SCIENTIFIC REPORTS
nature research

Check for updates

OPEN

On the Aggregation and Nucleation Mechanism of the Monoclonal Antibody Anti-CD20 Near Liquid-Liquid Phase Separation (LLPS)

Elvira Pantuso[2], Teresa F. Mastropietro[1], Maria L. Briuglia[3], Charline J. J. Gerard[3], Efrem Curcio[4,5], Joop H. ter Horst[3], Fiore P. Nicoletta[2] ✉ & Gianluca Di Profio[1,5] ✉

The crystallization of Anti-CD20, a full-length monoclonal antibody, has been studied in the PEG400/Na_2SO_4/Water system near Liquid-Liquid Phase Separation (LLPS) conditions by both sitting-drop vapour diffusion and batch methods. In order to understand the Anti-CD20 crystallization propensity in the solvent system of different compositions, we investigated some measurable parameters, normally used to assess protein conformational and colloidal stability in solution, with the aim to understand the aggregation mechanism of this complex biomacromolecule. We propose that under crystallization conditions a minor population of specifically aggregated protein molecules are present. While this minor species hardly contributes to the measured average solution behaviour, it induces and promotes crystal formation. The existence of this minor species is the result of the LLPS occurring concomitantly under crystallization conditions.

A thorough understanding of protein aggregation behaviour in solution is crucial for the controlled manufacture and formulation of therapeutics and for the development of crystallization-based purification strategies for biological drugs[1–4]. Unfortunately, while the state diagram of protein-solvent systems can provide indications about the phase behaviour of biomacromolecules and their stability in solution, the intimate mechanism underlying protein nucleation and crystal growth have not been fully understood yet and, in many cases, it remains elusive. Protein aggregation is directly related to protein-protein interactions and may be strongly related to protein crystallization behaviour[5]. Since different types of intermolecular forces are involved in protein interactions, such as electrostatic, van der Waals and hydrophobic[6], environmental factors like pH, ionic strength and the presence of additives, can drastically alter the intermolecular interaction profile of proteins in solution[7,8].

Colloidal and conformational stabilities are the two main factors that govern the existence of a protein aggregate in a solution[9,10]. Colloidal stability is a delicate balance between repulsive and attractive forces between proteins, while conformational stability is related to the free energy difference between two molecular states, folded and partially or totally unfolded. Several investigations have found that the melting temperature (T_m) and the osmotic second virial coefficient (B_{22}) can respectively account for the protein conformational[11] and colloidal[12] stability. Both T_m and B_{22} can be experimentally determined by using correspondingly Dynamic Light Scattering (DLS) and Static Light Scattering (SLS) methods[13]. Aggregation induced by thermal unfolding can be monitored by measuring the average hydrodynamic diameter value as a function of temperature, while B_{22} describes the first deviation from ideal behaviour of dilute colloidal solutions[14]. A positive or negative value of B_{22} accounts for a net repulsive or attractive interaction balance, respectively, and could therefore be used to predict stable, crystallizing or precipitating protein solution conditions. According to the experimental crystallization slot of George and

[1]National Research Council of Italy (CNR) - Institute on Membrane Technology (ITM), Via P. Bucci Cubo 17/C, 87036, Rende, CS, Italy. [2]Department of Pharmacy, Health and Nutritional Sciences, University of Calabria, Via P. Bucci Edificio Polifunzionale, 87036, Rende, CS, Italy. [3]EPSRC Centre for Innovative Manufacturing in Continuous Manufacturing and Crystallisation (CMAC), Strathclyde Institute of Pharmacy and Biomedical Sciences, Technology and Innovation Centre, University of Strathclyde, 99 George Street, Glasgow, G1 1RD, UK. [4]Department of Environmental Engineering (DIAm), University of Calabria, Via P. Bucci Cubo 45/A, 87036, Rende, CS, Italy. [5]Seligenda Membrane Technologies S.r.l., Via P. Bucci Cubo 45/A, 87036, Rende, CS, Italy. ✉e-mail: fiore.nicoletta@unical.it; g.diprofio@itm.cnr.it

Wilson[15], a protein solution would be: stable if $B_{22} > -0.8 \times 10^{-4}$ mL mol g^{-2}; crystallizing if $-8.4 \times 10^{-4} < B_{22} < -0.8 \times 10^{-4}$ mL mol g^{-2}; precipitating if $B_{22} < -8.4 \times 10^{-4}$ mL mol g^{-2}. Haas and Drenth[16] postulated a modified crystallization slot, ranging from -0.9×10^{-4} to -0.35×10^{-4} mol mL g^{-2} for protein with a molecular size as high as 140 kDa.

The Z-potential is another indicator of colloidal stability, which measures the magnitude of the electrostatic or charge repulsion/attraction between particles[17]. It is related to the surface charge of the molecules, the adsorbed layer at the interface and the nature and composition of the surrounding environment. A large Z-potential value ($|$Z-potential$|$ >30 mV) can be considered an indicator of colloidal stability. Recently, the use of the interaction parameter k_d has been also suggested as a more high-throughput means to quantify protein-protein interactions[18,19].

In this work, we focus on the relationships between the crystallization propensity of Anti-CD20 in the Na_2SO_4/PEG400 solvent system of different compositions and its relation with some measurable parameters normally used to assess protein conformational and colloidal stability in solution, with the aim to understand the aggregation mechanism of this complex biomacromolecule. Anti-CD20 is a full-length monoclonal antibody (mAb) and it is a therapeutic protein extensively used for treatment of chronic lymphocytic leukaemia and non-Hodgkin's lymphoma, which is commercialized with the brand name of Rituxan or Mabthera. Here, the crystallization of this biomolecule has been observed in a PEG400/Na_2SO_4/Water system near Liquid-Liquid Phase Separation (LLPS) conditions by both vapour diffusion and batch methods. LLPS is a very common phenomena for monoclonal antibodies[20] and is of great interest because the factors which induce the separation of protein solutions into coexisting protein-poor and protein-rich phases are observed to play a central role in protein interaction and crystallization[21]. For several protein/PEG or protein/salt systems both the LLPS and the aggregation propensity of proteins in solution have been shown to coincide with net attractive interactions between molecules[22–24]. Monte Carlo simulations performed on systems that undergo phase separation demonstrated that the interactions between the proteins fall into the short-range regime and are strongly anisotropic[25]. Over time, the anisotropic interactions may be responsible for the evolution of the system from the phase separation to reversible aggregation conditions. In this situation, the LLPS has been speculated to behave as a metastable state towards nucleation[26–28]. In fact, the pathway for the formation of crystal nuclei drastically changes near the metastable low density–high density liquid region coexistence, with the free-energy barrier for crystal nucleation being strongly reduced. In this scenario, the crystal nucleation rate increases by many orders of magnitude over that predicted from classical nucleation theory[28], thus generating a fast and diffuse crystallization that often results in poor-quality crystals.

The results of this investigation allowed us to gain further insights into the mechanism governing the nucleation-crystallization of a complex system such as the Anti-CD20 mAb.

Methods

Anti-CD20 crystallization experiments. Anti-CD20 monoclonal antibody (M_W = 144.5 kDa) was kindly provided by FUJIFILM Diosynth Biotechnologies (Billingham, UK) at high level of purity (>98% HPLC, >95% SDS-PAGE) in 0.025 M sodium citrate buffer at pH 6.5 and 0.154 M sodium chloride solution. Sodium sulphate anhydrous (purity ≥ 99.99%), PEG400, HEPES (purity ≥ 99.5%), TRIS (purity ≥ 99.5%), magnesium chloride hexahydrate (≥99%) and NaOH (purity ≥ 98%), from Sigma-Aldrich (Italy), were used without further purification. MilliQ water was used to prepare all solutions. Anti-CD20 was thawed for 2–3 hours in an 8 °C water bath before buffer exchange with HEPES 0.1 M and concentration by Amicon Ultra centrifugal filter tube (Ultra-4, cut-off 10 kDa, from Merck) at 6,000 g and 8 °C. Finally, the protein solution was diluted to the desired concentration by adding HEPES buffer 0.1 M. NaOH 1 M was used to adjust the pH to the desired value. The Anti-CD20 concentration was determined by measuring protein absorbance at 280 nm by UV-Vis spectrophotometer (extinction coefficient at 280 nm 237,380 M^{-1} cm^{-1}).

Protocols suitable for crystallizing Anti-CD20 reproducibly by vapour diffusion and batch methods were developed in previous works[29,30]. Precipitant solutions were made by dissolving sodium sulphate in the HEPES buffer solution containing PEG400 at the proper concentration. All solutions were filtered through a 0.22 µm Anotop 10 Filter Unit (Whatman). In sitting-drop vapour diffusion crystallization experiments, drops made of 10 µL of precipitant solution and 10 µL of protein solutions were equilibrated in a sealed well against 1 mL of a hypertonic solution of $MgCl_2{\cdot}6H_2O$ 30 wt.% used as reservoir. For the batch crystallization experiments, a volume of protein solutions having a certain initial concentration was mixed with the required volume of precipitant solution to obtain a supersaturated solution of known composition (final volume 1 mL). All crystallization experiments were performed in a vibration-free refrigerator at 20 °C.

Analysis of Anti-CD20 crystals. Crystals obtained from crystallization experiments were observed using an optical microscope (Eclipse LV 100ND, Nikon Instruments, Italy) equipped with a video camera, separated from the mother liquor by using a centrifuge (MiniSpin Plus Eppendorf model IVD) and centrifuge tubes (0.22 µm pore size) and extensively washed. Washing solution was made by 0.1 M HEPES, Na_2SO_4 and PEG400 whose concentration was optimized to avoid crystal solubilisation according to standard procedures. Washed crystals were dissolved by using 0.01 M TRIS-HCl, NaCl 0.15 M, pH 8.0 and analysed by UV spectrophotometry (Lambda EZ201, Perkin Elmer) to estimate protein concentration.

Circular dichroism (CD) measurements. The far-UV CD spectra of Anti-CD20 under initial conditions, after buffer exchange with HEPES 0.1 M, and after mixing with the precipitant solution, were recorded between 200 and 300 nm on a Chirascan CD spectrometer (Applied Photophysics, UK) at 20 °C with a 1 nm bandwidth resolution and current time-per-point of 3 s. 100 µL of mAbs solution were added in a 0.1 mm path-length quartz

cuvette that was placed in the spectrophotometer. Typically, three scans were recorded and baseline spectra were subtracted from each spectrum. Data were processed using Applied Photophysics Chirascan Viewer.

Melting point and Z-potential assessment by DLS. Thermal stability (melting temperature, T_m) of protein solutions was assessed by a Zetasizer Nano ZS instrument (Malvern) equipped with a 4 mW He–Ne laser at 632.8 nm and a detector placed at 173° in agreement to the proprietary NIBS (Non Invasive Back Scatter) technology. Samples were placed in quartz cuvettes (1 cm path-length), heated by a Peltier module and the average hydrodynamic diameter was measured as a function of temperature from 20 °C to 90 °C (temperature increment 0.5 °C). DLS can monitor the thermal denaturation of proteins, which leads to irreversible loss of their structure and function as the protein unfolds. When the average protein size is observed by DLS versus temperature, one will observe a significant increase in the average protein size for temperatures larger than the aggregation point, which is generally assumed as the onset of protein melting point.

For the Z-potential evaluations, the appropriate ratios of protein and buffer solutions were placed in disposable folded capillary cells (DTS 1070, Malvern), and then inserted into the Zetasizer instrument. The Z-potential was measured by laser Doppler micro-electrophoresis. Briefly, an electric field was applied to the protein solution and consequently each molecule moved with a velocity related to its Z-potential. The velocity was measured using a laser interferometric technique (M3-Phase analysis light scattering), which allowed the calculation of the electrophoretic mobility, and from this, the average Z-potential and Z-potential distribution.

Light scattering measurements. Static and dynamic light scattering measurements were performed on protein solution samples with a concentration between 0.1 and 5.0 mg mL^{-1} in different crystallization cocktails. The Anti-CD20 concentration in these solutions was checked by measurement of the absorbance at 280 nm. All the solutions were filtered on a 0.22 µm filter (Anotop 10, Whatman) before measurement. The samples were placed in quartz ultra-low volume cuvettes (ZEN 2112, optical path 3 mm, Malvern) to measure with the Zetasizer Nano ZS instrument (Malvern) in static mode for the second virial coefficient B_{22} or dynamic mode for the protein diffusion coefficient D_m. The light intensity and its time autocorrelation function were measured at 173° scattering angle. All measurements were performed at 20 °C after 2 min of equilibration using automatic time settings.

The Debye plots were generated by using Debye's light scattering equation:

$$\frac{KC}{R_\vartheta} = \frac{1}{M_W} + 2B_{22}C \tag{1}$$

where R_θ is the excess Rayleigh ratio of the protein in a solution with a protein concentration C and M_W is the average molecular weight of the protein. K is the optical constant and is defined as:

$$K = \frac{4\pi^2 n^2 \left(\frac{dn}{dc}\right)^2}{N_A \lambda^4} \tag{2}$$

where n is the solvent refractive index, dn/dc is the refractive index increment, λ is the wavelength of the incident light, and N_A is Avogadro's number. The second virial coefficient B_{22} [mL mol g^{-2}] was obtained from the slope of the linear Debye plot of KC/R_θ versus protein concentration C.

The static light scattering results for B_{22} were accepted if the following criteria were satisfied:

1. Signal to noise ratio >130%, where noise is solvent kcounts;
2. Increasing trend for signal kcounts vs. protein concentration.

The interaction parameter k_d [mL mg^{-1}] and the self-diffusion coefficient D_s [m^2 s^{-1}] were determined by using linear fitting of the protein diffusion coefficient D_m, determined from the dynamic light scattering measurements, plotted against the protein concentration C [mg mL^{-1}], using the following Eq. 3 [31]:

$$D_m = D_s(1 + k_d C) \tag{3}$$

The determined interaction parameter k_d can be interpreted as the difference between a thermodynamic term related to the product $2B_{22}{\cdot}M_W$ and a hydrodynamic term ($\xi_1 + \nu$):

$$k_d = 2B_{22}M_W - (\xi_1 + \upsilon) \tag{4}$$

where ξ_1 is obtained from the virial expansion of the concentration-dependent frictional coefficient and ν is the solvent viscosity.

Solutions that showed turbidity, indicating phase separation and crystallization/precipitation, were not used for measurements.

Protein partitioning measurements. Anti-CD20 repartition between aqueous and PEG phases has been estimated by the method described by Kress *et al.*[32]. A solution (5 mL) containing Na_2SO_4 1.1 M (13.54 wt.%) and PEG400 10.46% V/V (10.22 wt.%) in HEPES 0.1 M at pH 7.4, was added with 200 µL Anti-CD20 solution 10.3 mg mL^{-1} in the same buffer. After vigorous stirring for 5 minutes by vortex, the mixture has been left to rest overnight. The concentration of the protein in the two separated phases visually detected and in presence of excess of precipitated protein, was then assessed by UV spectrophotometry at 280 nm.

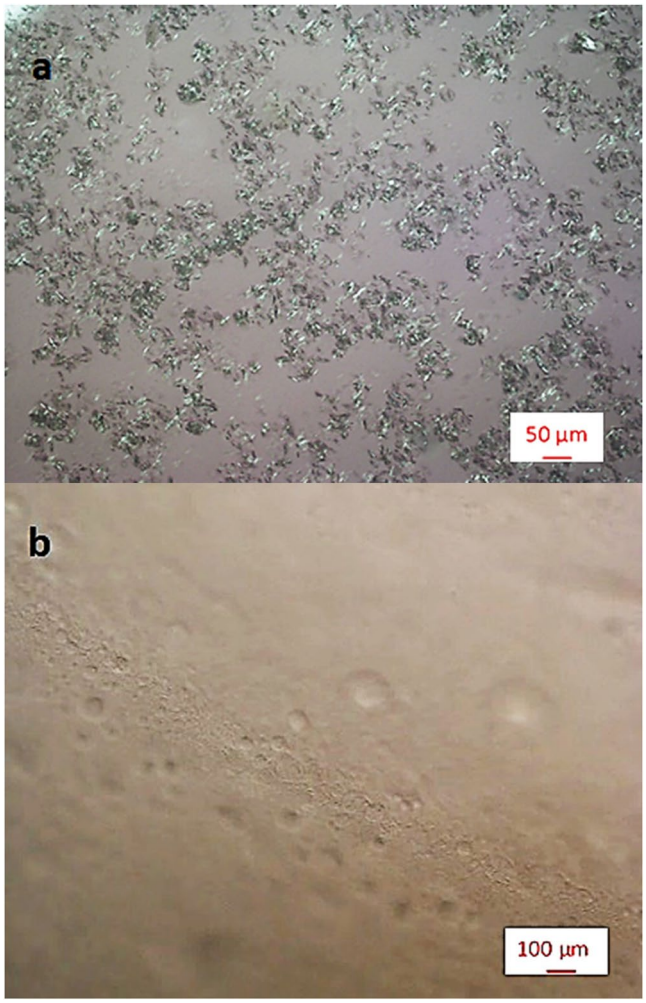

Figure 1. Optical microscope image of (**a**) crystalline Anti-CD20 monoclonal antibody obtained by sitting-drop vapour diffusion method and (**b**) Liquid-Liquid Phase Separation occurring just prior to crystallization. Solution composition near the crystallization point: 20 mg mL^{-1} Anti-CD20, 9.6% V/V PEG400, 0.86 M Na_2SO_4 in 0.1 M HEPES buffer at pH 7.4.

Results and discussion

Anti-CD20 crystallization and crystals analysis. Anti-CD20 needle-like crystals (Fig. 1a) appear by sitting-drop vapour diffusion crystallization experiments using a 1:1 volume ratio mixture of a 20 mg mL^{-1} protein solution in HEPES buffer 0.1 M at pH 7.4 and a precipitant solution containing PEG400 9.6% V/V (9.4 wt.%) and Na_2SO_4 0.86 M (10.6 wt.%) in the same buffer. Therefore, the initial composition of the drop in the vapour diffusion tests corresponds to 10 mg mL^{-1} Anti-CD20, 4.8% V/V PEG400, 0.43 M Na_2SO_4 in 0.1 M HEPES at a pH of 7.4 and is shown by the diamond symbol in Fig. 2. The crystals appear when the volume of the drop is nearly halved with respect to its initial volume. Accordingly, the final concentration of the components in the crystallizing drops is almost doubled that compared to the initial compositions and it is indicated by the pentagon symbol of Fig. 2.

Batch experiments in similar solution composition as that for crystallizing drops in the vapour diffusion tests (the term "similar" is used since the exact composition of the droplet at the crystallization point under equilibration with the reservoir solution it is not known), give also rise to crystals. Therefore, the batch working point coincides with the final conditions of the vapour diffusion tests. Since the batch composition does not vary with time, due to the absence of solvent evaporation, it is evident that the solution is supersaturated immediately upon mixing the protein and precipitant solutions, which results in a shorter crystallization time (12–18 h) than that observed with the sitting drop set up (36–48 h).

The UV-Vis and DLS measurements of washed and dissolved crystals confirmed their protein nature. All the analysed samples showed an intense band centred at 280 nm, typical of proteins absorption. The size distribution analysis by intensities show a single peak near 12 nm, size expected for Anti-CD20 protein molecules[33], and a polidispersity <20%, indicating a high monodispersity.

Liquid-Liquid Phase Separation (LLPS) is observed nearly simultaneously with, or just prior to, crystals appearance (Fig. 1b) both in the vapour diffusion and batch crystallization trials. In accordance with other mAbs

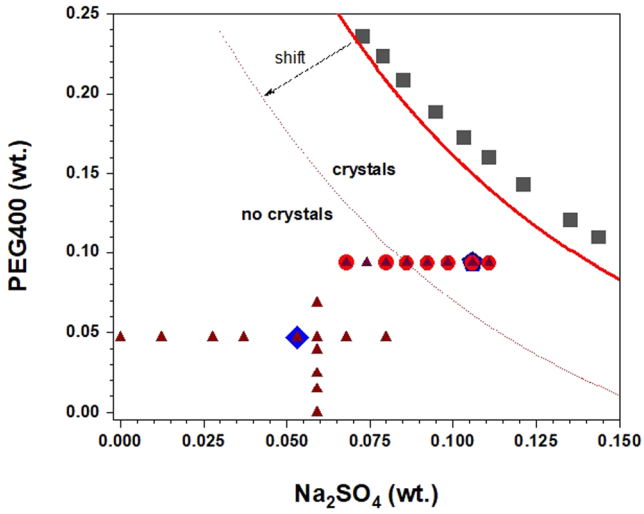

Figure 2. Experimental state diagram for the system PEG400/Na_2SO_4/H_2O at 293.15 K. Squares: experimental liquid-liquid equilibrium data from ref. [36]; solid line: asymptotic fitting to tie-lines data from ref. [36]; circles: batch crystallization conditions reported in Table 1; diamond: starting solution composition in vapour diffusion tests; pentagon: ending solution composition in vapour diffusion tests; triangles: measurements points for B_{22}; dotted line: shifted binodal curve in the presence of Anti-CD20 protein. Solution composition is expressed as weight fraction (wt.).

Na_2SO_4 [M]	Crystallization behaviour	Size [μm]	Shape
0.90	Opalescent solution, Instantaneous Precipitation	<1	gel beads
0.86	Precipitation/ crystallization in 12–24 hours	<10	spherulite
0.80	Crystallization in 18–36 hours	<20	needle
0.75	Crystallization in 24–48 hours	20	needle
0.70	Crystallization in 24–48 hours	20	needle
0.65	Clear and stable solution	—	—
0.55	Clear and stable solution	—	—

Table 1. Mixed solution compositions, crystallization behaviour and crystal appearance in the batch crystallization experiments using an Anti-CD20 concentration 20 mg mL^{-1}, 9.6% V/V PEG400 and 100 mM HEPES buffer at a pH of 7.4.

crystallization studies reported in literature[20,34,35], also here it appears that the LLPS is a key step in Anti-CD20 crystallization.

Crystallization trials, aiming to investigate the effect of precipitant solution composition on crystal quality have been also performed by changing the starting amount of PEG400 and/or Na_2SO_4. As expected, the shape, size and density of crystals (number of crystals per unity of volume) are found to depend on buffer composition, initial protein concentration and crystallization method (sitting drop or batch). All these parameters have also shown to impact on the time required to observe crystallization. Table 1 shows the effect of the salt concentration on the crystal appearance for the batch experiments. Needles with a size of nearly 20 μm are the largest crystals obtained for starting solutions with 9.6% V/V PEG400 and 0.70–0.80 M Na_2SO_4. A further increase in Na_2SO_4 concentration produces a faster and massive protein nucleation in less than 24 hours and crystals of smaller size (<10 μm) with a spherulite-like morphology appear. For a Na_2SO_4 concentration of 0.9 M or larger, upon mixing the solution immediately turns opalescent due to a macroscopic LLPS in the solution, followed by a rapid protein precipitation as gel beads. On the opposite side, for the lowest amount of salt ($Na_2SO_4 \leq 0.65$ M), the solutions remain clear for months.

Figure 2 shows the experimental state diagram for the system PEG400/Na_2SO_4/H_2O at 20 °C indicating the loci of the points related to solution compositions used for batch crystallization tests (circles) of Table 1, together with the starting (diamond) and ending (pentagon) solution compositions in vapour diffusion tests. The figure also displays the experimental liquid-liquid equilibrium data and the asymptotic fitting to tie-lines data reported in ref. [36].

Buffer	T_m [°C]	Z-potential [mV]	Z_{avg} [nm]	*PDI*
HEPES[a]	68.5 ± 0.5	2.2 ± 0.7	11.3 ± 0.1	0.10 ± 0.03
Sodium Citrate[b]	68.5 ± 0.5	1.1 ± 0.2	12.2 ± 0.2	0.04 ± 0.03
TRIS[c]	69.0 ± 0.5	7.0 ± 1.0	12.63 ± 0.03	0.12 ± 0.01
Sodium Citrate/HEPES (1:1 Vol)	69.0 ± 0.5	0.19 ± 0.04	13.6 ± 0.2	0.15 ± 0.03
TRIS/HEPES (1:1 Vol)	69.0 ± 0.5	2.3 ± 1.0	12.13 ± 0.03	0.06 ± 0.01

Table 2. Results for melting point (T_m), Z-potential, intensity weighted mean hydrodynamic size (Z_{avg}), and polydispersion index (*PDI*) from formulation buffers screening (Anti-CD20 concentration 20 mg mL^{-1}; [a]HEPES buffer 0.1 M, pH 7.4; [b]Sodium Citrate buffer 0.035 M, NaCl 0.15 M, pH 6.5; [c]TRIS buffer 0.1 M, NaCl 0.15 M, pH 8.0).

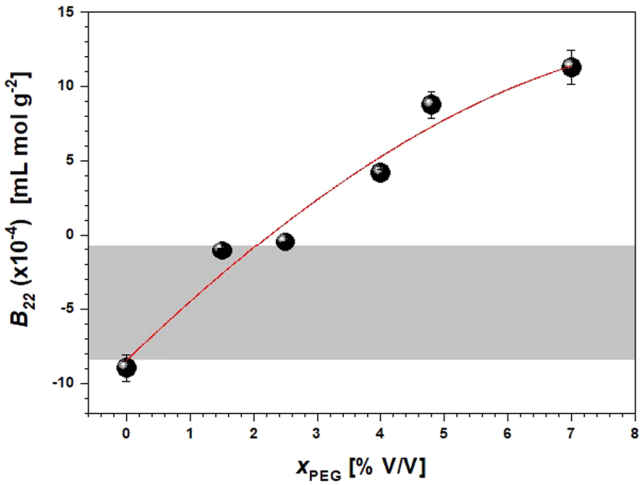

Figure 3. The second virial coefficient B_{22} as a function of PEG400 volume concentration (X_{PEG}) at a constant amount of 0.48 M Na_2SO_4 in HEPES 0.1 M buffer at pH 7.4. The crystallization slot according to ref. [15] is shown in light grey. The solid line is a guide for the eyes.

SLS/DLS analysis. The protein stability in several formulation buffers but without precipitating agents (PEG and/or Na_2SO_4), was assessed by DLS, resulting in values for the melting point and Z-potential. As reported in Table 2, all samples are characterized by similar values of melting point, Z-potential, intensity weighted mean hydrodynamic size Z_{avg} and polydispersion index (PDI). All protein samples show low positive values of the Z-potential, indicating a reduced colloidal stability in the various buffer solutions. Nevertheless, due to a PDI close to 0.1, all samples can be considered monodisperse. No significant variations were observed in the sample aggregation state over time in the range 12–24 h. Accordingly, further SLS/DLS studies were performed by using HEPES as formulation buffer since it was used in previously published crystallization protocols[30].

To study the aggregation propensity on Anti-CD20 in HEPES buffer and in various PEG400/Na_2SO_4 solution compositions, SLS and DLS measurements were performed on freshly prepared samples allowing to determine the second virial coefficient B_{22} and the interaction parameter k_d, respectively. The value of KC/R_θ was determined as a function of protein concentration, ranging from 0.1 to 5.0 mg mL^{-1}, by SLS. Results were used to evidence the combined effect of PEG400 and Na_2SO_4 on Anti-CD20 colloidal stability and to compare these results with effective crystallization outcomes.

It is well known that the addition of salts can cause an electrostatic double layer around the protein surface charges, which involves a shielding and the reduction of repulsive interactions among macromolecules. Furthermore, the salts ions compete with the protein for water molecules and dehydrate the protein (salting out effect[37]), thus inducing aggregation by mainly electrostatic and hydrophobic interactions. On the other side, polymers such as PEG bring the protein molecules together due to preferential interactions and osmotic potential[38]. Accordingly, the addition of PEG400 or Na_2SO_4 to Anti-CD20 solutions is expected to increase the aggregation aptitude of the protein and facilitate crystallization, as experimentally observed, e.g., in batch crystallization tests of Table 1. However, as displayed in Figs. 3 and 4, for increasing amounts of PEG or Na_2SO_4 when both components are present in solution, the B_{22} moves towards larger positive values, theoretically indicative of an increase in repulsive interactions. Furthermore, effective Anti-CD20 crystallization conditions seems to lay outside the crystallization slot defined by George and Wilson[15].

Both Figs. 3 and 4 show that the combined effect of PEG400 and Na_2SO_4 has a noteworthy consequence on the B_{22} value. For instance, for the solution containing 0.48 M Na_2SO_4 in absence of PEG, a B_{22} value of -8.9×10^{-4} mL mol g^{-2} is obtained (Fig. 3) that is slightly below the crystallization slot. However, when adding PEG400 to the solution while keeping salt concentration constant, B_{22} steeply increases through the crystallization window towards positive values already at 3% V/V of PEG.

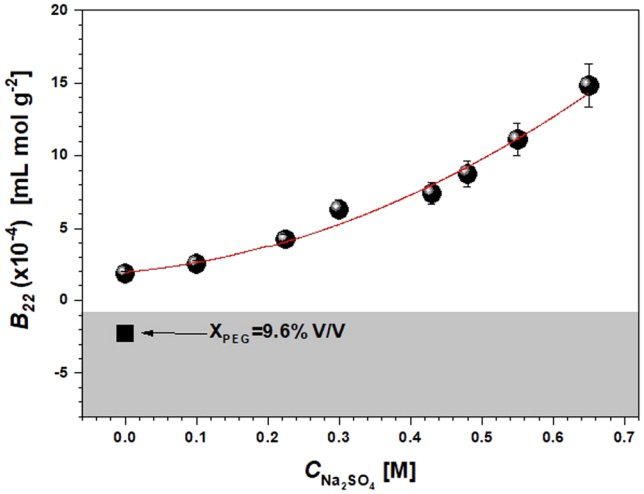

Figure 4. The second virial coefficient B_{22} as a function of Na_2SO_4 molar concentration ($C_{Na_2SO_4}$) for constant PEG400 volume concentration $x_{PEG} = 4.8\%$ V/V in HEPES 0.1 M, pH 7.4. The solid square symbol indicate B_{22} for a PEG400 concentration of 9.6% V/V without NaCl. The crystallization slot according to ref. [15] is shown in light grey. The solid line is a guide for the eyes.

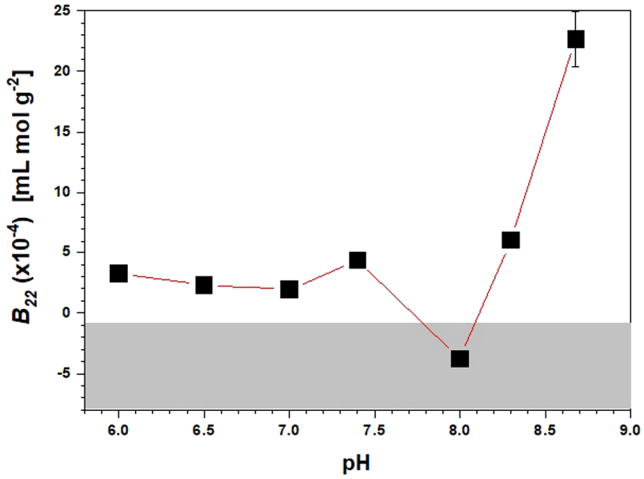

Figure 5. pH dependence of B_{22} for Anti-CD20 solutions with 4.8% V/V PEG400, 0.43 M Na_2SO_4, and HEPES 0.1 M. The crystallization slot according to ref. [15] is shown in light grey. The solid line is a guide for the eyes.

The B_{22} value for the solution containing 4.8% V/V PEG400 in absence of Na_2SO_4 is found to be slightly positive ($B_{22} = 1.8 \times 10^{-4}$ mL mol g^{-2}), becoming negative ($B_{22} = -2.26 \times 10^{-4}$ mL mol g^{-2}) when increasing PEG400 up to $x_{PEG} = 9.6\%$ V/V without Na_2SO_4 (Fig. 4). This is in agreement with results from Ahamed *et al.*[39], who observed the reduction in solubility for a mAb for increasing amounts of PEG400. Nevertheless, when salt is added to the solution ($C_{Na_2SO_4} \neq 0$), the virial coefficient rises steeply with the salt concentration. These observations clearly indicate that a different effect on the measured value of B_{22} takes place when PEG400 and Na_2SO_4 are combined together, compared to the cases when they are used separately.

Figure 5 shows the effect of pH on B_{22} for solutions containing PEG400 and Na_2SO_4. At pH $\leq$ 7.6, the B_{22} remains roughly unchanged ($2.28 < B_{22} < 4.35 \times 10^{-4}$ mL mol g^{-2}). These positive values are indicative of repulsive protein–protein interactions. At a pH of 8.0, the B_{22} becomes negative (-3.78×10^{-4} mL mol g^{-2}) and then sharply increases approaching pH = 8.8, which is the theoretical isoelectric point of Anti-CD20 (pI = 8.8)[40].

It is expected that protein-protein interactions at a pH substantially below or above the pI are repulsive, while at a pH near the pI they become increasingly attractive due to the overall neutral charge of the protein. However, the actual pI of biomolecules might change with ionic environment, and generally decreases at different extents with increasing ionic strength, depending on the nature of the counter-ion[41,42]. It is speculated that in the presence of 0.43 M Na_2SO_4, PEG400 4.8% V/V and 0.1 M HEPES, the effective pI of Anti-CD20 is around 8.0. Therefore, for pH < 8.0, the slightly positive B_{22} values are most likely due to partially screening of the net positive electrostatic charges of protein molecules, while at pH = 8.0, where the net charge becomes null, the net resultant

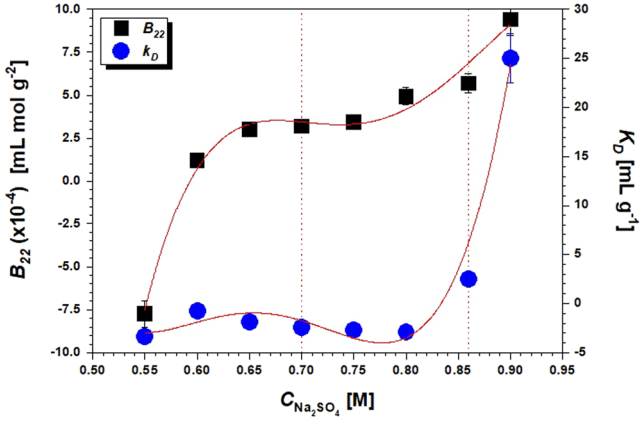

Figure 6. B_{22} and k_d as a function of Na_2SO_4 concentration for $x_{PEG} = 9.6\%$ V/V and pH 7.4. The experimental crystallization window boundaries (see Table 1) are highlighted by the dotted vertical lines. The solid lines are guides for the eyes.

interaction is attractive. Above pH 8.0, the B_{22} values become strongly positive, indicating an ineffective screening of the net negative charge of molecules that results in repulsive interactions.

Since solution composition strongly affects protein stability and aggregate formation, the Na_2SO_4 concentration dependence of B_{22} was also investigated at higher PEG and salt concentrations, thus reaching compositions close to that used for crystallization experiments of Table 1. Correlation between B_{22} and Na_2SO_4 concentration is shown in Fig. 6, confirming an increasing trend similar to that observed for lower concentrations of Fig. 4 also in this case.

In Fig. 6 it can be seen that the B_{22} values are negative (attractive interactions) for Na_2SO_4 concentration ranging from 0.55 to 0.6 M, corresponding to conditions in Table 1 that give clear solutions, while they become positive (repulsive interactions) when the salt concentration is further increased up to 0.9 M. Within the crystallization window, i.e. Na_2SO_4 ranging from 0.7 to 0.86 M, the B_{22} values range from 3.2×10^{-4} to 5.7×10^{-4} mL mol g^{-2}. This would be theoretically indicative of the absence of attractive colloidal interactions, in opposition to the results of the crystallization tests.

DLS measurements were then used to determine the interaction parameter k_d. Generally, k_d and B_{22} display a qualitatively similar trend[43]. However, looking at the behaviour of the interaction parameter k_d in the crystallization range of Fig. 6, negative values are found for $x_{PEG} = 9.6\%$ V/V and sodium sulfate 0.55–0.80 M, with slight variations within this range. Generally, a negative value of k_d implies the existence of attractive interactions between protein molecules which slow down their diffusion in solution[44]. Since under the same conditions B_{22} is negative only for $Na_2SO_4 < 0.6$ M and then increases up to $1.23–4.94 \times 10^{-4}$ mL mol g^{-2} in the salt concentration range 0.6–0.8 M, this indicates that underlying attractive interactions do not have a thermodynamic origin of colloidal nature. We can instead conclude that these interactions result from the hydrodynamic effects included in the second term of Eq. 4 [45], most likely due to the increase of solution viscosity (from 1.949 cP in the PEG400 9.6% V/V solution without salt to 2.546 cP for PEG400 9.6% V/V and Na_2SO_4 0.9 M). Nevertheless, as for B_{22}, the increasing trend of k_d with the concentration of Na_2SO_4 does not comply with the outcome of crystallization tests. In fact, the increase of Na_2SO_4 concentration at values larger than 0.8 M gives rise to a theoretically significant increase of the strength in protein-protein repulsions, according to measured B_{22}, values that are exacerbated near the LLPS region, while crystals are obtained in such conditions.

From the observed trends of B_{22} and k_d, we can conclude that neither PEG400 nor Na_2SO_4 act as conventional precipitants for Anti-CD20 when used together since, on average, Anti-CD20 monomers apparently repel one another. Nevertheless, this evidently contradicts the experimental outcomes of crystallization tests. The B_{22} values are negative (attractive interactions) in formulations that remain clears for months, whereas positive values, corresponding theoretically to net repulsive interactions, have been found for the best-performing crystallizing or either in precipitating conditions.

Considering the possible variations in their structure, morphology, size, surface charge and chemistry, proteins definitely represent one of the most complex colloidal systems. Several mechanisms could lead to molecular aggregation in solution and it is not excluded that more than one is active in certain conditions. Therefore, it is not surprising that B_{22} does not always reflect the aggregation propensities of proteins in a given formulation. Several papers can be found in literature reporting discrepancies between B_{22} and the susceptibility to aggregation of structurally complex macromolecules, like monoclonal antibodies[43,46]. Crystallizing conditions with positive B_{22} values were found for a mAb molecule with PEG400 by Rakel *et al.*[35]. This behaviour has been ascribed to the ambivalent character of PEG400 as precipitant and stabilizer[35,47]. Kress *et al.*[48] observed the decrease in B_{22} of an immunoglobulin G (IgG) when increasing PEG2000, sodium citrate and NaCl solute concentration, if used separately. However, the mutual effect of PEG and salts resulted in a shift of B_{22} from negative to almost neutral values when increasing the NaCl amount if used together[32]. This behaviour has been explained by a solubilizing effect of IgG in PEG2000 induced by NaCl until a maximum salt concentration. Exceeding this NaCl value, a solid-liquid equilibrium, due to protein precipitation, was observed. Herhut *et al.*[49] measured B_{22} for an IgG and

D-xylose ketol-isomerase (172.3 kDa) with ammonium sulphate and PEG12000 and PEG2000, respectively. They observed the initial decrease in B_{22} at low PEG concentration, passing through a minimum and then increasing for higher PEG concentrations. This behaviour has been explained by considering that polymer-induced interactions are not solely attractive over the whole polymer concentration range. Normally, polymer chains behave as coil-inducing attractive interactions between protein molecules at low-polymer concentration. However, at high polymer concentration, flexible coils elongate and enter space between the proteins, so that repulsive interaction are induced[50,51]. This explanation has been also confirmed by molecular dynamic simulations at different polymer concentration by Cao *et al.*[52].

It has also been reported that conformational stability plays an important role in aggregation propensity, with partially unfolded conformational intermediates being responsible of aggregate formation[9]. In this case, the B_{22} value is unlikely to correlate with long-term aggregation because the structurally perturbed state susceptible to aggregation could be present in a small fraction of molecules compared to the native species[46]. Aiming to shed light on this aspect, protein structure in solution has been investigated in the present work.

Protein structure in the solution. On the bases of the results above, measurements of melting temperature (T_m) for several solution compositions, including $x_{PEG} = 9.6\%$ V/V and Na_2SO_4 at various concentrations, and Circular Dichroism (CD) in selected formulations, were performed aiming to investigate conformational stability of Anti-CD20. Results show slight variations of the melting points: $T_m = 67.0$, 67.5, 68.0, and 67.5 ± 0.5 °C for 0.55, 0.65, 0.75, and 0.85 M Na_2SO_4, respectively, confirming that samples, on average, share similar water activity according to the Wyman-Tanford equation[53].

Circular Dichroism spectra of protein solutions give information on the secondary structure, folding properties and conformational variation of the proteins[54]. The general shape of the spectrum of the provided Anti-CD20 solution (Fig. 7, initial conditions) corresponds to the expected spectra for Anti-CD20, as it shows the typical shape of a monoclonal antibody with a minimum around 220 nm and a maximum around 200 nm, suggesting a secondary structure dominated by β-sheet motif[54]. Moreover, these spectra correspond to the CD-spectra of Anti-CD20 from literature[55].

These spectra show that the protein is well folded in every buffer, i.e. the 3D-structure of Anti-CD20 is preserved under all conditions tested. However, the maximum ellipticity (around 200 nm) is much higher after the buffer exchange (red and green spectra in Fig. 7) than in the initial buffer (blue spectrum), while ellipticity at other wavelengths is similar in every buffer. This suggest that the β-sheet motifs of the mAbs are favoured and stabilized in the new HEPES buffer, in comparison to the initial buffer (0.025 M sodium citrate at pH 6.5 and 0.154 M sodium chloride).

There is no significant difference in the far-UV CD spectra of Anti-CD20 before (Fig. 7, red spectrum) and after the addition of Na_2SO_4 and PEG400 as precipitating agents (Fig. 7, green spectrum), although a small shift can be observed, showing slight changes in the 3D-structure. The main difference between these spectra is around 230 nm, but the peak is quite narrow and this wavelength does not correspond to any typical feature of protein structure. Therefore, the difference could be due to Na_2SO_4 or PEG400 absorbance or might be due to some noise. It should be noted, though, that the crystallization solution had to be diluted in order to enable CD measurement. Since the spectra before and after the concentrated crystallization conditions are the same, it can be concluded that the concentrated crystallization conditions do not alter the secondary structure of the protein.

The Anti-CD20 secondary structure is preserved after centrifugation and addition of the precipitating agents, as the typical β-sheet motif appears clearly in the spectrum (Fig. 7, green spectrum). Moreover, this secondary structure seems to be favoured in the crystallization buffer, meaning that the 0.1 M HEPES at pH 7.4 buffer promotes Anti-CD20 crystallization by making the secondary structure less flexible. However, no clear effect of Na_2SO_4 or PEG400 on the Anti-CD20 secondary structure is visible.

The role of LLPS in Anti-CD20 crystallization. A complementary scenario can be envisioned when considering LLPS phenomena observed concurrently or just prior of Anti-CD20 crystallization.

For simple protein/salts or protein/PEG systems, it has been often reported that the B_{22} value can be used as a predictor of LLPS formation, since weak protein-protein interactions, in general hydrophobic in nature or even limited to specific protein chain sequences, are at the origin of the phase separation[22]. However, our results seem to suggest that the LLPS (observed just prior or at the point of protein crystallization) proceeds without attractive interactions, since B_{22} is negative when a mono-phase system is observed while it turns positive as approaching the liquid phase separation and crystals formation.

It could be argued that, in our complex mixture that includes several components, the B_{22} values fail to predict the experimentally observed crystallization since it may be rather related to the polymer-salt phase separation that can occur without protein. The liquid-liquid separation of PEG/salt aqueous solutions is a fundamentally different phenomenon from PEG/protein or salt/protein LLPS, which originate by a combination of hydrophobic and salting-out effects, whose origin has been attributed to the different conformations of the PEG polymer chain in aqueous solution and their respective hydration[24]. Actually, the PEG400/Na_2SO_4/H_2O system is known to generate aqueous two-phase systems[36,56,57]. Figure 2 displays the experimental binodal curve together with liquid-liquid equilibrium tie-lines data for PEG400 + Na_2SO_4 + H_2O system at 293.15 K (the temperature used in this work for crystallization experiments and SLS/DLS measurements)[36].

Concerning the specific system studied in this work, for a protein concentration of 20 mg mL^{-1} and a PEG400 content of 9.6% V/V (9.4 wt.%), phase separation is observed for salt concentrations $\geq$0.7 M ($\geq$8.6 wt.%) (crystallization points of Table 1 and dark circles in Fig. 2), that is quite far from both binodal curve. However, the presence of the protein can reduce the amount of salt needed for observing LLPS in the multicomponent PEG/salt/protein/water systems[24]. Consequently, it cannot be excluded that the effective binodal curve for PEG/salt

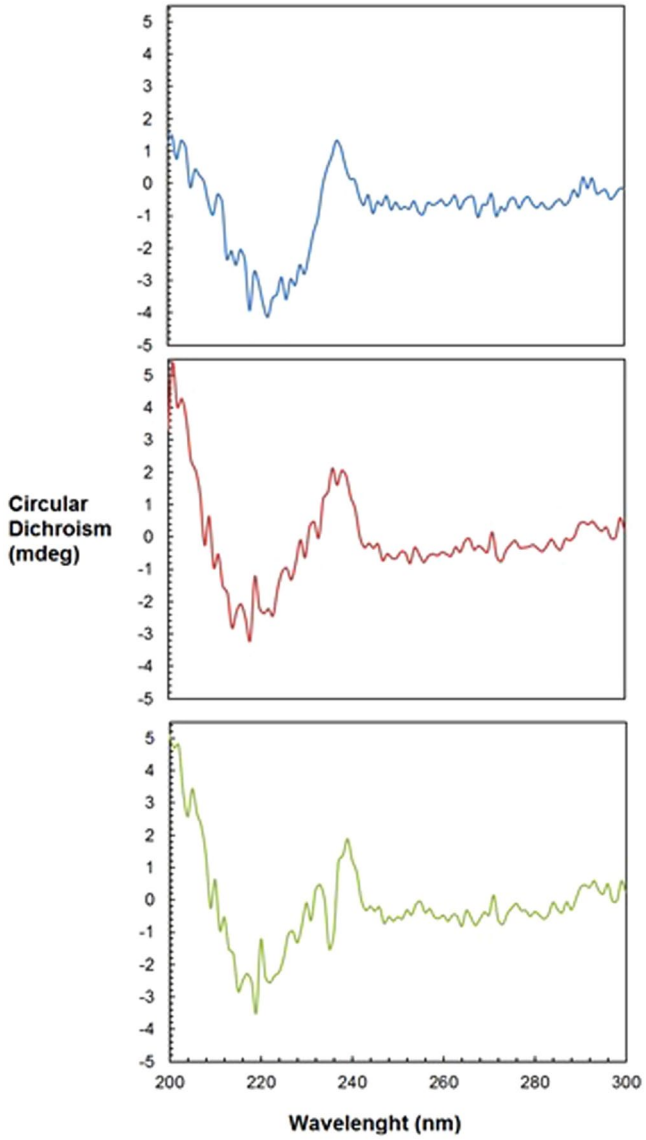

Figure 7. Far-UV CD spectra of Anti-CD20 under initial conditions (blue spectra, mAbs concentration 9.95 mg mL^{-1} in stock buffer), after buffer exchange with HEPES 0.1 M pH 7.4 (red spectra, mAb concentration 100 mg/mL in crystallization buffer) and after salt mixing (green spectra, mAb concentration 60 mg/mL in crystallization buffer and with 0.9 M Na_2SO_4 and 9.6% V/V PEG400).

LLPS is shifted down to conditions corresponding to the observed crystallization points, in a region of poor solvent conditions, thus giving a decisive contribute to the aggregation/crystallization behaviour of Anti-CD20.

Figure 8 shows the change in turbidity (Residual Intensity) for a solution containing PEG400 9.6% V/V in 0.1 M HEPES at pH 7.4 and increasing amount of Na_2SO_4. It is observed that, in absence of the Anti-CD20, a sharp increase in solution turbidity is observed for salt concentration >1.1 M (>13.5 wt.%), that is very close to the binodal curve reported in Fig. 2 for the PEG400 + Na_2SO_4 + H_2O system. Since in the presence of the protein turbidity due to liquid-liquid phase separation is observed for salt concentration larger than 0.65 M for the same amount of PEG, this proves the shift of LLPS region towards lower values of PEG and/or salt components when Anti-CD20 is present in the solution.

It is worth noting that B_{22} measurements reported in Figs. 3 and 4 have been made in a region of Fig. 2 where LLPS (and crystallization) are not observed (triangles in the left-bottom side of the figure). This means that PEG/salt LLPS is not responsible for the contradiction between B_{22} measurements and the outcome of crystallization trials, but rather the intrinsic complexity of the system makes ineffective the measurement of the osmotic coefficient as predictor parameter for crystallization conditions.

According to our results, it can be proposed that PEG/salt LLPS causes the preferential partitioning of proteins in a specific salt- or PEG-rich phase. In general, the affinity of the protein for the polymeric component, due to hydrophobic interactions, favours the protein concentration in the PEG-rich phase[39,58]. This is particularly true for PEG400, which is a low molecular weight polymer and, therefore, its size exclusion effects on proteins

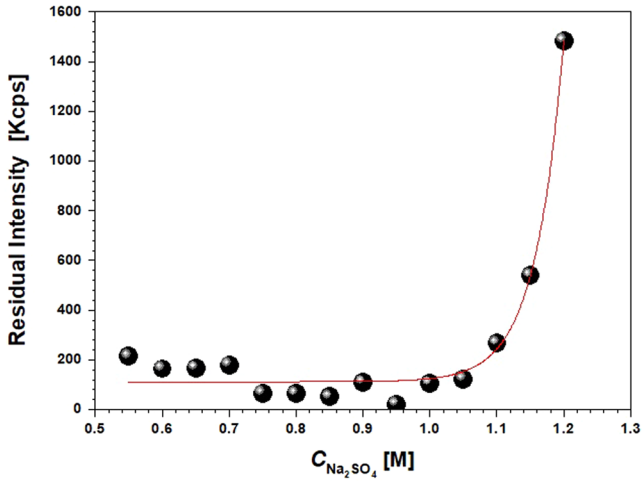

Figure 8. Residual Intensity measured by DLS at a fixed scattering angle of 173° for a solution containing PEG400 9.6% V/V in 0.1 M HEPES at pH 7.4 and increasing amount of Na_2SO_4, without protein. The solid line is a guide for the eyes.

are negligible while the dominant factor in the separation is the salting out effect, produced by the sulphate, which causes the protein to migrate to the PEG-rich phase. This effect has been confirmed by measurements on Anti-CD20 partitioning, showing that more than 90% of the solution protein migrates in the PEG-rich phase in the protein-saturated Anti-CD20/PEG400/Na_2SO_4/HEPES/H_2O mixture displaying phase separation.

In such conditions, during LLPS, the Anti-CD20 concentrates in droplets that have a density higher than the average density of the whole system thus generating a much higher supersaturation. When the protein concentration exceeds a certain limit, precipitation at the interface is observed and the precipitated protein is in equilibrium with the protein solubilised in each phase[59]. For particularly high Anti-CD20 ($>30\,mg\,mL^{-1}$ or PEG/salt concentrations (i.e. PEG400 9.4 wt.% and $Na_2SO_4 \geq 11.1$ wt.%), these droplets look not to be as liquid, having instead the appearance of gel beads. Protein crystals initiate to nucleate and grow near the boundary of the most protein-concentrated droplets (Fig. 9), suggesting that the droplet interface participates in the nucleation mechanism[60]. Subsequently, crystals appear within the bulk of the droplets, which spread and coalesce. Many droplets disappear as crystal growth proceeds, suggesting that a continuous exchange exists between the droplets and the surrounding medium.

Experimental observations clearly indicate the important role of LLPS and of the established liquid-liquid interface in inducing Anti-CD20 nucleation, even in absence of underlying net attractive interaction between proteins. It is likely that only protein molecules distributed at the droplets interface experience local short-range attractive interactions, due to the specific environment and/or conformational changes, not influencing the mean B_{22} value.

On the basis of our results, it could be hypothesized that the occurrence of LLPS in crystallizing/precipitating conditions results in a partially unfolded protein conformation or in a suitable variation of the secondary/tertiary protein structure, as already reported in other cases[27], which can be the species prone to aggregation through a mechanism driven by conformational change rather than colloidal interactions, and which initiate the crystallization process. The conformationally perturbed protein can be a minor species, which does not alter significantly the predominant native protein population and only slightly contributes to the net protein–protein interaction value measured by B_{22}. Indeed, SLS is an integrative technique, which provides only averaged information. If only specific short-range orientations or aggregation-prone structurally perturbed states can lead to attractive interactions, the remaining repulsive orientations or species dominate in the light scattering results.

It is important to note that the measurements of intermolecular interactions were performed under low protein concentration conditions ($\leq 5\,mg\,mL^{-1}$), while crystallization trials are generally conducted at higher concentrations ($\geq 20\,mg\,mL^{-1}$). Both B_{22} and k_d determined in dilute conditions are considered generally predictive of high-concentration behaviour. Nonetheless, high concentration behaviour can differ from that observed at dilute conditions, since "crowding" effects, i.e. the increased chance of molecular interactions, deviation from thermodynamic ideality and higher-order interactions (e.g., B_{23}, B_{222}) can significantly alter the net inter-protein interactions[55,61]. The contribution of these effects to aggregation propensity of structurally complex molecules such as mAbs, cannot be ruled out.

Conclusions

This work reports about the study of Anti-CD20 (a full-length monoclonal antibody) crystallization in a PEG400/Na_2SO_4/H_2O system occurring near LLPS conditions. The DLS/SLS studies performed to verify the correlation between the second virial coefficient and molecular diffusivity with the protein aggregation propensity in selected solution formulations indicates that B_{22} remains strongly positive (a sign of repulsive protein–protein interactions) for all conditions that have been instead ascertained to be effective for protein crystallization or precipitation. For all the crystallizing or precipitating conditions, the occurrence of LLPS phenomena was observed concomitantly or near the nucleation stage. Even in this case, the second virial coefficient does not correlate with

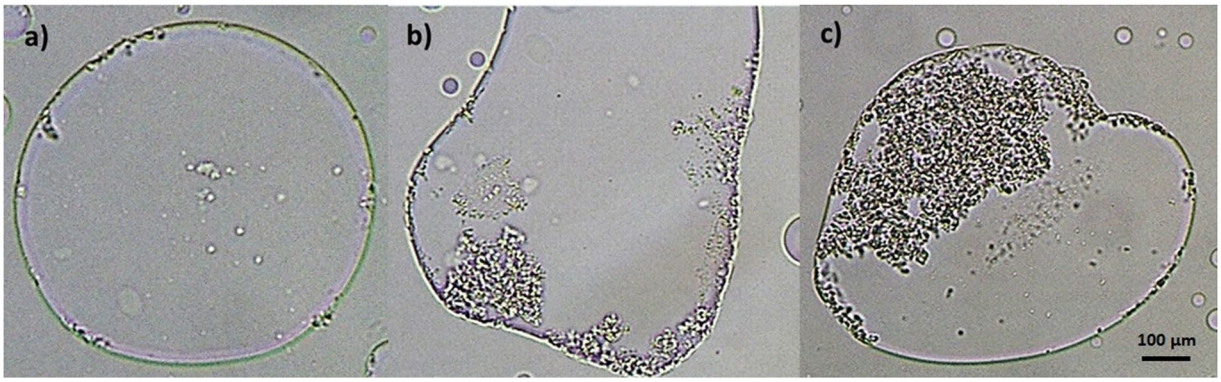

Figure 9. Protein crystals nucleating near the boundary of the droplets (**a**) and crystals forming within the bulk of the droplets (**b**,**c**).

the LLPS, which mainly proceeds without attractive protein interactions, and that may be rather related to the polymer-salt phase separation that can occur without protein. The possible interplay between the salt/PEG and protein/PEG phase separation, which can affect the solvent quality, can have a role on the polymer mediated forces and finally on Anti-CD20 aggregation. On the basis of these measurements and the experimental observation of LLPS near the nucleation zone, it is proposed that a minor population of protein molecules can experience valuable protein-protein attractive interactions, induced by local environmental factor, slightly affecting the averaged B_{22} parameter.

Received: 28 October 2019; Accepted: 30 April 2020;
Published online: 01 June 2020

References

1. Cromwell, M. E. M., Hilario, E. & Jacobson, F. Protein aggregation and bioprocessing. *AAPS J.* **8**, E572–E579 (2006).
2. Thömmes, J. & Etzel, M. Alternatives to chromatographic separations. *Biotechnol. Prog.* **23**, 42–45 (2007).
3. Low, D., O'Leary, R. & Pujar, N. S. Future of antibody purification. *J. Chromatogr. B* **848**, 48–63 (2007).
4. Dos Santos, R., Carvalho, A. L. & Roque, A. C. Renaissance of protein crystallization and precipitation in biopharmaceuticals purification. *Biotechnol. Adv.* **35**, 41–50 (2017).
5. Yin Chan, H. & Lubchenko, V. A mechanism for reversible mesoscopic aggregation in liquid solutions. *Nat. Commun.* **10**, 2381, https://doi.org/10.1038/s41467-019-10270-5 (2019).
6. Chari, R., Jerath, K., Badkar, A. V. & Kalonia, D. S. Long- and short-range electrostatic interactions affect the rheology of highly concentrated antibody solutions. *Pharm. Res.* **26**, 2607–2618 (2009).
7. Alford, J. R. *et al*. High concentration formulations of recombinant human Interleukin-1 receptor antagonist: I. Physical characterization. *J. Pharm. Sci.* **97**, 3035–3050 (2008).
8. Mahler, H. C., Friess, W., Grauschopf, U. & Kiese, S. Protein aggregation: pathways, induction factors and analysis. *J. Pharm. Sci.* **98**, 2909–2934 (2009).
9. Chi, E. Y., Krishnan, S., Randolph, T. W. & Carpenter, J. F. Physical stability of proteins in aqueous solution: mechanism and driving forces in nonnative protein aggregation. *Pharm. Res.* **20**, 1325–1336 (2003).
10. Wang, J. *et al*. Magic number colloidal clusters as minimum free energy structures. *Nat. Commun.* **9**, 5259, https://doi.org/10.1038/s41467-018-07600-4 (2018).
11. Yadav, S., Shire, S. J. & Kalonia, D. S. Factors affecting the viscosity in high concentration solutions of different monoclonal antibodies. *J. Pharm. Sci.* **99**, 4812–4829 (2010).
12. Roberts, C. J., Das, T. K. & Sahin, E. Predicting solution aggregation rates for therapeutic proteins: approaches and challenges. *Int. J. Pharm.* **418**, 318–333 (2011).
13. Velev, O. D., Kaler, E. W. & Lenhoff, A. M. Protein interactions in solution characterized by light and neutron scattering: comparison of lysozyme and chymotrypsinogen. *Biophys. J.* **75**, 2682–2697 (1998).
14. Neal, L. B., Asthagiri, D. & Lenhoff, A. M. Molecular origins of osmotic second virial coefficients of proteins. *Biophys. J.* **75**, 2469–2477 (1998).
15. George, A. & Wilson, W. W. Predicting protein crystallization from a dilute solution property. *Acta Crystallogr. D.* **50**, 361–365 (1994).
16. Haas, C. & Drenth, J. The protein-water phase diagram and the growth of protein crystals from aqueous solution. *J. Phys. Chem. B* **102**, 4226–4232 (1998).
17. Salgın, S., Salgın, U. & Bahadır, S. Zeta Potentials and Isoelectric Points of Biomolecules: The Effects of Ion Types and Ionic Strengths. *Int. J. Electrochem. Sci.* **7**, 12404–12414 (2012).
18. Hassan, P. A., Rana, S. & Verma, G. Making sense of Brownian motion: colloid characterization by dynamic light scattering. *Langmuir* **31**, 3–12 (2015).
19. Harding, S. E. & Johnson, P. The concentration-dependence of macromolecular parameters. *Biochem. J.* **231**, 543–547 (1985).
20. Trilisky, E., Gillespie, R., Osslund, T. D. & Vunnum, S. Crystallization and Liquid-Liquid Phase Separation of Monoclonal Antibodies and Fc-Fusion Proteins: Screening Results. *Biotechnol. Prog.* **27**, 1054–1067 (2011).
21. Muschol, M. & Rosenberger, F. Liquid–liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization. *J. Chem. Phys.* **107**, 1953, https://doi.org/10.1063/1.474547 (1997).
22. Wolf, M. *et al*. Effective interactions in protein–salt solutions approaching liquid–liquid phase separation. *J. Molec. Liq.* **200**, 20–27 (2014).
23. Lewus, R. A., Darcy, P. A., Lenhoff, A. M. & Sandler, S. I. Interactions and Phase Behavior of a Monoclonal Antibody. *Biotechnol. Prog.* **27**, 280–289 (2011).
24. Dumetz, A. C., Lewus, R. A., Lenhoff, A. M. & Kaler, E. W. Effects of ammonium sulfate and sodium chloride concentration on PEG/protein liquid-liquid phase separation. *Langmuir* **24**, 10345–10351 (2008).

25. Lomakin, A., Asherie, N. & Benedek, G. B. Monte Carlo study of phase separation in aqueous protein solutions. *J. Chem. Phys.* **104**, 1646, https://doi.org/10.1063/1.470751 (1996).
26. ten Wolde, P. R. & Frenkel, D. Enhancement of Protein Crystal Nucleation by Critical Density Fluctuations. *Science* **277**, 1975–1978 (1997).
27. Chen, Q., Vekilov, P. G., Nagel, R. L. & Hirsch, R. E. Liquid-Liquid Phase Separation in Hemoglobins: Distinct Aggregation Mechanisms of the b6 Mutants. *Biophys. J.* **86**, 1702–1712 (2004).
28. Wedekind, J. *et al*. Optimization of crystal nucleation close to a metastable fluid-fluid phase transition. *Sci. Rep.* **5**, 11260, https://doi.org/10.1038/srep11260 (2015).
29. Yang, M. X. *et al*. Crystalline monoclonal antibodies for subcutaneous delivery. *PNAS* **100**, 6934–6939 (2003).
30. Yang, H. *et al*. Optimization of Vapor Diffusion Conditions for Anti-CD20 Crystallization and Scale-Up to Meso Batch. *Crystals* **9**, 230, https://doi.org/10.3390/cryst9050230 (2019).
31. Zhang, J. & Liu, X. Y. Effect of protein-protein interactions on protein aggregation kinetics. *J. Chem. Phys.* **119**, 10972–10976 (2003).
32. Kress, C. & Brandenbusch, C. Osmotic virial coefficients as access to the protein partitioning in aqueous two-phase systems. *J. Pharm. Sci.* **104**, 3703–3709 (2015).
33. AMECRYS Project, Deliverable D5.1 Simulation code for thermodynamics of course-grained model of mAbs in confined geometry, http://www.amecrys-project.eu/images/documents/AMECRYS–Deliverable-D5.1.pdf, https://doi.org/10.5281/zenodo.1042207 (2017).
34. Hildebrandt, C., Mathaes, R., Saedler, R. & Winter, G. Origin of Aggregate Formation in Antibody Crystal Suspensions Containing PEG. *J. Pharm. Sci.* **105**, 1059–1065 (2016).
35. Rakel, N., Galm, L., Bauer, K. C. & Hubbuch, J. Influence of macromolecular precipitants on phase behavior of monoclonal antibodies. *Biotechnol. Prog.* **31**, 145–153 (2015).
36. de Araujo Sampaio, D. *et al*. Aqueous two-phase (polyethylene glycol + sodium sulfate) system for caffeine extraction: Equilibrium diagrams and partitioning study. *J. Chem. Thermodyn.* **98**, 86–94 (2016).
37. Roberts, D. *et al*. Specific Ion and Buffer Effects on Protein−Protein Interactions of a Monoclonal Antibody. *Mol. Pharmaceutics* **12**, 179–193 (2015).
38. Arakawa, T. & Timasheff, S. N. Mechanism of poly(ethylene glycol) interaction with proteins. *Biochemistry* **24**, 6756–6762 (1985).
39. Ahamed, T. *et al*. Phase Behavior of an Intact Monoclonal Antibody. *Biophys. J.* **93**, 610–619 (2007).
40. AMECRYS Project, Deliverable D2.2 HEL4 domain fragment & Anti-CD20 mAb process specification report, p. 20, http://www.amecrys-project.eu/images/documents/AMECRYS–Deliverable-D2.2.pdf.
41. Quigley, A. & Williams, D. R. The second virial coefficient as a predictor of protein aggregation propensity: A self-interaction chromatography study. *Europ. J. Pharmac. Biopharm.* **96**, 282–290 (2015).
42. Salgın, S., Salgın, U. & Bahadır, S. Zeta Potentials and Isoelectric Points of Biomolecules: The Effects of Ion Types and Ionic Strengths. *Int. J. Electrochem. Sci.* **7**, 12404–12414 (2012).
43. Saito, S. *et al*. Behavior of monoclonal antibodies: relation between the second virial coefficient (B(2)) at low concentrations and aggregation propensity and viscosity at high concentrations. *Pharm. Res.* **29**, 397–410 (2012).
44. Wu, G. *et al*. Elucidating the weak protein-protein interaction mechanisms behin d the liquid-liquid phase separation of a mAb solution by different types of additives. Euro. *J. Pharm. Biopharm.* **120**, 1–8 (2017).
45. Sorret, L. L., DeWinter, M. A., Schwartz, D. K. & Randolph, T. W. Challenges in Predicting Protein-Protein Interactions from Measurements of Molecular Diffusivity. *Biophys. J.* **111**, 1831–1842 (2016).
46. Bajaj, H. *et al*. Protein structural conformation and not second virial coefficient relates to long-term irreversible aggregation of a monoclonal antibody and ovalbumin in solution. *Pharm. Res.* **23**, 1382–1394 (2006).
47. Peat, T. S., Christopher, J. A. & Newman, J. Tapping protein data bank for crystallization information. *Acta Crystallogr. D.* **61**, 1662–1669 (2005).
48. Kress, C., Sadowski, G. & Brandenbusch, C. Solubilization of proteins in aqueous two-phase extraction through combinations of phase-formers and displacement agents. *Europ. J. Pharm. Biopharm.* **112**, 38–44 (2017).
49. Herhut, M., Brandenbusch, C. & Sadowski, G. Inclusion of mPRISM potential for polymer-induced protein interactions enables modeling of second osmotic virial coefficients in aqueous polymer-salt solutions. *Biotechnol. J.* **11**, 146–154 (2016).
50. Vivares, D., Belloni, L., Tardieu, A. & Bonneté, F. Catching the PEG induced attractive interaction between proteins. *Eur. Phys. J. E* **9**, 15–25 (2002).
51. Kozer, N., Kuttner, Y. Y., Haran, G. & Schreiber, G. Protein-protein association in polymer solutions: From dilute to semidilute to concentrated. *Biophys. J.* **92**, 2139–2149 (2007).
52. Cao, X. Z., Merlitz, H., Wu, C. X. & Sommer, J. U. Polymer-induced entropic depletion potential. *Phys. Rev. E* **84**, 041802, https://doi.org/10.1103/PhysRevE.84.041802 (2011).
53. Miyawaki, O., Dozen, M. & Nomura, K. Thermodynamic analysis of osmolyte effect on thermal stability of ribonuclease A in terms of water activity. *Biophys. Chem.* **185**, 19–24 (2014).
54. Manavalan, P. & Johnson, W. C. Sensitivity of circular dichroism to protein tertiary structure class. *Nature* **305**, 831–832 (1983).
55. Lee, H. K. *et al*. Analytical similarity assessment of rituximab biosimilar CT-P10 to reference medicinal product. *MAbs* **10**, 380–396 (2018).
56. Martins, J. P. *et al*. Liquid-Liquid Equilibrium of Aqueous Two-Phase System Composed of Poly(ethylene glycol) 400 and Sulfate Salts. *J. Chem. Eng. Data* **55**, 1247–1251 (2010).
57. Nascimento, G. R. *et al*. Liquid−Liquid Equilibrium of Two-Phase Aqueous Systems Composed of PEG 400, Na_2SO_4, and Water at Different Temperatures and pH Values: Correlation and Thermodynamic Modeling. *J. Chem. Eng. Data* **63**, 1352–1362 (2018).
58. Asenjo, J. A. & Andrews, B. A. Aqueous two-phase systems for protein separation: Phase separation and applications. *J. Chromatogr. A.* **1238**, 1–10 (2012).
59. Asenjo, J. A. & Andrews, B. A. Aqueous two-phase systems for protein separation: A perspective. *J. Chromatogr. A* **1218**, 8826–8835 (2011).
60. Kuznetsov, Y. G., Malkin, A. J. & McPherson, A. The liquid protein phase in crystallization: a case study - intact immunoglobulins. *J. Cryst. Growth* **232**, 30–39 (2001).
61. Blanco, M. A., Perevozchikova, T., Martorana, V., Manno, M. & Roberts, C. J. Protein–Protein Interactions in Dilute to Concentrated Solutions: α-Chymotrypsinogen in Acidic Conditions. *J. Phys. Chem. B* **118**, 5817–5831 (2014).

Acknowledgements

Authors would like to thank the European Union's Horizon 2020 FET-OPEN research and innovation program for funding this work within the AMECRYS project (http://www.amecrys-project.eu) under grant agreement n. 712965. Anti-CD20 monoclonal antibody was kindly provided by FUJIFILM Diosynth Biotechnologies (Billingham, UK). J.H.t.H, M.L.B. and C.J.J.G. would like to acknowledge that this work was carried out in the CMAC National Facility supported by UKRPIF (UK Research Partnership Fund) award from the Higher Education Funding Council for England (HEFCE) (Grant ref. HH13054).

Author contributions

E.P. and F.P.N. performed DLS/SLS measurements and analysed data. T.F.M., E.P. and M.L.B performed crystallization experiments. C.J.J.G., M.L.B. and J.H.t.H. performed CD measurements and analysed results. E.C. contributed to analyze experimental DLS/SLS data. T.F.M. and G.D.P. prepared the main manuscript text. E.P. and T.F.M. prepared figures. G.D.P. and F.P.N. conceived the work. G.D.P. coordinated the work. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to F.P.N. or G.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.