molecular pharmaceutics

Subscriber access provided by UCL Library Services

Article

Selective stabilisation and destabilisation of protein domains in tissue-type plasminogen activator using formulation excipients

Mathew J. Robinson, Paul Matejtschuk, Colin Longstaff, and Paul A. Dalby

Mol. Pharmaceutics, **Just Accepted Manuscript** • DOI: 10.1021/acs.molpharmaceut.8b01024 • Publication Date (Web): 19 Dec 2018

Downloaded from http://pubs.acs.org on January 4, 2019

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Published by American Chemical Society. Copyright © American Chemical Society. However, no copyright claim is made to original U.S. Government works, or works produced by employees of any Commonwealth realm Crown government in the course of their duties.

Selective stabilisation and destabilisation of protein domains in tissue-type plasminogen activator using formulation excipients

*Mathew J. Robinson[1], Paul Matejtschuk[2], Colin Longstaff[2], Paul A. Dalby[1]**

[1]Department of Biochemical Engineering, University College London, Gower Street, London, WC1E 6BT

[2]National Institute for Biological Standards and Control, South Mimms, Potters Bar, Hertfordshire, EN6 3QG

*Correspondence email: p.dalby@ucl.ac.uk

ABSTRACT

Multi-domain biotherapeutic proteins present additional behavioural and analytical challenges for the optimisation of their kinetic stability by formulation. Tissue-type plasminogen activator (tPA) comprises six protein domains that exhibit a complex and pH-dependent thermal unfolding profile, due to partially independent domain unfolding. Here we have used tPA as a model for evaluating the relationships between various thermal unfolding and aggregation parameters in multi-domain proteins. We show that changes in the thermal unfolding profile of tPA were parameterised by the overall thermal mid-point transition temperature T_m, and the Van't Hoff entropy for unfolding, ΔS_{vh}, which is a measure of unfolding cooperativity. The kinetics of degradation at 45 °C, leading to aggregation, were measured as rates of monomer and activity loss. These two rates were found to be coincident at all pH. Aggregation accelerated at pH 4 due to the early unfolding of the Serine Protease N-terminal domain (SP-N), whereas at pH 5-8, the fraction unfolded at 45 °C (f_{45}) was <1%, resulting in a baseline rate of aggregation from the native ensemble.

We used a Design of Experiments (DoE) approach to evaluate how formulation excipients impact and control the thermal unfolding profile for tPA, and found that the relative stability of each of the tPA domains was dependent on the formulation. Therefore, the optimisation of formulations for complex multi-domain proteins such as tPA may need to be multi-objective, with careful selection of the desired attributes that improve stability. As aggregation rates (ln v) correlated well to T_m ($R^2 = 0.77$), ΔS_{vh} ($R^2 = 0.71$), but not T_{agg} ($R^2 = 0.01$), we analysed how formulation excipients and pH

would be able to optimise T_m and ΔS_{vh}. Formulation excipient behaviours were found to group according to their combined impact on T_m and ΔS_{vh}. The effects of each excipient were often selectively stabilising or destabilising to specific tPA domains, and changed the stability of particular domains relative to the others. The types of mechanism by which this could occur might involve specific interactions with the protein surface, or otherwise effects that are mediated via the solvent as a result of the different surface hydrophobicities and polarities of each domain.

Keywords: Aggregation, cooperativity, entropy, formulations, stability, tissue-type plasminogen activator, multidomain protein

INTRODUCTION

As the number of biologics in clinical development continues to grow [1] there is an increased need to obtain stable formulations, on shortening time-scales. Aggregation is the principle route of degradation of protein biotherapeutics, and it is important to control in the formulation and storage phases, but also during upstream [2] and downstream [3] processing. Protein aggregates can be elicited through intrinsic and extrinsic factors [4], can have different morphologies [5, 6] and solubilities, and can arise from multiple pathways [7, 8] for a given protein. Aggregation will be defined in this work as the net irreversible formation of non-native species, larger than that of the native species [4, 9]. From a pharmaceutical standpoint, suppression of

aggregation is critical, given the possibility of aggregation-induced immunogenicity [10].

While the formulation of new therapeutic proteins must be achieved within increasingly shortened time-frames, the prediction of their long-term stability remains a challenge. Popular experimental approaches include the use of chemical denaturation to derive free-energies of unfolding (ΔG_{unf}), thermal scanning to derive thermal transition-midpoints (T_m) or aggregation temperatures (T_{agg}), or the use of forced degradation kinetics at elevated temperatures. These various surrogates for the long-term stability of the molecule, have been met with varying degrees of success [11-13]. The use of such measures is potentially confounded by multiple factors at play in aggregation mechanisms, and whose relative roles are dependent upon both the protein and the solution conditions. Key mechanisms thought to be important in protein aggregation include global and local folding events, the formation of unfolded intermediates, or native-like states, the exposure of hydrophobic surface patches, net surface charge, and the ability of specific sequences to form cross-beta sheets [4, 7, 12].

Our previous work on Fab A33 and GCSF found that T_m-values predict aggregation kinetics for formulations where the incubation temperature is below but close to the range of T_m-values [12, 13, 14]. While the proteins studied unfolded with relatively high cooperativity, it remains unknown whether T_m values can be used indirectly to optimise the aggregation kinetics of formulations for more complex multi domain proteins in which the domains unfold non-cooperatively. Multi-domain proteins, including antibodies, bring additional influences from inter-domain solution

conformations [15], and non-concerted domain unfolding events which can eventually resolve into separate T_m-values [16, 17], such that each state can have different aggregation propensities. The latter phenomenon would also be relevant to emerging classes of engineered fusion proteins as there would be no expectation for the respective protein domains to unfold cooperatively or have the same aggregation propensity.

The recombinant biotherapeutic tissue-type plasminogen activator (tPA) is a 65 kDa enzyme, consisting of five modules, and six domains, linked by disulphide bonds [18], and includes approx. 7 kDa from three carbohydrate moieties [19-21]. The five modules, containing six domains, are connected in the order:

FNT-I – EGF-like – K1 – K2 – SP-N – SP-C

where FNT-I is the fibronectin type I domain (also known as the finger domain), EGF-like is the epidermal growth-factor-like domain, K1 is the kringle-1 domain, K2 is the kringle-2 domain, SP-N is the serine protease N-terminal domain, and SP-C is the serine protease C-terminal domain.

Registered as Alteplase™, and marketed as Actilyse in Europe, human recombinant glycosylated tPA has treated over 2 million people for myocardial infarction and ischemic stroke [22]. Its complex structure, interactions between five domains [19], the hydrophobic clusters of the kringle-2 domain [23], and the hydrophobic regions at the active site [24] present a significant challenge to biophysical characterization, and

thus tPA is an interesting protein with which to investigate the roles of multi-domain unfolding and inter-domain interactions in stability, aggregation and formulation. The current clinical formulations for tPA marketed by Boehringer Ingelheim (Germany), and Genentech (CA, USA), are presented as freeze-dried solids that require reconstitution prior to parenteral administration. However, it remains of interest to study tPA in the context of liquid formulations to ensure stability prior to freeze-drying, or spray drying into nanoparticles [25], and liposomal delivery systems [26], where it would remain in the liquid state.

The thermal unfolding profile of tPA has been well characterised in the liquid state [19], and is known to be complex due to a degree of non-cooperative unfolding of its individual domains. It therefore provides a good model system to study the effects of formulation excipients on multi-domain unfolding. The commercial freeze-dried formulations for tPA contain arginine, Tween-80 and phosphate buffer [27]. In solution prior to freeze-drying or after reconstitution, 50-200 mM arginine enhances the solubility of tPA at neutral pH [27]. The kringle-2 domain is known to bind specifically to L-lysine with a 100 μM affinity, and to L-arginine with 21 mM affinity [28]. This demonstrates the possibility that some excipients may be able to act through the selective stabilisation of particular tPA domains.

The activity of the two-domain tPA serine-protease module, provides a characterisation of the functional state of that module, independent from measurements of total monomer loss. This could potentially enable the elucidation of the role of serine-protease domain unfolding in tPA aggregation. The serine protease domain is

known to uncouple from the concerted thermal unfolding of the other domains in a pH-dependent manner [19]. For example, during thermal unfolding at low pH, the N-terminal domain within the two-domain serine protease module (SP-N) unfolds first, whereas at physiological pH the protease domains remain coupled and the kringle-2 domain unfolds first. In addition, tPA can be cleaved by plasmin *in vivo* at the Arg_{275}-Ile_{276} peptide bond, to make a two-chain form that is held together by a Cys_{264}-Cys_{395} disulphide bridge [29], and that has an altered proteolytic activity [30, 31]. Thus, tPA exhibits a complex degradation pathway, that might severely affect the ability of simple T_m measurements to predict long-term stability, even at temperatures close to the T_m range of typical formulations.

In this work, we initially evaluated the relationships between degradation kinetics, and various thermodynamic parameters including T_m, T_{agg} and T_{onset} (onset of unfolding), and ΔS_{vh}. These were obtained simultaneously by intrinsic protein fluorescence and static light scattering, for tPA at a wide range of pH, enabling us to relate previously reported changes in tPA domain unfolding cooperativity, to aggregation kinetics. It was found that early unfolding of the SP-N domain at pH 4, also accelerated tPA aggregation. We also explored the formulation of tPA using DoE with 19 factors including pH and a range of common excipients. The rate of degradation was found to correlate well to T_m and ΔS_{vh}. Previously, for Fab A33 an increase in ΔS_{vh} due to rigidification of the native ensemble, was found to correlate to a decrease in aggregation rate, particularly in cases where the T_m remained the same [14].

The formulation excipients were found to differentially affect T_m and ΔS_{vh}, suggesting that optimisation of formulations by a DoE approach may need to be multi-objective. The excipient behaviours were found to group according to their combined impact on T_m and ΔS_{vh}. This study therefore provided a first insight into how different excipients could be used to optimise the conformational stability of tPA in terms of overall stability, cooperativity of unfolding, or with a multi-objective combination of the two. Furthermore, it demonstrated how different excipients were able to selectively stabilise or destabilise different domains within a complex multi-domain protein such as tPA.

MATERIALS AND METHODS

All formulation components were obtained from Sigma-Aldrich Co. (Poole, UK), except NaCl (Thermo Fisher Scientific Inc., Loughborough, UK).

Tissue Plasminogen Activator (tPA)

Expressed in Chinese Hamster Ovary (CHO) cells, glycosylated tissue plasminogen activator (tPA) (Actilyse), made by Boehringer Ingelheim, Germany. The lyophilised powder was reconstituted to 1 mg/ml tPA, aliquoted and snap frozen at -80 °C. Freezing tPA at -80 °C and thawing the samples has been previously shown to have little effect on t-PA activity [33].

Formulations

Reconstituted t-PA was dialysed at 4 °C using 10 kDa (Slide-a-Lyzer) dialysis cassettes (Thermo Fisher Scientific Inc., Loughborough, UK) into each buffered

formulation. For the pH studies, buffers from pH 3.5−10 (Table 1) were 50 mM sodium formate pH 3.5, 50 mM sodium acetate pH 4.0, 50 mM sodium acetate pH 5.0, 50 mM sodium phosphate-citrate pH 6.0, 50 mM sodium phosphate pH 7.0, 50 mM HEPES pH 8.0, and 50 mM CAPS pH 10.0. NaCl was added to all except the sodium phosphate buffer, to match the 110 mM ionic strength of that buffer. Arginine was added as arginine monohydrochloride to 75 mM at all pH, or to 0-200 mM arginine for the study at pH 8 only. For the DoE formulation study, buffers at pH 4.0 and pH 7.0 were produced at 110 mM ionic strength as above. Minitab 17 software (Minitab Inc, College Station, PA, USA) was used to construct a 2-level, resolution III, fractional factorial DoE study. The 32 formulations of 0.2 mg/ml tPA, (see Table S1 in SI), were generated at pH 4.0 or pH 7.0, with the following excipients: 0-50 mM glycine, glutamic acid, proline and methionine; 25-75 mM arginine; additional 0-50 mM NaCl; 0.1% (v/v) Tween-20 and Tween-80; 0-2.5% (w/v) sucrose, trehalose, mannitol, and glucose; and 0-3 mg/mL sorbitol, glycerol, ethanol, PEG 400, PEG 2000 and PEG 6000.

Protein Quantification

tPA was quantified through light absorbance; measured using the Cecil Aquarius CE 7500 spectrophotometer (Cecil Instruments Ltd, Cambridge, UK) with a Hellma cuvette (105.251-Qs). Wavelength scans (200-340 nm) were recorded, and the absorbance at 320 nm subtracted from that at 280 nm. Beer-Lambert's law ($A = \varepsilon.c.l$) was used to quantify the concentration of tPA, assuming an ε_{280} of 110395 M^{-1} cm^{-1} derived from the tPA sequence [34], and the pathlength of 1 cm.

Intrinsic Protein Fluorescence and Static Light Scattering Measurements

The Optim-1000 (Unchained Laboratories Ltd, West Drayton, UK) instrument was used to measure static light scattering (SLS) (at 266 nm) and intrinsic protein fluorescence (350:330 nm ratio) of 0.2 mg/ml t-PA in each formulation as the temperature was ramped (15-90 °C) at a rate of 1 °C/min. Each sample volume was 9 µl and each formulation was replicated five times. SLS and intrinsic fluorescence data from the five replicates were averaged before data fitting. Temperature dependent fluorescence intensity data (I_T) were fitted to two-state transitions with equation 1 as described previously [14],

$$I_T = \frac{(I_N + aT) + (I_D + bT)\exp[\frac{\Delta H_{vh}}{R}(\frac{1}{T_m} - \frac{1}{T})]}{1 + \exp[\frac{\Delta H_{vh}}{R}(\frac{1}{T_m} - \frac{1}{T})]} \quad \text{Equation 1}$$

using OriginPro 9.0 (OriginLab Corp, Northampton, MA, USA), to produce $T_{m,app}$ and ΔH_{vh} values for each formulation. I_N and I_D are the native and denatured state signals respectively, a and b are the slopes of the native and denatured state signal baselines respectively, R is the molar gas constant, ΔH_{vh} and ΔS_{vh} are the van't Hoff enthalpy and corresponding entropy, at the transition midpoint respectively. $T_{m,app}$ was the apparent thermal transition midpoint which represents an average across the known multiple domain unfolding events in tPA [19]. Unfolding is also known to be partially reversible, where the aggregation rate is slow relative to the unfolding rate during thermal ramping at 1 °C /min. ΔS_{vh} was then determined from equation 2, and the fraction unfolded (f_T) at any temperature T, determined using equation 3,

$$\Delta S_{\mathrm{vh}} = \frac{\Delta H_{\mathrm{vh}}}{T_{\mathrm{m}}} \qquad \text{Equation 2}$$

$$f_T = \frac{I_T - I_{\mathrm{N}} - \mathrm{a}T}{I_{\mathrm{D}} + \mathrm{b}T - I_{\mathrm{N}} - \mathrm{a}T} = \frac{\exp[\frac{\Delta H_{\mathrm{vh}}}{R}(\frac{1}{T_{\mathrm{m}}} - \frac{1}{T})]}{1 + \exp[\frac{\Delta H_{\mathrm{vh}}}{R}(\frac{1}{T_{\mathrm{m}}} - \frac{1}{T})]} \qquad \text{Equation 3}$$

T_{onset} was the temperature at which a mole fraction of 2% unfolded was present, and was calculated using equation 4 by setting f_T to 0.02.

$$T_{\mathrm{on}} = \frac{\Delta H_{\mathrm{vh}}}{\Delta S_{\mathrm{vh}} - R\ln\left(\frac{f_T}{1 - f_T}\right)} \quad (set\ f_T = 0.02) \qquad \text{Equation 4}$$

T_{agg} was the temperature for the first datapoint at the beginning of a transition, for which the static light scattering increases above the baseline data by at least 1 standard deviation.

Differential Scanning Fluorimetry (DSF)

DSF was carried out using a method similar to that previously reported [35]. 0.5 µL of SYPRO Orange stain (Invitrogen, Life Technologies Corp, Paisley, UK) was added to 100 µL of 0.2 mg/ml t-PA in pH 3.5-10.0. 20 µL of each sample was added to PCR tubes (4 replicates for each formulation) and analysed using a Corbett Rotor-Gene Thermocycler (Qiagen, Manchester, UK). The HRM filter (at gain 7) was applied and the samples were ramped (25-95°C) at 1 °C/min. Analysis was carried out by Rotor-Gene 6000 software (Qiagen, Manchester, UK) to obtain the fluorescence versus temperature profiles. DSF profiles were fitted to two-state transitions with equation 1 in OriginPro 9.0 to obtain $T_{\mathrm{m,DSF}}$. DSF profiles were also fitted to either two- or three-

Lorentzian peaks in OriginPro 9.0, taking the fits with the best residuals, and using the Lorentzian form:

$$y = y0 + (2A/\pi).(w/(4\,(T-T_{max})^2 + w^2)) \quad \text{Equation 5}$$

where T_{max} is the temperature at the peak maximum, w is the peak-width at half height, and A is the peak area. The baseline fluorescence y0 was a value shared by all Lorentzian peaks in each overall fit.

Isothermal Degradation

Into 2 ml screw-cap HPLC glass vials (Chromacol Ltd, Welwyn Garden City, UK) with 0.1 ml glass micro-inserts (VWR International Ltd, Leicester, UK), were placed 100 µL aliquots of 0.2 mg/ml t-PA within each formulation (Table 1 and Table S1), and then incubated at 45 °C. Forty vials of each formulation allowed for 4 replicates at each of 10 time-points. Mass measurements were taken before and after 45 °C incubation to evaluate for any potential evaporation.

Chromogenic Substrate Activity Assay

The S-2288 substrate (Ile-Pro-Arg-pNA, Chromogenix, Milan, Italy) was supplied by Quadratech Diagnostics Ltd (Epsom, UK) and used in the amidolytic activity assay for 0.2 mg/ml t-PA within the sacrificial time-point samples of given formulations subjected to 45 °C isothermal degradation. The assay buffer contained 100 mM Tris-base adjusted to pH 8.4 at 25 °C, and 106 mM NaCl. Assays were performed in 96-well plates, with 0.6 mM S-2288 as the final substrate concentration, and 667 ng/ml final enzyme concentration. Assays were conducted at 37 °C, with absorbance readings at

405 nm (A_{405}) taken using a FluoStar Optima plate reader (BMG Labtech GmbH, Ortenberg, Germany). Changes in A_{405} over time were fitted to linear decays in OriginPro 9.0 to give initial enzyme velocities (ΔA_{405} /min). Those initial enzyme velocities were plotted as a function of time (in days) incubated at 45 °C. These degradation kinetics were fitted where possible to equation 6 in OriginPro 9.0, to confirm that kinetics obeyed single exponential decays. All data were also fitted to a linear decay equation in the initial 20% of decay to enable a full comparison of the slowest with the fastest rates (in % day^{-1}).

$$y = y0 + Ae^{-kt} \qquad \text{Equation 6}$$

HPLC Size Exclusion Chromatography (SEC)

A TSK 3000 SWXL SEC column (dimensions 30 x 0.8 cm) plus guard column (dimensions 7 x 0.8 cm) on an Agilent 1200 HPLC machine was used to measure monomer and aggregation levels for t-PA throughout the 45 °C isothermal degradation. The mobile phase was 0.25 M sodium phosphate pH 7.0 and flow rate was 0.5 ml/min. Signal was detected at both 220 nm and 280 nm. Plots of % monomer versus time (in days at 45 °C) were used to obtain degradation rates as above.

RESULTS AND DISCUSSION

We first set out to map out conformational stability to unfolding, over a similar, though wider pH range to that reported previously for tPA using DSC [19]. In addition, we measured thermally-induced onset of aggregation, unfolding cooperativity, and aggregation kinetics at 45 °C. This aimed to confirm the complex

unfolding behaviour of tPA using orthogonal methods to DSC, relate tPA domain-unfolding to aggregation behaviour, and then also define two pH conditions to be explored further in complex formulations.

Effect of pH on thermal indices of global conformational stability and aggregation onset

Thermal ramping was used to compare the apparent thermal-transition midpoints ($T_{m,app}$), associated ΔS_{vh} for unfolding, and also the temperatures of unfolding onset (T_{onset}), and of aggregation onset (T_{agg}), for t-PA over a wide pH range (3.5-10.0), in the presence of 75 mM arginine (Table 1). $T_{m,app}$, ΔS_{vh} and T_{onset} were determined from the intrinsic protein fluorescence (350:330 nm emission ratio), which showed only one discernable transition in all cases. T_{agg} values were obtained simultaneously using static light-scattering at 266 nm, which is disproportionately more sensitive to larger aggregates than to the monomeric protein, and therefore provides a sensitive indication of the temperature at which the instrument first detects a small population of aggregates. All samples contained 75 mM arginine as this is a standard stabilising excipient for retaining t-PA solubility in commercial formulations [27].

Table 1 Effect of pH on thermal and kinetic parameters for 0.2 mg/ml t-PA, in the presence of 75 mM arginine.

Buffer	pH	T_{onset} (°C)	$T_{m,app}$ (°C)	T_{agg} (°C)	ΔS_{vh} (kJ mol^{-1} K^{-1})	f_{45}	ln v_{mon} (% day^{-1})	ln v_{act} (% day^{-1})
50mM Na Formate	3.5	22.3 (0.4)	47.5 (0.5)	44 (0.5)	0.37 (0.04)	0.41 (0.02)	*nd*	2.3 (0.2)
50mM Na Acetate	4.0	39.1 (0.3)	54.4 (0.4)	43 (0.5)	0.66 (0.06)	0.09 (0.02)	0.34 (0.3)	0.6 (0.3)

50mM Na Acetate	5.0	54.6 (0.3)	64.9 (0.3)	50 (0.5)	1.02 (0.09)	4.4 (3.0) x10^{-4}	0.03 (0.4)	-0.3 (0.3)
50mM Na Phosphate Citrate	6.0	62.0 (0.2)	68.6 (0.2)	55 (0.5)	1.6 (0.2)	4.6 (8.4) x10^{-7}	-0.7 (0.1)	-0.3 (0.4)
50mM Na Phosphate	7.0	59.2 (0.3)	67.9 (0.3)	63 (0.5)	1.2 (0.2)	2.2 (4.3) x10^{-5}	-0.39 (0.08)	-0.4 (0.5)
50mM HEPES	8.0	60.2 (0.3)	67.1 (0.3)	64 (0.5)	1.6 (0.3)	2.0 (5.4) x10^{-6}	-0.27 (0.09)	-0.7 (0.4)
50mM CAPS	10.0	61.0 (0.4)	65.1 (0.4)	65 (0.5)	2.7 (0.9)	1.7 (5.2) x10^{-9}	*nd*	-0.05 (0.1)

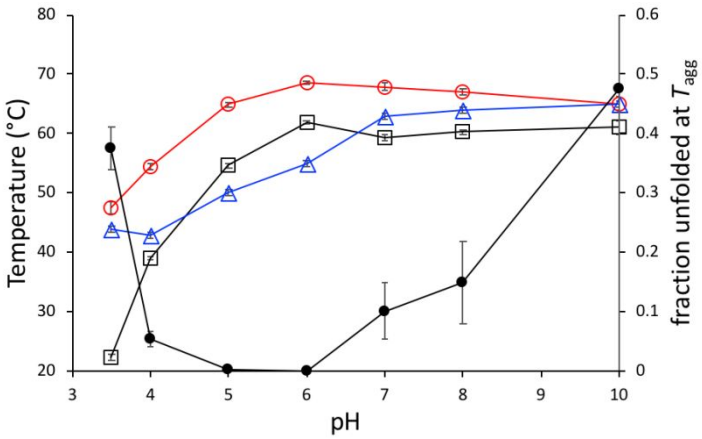
T_{onset}, $T_{m,app}$, and T_{agg} temperatures, ΔS_{vh} for unfolding, and fraction unfolded, at 45 °C (f_{45}). Initial rates (v) of monomer and activity loss (% day^{-1}) at 45 °C, expressed as ln v_{mon} and ln v_{act}, respectively. Standard errors of the mean are shown in parentheses.

tPA is known to undergo multiple overlapping domain-unfolding events [19]. As these are convolved into an apparent single transition, the data can be fitted to a two-state transition equation, and the $T_{m,app}$ obtained is the apparent thermal transition midpoint. Unfolding of tPA is also known to be partially reversible, where the aggregation rate is slow relative to the unfolding rate during the initial thermal ramping at 1 °C /min, but faster above the unfolding transition [19]. Hence, provided that the temperature ramping rate of 1 °C /min is used, $T_{m,app}$ will report on global conformational stability. The accompanying ΔS_{vh} for this transition measures the unfolding cooperativity, and would be expected to increase under conditions where the underlying multiple domain-unfolding events become more concerted. The fraction unfolded, f_T, can also be calculated from $T_{m,app}$ and ΔS_{vh}, at any temperature. T_{onset} values determined from intrinsic fluorescence occur below the $T_{m,app}$, are less

convoluted with aggregation events, and provide a robust indication of the temperature at which the instrument first detects protein unfolding (defined here as 2% unfolding).

It can be seen from Figure 1A, that *thermal-ramping induced* aggregation for tPA (with 75 mM arginine) was pH-dependent, as observed previously for many other proteins including GCSF [13], A33 Fab [12], α-chymotrypsinogen [8], β_2-microglobulin [36], γS-crystallin [37], and AS-IgG1 [38]. The tPA was most thermodynamically stable at pH 6, with the highest values of $T_{m,app}$ = 68.6 °C, and T_{onset} = 62 °C, and the lowest fraction unfolded at both T_{agg} ($2.9x10^{-4}$) and at 45 °C ($4.6x10^{-7}$) which was the temperature at which kinetic inactivation studies were carried out as reported below. $T_{m,app}$ decreased by 3.5 °C overall as the pH was increased from 6 to 10. These were consistent with T_m values in the literature [19], although slightly lower than DSC-derived T_m measurements of 71-73 °C.

A



B

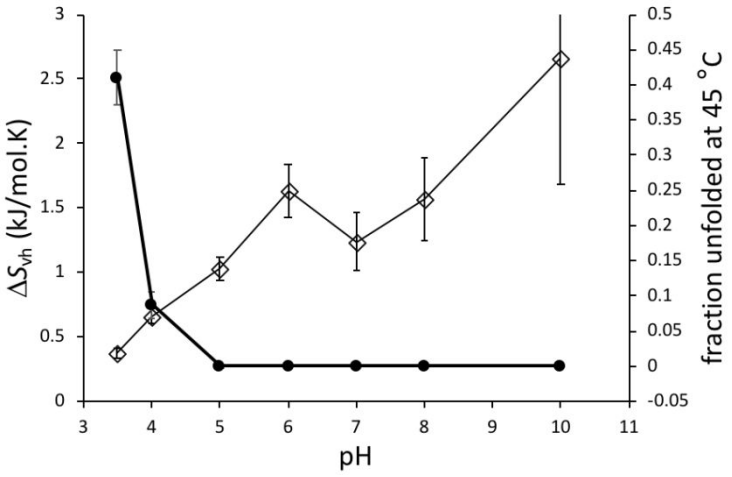


Figure 1 Dependence on thermal parameters on pH, for 0.2 mg/mL tPA in 75 mM Arginine. A) (Δ) T_{agg}, (○) $T_{m,app}$ and (□) T_{onset} temperatures, and (●) fraction unfolded at T_{agg}, and B) (●) fraction unfolded at 45 °C, and (◊) unfolding transition ΔS_{vh}.

T_{agg} increased above pH 4, reaching a plateau at pH 7-10, of 63-65 °C. At pH 4.0-8.0, T_{agg} remained clearly below $T_{m,app}$, with f_{Tagg} < 15%. At pH 5 and pH 6 in particular, T_{agg} was even below T_{onset}, with f_{Tagg} <0.3%, such that the *thermal ramping-induced* aggregation could be detected by light scattering prior to significant unfolding of the protein population, indicating that the initial onset of aggregation induced by thermal ramping was not driven by global protein unfolding under those conditions.

As the pH was decreased from 6.0 to 3.5, T_{agg}, $T_{m,app}$ and T_{onset} values all decreased sharply, although T_{agg} did not decrease further at pH 3.5, but almost converged with $T_{m,app}$ such that tPA was 37% unfolded at T_{agg}. Hence at low pH, tPA began to unfold at least 20 °C below the temperature at which aggregation was detected, allowing unfolded protein to accumulate as the temperature increased from 23 °C to 44 °C. The

effect at pH 10 was similar, with the accumulation of 48% unfolded protein prior to the detection of aggregates.

Overall, the data show that thermal ramping-induced aggregation of tPA initiates from a native population at pH 5-6. This implicates a low population of an aggregation-prone native-like state at equilibrium, or otherwise the frequent collision of native monomers which on rare occasions undergo conformational changes that induce aggregation. Such mechanisms have been implicated or observed previously for several other proteins [2, 12, 13, 39], and represents the key route to aggregation for therapeutic proteins that are stored under fully-native conditions. By contrast, tPA apparently aggregates from a substantially unfolded population at pH 3.5 and pH 10. Unfolding at the extremes of pH is expected. However, the higher net charge on the protein under these conditions, and particularly on the unfolded protein, would also be expected to increase their colloidal stabilisation, thus allowing more unfolding to occur prior to aggregation. The theoretical pI of tPA is 8.1 when calculated from the protein sequence using the ExPASy Compute pI tool [40], although the folded structure, microheterogeneity of glycosylation [41], and also interactions with 75 mM arginine and the ions in solution are each likely to change this value. Colloidal stabilisation is evident in the upturn in T_{agg} at pH 3.5, and also in the convergence of T_{agg} and $T_{m,app}$ at both pH 3.5 and pH 10. The T_{agg} was higher at pH 10 than at pH 3.5, most likely because the conformational stability of tPA was lower at pH 3.5 than at pH 10, as reflected in the $T_{m,app}$. Therefore, at both pH 3.5 and pH 10, colloidal stabilisation minimised aggregation from native-like states, and thus significant global unfolding

was also required before heat-induced aggregates could form at T_{agg}. Similar behaviours were observed previously for A33 Fab and GCSF [12, 13]. The role of unfolding in the kinetics of aggregation at 45 °C, will be investigated further below.

In Figure 1B it can be seen that the ΔS_{vh} for unfolding decreased at lower pH, indicative of a loss of unfolding cooperativity, and correlated with the decrease in $T_{m,app}$ at low pH. Cooperativity was greatest at pH 6.0-10.0. The decreased ΔS_{vh} indicated greater conformational flexibility in the native ensemble, consistent with differential scanning calorimetry (DSC) data [19] in which four unfolding events were deconvoluted at pH 7.4, and then five at pH 4.0 and pH 3.4, along with the same overall drop in $T_{m,app}$. The additional transition at pH 4.0, which had the lowest melting-temperature, was attributed previously to early unfolding of the N-terminal domain within the two-domain serine protease module by comparing the DSC melting profiles of individual domain fragments [19]. Such events may have direct implications on the aggregation propensity of tPA.

To confirm that the observed pH-dependence of tPA unfolding cooperativity in 75 mM arginine was correlated to independent domain unfolding, we used DSF, as an additional approach. The DSF thermograms for tPA at pH 3.5-10.0 are shown in Figure 2, and were fitted initially to 2-state sigmoidal transitions to obtain $T_{m,DSF}$ values for the first transition, but were then also fitted to up to three Lorentzian peaks to determine the contributions of different species populations to the overall signal, along with the temperatures (T_{max}) at which these reached their maximum. Addition of further Lorentzian peaks did not improve the overall fits, most likely due to the very

close proximity of the multiple transitions as deconvoluted by DSC [19]. At pH 5-8, there were two discernible populations, with maximum temperature of the second, T_{max2} typically 8 °C higher than that of the first peak, T_{max1}. These were equally populated and were most likely to correspond to dye-binding to at least two of the four native or near-native forms with one or more domains unfolded observed previously by DSC. The $T_{m,app}$ from intrinsic fluorescence measurements correlated strongly to $T_{m,DSF}$ (R^2>0.98, slope 1.27), and also to both T_{max1} (R^2>0.99, slope 1.35) and T_{max2} (R^2=0.88, slope 0.8), over the entire pH range 3.5-10.0 (see Fig S1 in SI). This indicated very good comparability between the orthogonal techniques.

At least one further fluorescence population emerged by DSF at pH 3.5, 4.0, and 10.0, indicating an additional domain unfolding event at these pHs, consistent with the additional state observed previously by DSC. At the same time, the relative signal from population 2 (with T_{max2}) decreased, and at low pH that of population 1 (with T_{max1}) increased, indicating that the relative population of the first two resolved native-like forms was pH-dependent. The curves at these pH-extremes also gave much wider peaks, indicative of multiple transitions and decreased cooperativity of unfolding, probably due to the two additional states resolved previously by DSC under all conditions. Thus, DSF profiles were consistent with the cooperativity changes determined above as ΔS_{vh} by intrinsic fluorescence, and demonstrated that the loss of unfolding cooperativity at the extremes of pH, was due to an increased number of unfolding transitions, that were not resolved directly by intrinsic protein fluorescence.

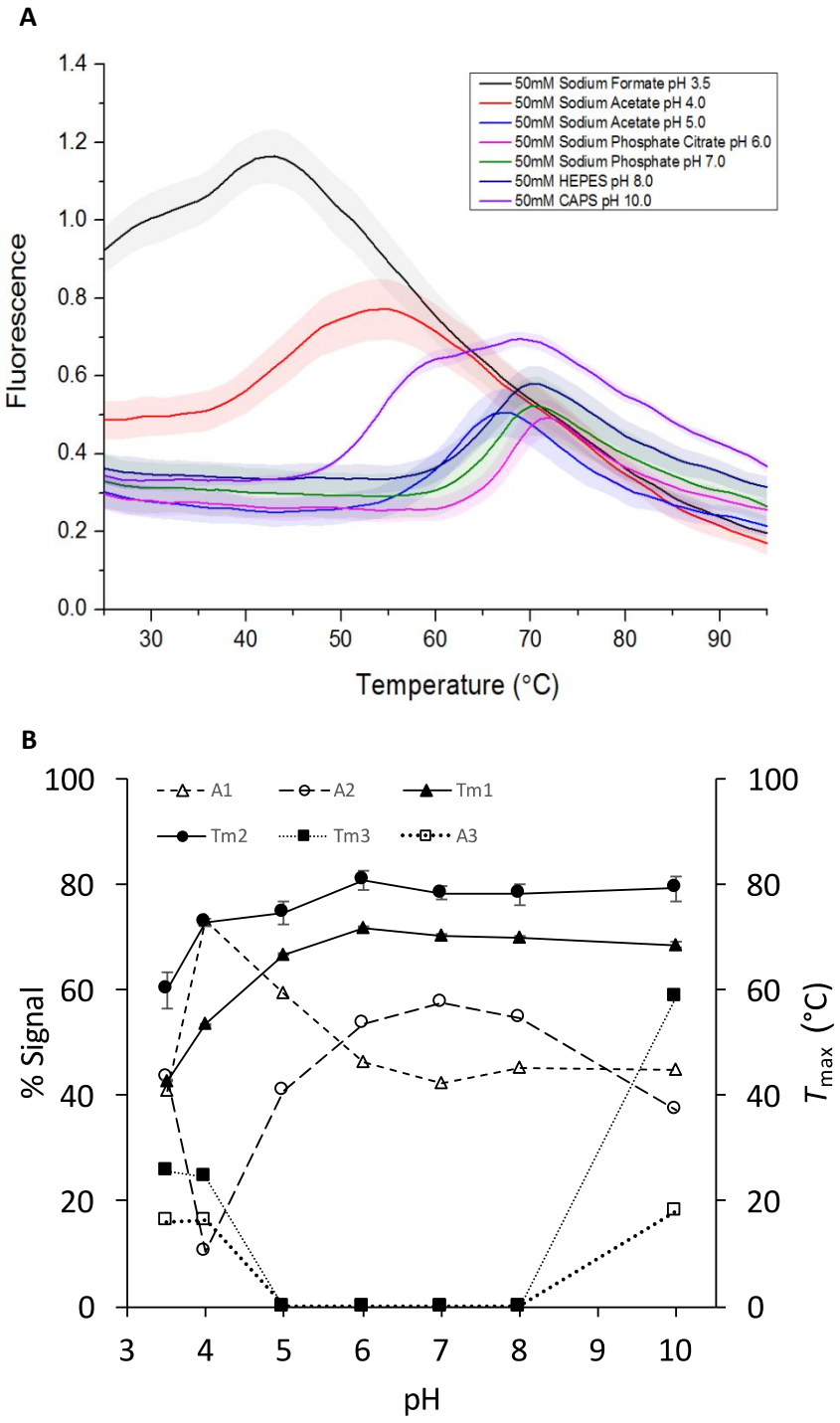


Figure 2 pH-dependence of DSF profiles for 0.2 mg/ml t-PA in 50 mM buffers containing 75 mM arginine. (A) Thermograms obtained at (black) pH 3.5, (red) pH 4.0, (neon blue) pH 5.0, (pink) pH 6.0, (green) pH 7.0, (dark blue) pH 8.0, (purple) pH 10.0; (B) T_{max} values for each population (▲) T_m1, (●) T_m2, (■) T_m3, and their

respective % contributions (Δ) A1, (○) A2 and (□) A3, to the total fluorescence, obtained by fitting DSF profiles to 2-3 Lorentzian peaks.

At pH 10, DSF also indicated an increase from two to at least three t-PA populations during unfolding, and yet the ΔS_{vh} from intrinsic fluorescence indicated an increased unfolding cooperativity, along with a 2 °C decrease in T_m, relative to those at pH 8.0. The different observations could be due to the lack of a tyrosine or tryptophan reporter probe in the region of structure that DSF found to unfold earlier at pH 10. Tryptophan is distributed throughout most of the tPA domains, with five in the serine protease module, three in each kringle domain, and two in the Fibronectin type-I domain (see Fig S2 in SI), and so intrinsic fluorescence reports strongly on the global unfolding of these domains. However, the EGF domain has no tryptophans and only one tyrosine residue, and so would not produce a significant change in intrinsic fluorescence upon early unfolding of this domain. Thus, decoupling of the EGF domain from the remaining structure at pH 10, and its early unfolding as observed by DSF, appears to have led to an increase in the unfolding cooperativity of the remaining structure as observed by intrinsic fluorescence, while the overall T_m decreased by 2 °C. The previous DSC results [19] suggested that at pH 3.5, the SP-N domain formed the first unfolding transition (T_m = 45.9 °C), but that at pH 7.4 the K2 domain unfolded first (T_m = 59.9 °C) with the SP-N domain unfolding at a higher temperature (T_m = 70.9°C) due to domain-domain stabilisation. Thus, it is conceivable that the EGF domain which was assumed to be the most stable at pH 7.4 in the absence of arginine [19], could

become the least stable at pH 10 with 75 mM arginine. Either way, the loss of cooperativity observed from intrinsic fluorescence at pH 3.5 but not at pH 10, strongly indicates that a different structure unfolds first at pH 3.5 (one containing tryptophan or tyrosines), to that which unfolds first at pH 10.

Impact of pH on aggregation/degradation kinetics

The kinetics of isothermal degradation for 0.2 mg/ml t-PA, at pH 3.5-10.0, with 75 mM arginine, were determined by storing t-PA at 45 °C and removing samples at different time points for HPLC-SE and bioactivity analyses. Monomer and activity loss each fitted well to single exponential kinetics in cases that were fast enough to reach zero in the time studied, indicating a monomolecular rate limiting step under these conditions. However, to allow a complete comparison, all data were fitted to linear equations to determine initial rates from slopes within the first 20% of monomer or activity loss. The pH-dependence of the rates for monomer and activity loss at 45 °C, are summarised in Table 1 and Figure 3.

The initial rates of monomer loss (ln v_{mon}) and activity loss (ln v_{act}) at 45 °C, were coincident within error over the range pH 4-8, indicating that unfolding of the activity-bearing SP-N domain at pH 4 was not faster than the rate limiting step for monomer loss and aggregation. One possibility is that aggregation was not caused by the SP-N domain unfolding, but that aggregation caused the loss of serine protease activity. However, the rate of monomer/activity loss was also found to correlate well to the T_m from intrinsic fluorescence (Figure 3B), and T_{max1} and $T_{m,DSF}$ from DSF, which all decreased at pH 4 due to the early unfolding of the N-terminal domain of the serine

protease observed previously [19]. Therefore, it is most likely that at pH 4 the unfolding of the SP-N domain led directly to aggregation, and the observed increase in the rate of aggregation at 45 °C.

The degradation rates were lowest at pH 6-8, and increased significantly at below pH 5, but less so at above pH 8. These observations were consistent with the decreased T_m (Figure 3B), and also the increased fraction unfolded f_{45} (Figure 3C), at low pH, leading to both loss of activity and aggregation of monomer. Colloidal stability arising from greater electrostatic repulsion at high pH did not seem to be a dominant factor under these conditions at 45 °C, as the rate of degradation did not decrease, but even increased slightly, despite retaining a folded structure at 45 °C as determined by both intrinsic fluorescence (Figure 1B) and DSF (Figure 2A). At low pH, colloidal stabilisation might have occurred, but any effects would be obscured by the more dominant and aggregation-accelerating impact of global unfolding, particularly with the early SP-N domain unfolding. Colloidal stabilisation effects did manifest at the T_{agg} for both low and high pH (Figure 1), as described above, whereby the fraction unfolded at T_{agg} (f_{Tagg}) was significantly increased.

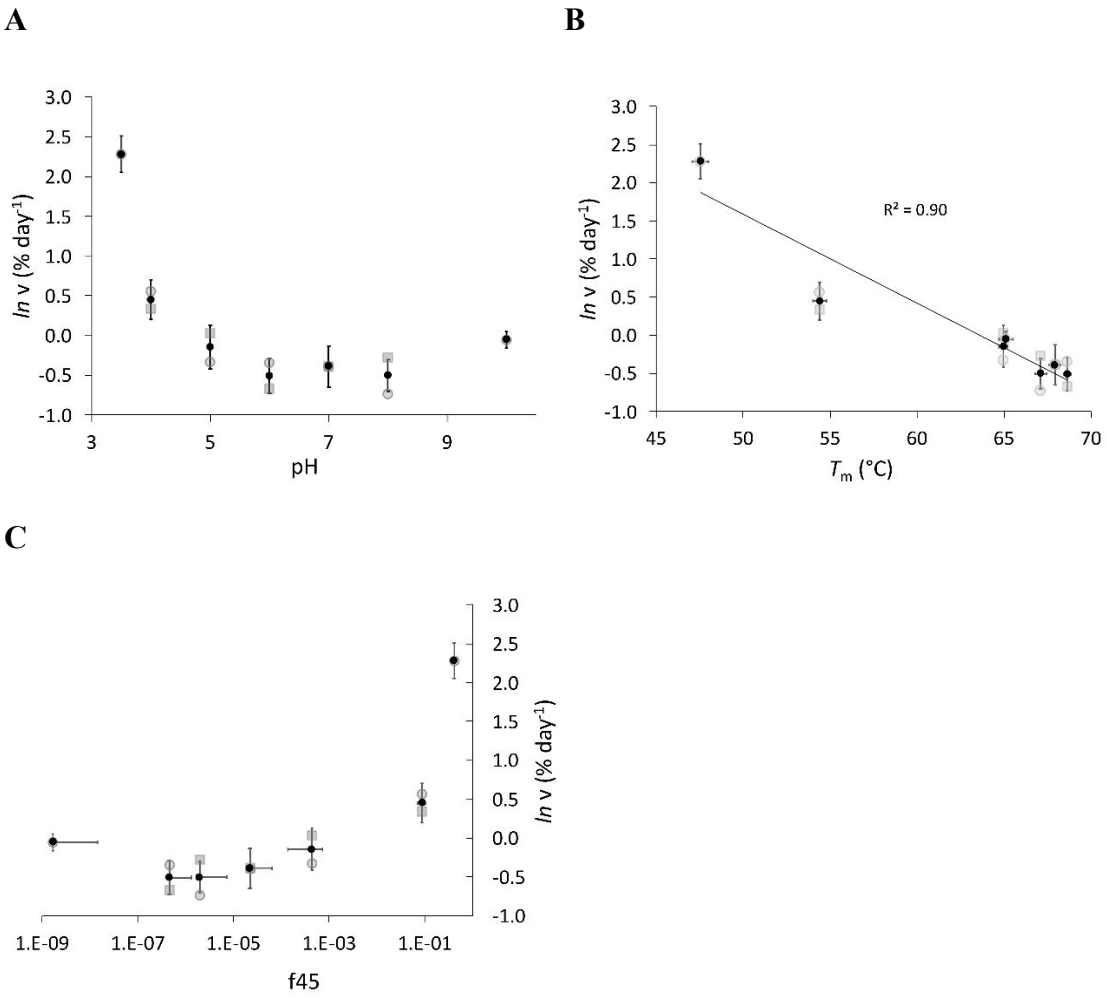


Figure 3 Rates of monomer and activity loss at 45 °C, for 0.2 mg/ml t-PA in 50 mM buffers, pH 3.5-10, containing 75 mM arginine. (□) Monomer loss, (o) activity loss, and (●) the average (ln v_{av}) from both measures. (A) Effect of pH on ln v. (B) Correlation between ln v and $T_{m,app}$. (C) Correlation between ln v and the fraction unfolded at 45 °C, f_{45}.

Impact of complex formulations on t-PA thermostability and aggregation kinetics

In earlier work with the more globular Fab, and single-domain protein GCSF, we found that the rate of monomer loss was dependent on the fraction of protein unfolded at the incubation temperature (f_T) until this value reached less than approximately 0.1% [12, 13]. At that point a baseline rate of monomer loss was reached, in which aggregation occurred directly from the native state or via local unfolding within the native ensemble. To date this has been found to hold true when varying the formulation conditions (pH or excipient addition), or when altering the global protein stability, or local flexibility through single mutations [12-14]. We therefore aimed to determine whether the same phenomenon existed for a wide range of formulations of tPA, which has the added complexity of non-cooperative unfolding between multiple domains, and at least one domain unfolding earlier at high and low pH conditions.

A DoE approach was used to construct a range of complex formulations from 19 factors including pH, arginine and other excipients. As commercial tPA is formulated in arginine, we first examined the influence of arginine concentration on the rate of activity loss at pH 8, in order to define the upper and lower concentrations to use in DoE. We found that the rate of activity loss was lowest at 50-75 mM, consistent with commercial formulations (see Fig S3 in SI), and so we used 75 mM in DoE to match commercial formulations. We also selected pH 4 and pH 7 for DoE as these produced very different behaviours in T_m, T_{agg}, ΔS_{vh}, and ln v, and yet had similar fractions unfolded at T_{agg}, while both avoided the most extreme pH values for which colloidal effects were observable on T_{agg}. A crucial difference was the early unfolding of the SP-N domain at pH 4 compared to pH 7.

Minitab 17 software (Minitab Inc, State College, PA, USA) was used to construct a 2-level, resolution III, fractional factorial DoE study. Thus 32 formulations of 0.2 mg/ml tPA, (see Table S1 in SI), were generated from the 19 factors of pH (4.0 or 7.0), 0-50 mM glycine, glutamic acid, proline, and methionine; 25-75 mM arginine; 110-160 mM ionic strength; 0.1% (v/v) Tween-20 and Tween-80; 0-2.5% (w/v) sucrose, trehalose, mannitol, and glucose; and 0-3 mg/mL sorbitol, glycerol, ethanol, PEG 400, PEG 2000 and PEG 6000. For all formulations, the T_m, T_{onset}, ΔS_{vh}, T_{agg} and f_{45} values, as well as the rates of activity loss at 45 °C, were obtained as above. Examples of two-state fitting to obtain T_m and ΔS_{vh} are shown in Figure SI8 (supplementary information).

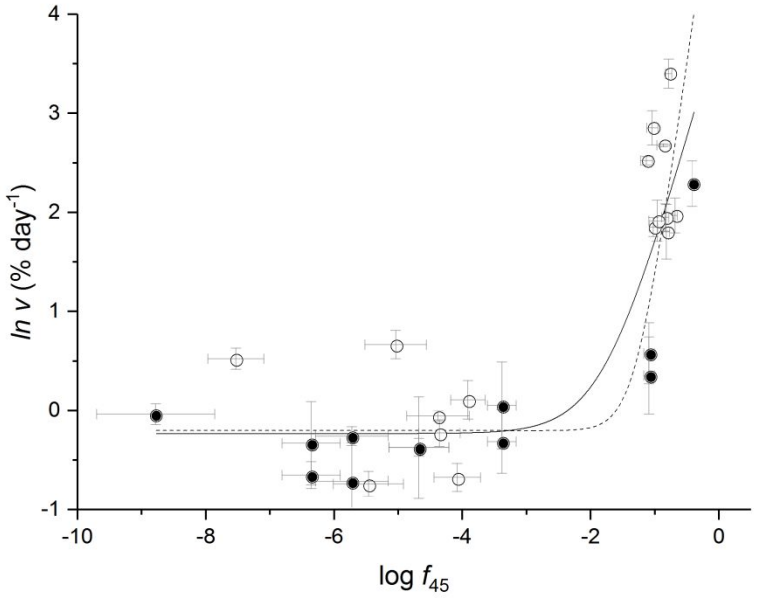


Figure 4 Effect of fraction unfolded on rates of activity/monomer loss at 45 °C, for 0.2 mg/ml t-PA in all formulations. Formulations at pH 4-8 in 50 mM buffers containing 75 mM arginine (●), DoE formulations at pH 4 and pH 7 (o). Fits to data

are shown for two models with (—) monomolecular, R^2 0.76, and (---) bimolecular, R^2 0.68, aggregation kinetics from the unfolded population. See SI for more details.

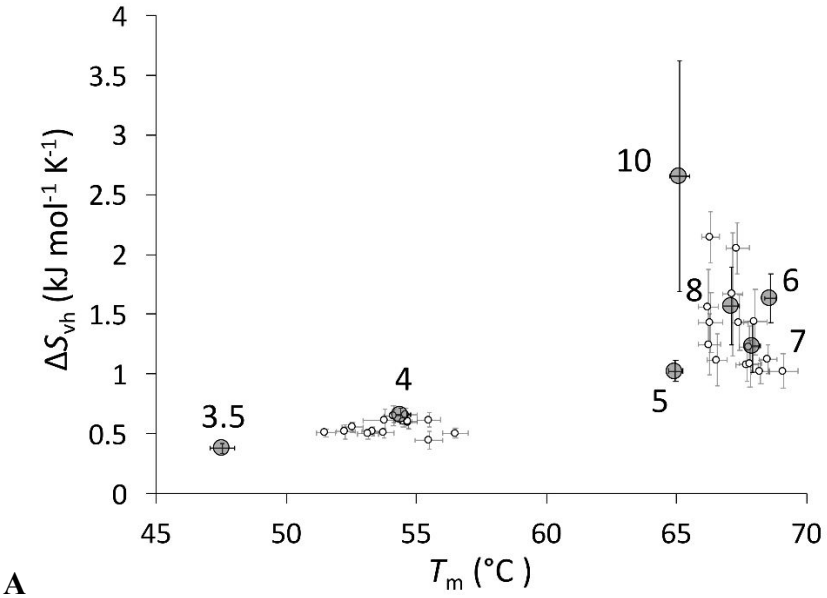
When the data from the DoE formulations were combined with the pH dependent data above, the rate of activity loss (as ln v) was found to correlate linearly with T_{onset} ($R^2 = 0.78$), T_m ($R^2 = 0.77$), and ΔS_{vh} ($R^2 = 0.71$), but not to T_{agg} ($R^2 = 0.01$). The rate of activity loss at 45 °C, was dependent on the fraction unfolded at 45 °C, but in a function where the rate of activity loss reached a baseline minimum value under conditions where less than 0.1 % was globally unfolded, *i.e.* $f_{45} < 0.001$ (Figure 4). Similar behaviour was observed recently for both A33 Fab and GCSF [12-14], and demonstrates that global unfolding does not contribute to differences in aggregation kinetics under conditions that favour the native ensemble. The data in Figure 4 were fitted to several kinetic models (See SI), as we have previously for Fab and GCSF [12, 13]. Models in which tPA aggregation under predominantly unfolding conditions ($f_{45} > 0.01$) was either monomolecular ($R^2 = 0.76$), or bimolecular diffusion-limited ($R^2 = 0.68$), both gave reasonable fits. Not surprisingly, all of the pH 7 formulations but none of the pH 4 formulations had $f_{45} < 0.01$. Therefore, for the complex multi-domain protein tPA, the early global unfolding of a single domain at low pH played a key role in accelerating aggregation.

The significantly increased fraction unfolded at T_{agg} (f_{Tagg}) at both low and high pH (Figure 1), and yet the lack of correlation between T_{agg} and ln v (at $T < T_{agg}$), was also observed previously for GCSF [13]. If T_{agg} (and f_{Tagg}) are indirect measures of

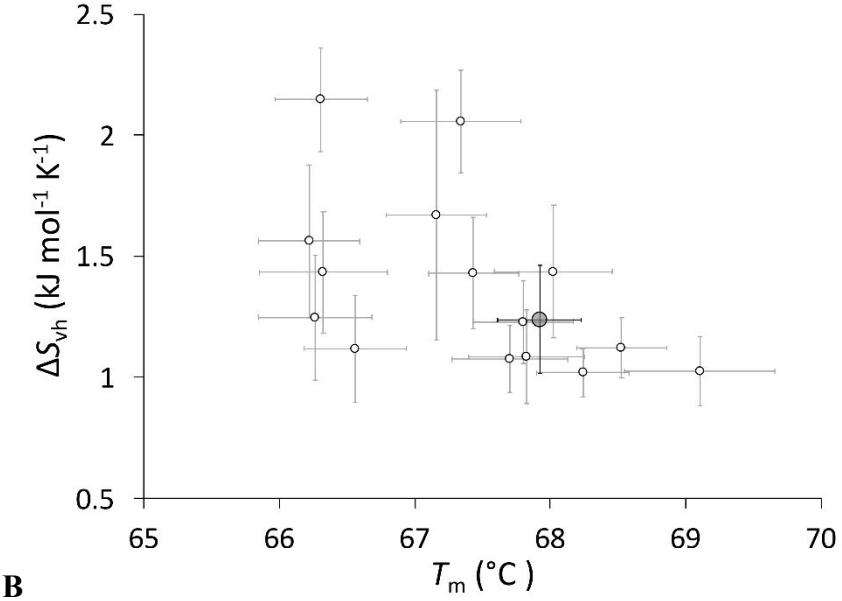
colloidal stability, then its lack of influence on the kinetics of aggregation at T below T_{agg} may be due to a more dominant role of global or partial unfolding as a rate-limiting step in aggregation.

Under predominantly native conditions (pH 7) the baseline aggregation rate in Figure 4 must be determined by a different mechanism than the early domain unfolding that occurred only at low and high pH. Previous work has shown that this is in part controlled colloidally by electrostatic and hydrophobic repulsions and attractions between native monomers [5, 40, 42, 43], and partly by local fluctuations within the native ensemble that potentially reveal aggregation-prone regions (APRs) [2, 14, 40, 43-46]. Interestingly, a plot of ΔS_{vh} vs T_m for all formulations of tPA at pH 7 (Figure 5), revealed that any increases in T_m were greatest when ΔS_{vh} was at the lower end (approx. 1 kJ mol^{-1} K^{-1}) of the range (1-2.1 kJ mol^{-1} K^{-1}) observed at pH 7. This implied a loss of unfolding cooperativity due to the selective *stabilisation* of at least one of the more stable domains, relative to the others. We have observed similar behaviour previously as a result of targeted mutations in the enzyme transketolase designed to stabilise local structural flexibility [47], but not previously in response to formulation excipients. This behaviour contrasts with the loss of ΔS_{vh} and T_m for tPA at pH 4, which was due to selective *destabilisation* of at least one domain. It was also different to the increase in ΔS_{vh} that accompanied the 2 °C decrease in T_m at pH 10 discussed above, which selectively destabilised the EGF domain, and resulted in an increase in unfolding cooperativity for the remaining structure upon decoupling from the less stable EGF domain. However, some formulations at pH 7 did follow this latter

pathway as seen by their increased ΔS_{vh} and decreased T_m and proximity to the pH 10 formulation in Figure 5. These various different effects relating T_m and ΔS_{vh} to the overall changes in the thermal denaturation profile of tPA, are outlined in Scheme 1.



A



B

Figure 5 Correlation between T_m and ΔS_{vh} for (A) all formulations of tPA, and (B) at pH 7 only. Formulations are plotted separately for (●) the pH 3.5-10 study containing 75 mM arginine, and (o) for the DoE-based formulations at pH 4 and pH 7.

None of the formulations at pH 7 achieved a T_m >70 °C and ΔS_{vh} >1.5 kJ mol^{-1} K^{-1}. Therefore, further optimisation at pH 7 should aim to improve T_m and ΔS_{vh} simultaneously, as this would lead to increased stability and ensure that no single domain is more susceptible to unfolding than the others. This contrasts with our previously reported strategy whereby an already highly stable Fab antibody fragment was engineered for slower aggregation kinetics by increasing only the ΔS_{vh} [14]. The Fab contained four domains with significant structural interactions, such that the domains themselves do not unfold independently. By contrast, the independent unfolding of one or more domains in the multi-domain architecture of tPA, makes the simultaneous improvement of T_m and ΔS_{vh} a more favourable strategy.

To gain more insight into which excipients cause the local stabilisation of one or more domains in tPA, we analysed the formulation data using DoE statistical tools, to determine the factors that most likely influenced the measured outputs T_m, T_{onset}, ΔS_{vh}, T_{agg}, f_{Tagg} and f_{45} (See full analysis in SI). It should be noted that as the DoE was at resolution III, the effects of each single factor could potentially be confounded by the effects of pairwise interactions.

T_{agg} was affected by PEG400 (positive) and glycerol (negative), and the influence of other polyols (PEG 2000, PEG 6000) decreased with their increasing molecular weight (Figure SI6). f_{Tagg} was primarily influenced by pH, and then, in order of priority by PEG400, glycerol and trehalose (Figure SI7). High T_{agg} was preferred by high PEG400, but low [glycerol]. Correspondingly, a low f_{Tagg} was preferred by low PEG400 and high glycerol. While pH affected f_{Tagg}, it did not impact T_{agg}. This was

interesting given that pH 4 and pH 7 were selected earlier in part as they had different T_{agg} but the same f_{Tagg}. In practice, the formulations introduced the same degree of variation and overall range in T_{agg} at either pH, and so pH was not a major influence on T_{agg}. By contrast the formulations created a wider distribution in f_{Tagg} at pH 4 than at pH 7, thus leading to an overall influence of pH on f_{Tagg} where the average was higher at pH 4. PEG400 increased the colloidal stability of tPA, whereby T_{agg} and f_{Tagg} increased, thus allowing more unfolding to occur before aggregation commenced. Glycerol had the opposite effect.

T_m, T_{onset}, and f_{45} were predominantly influenced by pH (Figure SI8), which confirms the behaviour observed in Figure 3. However, many other factors also contributed to the observed additional variability in T_m, T_{onset}, and f_{45} for the formulations at each pH, although at the concentrations explored, these contributions were small compared to that from pH. Most formulations at pH 7 lowered the T_m, but either maintained or increased ΔS_{vh} compared to the sample at pH 7 with 75 mM Arginine (Figure SI9). Specifically, the high pH, and addition of sorbitol and glucose each increased ΔS_{vh}, and hence unfolding cooperativity. However, these also decreased T_m slightly. Thus, at pH 7, sorbitol and glucose caused the selective *destabilisation* of one or more of the most stable domains, so that their stability became more similar to the others. Sorbitol (3 mg/ml) and glucose (25 mg/ml) are monosaccharides of similar molecular weight, and have the same type of impact, hence they appear to have acted via a similar mechanism. This mechanism does not appear to be simply a macromolecular crowding or preferential hydration effect, as often

observed with sugars [48], as the two disaccharides trehalose (25 mg/ml) and sucrose (25 mg/ml) were both ineffective at modifying ΔS_{vh}. Alternatively, this unusual finding could potentially be the result of a hidden interaction with one of the other excipients, as such pairwise interactions would not have been resolved using the resolution III DoE approach.

The few pH 7 formulations that increased T_m and yet retained lower ΔS_{vh} values, did not contain either sorbitol or glucose. The factors that increased T_m were, in order of significance, PEG400, Tween 80, mannitol, proline, PEG2000, and glycine. The factors that decreased ΔS_{vh} were, in order of significance, PEG6000, arginine, glycine, PEG400, sucrose, ethanol, glycerol, methionine, Tween 20, and mannitol. Thus, the only excipients that both increased T_m and decreased ΔS_{vh} were PEG400, glycine, and mannitol. Therefore, these three excipients each led to selective stabilisation of at least one domain relative to the others. As this also led to a decrease in unfolding cooperativity overall, that domain must have been one of those already the most stable, thus making an even wider range of temperatures over which the various tPA domains unfolded.

While increasing the T_m is a commonly used formulation strategy where the goal is to improve aggregation kinetics, we have previously observed with an antibody Fab fragment, that increasing ΔS_{vh} can also slow aggregation as it reflects less conformational flexibility in the native ensemble [14]. This is particularly relevant where $T_m >> T$, the temperature at which aggregation kinetics are tested. None of the formulations created in the DoE design for tPA led to increases in both T_m and ΔS_{vh}.

However, the DoE experiments can be used to determine which factors should be explored further to achieve this, or to predict specific combinations of the factors studied might achieve it. The excipients predicted to increase both T_m and ΔS_{vh} were proline and Tween 80. Setting the pH to 7, high values of 50 mM proline and 0.1% (w/v) Tween 80, and zero for all other excipients, predicted a T_m > 69.0 °C, and ΔS_{vh} > 1.53 kJ mol^{-1} K^{-1}. For comparison the unformulated tPA at pH 7 has a T_m of 67.9 °C and ΔS_{vh} = 1.24 kJ mol^{-1} K^{-1}. Thus, proline and Tween 80 were able to selectively stabilise the least stable domains relative to those already more stable. Further increasing the concentrations of both proline and Tween 80 beyond the highest values tested should continue to maximise T_m and ΔS_{vh}. Alternatively, an optimisation constraining all factors within the experimental ranges tested, predicted a formulation with the same T_m as with proline and Tween 80 at pH 7 (T_m = 69.1 °C), but with a higher ΔS_{vh} = 1.8 kJ mol^{-1} K^{-1} (see SI).

The different selective behaviours of each excipient appear to be specific to particular domains, which then changes the stability of those domains relative to the others. The types of mechanism by which this could occur might involve specific interactions with the protein surface, or otherwise effects that are mediated via the solvent as a result of the different surface hydrophobicities and polarities of each domain.

CONCLUSION

Despite the increased complexity of tPA due to non-cooperative unfolding of its multiple domains, it gave a good correlation between T_m-values and the aggregation

kinetics (at a temperature close to typical T_m values). The dependency of aggregation kinetics upon the fraction unfolded as measured under a two-state assumption, also followed the same type of profile observed previously with GCSF and Fab, which unfold with much higher cooperativity. Low pH conditions induced early unfolding of the SP-N domain, and this led to accelerated aggregation kinetics at 45 °C. Therefore, despite the greater complexity of tPA compared to Fab and GCSF, its formulations should also initially aim to reduce the globally unfolded fraction to <1%, so that the aggregation kinetics are in the baseline region dependent only on the native-state ensemble. For tPA, this was essentially achieved by simply adjusting the pH to > 4, where $f_{45} < 0.01$. As observed previously with Fab and GCSF, further formulation with excipients resulted in a variation spanning 1 log-order in the aggregation kinetics in the region where <1% was unfolded (Figure 4). Evaluation of more complex formulations containing a range of excipients determined that the unfolding cooperativity of tPA, measured via ΔS_{vh} was also found to vary under native conditions at pH 7, and to correlate well to aggregation kinetics, but was only partially correlated to T_m. In particular, the different formulation excipients were able to affect ΔS_{vh} and T_m in different ways, indicating that some were able to selectively stabilise, and others to selectively destabilising specific tPA domains. For example, sorbitol and glucose led to the selective destabilisation of the most stable domains, whereas PEG400, glycine and mannitol led to the selective stabilisation of the most stable domains. Only proline and Tween 80 increased both T_m and ΔS_{vh}, through selective stabilisation of the least stable domains.

The formulation of tPA and similar multi-domain proteins may potentially benefit from multi-objective optimisation, such as to maximise both T_m and ΔS_{vh} to maintain high stability and also high unfolding cooperativity. Mechanistically, this work opens up the potential to add excipients that specifically target local mechanisms involved in aggregation, such as the suppression of partial domain unfolding through excipients binding to surface hotspots within that domain. This could in future be rationally designed via molecular docking protocols as previously demonstrated for Fab [49].

ASSOCIATED CONTENT

Supporting Information

Correlations between T_m values derived by DSF and intrinsic-fluorescence; amino-acid sequence of tPA in Alteplase; effect of arginine concentration on rates of activity loss at 45 °C; fitting of aggregation kinetics to different models; DoE formulation design and analysis.

ACKNOWLEDGEMENT

The support of the BBSRC BRIC studentship (BB/J003824/1), and Engineering and Physical Sciences Research Council (EPSRC) Future Targeted Healthcare Manufacturing Hub (EP/P006485/1), is gratefully acknowledged. The Hub is part of the Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering, University College London. The authors thank Kiran Malik (NIBSC) for assistance with the DSF.

REFERENCE

(1) Aggarwal, S. What's fueling the biotech engine-2012 to 2013. *Nat. Biotechnol.* **2014**, *32* (1), 32-39.

(2) Espargaró, A.; Castillo, V.; de Groot, N. S.; Ventura, S. The *in Vivo* and *in Vitro* Aggregation Properties of Globular Proteins Correlate With Their Conformational Stability: The SH3 Case. *J. Mol. Biol.* **2008**, *378* (5), 1116-1131.

(3) Shukla, A. A.; Gupta, P.; Han, X. Protein aggregation kinetics during Protein A chromatography: Case study for an Fc fusion protein. *J. Chromatog. A* **2007**, *1171* (1–2), 22-28.

(4) Wang, W.; Nema, S.; Teagarden, D. Protein aggregation—Pathways and influencing factors. *Int. J. Pharma.* **2010**, *390* (2), 89-99.

(5) Jung, J.-M.; Savin G.; Pouzot, M.; Schmitt, C.; Mezzenga, R. Structure of Heat-Induced β-Lactoglobulin Aggregates and their Complexes with Sodium-Dodecyl Sulfate. *Biomacromolecules* **2008,** *9* (9), 2477-2486.

(6) Krebs, M. R. H.; Domike, K. R.; Donald, A. M. Protein aggregation: more than just fibrils. *Biochem. Soc. Transact.* **2009**, *37* (4), 682-686.

(7) Chi, E. Y.; Krishnan, S.; Kendrick, B. S.; Chang, B. S.; Carpenter, J. F.; Randolph, T. W. Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor. *Protein Sci.* **2003**, *12* (5),

903-913.

(8) Li, Y.; Ogunnaike, B. A.; Roberts, C. J. Multi-variate approach to global protein aggregation behavior and kinetics: Effects of pH, NaCl, and temperature for α-chymotrypsinogen A. *J. Pharm. Sci.* **2009**, *99* (2), 645-662.

(9) Roberts, C. J.; Das, T. K.; Sahin, E. Predicting solution aggregation rates for therapeutic proteins: Approaches and challenges. *Int. J. Pharm.* **2011**, *418* (2), 318-333.

(10) Moussa, E. M.; Panchal, J. P.; Moorthy, B. S.; Blum, J. S.; Joubert, M. K.; Narhi, L. O.; Topp, E. M. Immunogenicity of Therapeutic Protein Aggregates. *J. Pharm. Sci.* **2016**, *105* (2), 417–430.

(11) Maddux, N. R.; Iyer, V.; Cheng, W.; Youssef, A. M. K.; Joshi, S. B.; Volkin, D. B.; Ralston, J. P.; Winter, G.; Middaugh, R. High Throughput Prediction of the Long-Term Stability of Pharmaceutical Macromolecules from Short-Term Multi-Instrument Spectroscopic Data. *J. Pharm. Sci.* **2014**, *103* (3), 828-839.

(12) Chakroun, N.; Hilton, D.; Ahmad, S. S.; Platt, G. W.; Dalby, P. A. Mapping the Aggregation Kinetics of a Therapeutic Antibody Fragment. *Mol. Pharmaceutics* **2016**, *13* (2), 307–319.

(13) Robinson, M. J.; Matejtschuk, P.; Bristow, A. F.; Dalby, P. A. T_m-Values and Unfolded Fraction Can Predict Aggregation Rates for GCSF Variant Formulations, but Not under Predominantly Native Conditions. *Mol. Pharm.* **2018**, *15* (1), 256-267.

(14) Zhang, C.; Samad, M.; Yu, H.; Chakroun, N.; Dalby, P. A. Computational Design To

Reduce Conformational Flexibility and Aggregation Rates of an Antibody Fab Fragment. *Mol. Pharm.* **2018**, *15* (8), 3079-3092.

(15) Abe, Y.; Gor, J.; Bracewell, D. G.; Perkins, S. J.; Dalby, P. A. Masking of the Fc Region in Human IgG4 by Constrained X-ray Scattering Modelling: Implications for Antibody Function and Therapy. *Biochem. J.* **2010**, *432* (1), 101-114.

(16) Vermeer, A.W.; Norde, W. The Thermal Stability of Immunoglobulin: Unfolding and Aggregation of a Multi-domain Protein. *Biophys. J.* **2000**, *78* (1), 394-404.

(17) Andersen, C. B.; Manno, M.; Rischel, C.; Thorolfsson, M.; Martorana, V. Aggregation of a Multidomain Protein: A Coagulation Mechanism Governs Aggregation of a Model IgG1 Antibody Under Weak Thermal Stress. *Protein Sci.* **2010**, *19* (2), 279-290.

(18) Pennica, D.; Holmes, W. E.; Kohr, W. J.; Harkins, R. N.; Vehar, G. A.; Ward, C. A.; Bennett, W. F.; Yelverton, E.; Seeburg, P. H.; Heyneker, H. L.; Goeddel, D. V.; Collen, D. Cloning and Expression of Human Tissue-type Plasminogen Activator cDNA in *E. coli*. *Nature* **1983**, *301* (5897), 214-221.

(19) Novokhatny, V. V.; Ingham, K. C.; Medved, L. V. Domain Structure and Domain-Domain Interactions of Recombinant Tissue Plasminogen Activator. *J. Biol. Chem.* **1991**, *266* (20), 12994-13002.

(20) Pearlman, R.; Nguyen, T. Pharmaceutics of Protein Drugs. *J. Pharm. Pharmacol.* **1992**. *44* (Suppl 1), 178-185.

(21) Radek, J. T.; Castellino, F. J. A Differential Scanning Calorimetric Investigation of the

Domains of Recombinant Tissue Plasminogen Activator. *Arch. Biochem. Biophys.* **1988**, *267* (2), 776-786.

(22) Collen, D.; Lijnen, H. R. The Tissue-Type Plasminogen Activator Story. *Arterioscler. Thromb. Vasc. Biol.* **2009**, *29* (8), 1151-1155.

(23) Byeon, I. J.; Kelley, R. F.; Llinas, M. Kringle-2 Domain of the Tissue-Type Plasminogen Activator. *Eur. J. Biochem.* **1991**, *197* (1), 155-165.

(24) Renatus, M.; Bode, W.; Huber, R.; Sturzebecher, J.; Prasa, D.; Fischer, S.; Kohnert, U.; Stubbs, M. T. Structural Mapping of the Active Site Specificity Determinants of Human Tissue-type Plasminogen Activator. *J. Biol. Chem.* **1997**, *272* (35), 21713-21719.

(25) Korin, N.; Kanapathipillai, M.; Matthews, B. D.; Crescente, M.; Brill, A.; Mammoto, T.; Ghosh, K.; Jurek, S.; Bencherif, S. A.; Bhatta, D.; Coskun, A. U.; Feldman, C. L.; Wagner, D. D.; Ingber, D. E. Shear-Activated Nanotherapeutics for Drug Targeting to Obstructed Blood Vessels. *Science* **2012**, *337*, 738-742.

(26) Kim, J. Y.; Kim, J. K.; Park, J. S.; Byun, Y.; Kim, C. K.; The use of PEGylated Liposomes to Prolong Circulation Lifetimes of Tissue Plasminogen Activator. *Biomaterials* **2009**, *30*, (29), 5751-5756.

(27) Bennett, W. F.; Builder, S. E.; Gatlin, L. A. Stabilized human tissue plasminogen activator compositions. US 5,034,225.

(28) Cleary, S.; Mulkerrin, M. G.; Kelley, R. F. Purification and Characterization of Tissue Plasminogen Activator Kringle-2 Domain Expressed in *Escherichia coli. Biochemistry*

1989, *28* (4), 1884-1891.

(29) Rijken, D. C.; Collen, D. Purification and Characterization of the Plasminogen Activator Secreted by Human Melanoma Cells in Culture. *J. Biol. Chem.* **1981**, *256* (13), 7035-7041.

(30) Rijken, D. C.; Hoylaerts, M.; Collen, D. Fibrinolytic Properties of One-Chain and Two-Chain Human Extrinsic (Tissue-Type) Plasminogen Activator. *J. Biol. Chem.* **1982**, *257* (6), 2920-2925.

(31) Wallén, P.; Bergsdorf, N.; Rånby, M. Purification and Identification of Two Structural Variants of Porcine Tissue Plasminogen Activator by Affinity Adsorption on Fibrin. *Biochim. Biophys. Acta* **1982**, *719* (2), 318-328.

(32) Dagan, S.; Hagai, T.; Gavrilov, Y.; Kapon, R.; Levy, Y.; Reich, Z. Stabilization of a Protein Conferred by an Increase in Folded State Entropy. *Proc. Natl. Acad. Sci.* **2013**, *110* (26), 10628–10633.

(33) Shaw, G. J.; Sperling, M.; Meunier, J. M. Long-term Stability of Recombinant Tissue Plasminogen Activator at -80 C. *BMC Res. Notes* **2009**, *2*, 117.

(34) Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. How to Measure and Predict the Molar Absorption Coefficient of a Protein. *Protein Sci.* **1995**, *4*, 2411-2423.

(35) Malik, K.; Matejtschuk, P.; Thelwell, C.; Burns, C. J. Differential Scanning Fluorimetry: Rapid Screening of Formulations that Promote the Stability of Reference Preparations. *J. Pharm. Biomed. Anal.* **2013**, *77*, 163-166.

(36) Gosal, W. S.; Morten, I. J.; Hewitt, E. W.; Smith, D. A.; Thomson, N. H.; Radford, S. E. Competing Pathways Determine Fibril Morphology in the Self-assembly of β_2-Microglobulin into Amyloid. *J. Mol. Biol.* **2005**, *351* (4), 850−64.

(37) Roskamp, K. W.; Montelongo, D. M.; Anorma, C. D.; Bandak, D. N.; Chua, J. A.; Malecha, K. T.; Martin, R. W. Multiple Aggregation Pathways in Human γS-Crystallin and its Aggregation-Prone G18V Variant. *Invest Ophtalmol. Vis. Sci.* **2017**, *58* (4), 2397-2405.

(38) Barnett, G. V.; Razinkov, V. I.; Kerwin, B. A.; Laue, T. M.; Woodka, A. H.; Butler, P. D.; Perevozchikova, T.; Roberts, C. J. Specific-Ion Effects on the Aggregation Mechanisms and Protein-Protein Interactions for Anti-streptavidin Immunoglobulin Gamma-1. *J. Phys. Chem.* **2015**, 119, 5793-5804.

(39) Krishnan, S.; Chi, E. Y.; Webb, J. N.; Chang, B. S.; Shan, D.; Goldenberg, M.; Manning, M. C.; Randolph, T. W.; Carpenter, J. F. Aggregation of Granulocyte Colony Stimulating Factor under Physiological Conditions: Characterization and Thermodynamic Inhibition. *Biochemistry* **2002**, *41* (20), 6422–6431.

(40) Gasteiger E.; Hoogland C.; Gattiker A.; Duvaud S.; Wilkins M. R.; Appel R. D.; Bairoch A.

Protein Identification and Analysis Tools on the ExPASy Server.

In *The Proteomics Protocols Handbook*, Walker, J. M., Ed.; Humana Press: New York, **2005**; pp 571-607.

(41) 01/2008:1170 Alteplase for Injection. In *Ph. Eur.*, 6th ed.; EDQM: Strasbourg, **2008**; pp 1145-1149.

(42) Roberts, C. J. Therapeutic Protein Aggregation: Mechanisms, Design, and Control. *Trends Biochem. Technol.* **2014**, *32* (7), 372-380.

(43) Banks, D. D.; Latypov, R. F.; Ketchem, R. R.; Woodard, J.; Scavezze, J. L.; Siska, C. C.; Razinkow, V. I. Native-state Solubility and Transfer Free Energy as Predictive Tools for Selecting Excipients to Include in Protein Formulation Development Studies. *J. Pharm. Sci.* **2012**, *101*, 2720-2732.

(44) Chiti, F.; Dobson, C. M. Amyloid Formation by Globular Proteins Under Native Conditions. *Nat. Chem. Biol.* **2009**, *5* (1), 15–22.

(45) Costanzo, J. A.; O'Brien, C. J.; Tiller, K.; Tamargo, E.; Robinson, A. S.; Roberts, C. J.; Fernandez, E. J. Conformational Stability as a Design Target to Control Protein Aggregation. *Protein Eng. Des. Sel.* **2014**, *27* (5), 157–167.

(46) Neudecker, P.; Robustelli, P.; Cavalli, A.; Walsh, P.; Lundström, P.; Zarrine-Afsar, A.; Sharpe, S.; Vendruscolo, M.; Kay, L. E. Structure of an Intermediate State in Protein Folding and Aggregation. *Science* **2012**, *336* (6079), 362-366.

(47) Yu, H.; Yihan, Y.; Zhang, C.; Dalby, P. A. Two Strategies to Engineer Flexible Loops for Improved Enzyme Thermostability. *Sci. Reps.* **2017**, *7*, 41212.

(48) Benton, L. A.; Smith, A. E.; Young, G. B.; Pielak, G. J. Unexpected Effects of Macromolecular Crowding on Protein Stability. *Biochem.* **2012**, *51*, 9773-9775.

(49) Barata, T. S.; Zhang, C.; Dalby, P. A.; Brocchini, S.; Zloh, M. Identification of Protein–Excipient Interaction Hotspots Using Computational Approaches. *Int. J. Mol. Sci.* **2016**, *17* (6), 853

Selective stabilisation and destabilisation of protein domains in tissue-type plasminogen activator using formulation excipients

*Mathew J. Robinson, Paul Matejtschuk, Colin Longstaff, Paul A. Dalby**