connected through loops (Supporting Information, Figure S1A). IFN $\alpha$ -2a belongs to the type I interferon cytokines, which are part of the innate immune, system and is used in the

treatment of, for example, hepatitis, carcinoma, leukemia, and lymphoma.<sup>12–14</sup> Type I interferons show a high structural identity (Figure S1, Table S1).

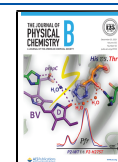
IFN $\alpha$  was the first cytokine approved for therapeutic use by the FDA in 1986. Two IFN $\alpha$ -based drugs are commercially available: Roferon-A (IFN $\alpha$ -2a) and Intron-A1 (IFN $\alpha$ -2b), which only differ in one amino acid in position 23 and have similar biological activity. While Roferon-A is formulated in acetate buffer at pH 5, Intron-A is formulated in phosphate buffer at pH 7.<sup>15</sup>

Several studies on the stability of IFN $\alpha$ -2a are already published. IFN $\alpha$ -2a oligomerizes at pH > 5.<sup>16</sup> Isothermal titration calorimetry (ITC)<sup>17</sup> and stability studies using the intrinsic fluorescence of the two tryptophans in IFN $\alpha$ -2a<sup>18</sup> have shown that temperature plays a major role in IFN $\alpha$ -2a's stability. At pH 3, IFN $\alpha$ -2a has shown partial unfolding,<sup>18</sup> while at pH 4 it seemed to be most stable.<sup>18,19</sup> Aggregation and

Received: August 10, 2021

Revised: November 28, 2021

Published: December 13, 2021



oligomerization have been observed around pH 6, which correlates with the protein's isoelectric point (pI) of 5.98. It was shown that the addition of salt decreased IFN $\alpha$ -2a's thermal stability and induced the formation of large aggregates in certain conditions.<sup>17,19</sup> The addition of protein stabilizers or surfactants is a common approach to increase stability and decrease aggregation.<sup>20,21</sup> For Interferon, the formulation with albumin as excipient used to be common, but a preference for albumin-free formulations due to increased aggregation<sup>22</sup> led to the addition of polysorbate 80 for better storage and handling.<sup>23</sup> An additional approach was the development of polyethylene glycol (PEG) conjugated IFNs, which have displayed an increased half-life.<sup>24</sup> As shown from these studies, the formulation of a protein drug plays a major role in its stability. Even though many studies were performed on IFN $\alpha$ -2a's stability, the correlation between protein structure, electrostatics, and aggregation propensity has, to our knowledge, not been studied so far.

IFN $\alpha$ -2a was selected as a representative, aggregation-prone protein to study protein self-association as a function of formulation condition. We used a combination of experimental techniques and a computational model to characterize and identify the cause of self-association. We performed high throughput screening using dynamic light scattering (DLS), isothermal chemical denaturation (ICD), and nano differential scanning fluorimetry (nanoDSF) in a pH range where IFN $\alpha$ -2a is known to self-associate (pH 5–9). We performed a detailed analysis and compared the oligomerization of IFN $\alpha$ -2a using analytical ultracentrifugation (AUC) and small-angle X-ray scattering (SAXS) on two selected pH values, below and above IFN $\alpha$ -2a's pI. Additionally, we analyzed the effect of salt on oligomerization and aggregation. We performed coarse-grained Monte Carlo simulations to gain a molecular understanding of solution behavior observed experimentally. While there have been many studies to find additives or conditions to prevent aggregation, studies characterizing the aggregates formed are limited so far. This study aims to characterize the oligomers and aggregates formed by IFN $\alpha$ -2a and analyze the protein–protein interactions. This enables new approaches to avoid protein aggregation, such as through directed mutation of specific sites involved in protein–protein interaction or the development of PEG conjugates, which disrupt these interactions. Due to the structural similarities within this protein family (Figure S1), the results from our study may be applied to other interferons and cytokines.

## MATERIALS AND METHODS

**Sample Preparation.** IFN $\alpha$ -2a ( $c = 1.35$  mg/mL) was kindly provided by Roche Diagnostics GmbH, formulated in 25 mM ammonium acetate pH 5 with 120 mM sodium chloride. If not stated otherwise, IFN $\alpha$ -2a samples were dialyzed into the desired condition using Slide-A-Lyzer 3500 MWCO dialysis cassettes (Thermo Fisher) as described in Pohl et al.<sup>25</sup> The protein concentration after dialysis was measured using a NanoDrop 8000 spectrophotometer (Thermo Fisher). If not stated otherwise, all measurements were performed at a protein concentration of 1 mg/mL obtained by dilution into the final solution condition. Filtration of the sample and buffer was performed using Luer-Lok syringes (BD) and 0.22  $\mu$ m Millex-GV filter (Merck) or 0.2  $\mu$ m Whatman Anotop 10 filter (GE Healthcare Life Science) syringe filters.

**DLS Measurements.** IFN $\alpha$ -2a samples were concentrated by centrifugation to a concentration of approximately 20 mg/mL and dialyzed in 10 mM His pH 5.5, 10 mM His pH 7, and 10 mM Tris pH 8.5 as described above. The material was filtered (0.22  $\mu$ m), and the concentration was measured using a NanoDrop 8000 spectrophotometer (Thermo Fisher). A protein stock solution of 20 mg/mL was obtained by dilution with the filtered dialysis buffer (0.2  $\mu$ m). The respective formulations were obtained by a 20 times dilution. The measurement was performed with a DynaPro Plate Reader II (Wyatt Technology) using Aurora 384 LV/EB plates (Brookes Life Science Systems). Silicone oil (Sigma-Aldrich) was used for sealing the wells. All measurements were performed isothermally at 25 °C with 5 s acquisition time and 20 acquisitions per well. All formulations were measured in technical triplicates. Analysis was performed using DYNAMICS version 7.8.1.3.

**NanoDSF Measurements.** Sample preparation was performed likewise for DLS measurements. The measurements were performed with Prometheus NT.48 (NanoTemper Technologies). A constant temperature ramp was applied from 20 to 95 °C with 1 °C/min. The analysis was performed with PR.ThermControl (NanoTemper Technologies).

**ICD Measurements.** Sample preparation was performed as described above. A protein stock solution of 1 mg/mL was prepared. Formulation buffer for the desired conditions was prepared. Denaturation buffer included 6 M GnHCl in the formulation buffer. Measurements were performed using the HUNK system (Unchained Laboratories) with no additional incubation time and a gain of 100 for fluorescence detection. The analysis was performed with Formulator (Unchained Laboratories).

**SAXS Measurements.** IFN $\alpha$ -2a samples were concentrated to a concentration of approximately 30 mg/mL and dialyzed as described above. All samples and dialysis buffers were filtered (0.22  $\mu$ m). The absorbance of all samples was measured using a NanoDrop 8000 spectrophotometer (ThermoFisher). A concentration series was obtained by dilution with the dialysis buffer. Measurements were performed at the German Electron Synchrotron DESY at the P12 EMBL BioSAXS beamline and the European Synchrotron Radiation Facility ESRF at the BM29 BioSAXS beamline. Data analysis was performed using the ATSAS software package version 2.8.4.<sup>26</sup> A summary of the SAXS data collection is shown in Table S3.

**AUC Measurements.** All samples were prepared as described above. The absorbance of all samples was measured using a NanoDrop 8000 spectrophotometer (ThermoFisher). All samples were diluted with dialysis buffers to an absorbance of  $A_{280} = 1$ ,  $A_{280} = 0.5$ , and  $A_{280} = 0.25$ , measured with a NanoDrop 8000 spectrophotometer (ThermoFisher). The measurements were performed as a sedimentation velocity run with an Optima XL-1 ultracentrifuge (Beckman-Coulter) at 50000 rpm. Data analysis was performed using SEDFIT version 16.1c.<sup>27,28</sup> The RMSD was used to determine the goodness of fit (see Table 1).

Final graphs of experimental results were prepared using Origin 2019 (OriginLabs) and MATLAB (MathWorks).

**Monte Carlo Simulations.** Starting from the all-atom PDB structure (PDB ID: 1ITF) a coarse-grained model of the protein, where each amino acid was replaced by a spherical bead centered in its center of mass, was constructed. Metropolis Monte Carlo simulations<sup>29</sup> (MC) were performed

**Table 1.** Root Mean Square Deviation (RMSD) to Determine the Goodness of the Fit of AUC Data

|                     | RMSD          |                 |                  |
|---------------------|---------------|-----------------|------------------|
|                     | $A_{280} = 1$ | $A_{280} = 0.5$ | $A_{280} = 0.25$ |
| pH 5                | 0.0086        | 0.0060          | 0.0061           |
| pH 5, 140 mM NaCl   | 0.0089        | 0.0060          | 0.0055           |
| pH 7.5              | 0.0096        | 0.0073          | 0.0064           |
| pH 7.5, 140 mM NaCl | 0.0109        | 0.0074          | 0.0063           |
| pH 7.5 concentrated | 0.0115        | 0.0077          | 0.0065           |

using Faunus<sup>30</sup> (v2.4.2 git revision bbd3545c) which allows for different MC moves such as protein translation and rotation, amino acid charge titration, and a cluster move (see below for more description). The potential energy function of the system was defined as

$$\beta U = \underbrace{\sum_i^{n-i} \sum_{j=i+1}^n \frac{e^2 z_i z_j}{4\pi\epsilon_0\epsilon_r r_{ij}} e^{-r_{ij}/\lambda}}_{\text{electrostatic energy}} + \underbrace{4\epsilon_{ij} \sum_i^{n-i} \sum_{j=i+1}^n \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]}_{\text{vdW and excluded volume energy}} - \underbrace{\sum_i \ln \left( \frac{N_i!}{(N_i + \nu_i)!} V^{\nu_i} \right) - \ln \prod_i a_i^{\nu_i}}_{\text{titration energy}}$$

Sums run over each  $i$ - $j$  amino acid pair for electrostatic, van der Waals (vdW), and excluded volume energy terms, while the titration energy was taken into account for charged amino acids.  $e$  is the elementary unit charge,  $z_i$  and  $z_j$  are the charges corresponding to the  $i^{\text{th}}$  and  $j^{\text{th}}$  amino acid.  $r_{ij}$  is the distance between  $i$  and  $j$ ,  $\epsilon_0$  is the vacuum permittivity, and  $\epsilon_r = 80$  is the water relative dielectric constant (i.e., the solvent is considered as a continuum homogeneous medium). Salt was included using Debye–Hückel theory through an appropriate Debye length,  $\lambda$ , according to the relation:  $\lambda/\text{\AA} = 3.04/I(M)$ , where  $I$  is the ionic strength in molar units. In the VdW term,  $\sigma_{ij}$  and  $\epsilon_{ij}$  represent the minimum distance between the  $i^{\text{th}}$  and  $j^{\text{th}}$  amino acid and the energy depth which characterize their interactions. They were obtained from the Lorentz–Berthelot mixing rules,<sup>31</sup>  $\sigma_{ij} = (\sigma_i + \sigma_j)/2$  and  $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$ , where  $ii$  and  $jj$  subscripts indicate the self-interaction parameters. The titration energy term accounts for the energy change due to the charge titration of the lateral chain of charged amino acids. For each charged amino acid species, the reaction  $HA \rightleftharpoons H + A$ , where  $A$  is a charged amino acid in its deprotonated form and  $H$  is a proton, is propagated back and forward using a reactive Monte Carlo scheme.<sup>32</sup> To each reaction, a  $pK_a$  value is assigned based on the amino acid species (Table 2).<sup>33</sup> In the equation,  $N_i$  is the number of amino acids of type  $i$ ,  $\nu_i$  is the stoichiometric coefficient (positive for the products and negative for the reagents),  $V$  is the volume of the system, and  $a_i$  is the activity of the amino acid species.

The solution properties were sampled using two-body and many-body models (Figure 1). A two-body model consisted of two identical proteins in a spherical cell (Figure 1 left). One protein was placed in the middle of the cell and could only rotate around its center of mass while a second one could also

**Table 2.** Reaction List for Amino Acid Titration and Relative  $pK_a$  Values<sup>33</sup>

| reaction                          | $pK_a$ |
|-----------------------------------|--------|
| HCTR $\rightleftharpoons$ H + CTR | 3.67   |
| HASP $\rightleftharpoons$ H + ASP | 3.67   |
| HGLU $\rightleftharpoons$ H + GLU | 4.25   |
| HHIS $\rightleftharpoons$ H + HIS | 6.54   |
| HNTR $\rightleftharpoons$ H + NTR | 8.0    |
| HCYS $\rightleftharpoons$ H + CYS | 8.55   |
| HTYR $\rightleftharpoons$ H + TYR | 9.84   |
| HLYS $\rightleftharpoons$ H + LYS | 10.4   |
| HARG $\rightleftharpoons$ H + ARG | 12.0   |

translate along the  $z$ -axis. In two-body simulations,  $\epsilon_{ij}$  was set equal to  $0.05 k_B T$  while  $\sigma_{ii} = \sqrt[3]{6M_{w,i}/4\pi\rho}$ , where  $M_{w,i}$  is the molecular weight of the  $i^{\text{th}}$  amino acid and  $\rho = 1 \text{ g/mol}/\text{\AA}^3$  is an average amino acid density. For each protein–protein center of mass separation along the  $z$ -axis, a virtual displacement perturbation,  $dL$ , of the second protein was performed and the corresponding energy perturbation,  $du$ , was measured. The force profile,  $F(z)$ , was then calculated as

$$F(z) = \frac{k_B T \ln \langle e^{-\beta du(z)} \rangle}{dL}$$

where the angle brackets,  $\langle \rangle$ , indicate an ensemble average. The integration of the force profile led to the potential of the mean force,  $U(z)$ , which allows for the calculation of the osmotic second virial coefficient,<sup>34</sup>  $B_2$  as

$$B_2 = -2\pi \int_{\sigma}^{\infty} [e^{-\beta U(z)} - 1] z^2 dz$$

where  $\sigma$  is the minimum contact distance between the proteins. In this paper, we report  $B_2$  in its reduced form as

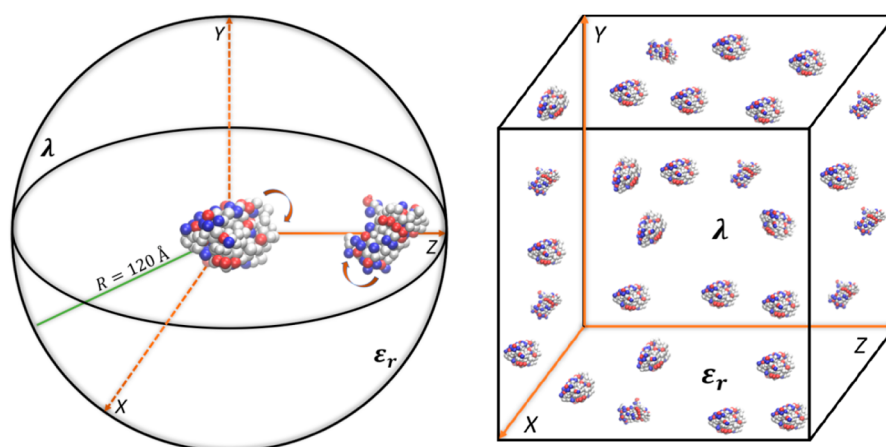
$$B_2^* = \frac{B_2}{B_2^{\text{hs}}}$$

Where  $B_2^{\text{hs}} = 2\pi(2R_h^3)/3$  is the hard sphere contribution of a sphere of radius equal to the protein hydrodynamics radius,  $R_h = 22.71 \text{ \AA}$ .  $B_2^*$  indicates a net repulsion between the proteins while a negative  $B_2^*$  indicates a net attraction.<sup>35</sup>

The proteins were treated as charge distributions, and the thermally averaged electrostatic energy,  $u(z)$ , was analyzed as a function of their mass center separation,  $z$ . Each protein was treated as either a monopole, a dipole, or a quadrupole, and the total multipolar interaction energy was calculated as a series of multipole terms:<sup>36</sup>

$$u_{\text{total}} = u_{\text{ion-ion}} + u_{\text{ion-dipole}} + u_{\text{dipole-dipole}} + u_{\text{ion-quadrupole}} + \dots \approx u_{\text{exact}} = \sum_i \sum_j \frac{q_i q_j}{r_{ij}}$$

where  $i$  and  $j$  run over the charges of the first and the second protein



**Figure 1.** Illustration of Metropolis Monte Carlo simulations. (Left) Two-body model with one protein fixed in the center of a sphere (radius = 120 Å) which is only allowed to rotate around its center of mass. The second protein was allowed to rotate and translate along the z-axis. Amino acid charges could titrate according to the solution conditions (pH and ionic strength). (Right) Many-body model including 30 proteins that were allowed to translate and rotate in a periodic cube-sized to reproduce three measured protein concentrations: 1 mg/mL, 10 mg/mL, 30 mg/mL. A fixed average charge was set for each amino acid.

$$u_{\text{ion-ion}} = \frac{Q_1 Q_2}{R}$$

$$u_{\text{ion-dipole}} = \frac{Q \mu_2 \mathbf{R}}{R^3}$$

$$u_{\text{dipole-dipole}} = \frac{\mu_1 \mu_2}{R^3} - \frac{3(\mu_1 \cdot \mathbf{R})(\mu_2 \cdot \mathbf{R})}{R^5}$$

$$u_{\text{ion-quadrupole}} = \frac{q_1 \mathbf{R}^T \mathbf{Q}_2 \mathbf{R}}{R^5} - \frac{q_1 \text{tr}(\mathbf{Q}_2)}{R^3}$$

$Q_{1/2}$  and  $\mu_{1/2}$  are the total protein charges and dipole moments of the first and the second protein and  $\mathbf{Q}_2 = \frac{1}{2} \sum_i q_i \mathbf{r}_i \mathbf{r}_i^T$ . To understand what contributed more to the interaction mechanism, the ensemble average of each multipole term,  $u_x$ , was calculated as

$$u_x = \frac{\int_0^\infty u_x e^{-\beta U(z)} z^2 dz}{\int_0^\infty e^{-\beta U(z)} z^2 dz}$$

Using the two-body model, we screened a pH range from 5 to 9 (with an increment of 0.5) at 0 and 140 mM NaCl. For each solution conditions, we performed five replicas with  $10^5$  equilibration steps, for charge and energy equilibration, followed by  $10^6$  steps for the production runs.

A many-body model consisted of 30 proteins in a cubic box (Figure 1 right). Each protein was allowed to translate in any direction and rotate around its center of mass. For each condition, parallel tempering was performed with 10 replicas which differed from each other by the  $\epsilon_{ij}$  being increased by  $0.01 k_B T$ . The amino acid charges were equilibrated and fixed. A cluster move was performed. Two proteins were considered to start a cluster if their centers of mass were placed within a distance threshold. A third protein was considered part of the cluster if the center of mass distances between the second and the third was less than the threshold and so on. The cluster growing stopped if no further protein satisfied the threshold condition. Once the cluster was formed, it was rotated and translated as an individual object in solution using a cluster

move.<sup>37</sup> The cluster shape was analyzed in terms of the relative shape anisotropy,<sup>38</sup>  $k^2$ , defined as

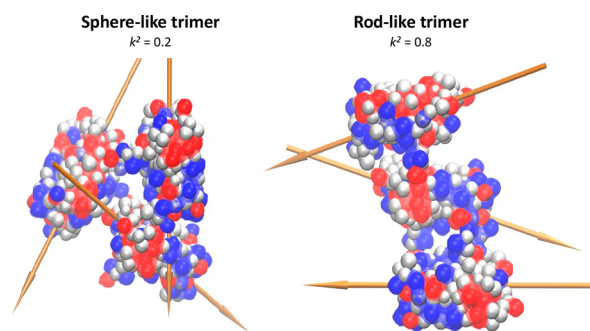
$$k^2 = 1 - 3 \frac{\lambda_1 \lambda_2 + \lambda_2 \lambda_3 + \lambda_3 \lambda_1}{(\lambda_1 + \lambda_2 + \lambda_3)^2}$$

$$= 1 - 3 \frac{\lambda_1 \lambda_2 + \lambda_2 \lambda_3 + \lambda_3 \lambda_1}{R_g^2}$$

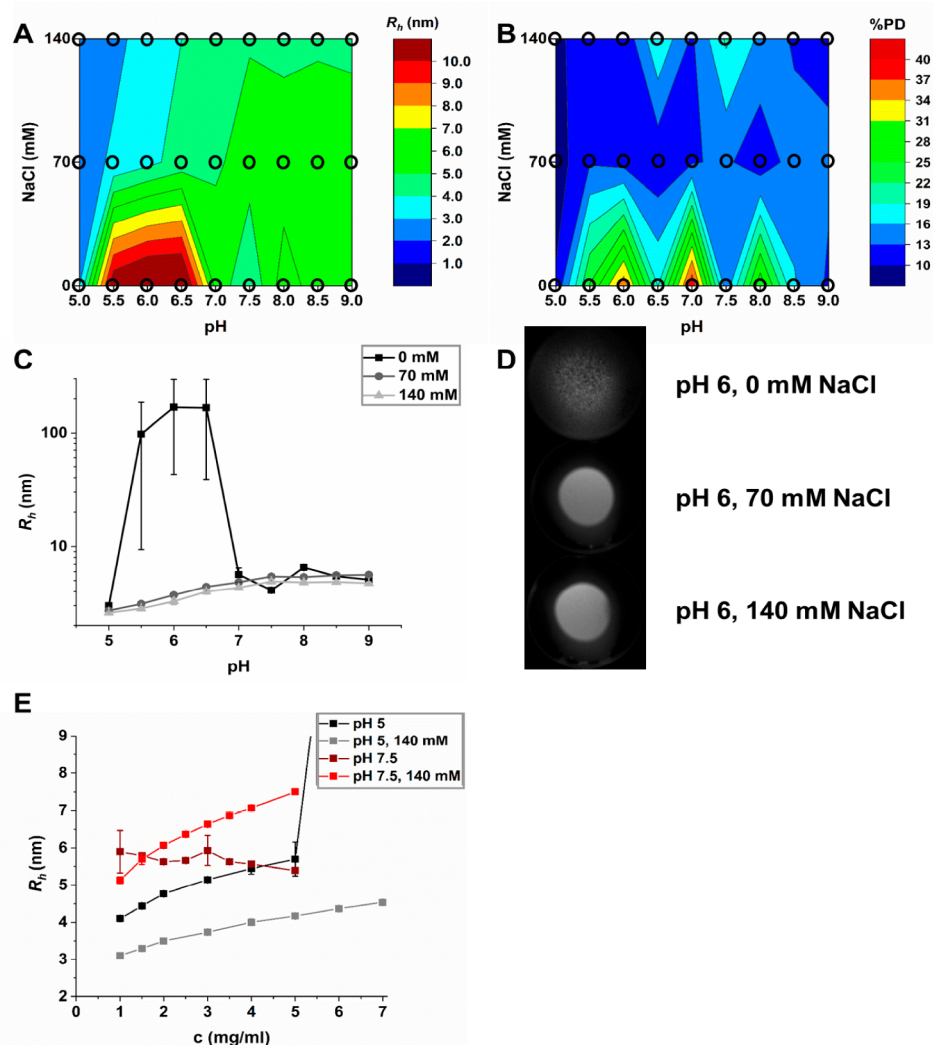
where  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  are the eigenvalues of the gyration tensor,  $S$ , sorted by  $\lambda_1 \geq \lambda_2 \geq \lambda_3$ ,  $R_g$  is the protein radius of gyration.  $k^2$  varies between 0 and 1, where 0 indicates a spherical symmetric cluster shape, while 1 indicates a linear chain.  $k^2$  is defined in respect to the center of mass distances, then a dimer will always have a  $k^2 = 1$ . Therefore, we calculated  $k^2$  for clusters bigger than dimers. To illustrate the meaning of  $k^2$  values, the cluster shape snapshots of IFN $\alpha$ -2a trimers with  $k^2 = 0.2$  and  $k^2 = 0.8$  are shown in Figure 2.

## RESULTS

**Colloidal Stability of IFN $\alpha$ -2a.** The colloidal stability of IFN $\alpha$ -2a was investigated using DLS (Figure 3) at different solution conditions (pH, added salt, and protein concentration). With no salt, the measurements showed a dramatic



**Figure 2.** Snapshots from many-body Monte Carlo simulations showing representative structures of trimers. Left: sphere-like trimer,  $k^2 = 0.2$  (left). Right: rod-like trimer,  $k^2 = 0.8$ . The golden-colored arrows represent molecular dipole moments for each of the participating proteins.



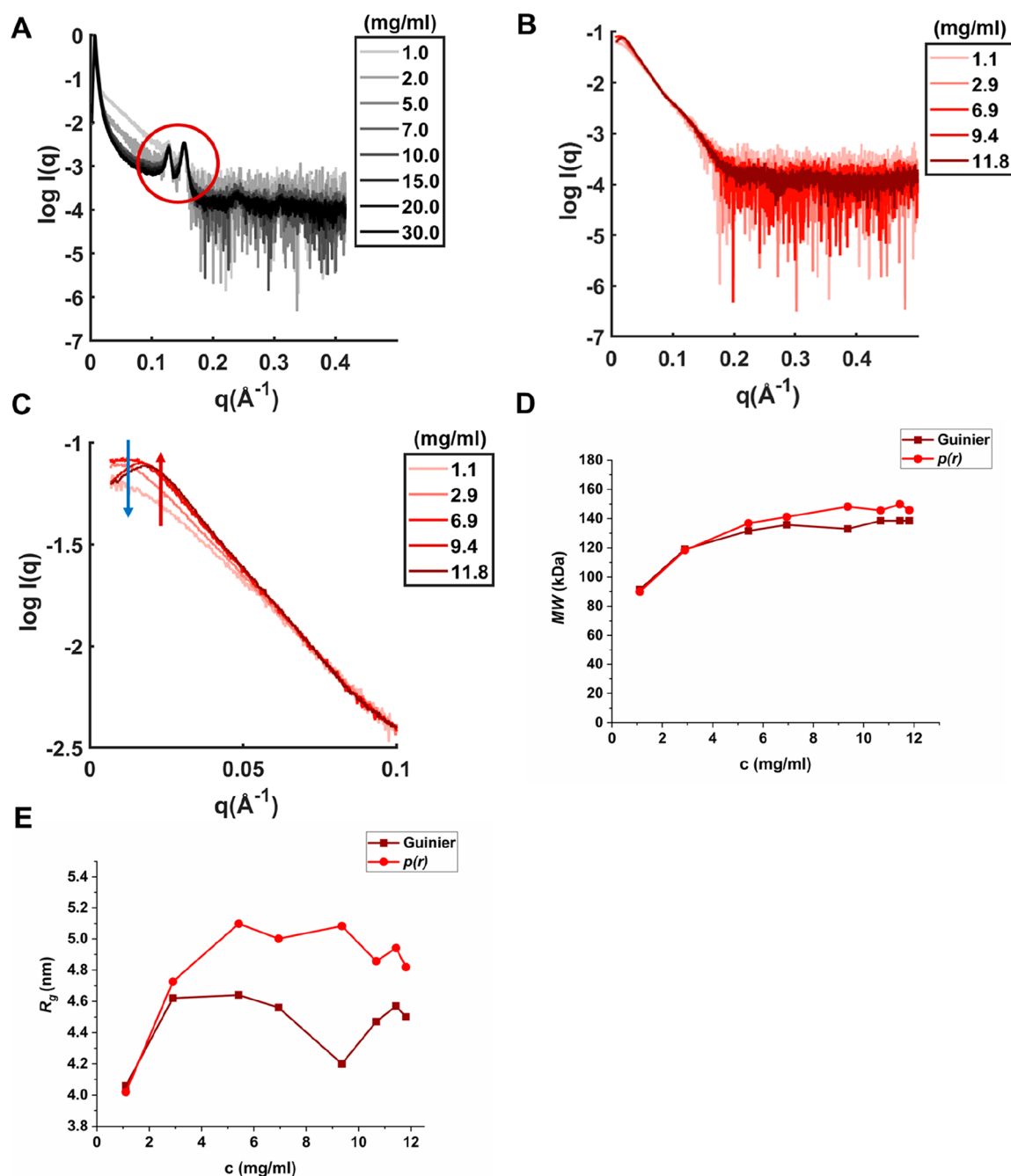
**Figure 3.** (A) Contour plot of the hydrodynamic radius  $R_h$  (nm) of IFN $\alpha$ -2a measured with DLS as a function of pH and salt (NaCl) concentration. (B) %Polydispersity. Formulations measured are indicated as dots. (C) Line plots of  $R_h$  measurements. Data are mean  $\pm$  SD for 3 replicates. (D) Visual inspection of IFN $\alpha$ -2a sample taken at pH 6 with 0, 70, and 140 mM NaCl. (E)  $R_h$  of IFN $\alpha$ -2a at pH 5 and pH 7.5 in the presence and absence of salt at various protein concentrations (see also Figure S4). Measured conditions are indicated by black circles.

increase and high standard deviation of the derived hydrodynamic radius  $R_h$  around pH 6, indicating the formation of large aggregates (Figure 3A,C). These aggregates were also observed upon visual inspection (Figure 3D). With salt (NaCl), we observed no visible aggregation. The polydispersity within the sample showed lower variation in the presence of salt (Figure 3B) resulting in more homogeneous results for  $R_h$  (Figure 3C). This indicated higher colloidal stability in the presence of salt. Over the whole pH range, IFN $\alpha$ -2a showed an increasing  $R_h$  upon increasing pH (Figure 3C), indicating an increased propensity to form reversible soluble oligomers. Under all measured conditions, the derived  $R_h$  was higher than the theoretical  $R_h$  of IFN $\alpha$ -2a of 2.27, calculated from the monomeric structure<sup>39</sup> (Table S4). We observed no stabilizing effect of different buffer systems (acetate and phosphate) or by the addition of excipients (280 mM sucrose, 140 mM arginine, 280 mM proline) (Figure S3). Concentration by centrifugation and subsequent dilution had no effect on  $R_h$ .

This high throughput screening was used to identify conditions to further investigate the colloidal stability in detail at pH 5 and pH 7.5 (below and above pI) with and without

salt. We investigated IFN $\alpha$ -2a at various protein concentrations using DLS (Figure 3E). At pH 5 the  $R_h$  of IFN $\alpha$ -2a showed an increasing trend toward higher protein concentration indicating the formation of soluble oligomers. At pH 7.5 in the presence of salt, we observed a similar increasing trend. Without salt,  $R_h$  did not increase at higher protein concentration but showed higher polydispersity and standard deviation (see also Figure S4). This indicates the presence of a small fraction of insoluble aggregates. Generally,  $R_h$  was lower when salt was present. At higher protein concentrations large insoluble aggregates formed in the absence of salt (see also Figure S4).

**Structural Analysis of IFN $\alpha$ -2a Using SAXS.** We measured SAXS at pH 5 and pH 7.5 to investigate the colloidal stability further and structurally analyze the oligomers formed by IFN $\alpha$ -2a. With no salt present, the SAXS measurements of IFN $\alpha$ -2a at pH 5 showed Bragg peaks in the SAXS curve (Figure 4A), indicating the formation of protein crystals, which were confirmed upon visual inspection under a microscope. The protein crystal formation appeared to be independent of the buffer system (Figure S5). At pH 7.5 the

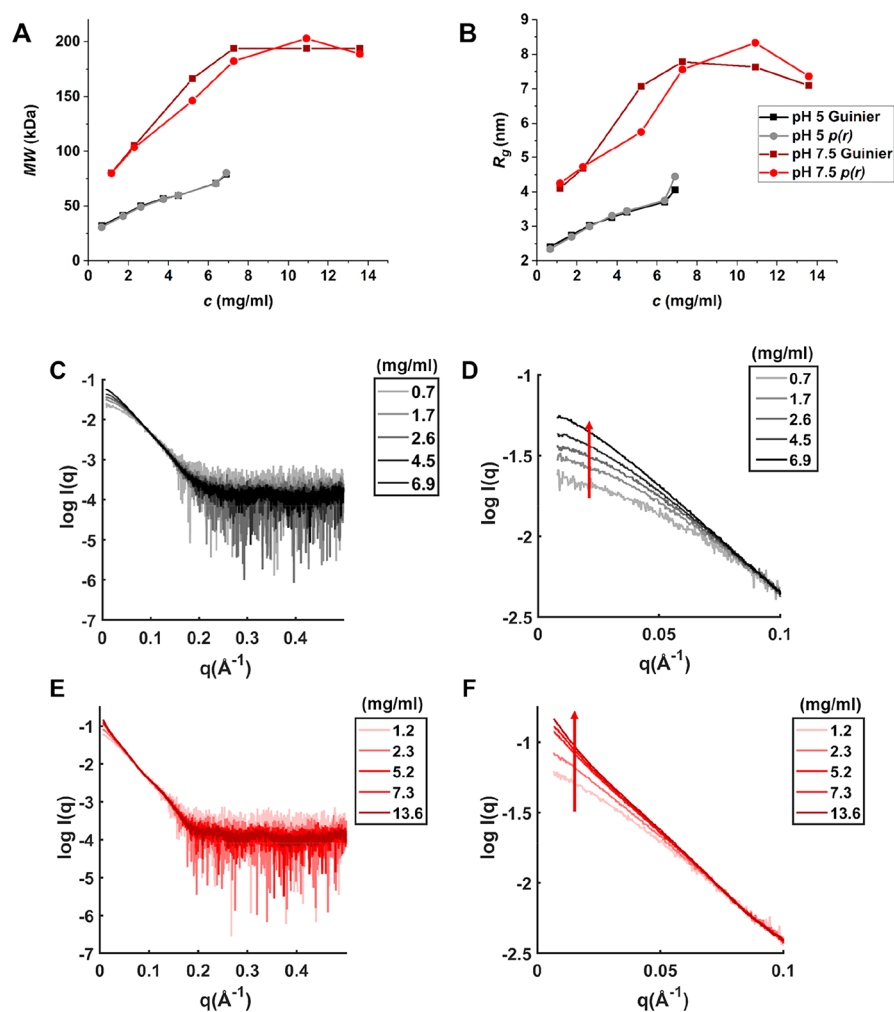


**Figure 4.** SAXS measurements of IFN $\alpha$ -2a in histidine at pH 5 and pH 7.5. (A) SAXS curves in histidine pH 5. Bragg peaks are indicated by a red circle. (B) SAXS scattering curves in histidine pH 7.5. (C) Scattering intensities at low  $q(\text{\AA}^{-1})$  at pH 7.5. Oligomerization or aggregation is indicated by the red arrow; repulsion is indicated by the blue arrow. (D) Derived apparent molecular weight MW from Guinier analysis and  $p(r)$  analysis at pH 7.5. (E) Derived radius of gyration  $R_g$  from Guinier analysis (dark red) and  $p(r)$  analysis (red) at pH 7.5. For distance distribution functions,  $p(r)$ , see Figure S6.

derived apparent molecular weight (MW) and radius of gyration ( $R_g$ ) increased with protein concentration (Figure 4). This was contradictory to the DLS measurements but might be explained by the high sensitivity of DLS toward aggregates present in the sample, introducing higher standard deviation and polydispersity.<sup>40</sup> Our SAXS measurements revealed the simultaneous presence of different-sized soluble oligomers, which is in agreement with previous studies on IFN $\alpha$ -2b.<sup>41</sup> The derived  $R_g$  showed a deviation between Guinier analysis and  $p(r)$ , most likely due to present insoluble aggregates, which have a larger impact on the analysis when the Guinier

approximation is used. The upper part of the SAXS curve (low  $q(\text{\AA}^{-1})$ ), which indicates protein–protein interactions, showed increased scattering intensity (red arrow) but also a decrease in scattering intensity at higher protein concentrations (blue arrows) (Figure 4C). This might be explained by increased self-association propensity at higher protein concentration, but repulsion is still present in the sample. A change in the buffer system did not change the overall shape of the SAXS curve of IFN $\alpha$ -2a at pH 7.5 (Figure S5).

In the presence of salt (140 mM NaCl), no IFN $\alpha$ -2a protein crystals formed. IFN $\alpha$ -2a showed a significantly lower apparent



**Figure 5.** SAXS measurements of IFN $\alpha$ -2a in histidine at pH 5 and pH 7.5 with 140 mM NaCl. (A) Derived apparent MW from Guinier analysis and  $p(r)$  analysis. (B) Derived radius of gyration  $R_g$  from Guinier analysis and  $p(r)$  analysis. (C) SAXS scattering curves at pH 5. (D) Scattering intensities at low  $q$  ( $\text{\AA}^{-1}$ ) at pH 5. Oligomerization/aggregation is indicated by the red arrow. (E) SAXS scattering curves at pH 7.5. (F) Scattering intensities at low  $q$  ( $\text{\AA}^{-1}$ ) at pH 7.5. Oligomerization/aggregation is indicated by the red arrow.

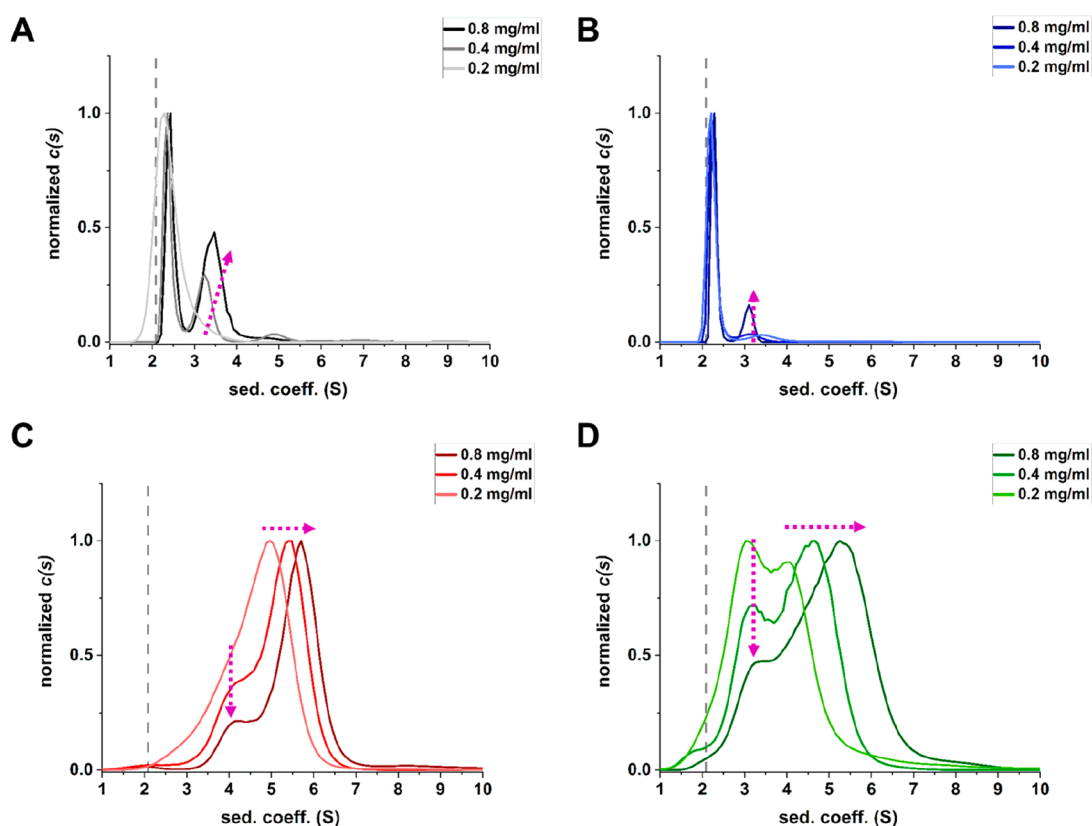
MW and smaller  $R_g$  at pH 5 than at pH 7.5 (Figure 5A,B), which agrees with the results from DLS. Similar to the measurements in the absence of salt, we observed an increase of apparent MW and  $R_g$  with increasing protein concentration. The scattering intensities at low  $q$  ( $\text{\AA}^{-1}$ ) indicated increased self-association at higher protein concentrations (Figure 5D,F). This was more pronounced at pH 7.5. With added salt, we observed a reduction of repulsion, but the overall shape of the SAXS scattering curves was unaffected by the presence of salt (Figure 5F, Figure S7), indicating the formation of similar soluble oligomers.

To characterize the structure of the oligomer formed, we attempted to use *ab initio* modeling using DAMMIF<sup>42</sup> and GASBOR.<sup>43</sup> *Ab initio* modeling determines the average particle in solution. We observed an elongated screw-shaped form as a common feature for the oligomers (Figure S8), but due to the simultaneous presence of different sized oligomers the exact assembly of IFN $\alpha$ -2a could not be modeled (see also supplementary discussion).

**Oligomeric Species of IFN $\alpha$ -2a Analyzed by AUC.** To examine which oligomeric species of IFN $\alpha$ -2a were present, we performed AUC. IFN $\alpha$ -2a was measured in histidine buffer pH 5 and pH 7.5 at different protein concentrations ( $c = 0.8$  mg/

mL ( $A_{280} = 1$ ),  $c = 0.4$  mg/mL ( $A_{280} = 0.5$ )  $c = 0.2$  mg/mL ( $A_{280} = 0.25$ )) in the presence or absence of salt (Figure 6). We calculated the sedimentation coefficient from monomeric IFN $\alpha$ -2a to  $s = 2.08$ <sup>39</sup> (dashed line). Soluble oligomeric species were present in all conditions but to different extents (Table S5). At pH 5, IFN $\alpha$ -2a was mainly monomeric. A second oligomeric species was present at pH 5, which showed a concentration-dependent increase. This was more pronounced in the salt-free case. Remarkably, with higher protein concentration this fraction showed a higher sedimentation coefficient, indicating a fast equilibrium between different soluble oligomeric species. This equilibrium shifts toward larger species at higher protein concentrations, which agreed with the SAXS measurements. At pH 5, this fraction corresponded to a dimer in equilibrium with higher oligomeric species. At pH 7.5, IFN $\alpha$ -2a was mainly present as soluble oligomeric species of different sizes and only a small fraction was monomeric. Two fractions of different oligomeric species were separated. The smaller oligomeric species corresponded most likely to dimers, and its proportion decreased at higher protein concentration. The proportion of larger oligomers increased with concentration and showed a concentration-dependent equilibrium toward larger oligomeric species. The





**Figure 6.** Analysis of IFN $\alpha$ -2a oligomers present in different conditions measured with AUC. (A) histidine pH 5, 0 mM NaCl. (B) histidine pH 5, 140 mM NaCl. (C) histidine pH 7.5, 0 mM NaCl. (D) histidine pH 7.5, 140 mM NaCl. All measurements were performed at  $c = 0.8$  mg/mL ( $A_{280} = 1$ ),  $c = 0.4$  mg/mL ( $A_{280} = 0.5$ ) and  $c = 0.2$  mg/mL ( $A_{280} = 0.25$ ) and analyzed with sedfit 16.1c.<sup>28</sup> The dashed line indicates the theoretical sedimentation coefficient of the monomer ( $s = 2.08$ ) based on alignment to the crystal structure of IFN $\beta$  (PDB:1AU1)<sup>44</sup> calculated with HULLRAD.<sup>39</sup>

proportion of larger oligomers appeared to be higher without salt than with salt. A comparison of different formulated samples indicated no significant difference in oligomers present in the same condition (Figure S9). When no salt was present, a small fraction of very large species corresponding to insoluble aggregates were present, which is in agreement with our DLS measurements.

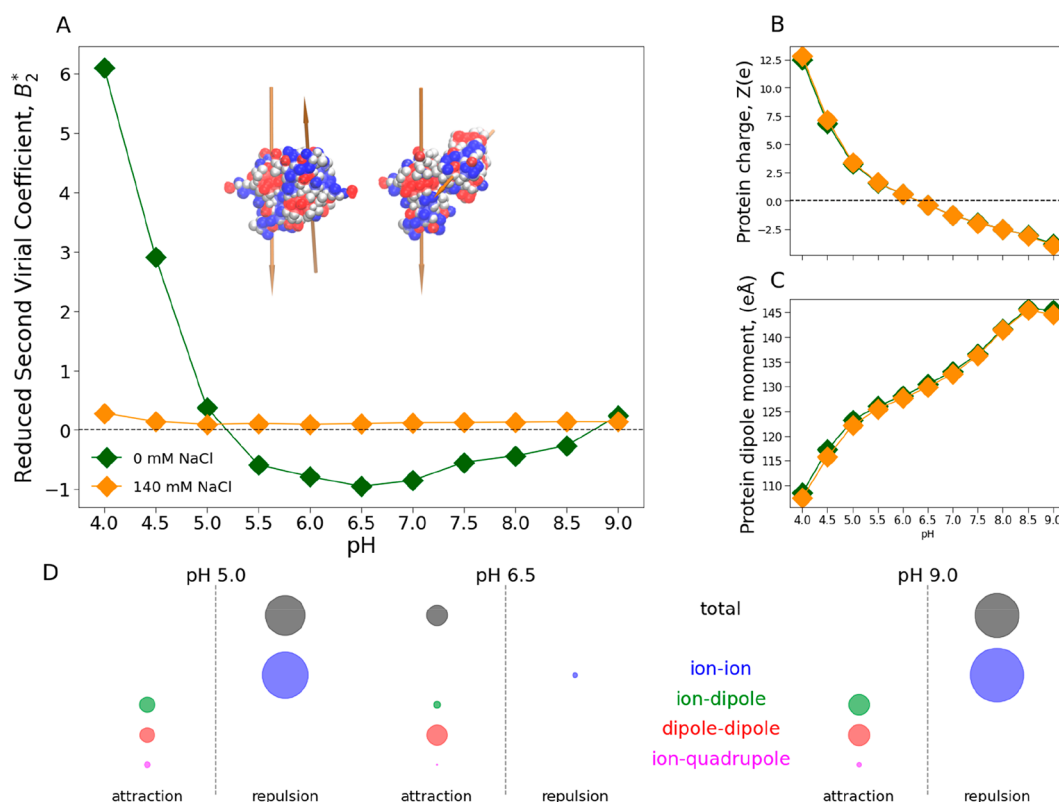
**Metropolis Monte Carlo Simulations.** To obtain insight into the molecular mechanism driving the aggregation/oligomer formation of the IFN $\alpha$ -2a solution, we applied a combination of Metropolis Monte Carlo simulations and coarse-grained modeling. Details regarding the simulations and models used are described in the Materials and Methods section. Figure 7 shows the calculated reduced second osmotic virial coefficient,  $B_2^* = B_2/B_2^{\text{hs}}$ , the net protein charge,  $Z$ , and the protein dipole,  $\mu$ , as a function of pH, with no added salt and with 140 mM of NaCl. The simulations were performed using a two-body model, that is, where only two proteins were embedded in the aqueous salt solution.

Our simulations showed that without salt,  $B_2^*$  is strongly pH-dependent. At low and high pH IFN $\alpha$ -2a showed net repulsion in solution, indicated by positive  $B_2^*$ . At neutral pH,  $B_2^*$  was negative indicating attraction, with the highest attraction at pH 6.5, corresponding approximately to the protein's pI. With added salt, IFN $\alpha$ -2a showed constant net repulsion in solution independent of pH. These results were in very good agreement with the colloidal stability of IFN $\alpha$ -2a measured with DLS,

indicating that our model captured the main molecular mechanisms leading to self-association of IFN $\alpha$ -2a.

Treating protein charge distribution as a set of electrostatic multipoles, we first investigated the protein net charge,  $Z$ , and the protein dipole,  $\mu$  (Figure 7B,C). As expected, the absolute value of  $Z$  was high at pH far from the protein's pI ( $\text{pH} \approx 6$ ) and zero at the pI.  $\mu$  showed pH dependence and increases with pH. We analyzed the multipole terms in the absence of salt to investigate their contribution to protein–protein interaction of IFN $\alpha$ -2a (Figure 7D). We saw that repulsion was driven by ion–ion contributions. The attraction was driven by ion–dipole, dipole–dipole, and ion–quadrupole contributions. At low and high pH, where IFN $\alpha$ -2a was highly charged, ion–ion interactions dominated the system, leading to overall repulsion. Around the protein's isoelectric point with little or no protein charge, dipole–dipole contributions dominated, leading to overall attraction. With added salt, electrostatic contributions were effectively screened, independent of pH. This electrostatic behavior agreed with our SAXS measurements where we observed a reduction in repulsion when adding salt. Using the graphical software VMD,<sup>45</sup> we analyzed the dipole–dipole orientation during the simulations. The energy configurations were minimized at antiparallel or head-to-tail orientation (Figure 7A), indicating these as preferred orientations.

To investigate the formation of clusters (or soluble oligomers), we used a many-body model where multiple proteins are allowed to interact. We analyzed the average



**Figure 7.** Two-body coarse-grained Metropolis Monte Carlo simulations as a function of pH with 0 mM (green) and 140 mM NaCl (orange). (A) Reduced second osmotic virial coefficient,  $B_2^*$ , and snapshots of the dipole orientations corresponding to the minimum energy configurations at pH = pI. The arrows, pointing from the negative to the positive charge, represent the dipole magnitudes ( $\approx 122$  eÅ or  $\approx 25$  Debye). (B) Total protein charge,  $Z$ . (C) Net protein dipole moment,  $\mu$ . (D) Electric multipole contributions to the protein–protein interaction for pH 5, 6.5, and 9 (0 mM NaCl). The circle area illustrates the magnitude of the contribution. Negative (attractive) and positive (repulsive) contributions are separated by the dashed line. The sum of all multipole contributions is represented by the gray circle (“total”).

cluster size and shape as a function of pH with and without salt. The strength of the Lennard-Jones interactions, represented by the  $\epsilon_{ij}$  parameter, was also explored using the parallel tempering method (see [Materials and Method](#) section). Increasing  $\epsilon_{ij}$  can be approximately interpreted as a lowering of the system temperature.

At low protein concentration, the average cluster size is close to 1, meaning that only monomers were observed ([Figure 8A](#), left). At higher protein concentration, the average cluster size increases with increasing pH ([Figure 8A](#), right). The largest clusters are formed around the protein’s pI and at a lower temperature (high  $\epsilon_{ij}$ ). This effect was less pronounced in the presence of salt, due to screening of the electrostatics. These results correlate very well with our experimental observations where we found increased oligomerization with increasing pH and a pronounced oligomer formation around the protein pI.

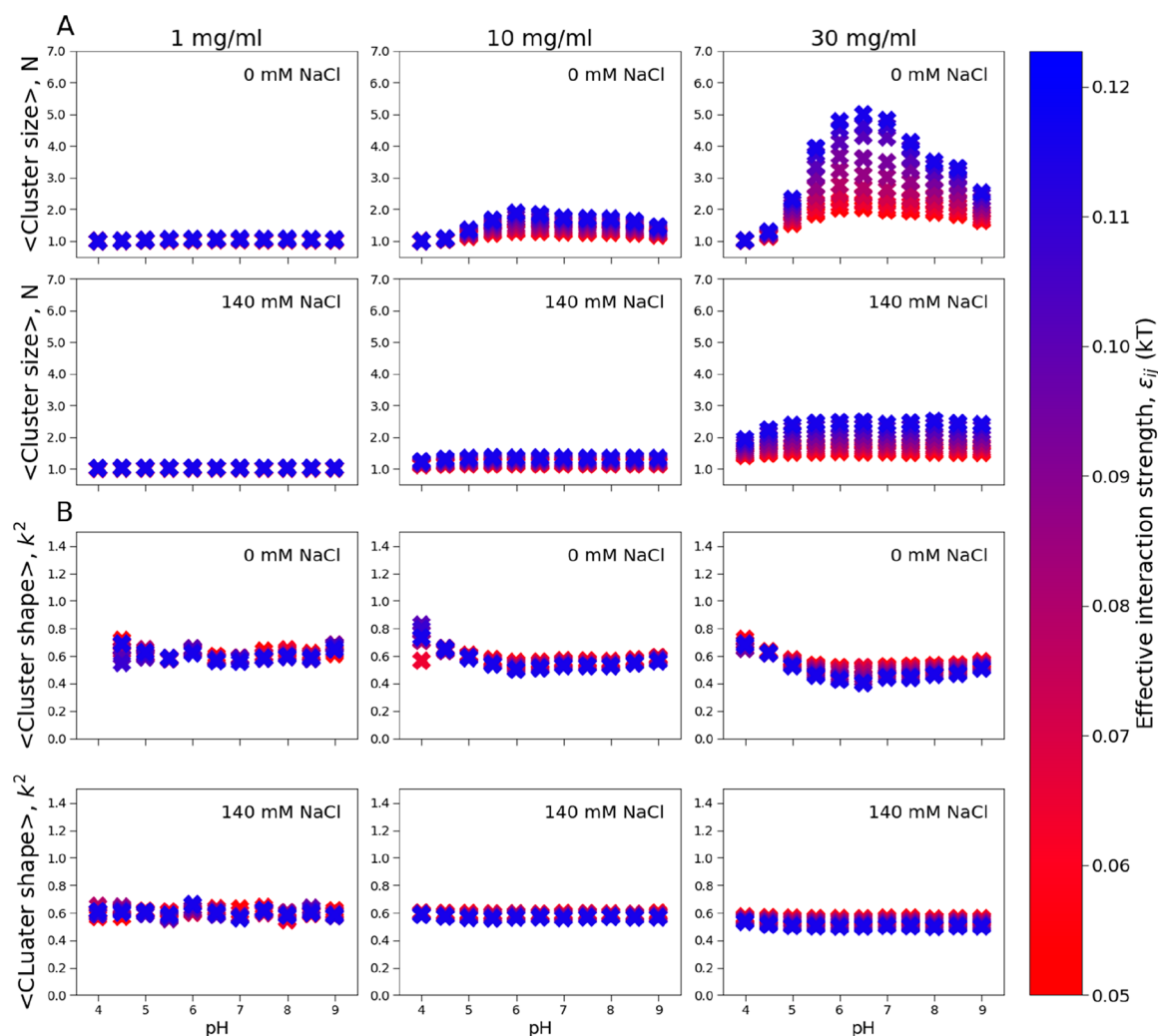
We additionally analyzed the average cluster shape by determining the shape anisotropy parameter,  $k^2$ , a method that is commonly applied to polymers.<sup>46</sup> We observed more spherical cluster shapes ( $k^2$ ) with increasing cluster size, indicating a less defined orientation for larger oligomers ([Figure 8B](#)). Snapshots of IFN $\alpha$ -2a trimers, corresponding to different  $k^2$  values, are shown in the [Materials and Methods](#) section.

**Stability toward Thermal and Chemical Denaturation of IFN $\alpha$ -2a.** In addition to the colloidal stability of IFN $\alpha$ -2a, we studied the conformational stability as a function of pH at 0, 70, and 140 mM NaCl. The thermally induced unfolding ( $T_{1/2}$ ) of IFN $\alpha$ -2a was measured using nanoDSF ([Figure 9A](#)).

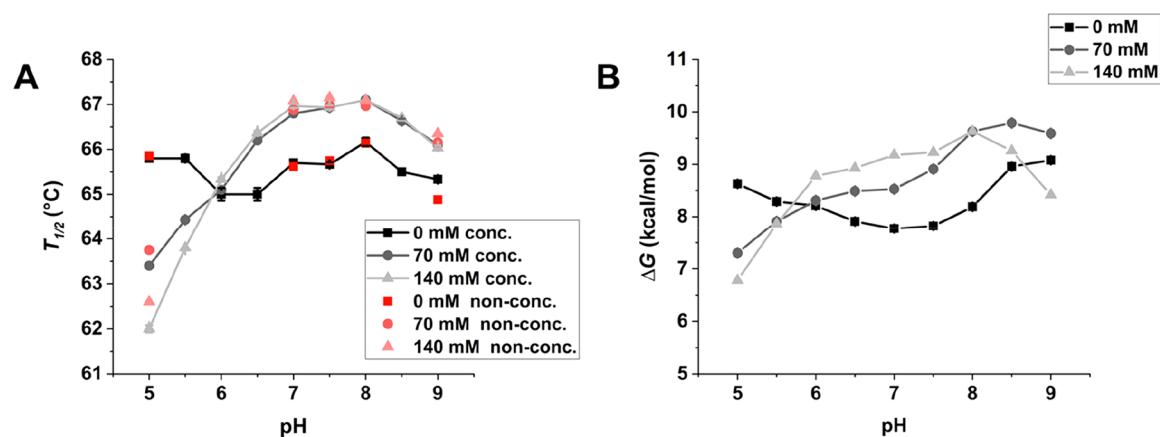
With no salt present, IFN $\alpha$ -2a unfolded at around 65 °C, independent of pH. With added salt (70 and 140 mM NaCl),  $T_{1/2}$  of IFN $\alpha$ -2a was highly pH-dependent. NaCl showed a destabilizing effect on the thermal stability below pH 6 and a stabilizing effect above pH 6. We tested the influence of different formulation procedures of IFN $\alpha$ -2a.  $T_{1/2}$  of the sample concentrated before dialysis (black and gray) was in good agreement with the nonconcentrated sample (red). We determined the free energy changes by denaturant-induced unfolding ( $\Delta G$ ) with ICD ([Figure 9B](#)). The conformational stability toward chemically induced denaturation showed a similar trend as seen for thermal-induced denaturation. Remarkably, the curves for both  $T_{1/2}$  and  $\Delta G$  cross at pH 6, correlating with the protein’s pI of 5.98. The conformational stability of IFN $\alpha$ -2a was additionally investigated in different buffer systems (acetate pH 5 and phosphate pH 7.5) and in addition of selected excipients (sucrose (280 mM), arginine-HCl (140 mM) and proline (280 mM), but no significant changes could be seen ([Figure S11](#)).

## DISCUSSION

IFN $\alpha$ -2a is widely used as a therapeutic drug against various diseases<sup>12–14</sup> but has shown difficulties in formulation due to aggregation. We assessed the colloidal stability of IFN $\alpha$ -2a as a function of pH experimentally and through molecular simulations. Our experimental results were in excellent agreement with our simulations, indicating that the main factors leading to IFN $\alpha$ -2a aggregation were captured by the



**Figure 8.** Cluster formation of IFN $\alpha$ -2a in many-body coarse-grained Metropolis Monte Carlo simulations. (A) Average cluster size.  $N$  indicates the average number of monomers in the cluster observed during the simulations. (B) Average cluster shape.  $k^2$  is the relative shape anisotropy.  $k^2$  close to 0 indicates a spherical cluster while a  $k^2$  close to 1 indicates an elongated, rod-like structure. By definition,  $k^2$  is equal to 1 for dimers, so the analysis is reported only for trimers or larger. Clustering is shown as a function of pH at three protein concentrations (1, 10, and 30 mg/mL), and at 0 mM and 140 mM NaCl. For each condition replica exchange simulations are reported for  $\epsilon_{ij}$  values in the range of 0.05 kT (red) to 0.012  $k_B T$  (blue).



**Figure 9.** (A) Thermal denaturation point  $T_{1/2}$  of IFN $\alpha$ -2a measured with nanoDSF at different formulation procedures. A sample concentrated before dialysis (conc., black) and a nonconcentrated sample (non-conc., red) were measured. Data are mean  $\pm$ SD for three replicates (error bars obscured by points). (B) Stability of IFN $\alpha$ -2a toward chemical denaturation measured with ICD in the presence of 0, 70, and 140 mM NaCl as a function of pH.

simulation models. We observed increased formation of soluble oligomers with increasing pH and the formation of large insoluble aggregates around the protein's pI of 5.98 experimentally as well as in our simulations, which is in agreement with previous studies.<sup>16</sup> The addition of salt prevented the formation of insoluble aggregates and shifted the equilibrium of soluble oligomers slightly toward smaller species. The effect of salt on protein stability is poorly understood, but molecules with significant dipole moment were shown to be stabilized by the addition of salt through screening of the anisotropic charge distribution within the protein.<sup>47,48</sup> With added salt, a positive  $B_2$  indicated repulsion independent of pH, thus supporting the idea that nonspecific ion binding and effective screening of charges stabilized IFN $\alpha$ -2a. In our simulations, the only electrostatic contribution for repulsive behavior was pH dependent due to ion–ion interactions. Attraction instead came mainly from dipole–dipole interactions. Further, the dipole moment of IFN $\alpha$ -2a showed a pH-dependent increase, which might explain increased formation of soluble oligomers seen in our experiments. Huang et al. showed that the addition of salt led to increased attraction.<sup>17</sup> However, in our study, we showed that  $B_2$  is highly pH-dependent in the absence of salt, and salt-induced attraction thus cannot be generalized.

Our experimental results revealed the simultaneous presence of different species of soluble oligomers that seemed to be in a fast, concentration-dependent equilibrium with each other, which is in agreement with previous studies on IFN $\alpha$ -2b.<sup>41</sup> Our experiments showed a clear correlation between pH and the formation of soluble oligomers. Our many-body simulations indicated a general trend toward larger average cluster size with increasing pH, which was in good agreement with experimental results. Analysis of the cluster shape in our simulations showed more spherical, less defined clusters when the cluster formation is increased. Our SAXS measurements, however, showed elongated oligomers. Even though we believe the main driving forces for IFN $\alpha$ -2a oligomerization and aggregation are electrostatic interactions, hydrophobic interactions, which are unaccounted for in our simulations, might influence the self-association of IFN $\alpha$ -2a.

Our experimental measurements in combination with the Monte Carlo simulations point toward different types of protein–protein interaction for IFN $\alpha$ -2a. The formation of large insoluble aggregates and protein crystals could be diminished by the presence of salt. The formation of smaller soluble oligomers appeared to be less influenced by the presence of salt but very much pH and protein concentration dependent. From a pharmaceutical point of view, irreversible aggregation can lead to compromised activity and immune response in patients. We found the least number of insoluble aggregates at low pH (pH 5) in the presence of salt (140 mM NaCl). This formulation resembles the commercial formulation of IFN $\alpha$ -2a (Roferon-A; 10 mM ammonium acetate buffer pH 5, 120 mM NaCl) closely. As IFN $\alpha$ -2a has been reported to be acid labile and undergoes partial unfolding at pH 4,<sup>18</sup> this formulation likely represents the optimal conditions to prevent aggregation. However, other IFN $\alpha$  products (IFN $\alpha$ -2b (Intron-A), Interferon aphacon-1 (Infergen) and IFN $\alpha$ -n3 (Alferon)) are formulated in phosphate buffer at pH 7 despite studies that show self-association under this condition.<sup>41,49</sup> This might be explained by the low protein concentration in commercial formulation, which shifts the equilibrium toward the monomeric form as we could show in our study.

IFN $\alpha$ -2a only formed protein crystals when no salt was present, indicating that the crystal interactions are most likely formed by ionic interactions. The crystal formation has not been reported in any previous studies of IFN $\alpha$ -2a.<sup>16–19</sup> SAXS is currently the only technique applied for IFN $\alpha$ -2a characterization that can show the presence of protein crystals apart from visual inspection in a microscope. Remarkably, these IFN $\alpha$ -2a protein crystals formed without zinc(II), which raises the question of whether the zinc being a mediator, as reported for the dimerization of IFN $\alpha$ -2b and IFN $\beta$ ,<sup>44,50</sup> is necessary to form interferon crystals. The pH of IFN $\alpha$ -2b and IFN $\beta$  crystal formation is close to pH 5, indicating that some protein–protein interactions are similar within the group of interferons. Solving the crystal structure of IFN $\alpha$ -2a would highly benefit this study but has not been possible to date as the obtained crystals were too fragile to handle.

The measurement of the stability of IFN $\alpha$ -2a toward thermal and chemical denaturation showed a lower stability in the presence of salt at pH 5, which is in agreement with previous studies.<sup>17–19</sup> The contrast between IFN $\alpha$ -2a conformational stability below and above its pI is remarkable. These measurements indicate, that in the case of a predominantly negatively charged protein, the presence of salt enhances IFN $\alpha$ -2a's conformational stability, while it is lower in the case of a predominantly positively charged protein. This is in agreement with previous studies on IFN $\alpha$ -2b.<sup>49</sup> The conformational stability of IFN $\alpha$ -2a was independent of the colloidal stability and shows that a combination of different variables in a screening for stability is important and cannot be generalized. The concentration of IFN $\alpha$ -2a had no influence on IFN $\alpha$ -2a's colloidal and conformational stability, which was indicated in previous studies.<sup>16</sup>

## CONCLUSION

Through systematic biophysical and computational characterization of IFN $\alpha$ -2a, we gained molecular insights into the driving forces of the oligomerization and aggregation. With this combination of methods, we could assess the challenge of characterizing a polydisperse system as IFN $\alpha$ -2a in a new way. Because of the sequence and structural similarity among cytokines, our results benefit the study of these powerful protein drugs. Formulation and reformulation have no impact on the results, which could lead to new approaches to the formulation of protein drugs. They could, for example, be stored at low pH, a condition where they show high stability as indicated in our study, and reformulated just before treatment into conditions where they show lower stability but are more beneficial for the patient.

The data presented in this study will be connected with similar data sets on other types of protein drugs in the context of the EU international training network project on 'protein excipient and protein–protein interaction in formulation (PIPPPI)'.<sup>51</sup> This will finally result in a comprehensive data set characterizing different kinds of protein drugs as a function of formulation conditions and facilitate a better understanding of the solution behavior of protein drugs.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.1c07090>.

Sequence and structural alignment within the Interferon class; additional SAXS parameters and analysis (full SAXS curves,  $p(r)$  functions, *ab initio* modeling); calculated biophysical parameters; additional DLS and ICD measurements in addition of excipients; quantitative AUC analysis; supplementary discussion regarding dimer assembly (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

**Christin Pohl** – *Novozymes A/S, Bagsvaerd 2880, Denmark; Department of Chemistry, Technical University of Denmark, Kongens Lyngby 2800, Denmark; Present Address: Lund University, Division of Biochemistry and Structural Biology, Department of Chemistry, 221 00 Lund, Sweden; [orcid.org/0000-0001-6193-4016](https://orcid.org/0000-0001-6193-4016); Email: [christin.pohl@biochemistry.lu.se](mailto:christin.pohl@biochemistry.lu.se)*

**Pernille Harris** – *Department of Chemistry, Technical University of Denmark, Kongens Lyngby 2800, Denmark; Present Address: Department of Chemistry, University of Copenhagen, 2100 Copenhagen, Denmark; [orcid.org/0000-0002-6806-4903](https://orcid.org/0000-0002-6806-4903); Email: [phharris@chem.ku.dk](mailto:phharris@chem.ku.dk)*

### Authors

**Marco Polimeni** – *Division of Theoretical Chemistry, Department of Chemistry, Lund University, 221 00 Lund, Sweden; [orcid.org/0000-0002-9670-7232](https://orcid.org/0000-0002-9670-7232)*

**Sowmya Indrakumar** – *Department of Chemistry, Technical University of Denmark, Kongens Lyngby 2800, Denmark; Present Address: Department of Biology, University of Copenhagen, 2100 Copenhagen, Denmark; [orcid.org/0000-0002-5144-3261](https://orcid.org/0000-0002-5144-3261)*

**Werner Streicher** – *Novozymes A/S, Bagsvaerd 2880, Denmark; Present Address: NanoTemper Technologies GmbH, Muenchen, Germany.*

**Günther H.J. Peters** – *Department of Chemistry, Technical University of Denmark, Kongens Lyngby 2800, Denmark; [orcid.org/0000-0001-9754-2663](https://orcid.org/0000-0001-9754-2663)*

**Allan Nørgaard** – *Novozymes A/S, Bagsvaerd 2880, Denmark*  
**Mikael Lund** – *Division of Theoretical Chemistry, Department of Chemistry, Lund University, 221 00 Lund, Sweden; Present Address: Lund Institute of Advanced Neutron and X-ray Science (LINXS), 22370 Lund, Sweden; [orcid.org/0000-0001-8178-8175](https://orcid.org/0000-0001-8178-8175)*

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jpcc.1c07090>

### Author Contributions

#C.P. and M.P. contributed equally. C.P., M.P., A.N., M.L., and P.H. designed research; C.P. performed experimental research; M.P. performed molecular Metropolis Monte Carlo simulations; S.I. performed charge distribution on dimers; C.P. and M.P. analyzed data; W.S., G.H.J.P., A.N., M.L., and P.H. supervised the study. C.P. and M.P. wrote the paper with support from A.N., M.L., and P.H.

### Notes

The authors declare no competing financial interest. Electronic notebooks for reproducing the presented simulations and analysis are available at <https://github.com/mlund/SI-interferon-oligomerization>

## ACKNOWLEDGMENTS

The synchrotron SAXS data were collected at beamline P12 operated by EMBL Hamburg at the Petra III storage ring (DESY, Hamburg, Germany), and at the BioSAXS beamline, BM29, at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We thank Melissa Gräwert, Stefano Da Vela, Andrey Gruzinov, Karen Manalastas (all: EMBL, Hamburg) and Martha Brennich (ESRF, Grenoble) for local support. This work was funded by European Union's Horizon 2020 research and innovation program (Grant Agreement No. 675074). We thank the Danish Agency for Science, Technology, and Innovation for funding the instrument center DanScatt. Simulations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at the center for scientific and technical computing at Lund University (LUNARC), at the high-performance computing (HPC) services at DTU, and in-house CPU/GPU cluster facilities at DTU chemistry. M.L. thanks the Swedish Research Council for financial support.

## REFERENCES

- (1) Walsh, G. *Biopharmaceutical benchmarks 2018*; Nature Publishing Group, 2018; Vol. 36, pp 1136–1145.
- (2) Mullard, A. 2018 FDA drug approvals. *Nat. Rev. Drug Discovery* **2019**, *18*, 85–89.
- (3) Revers, L.; Furczon, E. An introduction to biologics and biosimilars. Part I: Biologics: What are they and where do they come from? *Canadian Pharmacists Journal* **2010**, *143*, 134–139.
- (4) Manning, M. C.; Patel, K.; Borchardt, R. T. Stability of Protein Pharmaceuticals. *Pharm. Res.* **1989**, *06*, 903–918.
- (5) Manning, M. C.; Chou, D. K.; Murphy, B. M.; Payne, R. W.; Katayama, D. S. Stability of protein pharmaceuticals: An update. *Pharm. Res.* **2010**, *27*, 544–575.
- (6) Wang, W. Instability, stabilization, and formulation of liquid protein pharmaceuticals. *Int. J. Pharm.* **1999**, *185*, 129–188.
- (7) Nahri, L. O.; Schmit, J.; Bechtold-Peters, K.; Sharma, D. Classification of Protein Aggregates. *J. Pharm. Sci.* **2012**, *101*, 493–498.
- (8) Roberts, C. J. Non-native protein aggregation kinetics. *Biotechnol. Bioeng.* **2007**, *98*, 927–938.
- (9) Roberts, C. J.; Das, T. K.; Sahin, E. Predicting solution aggregation rates for therapeutic proteins: Approaches and challenges. *Int. J. Pharm.* **2011**, *418*, 318–333.
- (10) Ruiz, L.; Aroche, K.; Reyes, N. Aggregation of recombinant human interferon alpha 2b in solution: Technical note. *AAPS PharmSciTech* **2006**, *7*, 5–9.
- (11) Klaus, W.; Gsell, B.; Labhardt, A. M.; Wipf, B.; Senn, H. The three-dimensional high-resolution structure of human interferon  $\alpha$ -2a determined by heteronuclear NMR spectroscopy in solution. *J. Mol. Biol.* **1997**, *274*, 661–675.
- (12) Dragomiretskaya, N.; Izha, A.; Kalinichenko, N.; Szark-Eckardt, M. I.; Klimczyk, M.; Cieslicka, M. I.; Muszkieta, R. I.; Prusik, K.; Napierala, M.; Zukowska, H.; Zukow, W. Use of antiviral therapy in patients with chronic hepatitis C. *Open Medicine (Poland)* **2015**, *10*, 209–215.
- (13) Escudier, B.; et al. Bevacizumab plus interferon alfa-2a for treatment of metastatic renal cell carcinoma: a randomised, double-blind phase III trial. *Lancet* **2007**, *370*, 2103–2111.
- (14) Jonasch, E.; Haluska, F. G. Interferon in Oncological Practice: Review of Interferon Biology, Clinical Applications, and Toxicities. *Oncologist* **2001**, *6*, 34–55.
- (15) Lipiäinen, T.; et al. Formulation and stability of cytokine therapeutics. *J. Pharm. Sci.* **2015**, *104*, 307–326.
- (16) Panjwani, N.; Hodgson, D. J.; Sauv e, S.; Aubin, Y. Assessment of the effects of pH, formulation and deformulation on the

conformation of interferon alpha-2 by NMR. *J. Pharm. Sci.* **2010**, *99*, 3334–3342.

(17) Huang, S. L.; Lin, F. Y.; Yang, C. P. Microcalorimetric studies of the effects on the interactions of human recombinant interferon- $\alpha$ 2a. *Eur. J. Pharm. Sci.* **2005**, *24*, 545–552.

(18) Sharma, V. K.; Kalonia, D. S. *Pharm. Res.* **2003**, *20*, 1721–1729.

(19) Svilenov, H.; Winter, G. Rapid sample-saving biophysical characterisation and long-term storage stability of liquid interferon alpha2a formulations: Is there a correlation? *Int. J. Pharm.* **2019**, *562*, 42–50.

(20) Schiffter, H. A. Pharmaceutical Proteins - Structure, Stability, and Formulation. *Comprehensive Biotechnology*, 2nd ed.; Elsevier B.V., 2011; Vol. 5.

(21) Jorgensen, L.; Hostrup, S.; Moeller, E. H.; Grohgan, H. Recent trends in stabilising peptides and proteins in pharmaceutical formulation - Considerations in the choice of excipients. *Expert Opin. Drug Delivery* **2009**, *6*, 1219–1230.

(22) Hochuli, E. Interferon immunogenicity: technical evaluation of interferon-alpha 2a. *J. Interferon Cytokine Res.* **1997**, *17* (1), S15–21.

(23) Braun, A.; Kwee, L.; Labow, M. A.; Alsenz, J. Protein aggregates seem to play a key role among the parameters influencing the antigenicity of interferon alpha in normal and transgenic mice. *Pharm. Res.* **1997**, *14*, 1472–1478.

(24) Foser, S.; et al. Isolation, structural characterization, and antiviral activity of positional isomers of monopegylated interferon  $\alpha$ -2a (PEGASYS). *Protein Expression Purif.* **2003**, *30*, 78–87.

(25) Pohl, C.; et al. The effect of point mutations on the biophysical properties of an anti-microbial peptide: development of a screening protocol for peptide stability screening. *Mol. Pharmaceutics* **2020**, *17*, 3298–3313.

(26) Franke, D.; et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **2017**, *50*, 1212–1225.

(27) Schuck, P.; Perugini, M. A.; Gonzales, N. R.; Howlett, G. J.; Schubert, D. Size-distribution analysis of proteins by analytical ultracentrifugation: Strategies and application to model systems. *Biophys. J.* **2002**, *82*, 1096–1111.

(28) Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **2000**, *78*, 1606–1619.

(29) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087.

(30) Stenqvist, B.; Thuresson, A.; Kurut, A.; Vácha, R.; Lund, M. Faunus-A flexible framework for Monte Carlo simulation. *Mol. Simul.* **2013**, *39*, 1233.

(31) Hagler, A. T.; Huler, E.; Lifson, S. Energy Functions for Peptides and Proteins. I. Derivation of a Consistent Force Field Including the Hydrogen Bond from Amide Crystals. *J. Am. Chem. Soc.* **1974**, *96*, 5319.

(32) Johnson, J. K.; Panagiotopoulos, A. Z.; Gubbins, K. E. Reactive canonical Monte Carlo. *Mol. Phys.* **1994**, *81*, 717.

(33) Thurlkill, R. L.; Grimsley, G. R.; Scholtz, J. M.; Pace, C. N. pK values of the ionizable groups of proteins. *Protein Sci.* **2006**, *15*, 1214.

(34) Moon, Y. U.; Anderson, C. O.; Blanch, H. W.; Prausnitz, J. M. Osmotic pressures and second virial coefficients for aqueous saline solutions of lysozyme. *Fluid Phase Equilib.* **2000**, *168*, 229.

(35) Quigley, A.; Williams, D. R. The second virial coefficient as a predictor of protein aggregation propensity: A self-interaction chromatography study. *Eur. J. Pharm. Biopharm.* **2015**, *96*, 282.

(36) Lacava, F. Classical Electrodynamics: A Short Review. *Classical Electrodynamics*; Springer, 2016; Chapter 1. DOI: 10.1007/978-3-319-39474-9\_1.

(37) Wu, D.; Chandler, D.; Smit, B. Electrostatic analogy for surfactant assemblies. *J. Phys. Chem.* **1992**, *96*, 4077–4083.

(38) Arkin, H.; Janke, W. Gyration tensor based analysis of the shapes of polymer chains in an attractive spherical cage. *J. Chem. Phys.* **2013**, *138*, 054904.

(39) Fleming, P. J.; Fleming, K. G. HullRad: Fast Calculations of Folded and Disordered Protein and Nucleic Acid Hydrodynamic Properties. *Biophys. J.* **2018**, *114*, 856–869.

(40) Pohl, C.; Mahapatra, S.; Kulakova, A.; Streicher, W.; Peters, G. H.J.; Nørgaard, A.; Harris, P. Combination of high throughput and structural screening to assess protein stability - a screening perspective. *European Journal of Pharmaceutical Sciences* **2021**, *1* DOI: 10.1016/j.ejpb.2021.08.018.

(41) Li, Y.; et al. Characterization of the Self-Association of Human Interferon- $\alpha$ 2b, Albinterferon- $\alpha$ 2b, and Pegasys. *J. Pharm. Sci.* **2012**, *101*, 68–80.

(42) Franke, D.; Svergun, D. I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **2009**, *42*, 342–346.

(43) Svergun, D. I.; Petoukhov, M. V.; Koch, M. H. J. Determination of domain structure of proteins from x-ray solution scattering. *Biophys. J.* **2001**, *80*, 2946–2953.

(44) Karpusas, M.; et al. The crystal structure of human interferon  $\beta$  at 2.2-Å resolution. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 11813–11818.

(45) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33.

(46) Theodorou, D. N.; Suter, U. W. Shape of unperturbed linear polymers: polypropylene. *Macromolecules* **1985**, *18*, 1206–1214.

(47) Hamabata, A.; Chang, S.; von Hippel, P. H. Model Studies on the Effects of Neutral Salts on the Conformational Stability of Biological Macromolecules. III. Solubility of Fatty Acid Amides in Ionic Solutions. *Biochemistry* **1973**, *12*, 1271–1278.

(48) Kirkwood, J. G. *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*; Reinhold Publishing Corporation, 1943.

(49) Beldarraín, A.; Cruz, Y.; Cruz, O.; Navarro, M.; Gil, M. Purification and conformational properties of a human interferon  $\alpha$ 2b produced in *Escherichia coli*. *Biotechnol. Appl. Biochem.* **2001**, *33*, 173.

(50) Radhakrishnan, R.; et al. Zinc mediated dimer of human interferon- $\alpha$ (2b) revealed by X-ray crystallography. *Structure* **1996**, *4*, 1453–1463.

(51) Gentiluomo, L.; Svilenov, H. L.; Augustijn, D.; El Bialy, I.; Greco, M. L.; Kulakova, A.; Indrakumar, S.; Mahapatra, S.; Morales, M. M.; Pohl, C.; et al. Advancing Therapeutic Protein Discovery and Development through Comprehensive Computational and Biophysical Characterization. *Mol. Pharmaceutics* **2020**, *426* DOI: 10.1021/acs.molpharmaceut.9b00852.