General role of amorphous aggregates in crystal nucleation

Zhiyu Liao,[1] Ankita Das,[1] Christina Robb,[2] Rebecca Beveridge,[2] and Klaas Wynne[1]

[1]School of Chemistry, University of Glasgow, G12 8QQ, UK

[2]Dept. of Pure and Applied Chemistry, University of Strathclyde, Glasgow G1 1BX, UK

There is mounting evidence that crystal nucleation from supersaturated solution involves the formation and reorganisation of pre-nucleation clusters contradicting classical nucleation theory. Here, a wide range of amino acids and peptides is investigated using light scattering, mass spectrometry, and *in-situ* terahertz Raman spectroscopy to demonstrate that the presence of amorphous aggregates with a wide range of sizes is a general phenomenon in supersaturated solutions. Additionally, these amorphous aggregates act as intermediates for laser-induced crystal nucleation. These observations are inconsistent not only with classical nucleation theory, but also non-classical theories involving liquid-liquid phase separation, requiring a new approach.

Introduction

Nucleation of crystals from solution is traditionally described in the framework of classical nucleation theory, in which the critical size of a nucleus that grows by attachment of solute molecules is the key criterion. However, classical nucleation theory has been challenged by observations of nanoscale and mesoscale metastable solute species in super- and even undersaturated solutions without initiating crystal growth.[1–3] Accordingly, an alternative mechanism, called non-classical or two-step nucleation theory, was proposed in which nucleation involves formation and reorganisation of pre-nucleation clusters.[1,2,4,5] Such pathways have been discovered in systems that are close to a liquid-liquid phase transition.[6–8] Subsequently, many more examples have been found of apparently amorphous intermediates or mesoscale aggregates in solution playing a role in crystal nucleation of organic[3,9,10] and inorganic molecules,[11–16] proteins,[17] and laser-induced crystal nucleation.[18–20] Such clusters may range in scale from 10 to 250 nm[21] or even 500 to 1,400 nm.[20] The relative stability of such aggregates (and lack of macroscopic phase separation) would seem to be inconsistent with liquid-liquid phase separation. Hence the investigation of the nature and behaviour of such aggregates is of great importance to the fundamental understanding of crystal nucleation.

The crystallisation and amorphous phases of amino acids and small peptides are of particular interest due their application in peptide-based supramolecular materials,[22] engineered amino acid crystals with special physical properties,[23] and peptide-based drugs[24] where crystallisation is key to unlock pharmaceutical applications.[25]

Here, we will show that a wide range of amino acids as well as di- and tripeptides in supersaturated aqueous solution form aggregates and investigate their role in laser-induced nucleation and crystallisation. Using light scattering, we can demonstrate that these aggregates are far from monodisperse but have a wide range of sizes. They form on a timescale of about a day and redissolve on a timescale of hours, while *in situ* Raman spectroscopy confirms their amorphous nature. Mass spectrometry is used to confirm that the solute molecules cluster over a very wide range of sizes. Nine out of twelve of the samples investigated shows aggregate-assisted laser-induced nucleation. These results suggest a general role of amorphous aggregates in crystal nucleation and a universal role in laser-induced nucleation. The wide range of observed sizes of the aggregates is inconsistent with both classical nucleation theory as well as non-classical theories involving liquid–liquid phase separation, establishing the need for a new theory of crystal nucleation.

Results

Formation and dissolution of aggregates. Numerous studies have described the formation of chainlike structures, mesoscale clusters, amorphous nanoparticles, transient liquid droplets, *etc.* in supersaturated solutions. However, due to the slow rate of formation and nonequilibrium nature, it is unclear whether these aggregates have the expected thermodynamic behaviour of nucleation and redissolution. Therefore, we carried out dynamic light scattering studies (DLS, see Methods for details) on a range of amino acids and peptides to elucidate the formation and redissolution of aggregates in solution on ultraslow timescales.

Figure 1 shows an example of the measurement by DLS of the formation and dissolution of aggregates of glycine tripeptide (Gly-Gly-Gly) in aqueous solution as function of temperature. The solution was prepared at 80°C at a concentration of 0.1 g/mL (see Methods for sample preparation). The intensity autocorrelation curves show that there are two components in the decays: a fast decay due to concentration fluctuations of solute molecules and a much slower decay due to the

1

aggregates. On slowly raising the temperature from 20°C to 70°C, the slow decay gradually weakens. On subsequent cooling back to 20°C at the same rate, the slow decay restored gradually, but not to the starting level. Only after ageing the solution for another two days is the slow decay amplitude fully restored. This demonstrates that the aggregates form extremely slowly in supersaturated solution (consistent with literature reports) but otherwise redissolve normally. In general, fresh solutions do not tend to nucleate crystals while aged solutions, having formed aggregates, eventually nucleate crystals after a day or days. This shows that the aggregates are thermodynamically more stable than the supersaturated solution but metastable with respect to crystallisation.
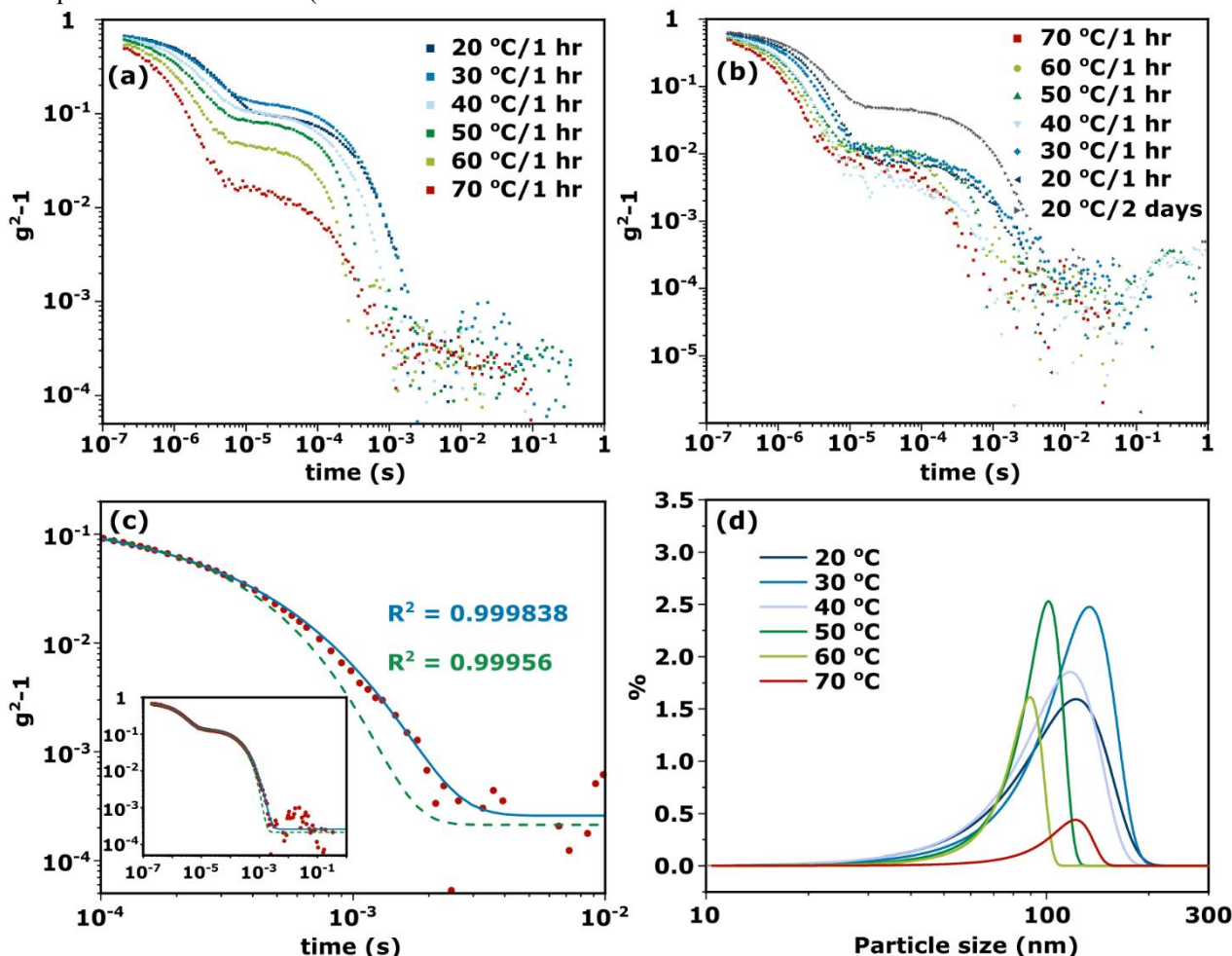

Figure 1. Dynamic light scattering showing the presence of nanometre aggregates in supersaturated solution, dissolution on heating, and reappearance on cooling. Experimental dynamic light scattering intensity autocorrelation functions with baseline subtracted of Gly-Gly-Gly (0.1g/mL) as function of temperature. (a) Solution aged for 1 day slowly warmed up (1 hour equilibration at each temperature) to induce dissolution. (b) Same sample cooled down after the heating experiment shown in (a). (c) Comparison of standard exponential fitting and stretched exponential fitting to the intensity autocorrelation function at 20°C shown in (a). Red dots: experimental intensity autocorrelation function; green dashed line: fit with an exponential function; blue solid line: fit with a stretched exponential function. (d) Particle size distribution of the slower components in the decays shown in (a), based on stretched exponential fitting.

The hydrodynamic diameters and size distributions obtained by analysing the intensity autocorrelation functions using cumulant analysis (using software that comes with the instrument) are shown in Supplementary Figure 1. The first peak at 0.97 nm is consistent with the size of Gly-Gly-Gly. The second peak near 100-200 nm is more variable in position and with a height falling with increasing temperature as expected. These results are broadly consistent with previous studies on similar solutions.

However, it is highly surprising that aggregates—which are thermodynamically more stable than the supersaturated solution—would grow to a size of 100-200 nm and then stop, especially since the component parts—molecules—are only ~1 nm. This suggests that the solutions should contain a range of aggregate sizes that is missed by a simple DLS experiment.

To test this idea, the DLS intensity correlation functions were instead analysed using a stretched exponential function (often referred to as a Kohlrausch-Williams-Watts function), $e^{-(t/\tau)^\beta}$. Much better fits could be obtained using this function (see Figure 1(c)). Figure 1(d) shows the particle size distributions (see Methods) obtained by fitting the temperature-dependent data with the stretched exponential function. These distributions are very broad at low temperature while sharpening up

and moving to lower particle sizes at higher temperature. This unambiguously demonstrates the redissolution of the aggregates at higher temperatures.

Results from nanoelectrospray ionisation-mass spectrometry also support the idea that aggregates assume a large range of sizes ranging from dimers to oligomers. As an example, Figure 2 shows the size distribution of aggregates formed in Gly–Gly–Gly solution (0.1 g/mL). The detailed mass spectra reveal that dimer, trimer, and up to 20th oligomers can be detected within the signal-to-noise, while the intensity (in other words, concentration) decreases approximately exponentially with aggregate size. Peaks with m/z smaller than monomer and in between oligomers are likely due to multiple charging, which is not included in the intensity histogram. Similar results have been obtained in several of the solutions studied here.
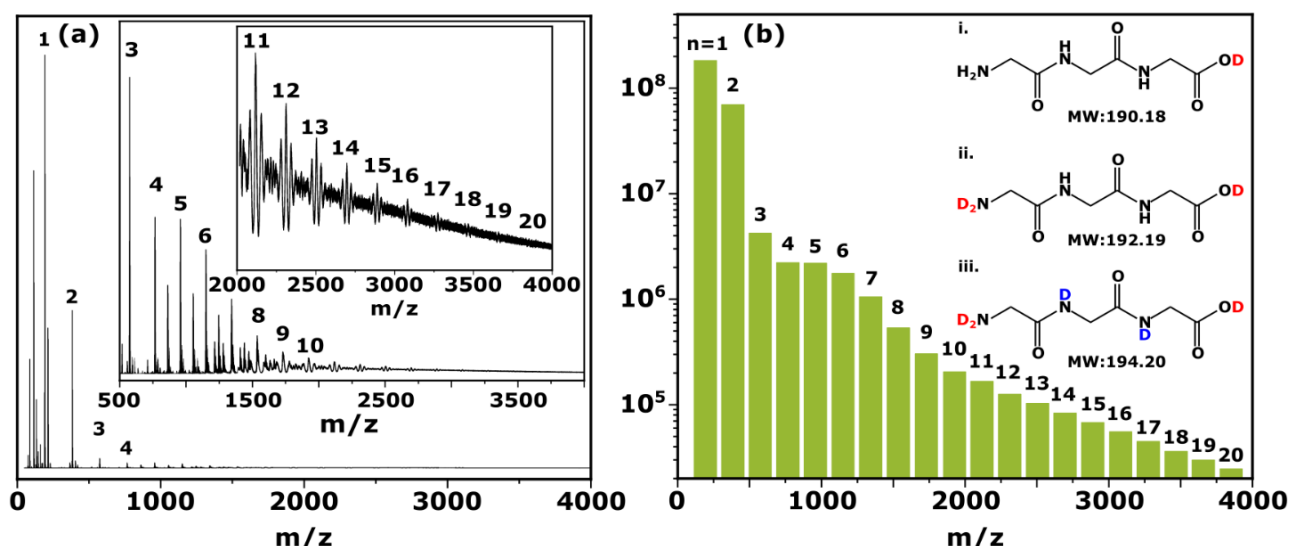

Figure 2. Characterisation of small Gly–Gly–Gly aggregates using mass spectrometry. Samples were prepared in D_2O (0.1 g/mL) and characterised 4 hrs after preparation at 70°C using mass spectroscopy. (a) Detailed spectra showing mass from 0 – 4000. Numbers above peaks indicate the size (number of monomers in the oligomer) of oligomers within aggregates; (b) Intensity histogram of the main peaks in the spectrum on a logarithmic scale. Insert show the structures of Gly-Gly-Gly in D_2O with possible deuteration sites.

The amorphous character of the aggregates. Raman spectroscopy was employed to characterise the aggregates in solution using a set-up that allows optical tweezing and *in situ* confocal Raman microscopy as described previously.[18] In supersaturated Gly-Gly-Gly solutions, aggregates formed that clustered together over time to form larger aggregates assisted by optical tweezing by the Raman excitation laser (see insert of Figure 3(a)). The *in-situ* Raman spectra of the (supersaturated) solution, the aggregates trapped in the laser beam, and the crystal are shown in Figure 3(a). The spectra of the aggregates and solution are very similar but with the former having about double the intensity. This shows that the aggregates are solute rich and are not foreign particles such as dust. At very low frequencies (0-200 cm^{-1}), the aggregate spectrum is more well defined than that of the solution, however, it does not show the three sharp phonon peaks of the crystal. This indicates a degree of order in between solution and crystal.

A number of bands in the fingerprint region that are prominent in the crystal are weak in the aggregates. For example, the CH_2 rocking band (970 cm^{-1}) and the CO_2 symmetric and asymmetric stretching modes (1407 cm^{-1} and 1638 cm^{-1}). The CO_2 asymmetric stretching peak (1638 cm^{-1}) is absent or very weak in solution but much more intense in the aggregate while dominating in the spectrum of the crystal. Additional, less prominent changes are the splitting of the ND_3 rocking band (993.3 to 993.3/1006.2 cm^{-1}), the CH_2 twisting (1269 to 1256.6/1271.6 cm^{-1}) and amide I vibration (at 1679 to 1669.6/1683.8 cm^{-1}) on going from aggregate to crystal.

Similar spectral changes have been found in other amino acids and peptides. For example, Figure 3(b) shows spectra of aggregates formed in Ala-Ala solution. Here, the aggregates show much stronger Raman intensity and some distinct peaks (*e.g.*, 984.2 cm^{-1}) compared to solution, as well as peak coalescence and shifting compared to the crystal (*e.g.*, 242.4, 878, 1096.7, and 1665.5 cm^{-1}) is observed. The very low frequency spectrum (0-200 cm^{-1}) shows a broad featureless band suggesting even less ordering in Ala-Ala compared to Gly-Gly-Gly aggregates. Another example is that of threonine aggregates (Supplementary Figure 2), which show prominent and distinctive fingerprint peaks compared to solution.
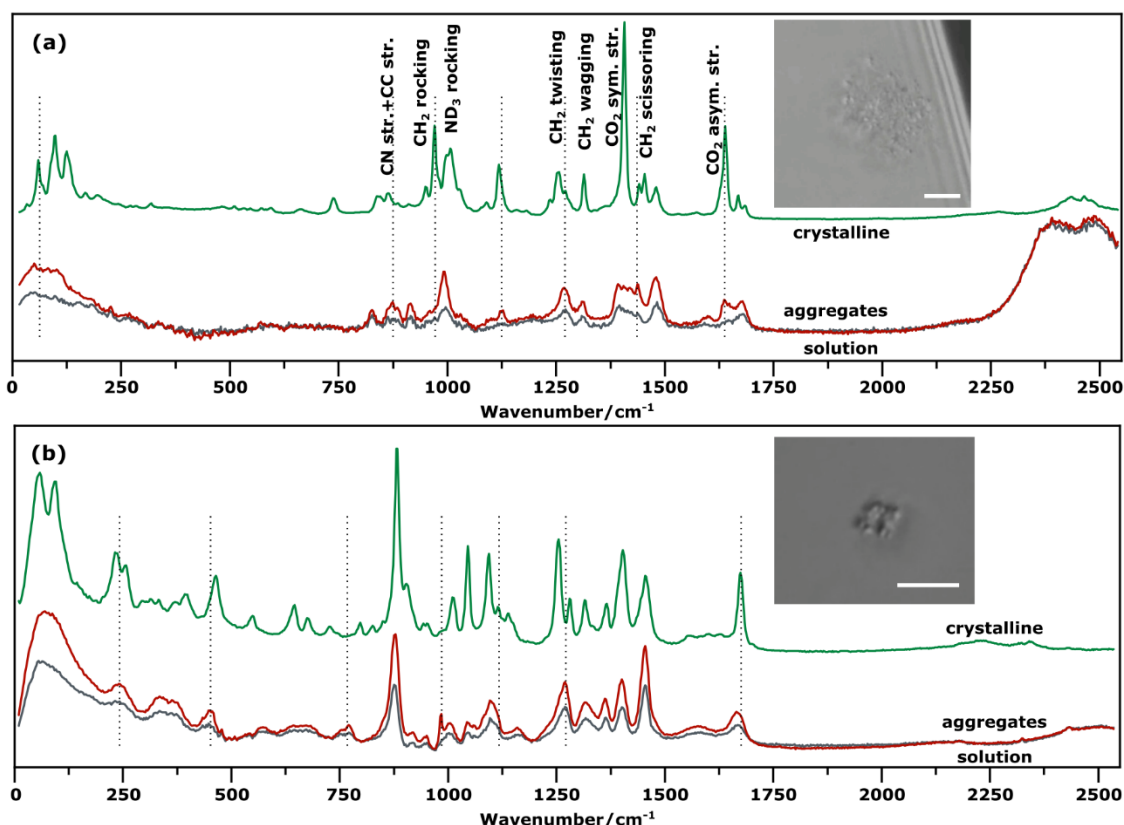
Figure 3. *In-situ* Raman spectroscopic characterisation of Gly-Gly-Gly and Ala-Ala aggregates in solution demonstrating their amorphous nature. Shown are the Raman spectra of the solution, aggregates, and a crystal of (a) Gly-Gly-Gly and (b) Ala-Ala. (insets) Aggregates in solution observed by phase-contrast microscopy (scale bars 5 μm).

The role of aggregates in laser-induced nucleation. In previous work, we have shown that aggregates in supersaturated glycine solution are an intermediate to laser-induced crystal nucleation[18] but could not demonstrate the generality of the effect. Here, we studied nine amino acids and three peptides (see Table 1), found aggregates in all but one, and observed laser-induced nucleation in nine of these.

A few examples will be presented here, where a 50-mW 532-nm CW laser was employed for laser-induced nucleation, optical trapping, and as the Raman excitation source. Figure 4(a) shows laser-induced nucleation of alanine from supersaturated aqueous solution that has been aged for 2 days (see also Supplementary Movie S1) and followed using *in-situ* Raman spectroscopy. Due to the small size of the aggregates in this case, the Raman spectrum of the aggregate ($t = 2$ s) is weak and difficult to distinguish from that of the solution. Hence, the amorphous nature cannot be confirmed. However, once laser-induced nucleation takes place ($t = 4$ s), a burst of emission is observed (peaking at ~580 nm) that is also observable in microscopy. This effect is likely due to crystalloluminescence in which dopant metal cations trapped in the growing crystal lattice relax through light emission.[26]

The crystal tumbled at the fluence of laser irradiation (and possibly free energy released from crystal growth) and went through fast conformational changes, which indicates the supersaturation of the solution was not high enough to stabilise the crystals. This was confirmed later by blocking the laser, which saw the crystals dissolve again. A back-of-the-envelope calculation shows that the formation of a ~3 μm³ crystal is followed by heat release of ~ 0.968 ± 0.22 nJ (given the heat of fusion of alanine is 22 ± 5 kJ/mol[27]) elevating the temperature of the surrounding solution (assuming the same volume as the crystal) by 8.27 ± 1.88 K. This explains the almost instant dissolution of the crystal once the laser is blocked.

Another example of laser-induced nucleation of a supersaturated Gly-Gly solution is shown in Figure 3(b), as well as Supplementary Movie S2. Despite the small size of the aggregate formed (< 1 μm, indicated by the arrow in the micrograph), the intensity of the Raman spectrum over the entire range is visibly stronger than that of the surrounding solution (~×1.6, see the comparison of spectra at $t = 3$ s and solution) without any detectable difference in peak positions or heights. When brought into the laser focus, nucleation from the aggregate starts almost immediately with a sizable crystal forming within seconds. The transition from aggregate to crystal is accompanied by peak shifts of the C-C-stretch (879 cm⁻¹), amide III (1274.4 cm⁻¹), and amide I (1687.4 cm⁻¹). Unlike the laser-induced nucleation of glycine we reported previously,[18] no intermediate states could be detected during the transition here despite the good signal strength. Similar to the case of alanine, blocking the laser would cause the formed crystal to dissolve again.
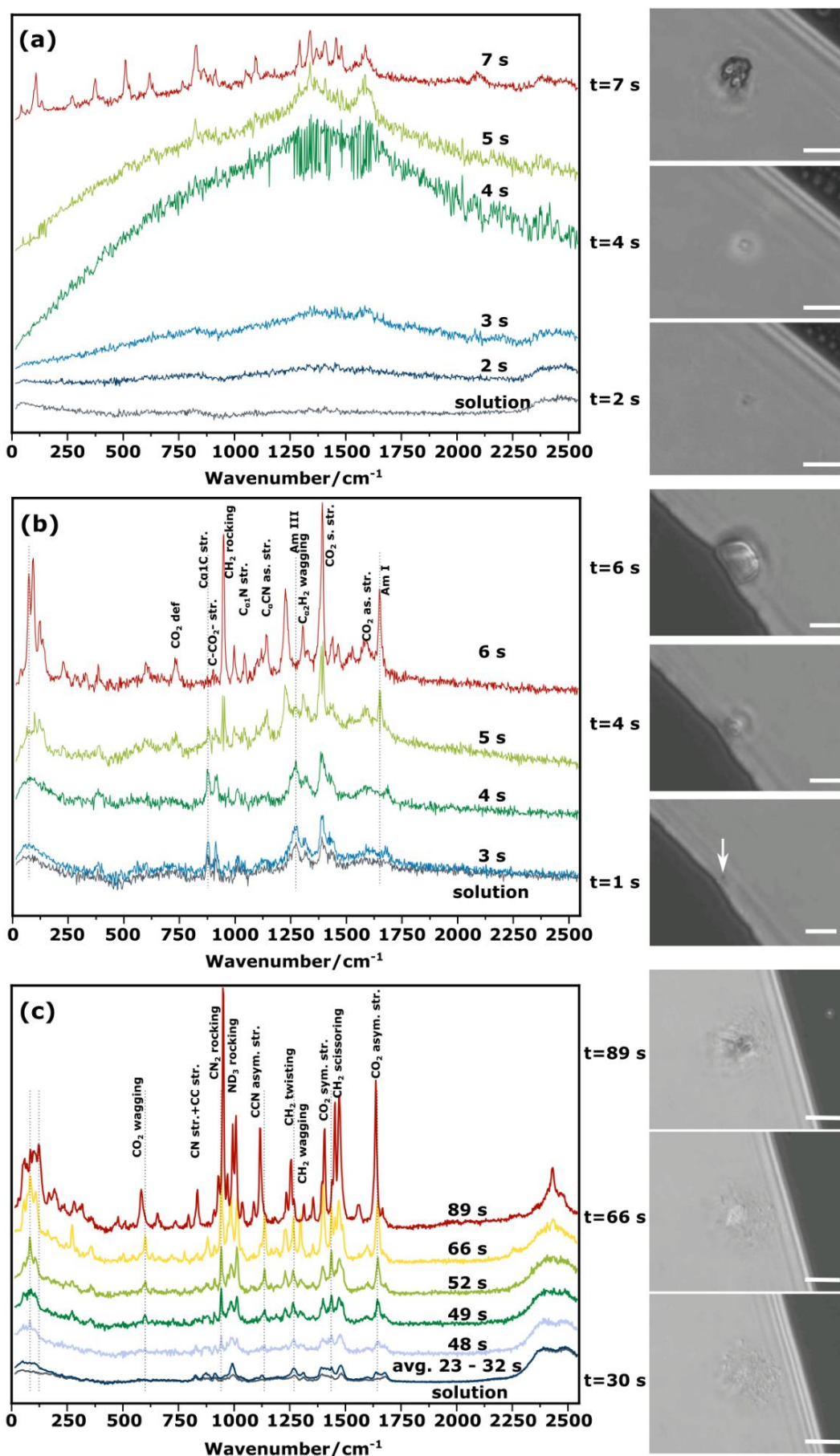
4

Figure 4. The amorphous aggregates are essential to laser-induced nucleation in amino-acid and peptide solutions. Aggregate-assisted laser-induced nucleation from solution: (a) alanine, (b) Gly-Gly and (c) Gly-Gly-Gly. Shown are time-dependent Raman and luminescence spectra on the left, along with phase-contrast microscopy images at selected times on the right (scale bars, 5 μm). The burst of luminescence in (a) at *circa* 4 s and in (c) at *circa* 66 s is likely due to crystalloluminescence.[26]

5

In most of the investigated amino-acid and peptide solutions, the aggregates visible in microscopy are small (< 1 μm) and rare. In contrast, Gly-Gly-Gly and Ala-Ala form many aggregates after aging, which cluster to form bigger aggregates. As before, *in-situ* Raman spectroscopy shows that the aggregates are amorphous and similar to the solution with subtle differences while being distinct from the crystal (see Figure 3(a) for Gly-Gly-Gly and Figure 3(b) for Ala-Ala). Figure 3(c) shows laser-induced nucleation in a supersaturated solution of Gly-Gly-Gly (Supplementary Movie S3), where again crystalloluminescence is observed (in microscopy only, showing that the emission is outside the Raman spectral range). The gradual appearance and sharpening of phonon peaks in the low frequency region (0-400 cm^{-1}) shows the process of transition from disordered aggregate to partially ordered intermediate and, finally, to the crystal at $t = 89$ s. Before this point, the peaks in the fingerprint region increase in intensity but otherwise do not change in position or relative intensity. At $t = 89$ s, the morphology changes in microscopy, and more drastic changes in the spectrum are observed.

Table 1. *Summary of aggregate formation and laser-induced nucleation in the investigated amino acids and peptides demonstrating the general nature of the formation of aggregates and their role in laser-induced nucleation.*

	Solubility in H$_2$O at 298 K[28]	Prepared in D$_2$O/H$_2$O at 293 K	Aggregate seen under microscope	Laser-induced nucleation
L-histidine	0.0436 g/mL	0.13 g/mL	Y	Y
L- phenylala-	0.028 g/mL	0.05 g/mL	Y	Y
L-threonine	0.0979 g/mL	0.2 g/mL	Y	N
L-arginine	0.1959 g/mL	0.3 g/mL	N	N
L-alanine	0.1663 g/mL	0.18 g/mL	Y	Y
L-lysine	0.2466 g/mL	0.5 g/mL	Y	Y
L-serine	0.3657 g/mL	0.4 g/mL	Y	Y
L-proline	1.3 g /mL	1.7 g/mL	Y	N
L-glutamic acid	0.0425 g/mL	0.13 g/mL	Y	Y
Ala-Ala	0.29 g/mL[29]	0.3 g/mL	Y	Y
Gly-Gly	0.195 g/mL[29]	0.2 g/mL	Y	Y
Gly-Gly-Gly	0.061 g/mL[29]	0.1 g/mL	Y	Y

Discussion

In dynamic light scattering (DLS), we find that the intensity-correlation function data are modelled much better by fitting to a stretched exponential function, implying a distribution of aggregate sizes as shown in Figure 1(d). However, light scattering scales with particle size to the 6th power and hence light scattering hugely overemphasises the larger aggregates. The true aggregate-size probability distribution therefore would strongly peak near zero (monomers) and fall off approximately exponentially. The signal-to-noise ratio of DLS is of course not nearly sufficient for a reliable extrapolation to zero.

However, there are other published experiments that support the idea of a wide distribution of aggregate sizes. Small oligomers (n ≤ 3) have been found to form in undersaturated aqueous solutions of the whole range of DL-amino acids,[30] in a number of undersaturated solutions of simple aminosulfonic acids (n ≤ 8),[31] concentrated but undersaturated DL-glutamic acid (S = 0.5, n ≤ 10),[32] and sodium metasilicate solutions of different concentrations (n ≤ ~10)[33] using electrospray ionisation mass spectrometry (ESI-MS). Indeed, oligomers (n ≤ 20) are also observed by mass spectrometry in our studies. Just as DLS is sensitive to the largest aggregates, mass spectrometry is prone to emphasise the smallest aggregates due to their much higher concentration. Therefore, these results strongly suggest that the aggregate distribution is continuous from dimers to the largest aggregates.

Phase-contrast microscopy as used here places an even stronger emphasis on the largest aggregates. Thus, the sub-micrometre scale aggregates observed by microscopy here must be at the very tail end of this wide distribution. These, comparatively huge aggregates are always observed near the edge of droplets deposited there through the coffee-ring effect.[34]

In our previous work[18], we demonstrated that metastable aggregates formed in glycine solution are solute-rich and amorphous to partially ordered. Here we have shown that this is more generally true for aggregates formed after 1 to 5 days aging of supersaturated solutions of a range of amino acids and peptides. The increased Raman scattering intensity of the aggregates over the

6

solution proves the solute-rich character of aggregates in solutions of threonine, Gly-Gly, Ala-Ala and Gly-Gly-Gly. The broad unstructured bands in the lattice phonon region (0-400 cm⁻¹) are typically different from that in the solutions, suggesting a different but still amorphous phase for the aggregates. The aggregates of the remaining amino acids are either too small or not dense enough to show a difference above the signal-to-noise. In the case of proline and alanine, the aggregates are not optically trapped but pushed away instead making it impossible to take a Raman spectrum. Finally, aggregates of phenylalanine, histidine, lysine, and serine undergo instantaneous laser-induced nucleation leaving no time for integration of a Raman spectrum with distinguishable features.

The amorphous aggregates have two possible roles in laser-induced nucleation: provide an effective local supersaturation greater than the starting solution or act as a heterogeneous nucleation site. In the former scenario, the laser either induces greater order through the Kerr effect (however, known to be a small effect[35]) or simply provides enough heat for the aggregate to undergo homogeneous nucleation at the much greater effective supersaturation.

A number of recent publications have claimed that laser-induced nucleation is caused by the optical tweezing of solute molecules or solute clusters.[36,37] However, the optical gradient force cannot trap single molecules or clusters with a radius <1 μm as the trap depth is insufficient to overcome Brownian motion. A possible exception would be laser-induced phase separation (LIPS) near a liquid–liquid critical point[8] but there is no evidence for liquid–liquid phase separation in any of the solutions studied. However, these results can be easily understood considering the results presented here. Many (aqueous) supersaturated solutions form amorphous aggregates with a wide distribution of length scales, including some aggregates large enough to be optically trapped if their refractive index is higher than the surrounding liquid. Thus, it is highly likely that so-called optical trapping induced crystallisation (OTIC) is caused by the effects described here rather than any actual optical trapping.

There is mounting evidence to support the idea of the existence of amorphous aggregates as being intermediate (on-path) or inhibiting (off-path) for nucleation of a crystal from a supersaturated solution with or without a laser to trigger the nucleation.[1,38–40] Here, we have demonstrated that such amorphous intermediates are found in aqueous solutions of many amino acids and some peptides, supporting the idea that it is a general phenomenon. The key remaining questions are of course how these amorphous aggregates nucleate themselves, why they do not grow through Ostwald ripening, and how they can be manipulated to produce crystals of the desired polymorph.

Methods

Sample preparation. All the amino acids (L-histidine, ≥99.5%, L-phenylalanine, ≥99.0 %, L-proline, ≥99.0%, L-serine, ≥99.5%, L-glutamic acid, ≥99.0%, L-lysine, ≥98.0%, L-alanine, ≥99.0%, L-threonine, ≥98%, L-arginine, ≥98.0%) and peptides (Ala-Ala, ≥98.0% and Gly-Gly, ≥99.0%, Gly-Gly-Gly, ≥98.0%) were purchased from Sigma Aldrich and used as received. Supersaturated solutions were prepared by dissolving the amino acids and peptides either in H_2O (Fisher Scientific, HPLC grade) or D_2O (Sigma-Aldrich, 99.9%) in clean glass vials at 80°C on a thermal shaker for 8 hours at a speed of 500 rpm, and then gradually cooled down to room temperature. All solutions were then filtered with Millex® PVDF filters with pore size of 0.22 μm before DLS measurements and laser-induced nucleation investigation.

Dynamic Light Scattering (DLS). Particle size analysis was carried out using a 647-nm laser and under precise temperature control using a particle size analyser (Anton Paar Litesizer 500) using a backscattering configuration ($\theta = 175°$). Intensity autocorrelation traces were recorded and initially the standard cumulant method (Kalliope software) was used to analyse the autocorrelation decay curves and estimate a mean hydrodynamic diameter of the aggregates.

As the analysis with standard software was not fitting the data well, we carried out nonlinear curve fitting (using Mathematica) with a stretched-exponential function, $e^{-(t/\tau)^\beta}$. The stretched-exponential function can be written as a distribution of exponentials as

$$e^{-(t/\tau_0)^\beta} = \int_0^\infty d\tau \, \rho_{\tau_0,\beta}(\tau) e^{-t/\tau},$$

where $\rho_{\tau,\beta}(k)$ is the relaxation-rate distribution function, which can be evaluated from the integral

$$\rho_{\tau_0,\beta}(\tau) = \frac{\tau_0}{\pi\tau^2} \int_0^\infty$$
$$\exp\left[-u^\beta \cos\left(\frac{\beta\pi}{2}\right)\right] \cos\left[u^\beta \sin\left(\frac{\beta\pi}{2}\right) - \frac{\tau_0}{\tau} u\right] du,$$

where β the stretching parameter.[41]

The relaxation times, τ, may be related to a hydrodynamic radius using the Stokes-Einstein relation, that is,

$$r = \frac{k_B T q^2}{3\pi\eta} \tau,$$

where η is the shear viscosity and $q = 4\pi\eta \sin\left(\frac{\theta}{2}\right)/\lambda$ the scattering vector. The particle-size probability distribution functions, $\rho_{\tau_0,\beta}(r)$, have not been corrected for the size dependent scattering.

Mass Spectrometry. Native mass spectrometry was carried out on a Synapt G2Si instrument (Waters) with a nanoelectrospray ionisation source (nESI). Mass calibration was conducted by infusing NaI cluster ions separately. Solutions were ionised from a thin-walled borosilicate glass capillary (i.d. 0.78 mm, o.d. 1.0 mm, Sutter Instrument) pulled in-house to an nESI tip with a

Flaming/Brown micropipette puller (Sutter Instrument). A potential of 0.8 kV was applied to the solution using a thin platinum wire (diameter 0.125 mm, Goodfellow). The following instrument parameters were used for the Gly-Gly-Gly solution in D_2O: capillary voltage 0.8 kV, sample cone voltage 40 V, source offset 60 V, source temperature 40°C, trap collision energy 4.0 V, and trap gas 3 mL/min. Data were processed using Masslynx V4.2 and OriginPro 2021.

Microscopy and Raman setup. A home-built setup for microscopy and Raman spectroscopy was employed on a double-deck inverted microscope (Olympus IX73). Phase contrast microscopy images and videos were taken using a 60×/0.7 N.A. objective (Olympus, UCPlanFL N Ph2) and a CMOS camera (Teledyne Dalsa, Genie Nano-1GigE). Two laser sources were simultaneously aligned into the microscope objective: a high-power pulsed 1040 nm laser (Spectra-Physics, Spirit One, 8 W) intended for optical tweezing and trapping, and a single-frequency 532 nm laser with linear polarization (Laser Quantum, gem 532, 500 mW) for Raman excitation. Ultra-low frequency Raman spectroscopy was achieved using BragGrate^{TM} bandpass and notch filters (OptiGrate), enabling detection of Raman scattering to frequencies as low as 10 cm^{-1}. A spectrometer (Andor, Shamrock 500i with 600 groove/mm grating) and a CCD camera (Andor, iDUS 401) were utilized for detection. A confocal Raman collection employed an optical fibre with 50 μm core size, and the spectral resolution of the system was around 2 cm^{-1}.

Data processing. All raw Raman scattering spectra were corrected (using MATLAB by MathWorks) with the Bose thermal-occupation factor, $I(\omega)/(1+n(\omega))$, after background subtraction (dark counts of CCD detector), where $n(\omega) = (exp(-\hbar\omega/k_BT) - 1)^{-1}$, ω is the angular frequency, and I the Raman amplitude at ω, to obtain a reduced Raman spectrum.

As the Raman scattering signal strength from the small aggregates is weak, there is a relatively strong background from the glass microscope slides (as shown in Supplementary Figure 3). All the spectra shown here have this background caused by glass subtracted.

Data availability
The data that support the findings of this study are available in Enlighten: Research Data Repository (University of Glasgow) with the identifier: http://dx.doi.org/10.5525/gla.researchdata.@@@.

ACKNOWLEDGMENT

References
1. Gebauer, D., Völkel, A. & Cölfen, H. Stable Prenucleation Calcium Carbonate Clusters. *Science* **322**, 1819–1822 (2008).
2. Gebauer, D., Kellermeier, M., Gale, J. D., Bergström, L. & Cölfen, H. Pre-nucleation clusters as solute precursors in crystallisation. *Chem. Soc. Rev.* **43**, 2348–2371 (2014).
3. Svärd, M. Mesoscale clusters of organic solutes in solution and their role in crystal nucleation. *CrystEngComm* (2022) doi:10.1039/D2CE00718E.
4. Gebauer, D. & Wolf, S. E. Designing Solid Materials from Their Solute State: A Shift in Paradigms toward a Holistic Approach in Functional Materials Chemistry. *J. Am. Chem. Soc.* **141**, 4490–4504 (2019).
5. De Yoreo, J. J. *et al.* Crystallization by particle attachment in synthetic, biogenic, and geologic environments. *Science* **349**, aaa6760 (2015).
6. Wallace, A. F. *et al.* Microscopic Evidence for Liquid-Liquid Separation in Supersaturated $CaCO_3$ Solutions. *Science* **341**, 885–889 (2013).
7. Mosses, J., Turton, D. A., Lue, L., Sefcik, J. & Wynne, K. Crystal templating through liquid–liquid phase separation. *Chem. Commun.* **51**, 1139–1142 (2015).
8. Walton, F. & Wynne, K. Control over phase separation and nucleation using a laser-tweezing potential. *Nat. Chem.* **10**, 506–510 (2018).
9. Tsarfati, Y. *et al.* Crystallization of Organic Molecules: Nonclassical Mechanism Revealed by Direct Imaging. *ACS Cent. Sci.* **4**, 1031–1036 (2018).
10. Tsarfati, Y. *et al.* Continuum Crystallization Model Derived from Pharmaceutical Crystallization Mechanisms. *ACS Cent. Sci.* **7**, 900–908 (2021).
11. Yang, S., Guo, Z., Bian, B., Du, J. & Hu, Y. Dynamic Observation of Anisotropic Chainlike Structures during Nonclassical Two-Step Nucleation in Solid-State Iron Oxide Crystallization. *J. Phys. Chem. Lett.* **13**, 8352–8358 (2022).
12. Durelle, M. *et al.* Coexistence of Transient Liquid Droplets and Amorphous Solid Particles in Nonclassical Crystallization of Cerium Oxalate. *J. Phys. Chem. Lett.* **13**, 8502–8508 (2022).
13. Schreiber, R. E. *et al.* Real-time molecular scale observation of crystal formation. *Nat. Chem.* **9**, 369–373 (2017).
14. Dey, A. *et al.* The role of prenucleation clusters in surface-induced calcium phosphate crystallization. *Nat. Mater.* **9**, 1010–1014 (2010).
15. Parmar, D., Niu, Z., Liang, Y., Dai, H. & Rimer, J. D. Manipulation of amorphous precursors to enhance zeolite nucleation. *Faraday Discuss.* (2022) doi:10.1039/D1FD00096A.
16. Avaro, J., Moon, E. M., Schulz, K. G. & Rose, A. L. Calcium Carbonate Prenucleation Cluster Pathway Observed via In Situ Small-Angle X-ray Scattering. *J. Phys. Chem. Lett.* **14**, 4517–4523 (2023).

17. Yamazaki, T. *et al.* Two types of amorphous protein particles facilitate crystal nucleation. *Proc. Natl. Acad. Sci.* **114**, 2154–2159 (2017).

18. Liao, Z. & Wynne, K. A Metastable Amorphous Intermediate Is Responsible for Laser-Induced Nucleation of Glycine. *J. Am. Chem. Soc.* **144**, 6727–6733 (2022).

19. Zhiyu Liao & Klaas Wynne. Mesoscopic amorphous particles rather than oligomeric molecular aggregates are the cause of laser-induced crystal nucleation | PNAS. *Proc. Natl. Acad. Sci.* **119**, e2207173119 (2022).

20. Gowayed, O. Y. *et al.* Dynamic Light Scattering Study of a Laser-Induced Phase-Separated Droplet of Aqueous Glycine. *J. Phys. Chem. B* **125**, 7828–7839 (2021).

21. Zimbitas, G. *et al.* Investigation of molecular and mesoscale clusters in undersaturated glycine aqueous solutions. *Colloids Surf. Physicochem. Eng. Asp.* **579**, 123633 (2019).

22. Sheehan, F. *et al.* Peptide-Based Supramolecular Systems Chemistry. *Chem. Rev.* **121**, 13869–13914 (2021).

23. Yuan, H. *et al.* The engineering of molecular packing in amino acid crystals for the enhanced triboelectric effect. *Nano Energy* **110**, 108375 (2023).

24. Muttenthaler, M., King, G. F., Adams, D. J. & Alewood, P. F. Trends in peptide drug discovery. *Nat. Rev. Drug Discov.* **20**, 309–325 (2021).

25. Chen, J., Sarma, B., Evans, J. M. B. & Myerson, A. S. Pharmaceutical Crystallization. *Cryst. Growth Des.* **11**, 887–895 (2011).

26. Alexander, A. J. Deep ultraviolet and visible crystalloluminescence of sodium chloride. *J. Chem. Phys.* **136**, 064512 (2012).

27. Zen Chua, Y., Tam Do, H., Schick, C., Zaitsau, D. & Held, C. New experimental melting properties as access for predicting amino-acid solubility. *RSC Adv.* **8**, 6365–6372 (2018).

28. Michel Fleck , Aram M. Petrosyan. *Salts of amino acids. Crystallization, Structure and Properties*. (Cham, Switzerland: Springer International Publishing, 2014).

29. Tam Do, H. *et al.* Melting properties of peptides and their solubility in water. Part 1: dipeptides based on glycine or alanine. *RSC Adv.* **9**, 32722–32734 (2019).

30. Kellermeier, M. *et al.* Amino acids form prenucleation clusters: ESI-MS as a fast detection method in comparison to analytical ultracentrifugation. *Faraday Discuss.* **159**, 23–45 (2013).

31. Haupa, K., Szewczuk, Z. & Mielke, Z. Clustering of simple aminosulfonic acids – electrospray ionization mass spectrometric study. *Rapid Commun. Mass Spectrom.* **27**, 1993–1998 (2013).

32. Jiang, Y. *et al.* Growth of organic crystals via attachment and transformation of nanoscopic precursors. *Nat. Commun.* **8**, 15933 (2017).

33. Marsiske, M. R., Köser, R., Bäumle, B. & Ruiz-Agudo, C. Uncovering the Early Stages of Magnesium Silicate Hydrate Formation: A Nonclassical Multistep Pathway. *ACS Appl. Eng. Mater.* **1**, 696–707 (2023).

34. Deegan, R. D. *et al.* Capillary flow as the cause of ring stains from dried liquid drops. *Nature* **389**, 827–829 (1997).

35. Knott, B. C., Doherty, M. F. & Peters, B. A simulation test of the optical Kerr mechanism for laser-induced nucleation. *J. Chem. Phys.* **134**, 154501 (2011).

36. Urquidi, O., Brazard, J., LeMessurier, N., Simine, L. & Adachi, T. B. M. In situ optical spectroscopy of crystallization: One crystal nucleation at a time. *Proc. Natl. Acad. Sci.* **119**, e2122990119 (2022).

37. Yuyama, K., Islam, M. J., Takahashi, K., Nakamura, T. & Biju, V. Crystallization of Methylammonium Lead Halide Perovskites by Optical Trapping. *Angew. Chem. Int. Ed.* **57**, 13424–13428 (2018).

38. Jacobson, L. C., Hujo, W. & Molinero, V. Amorphous Precursors in the Nucleation of Clathrate Hydrates. *J. Am. Chem. Soc.* **132**, 11806–11811 (2010).

39. Yamazaki, T. *et al.* Two types of amorphous protein particles facilitate crystal nucleation. *Proc. Natl. Acad. Sci.* **114**, 2154–2159 (2017).

40. Reichenbach, J. & Wynne, K. Frustration vs Prenucleation: Understanding the Surprising Stability of Supersaturated Sodium Thiosulfate Solutions. *J. Phys. Chem. B* **122**, 7590–7596 (2018).

41. Berberan-Santos, M. N., Bodunov, E. N. & Valeur, B. Mathematical functions for the analysis of luminescence decays with underlying distributions 1. Kohlrausch decay function (stretched exponential). *Chem. Phys.* **315**, 171–182 (2005).

Supplementary information—General role of sub-micron amorphous aggregates in crystal nucleation
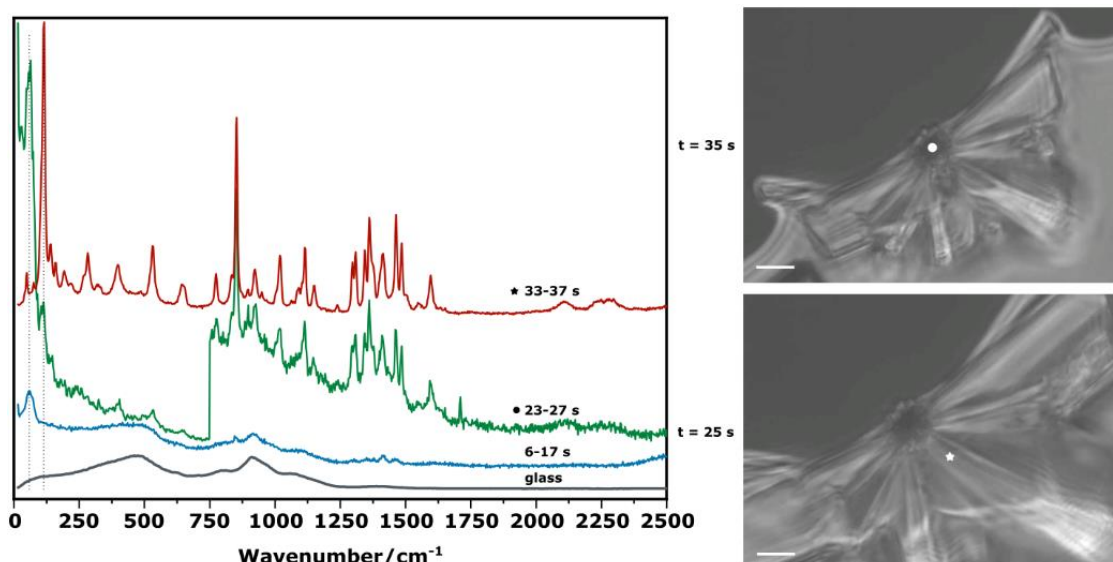
Supplementary figures
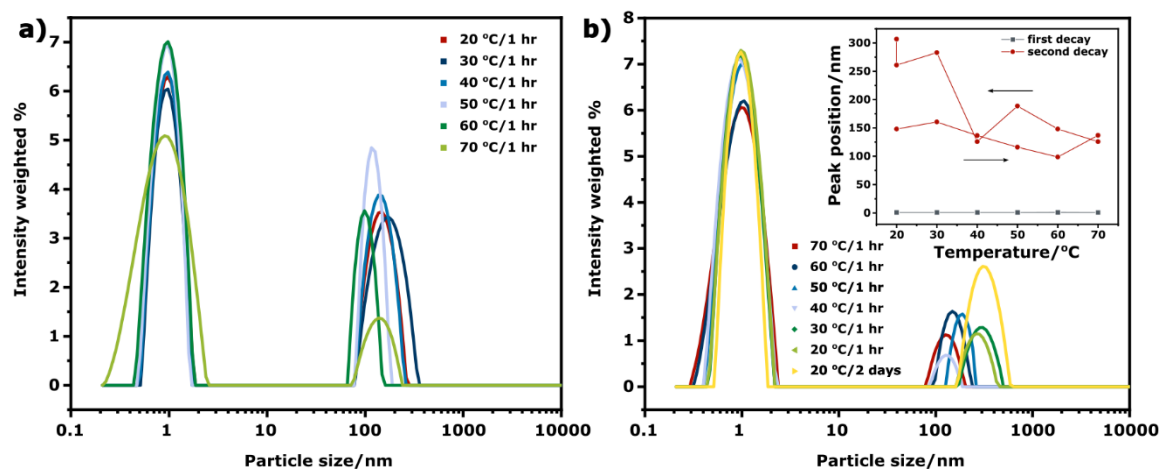
Supplementary Figure 1. Intensity weighted size distribution of the intensity correlation functions shown in Figure 1. (a) Warm up from 20°C to 70°C; (b) Cooling down from 70°C to 20°C. The inset shows the changes of peak position as function of temperature.
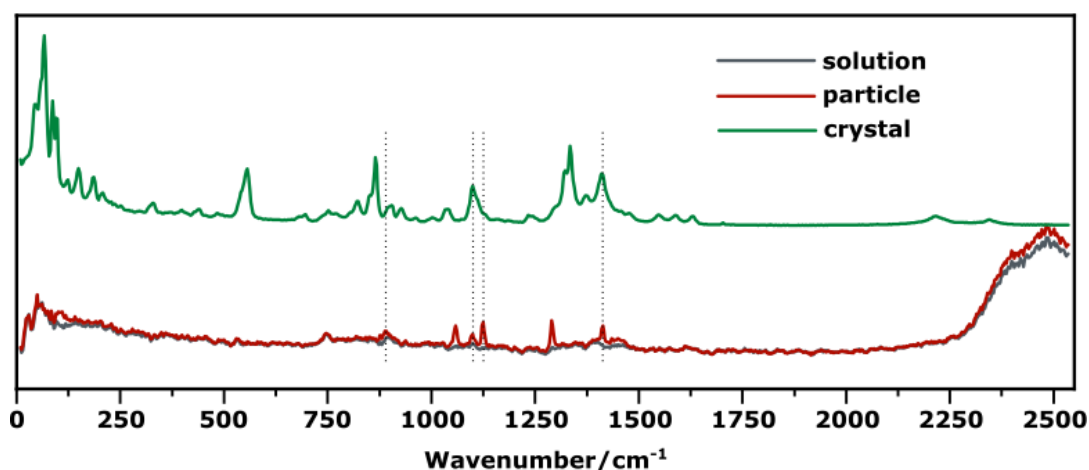
Supplementary Figure 2. Raman spectrum of threonine aggregate in solution compared to surrounding solution and crystal.

Supplementary Figure 3. Aggregate-assisted laser-induced nucleation of alanine from solution. Averaged Raman spectra are shown on the left, along with microscopy images at selected times on the right; The green spectrum is 5× magnified from 750 – 2500 cm^{-1} for comparison with better clarity; dot and star signs indicate the locations where spectra are taken. scale bar, 5 μm.