Supporting Text for

Does solution viscosity scale the rate of aggregation of folded proteins?

Mike Sleutel, Alexander E.S. Van Driessche, Weichun Pan, Erwin K. Reichel, Dominique Maes, Peter G. Vekilov

The enzyme

Protein function. Xylose isomerase (EC:5.3.1.5) is an enzyme which catalyzes the interconversion of D-xylose to D-xylulose *in vivo*. *In vitro* it also catalyzes the commercially attractive conversion of glucose to fructose and is therefore commonly referred to glucose isomerase (gluci) [1,2]. These functions of the enzyme are illustrated in Fig. S1.

Gluci is widely distributed in prokaryotes and the interconversion of xylose to xylulose serves a nutritional requirement in bacteria that thrive on decaying plant material. Gluci also exists in plants and its presence was also reported in certain animals [4]. D-xylose is a necessary sugar for animals. After its conversion, D-xylulose is an important intermediate in the pentose metabolism cycle.

The main industrial use of this enzyme is in the manufacture of high-fructose corn syrup to satisfy the demand of cost-effective production of sweeteners. In this process starch undergoes enzymatic liquefaction via an α-amylase, subsequent saccharification by a glucoamylase and finally the isomerisation of glucose into the much sweeter fructose. At present more than ten million tons of high-fructose corn syrup are produced annually.

Because xylulose can be fermented to ethanol by conventional yeasts, gluci has also gained some industrial attention for its potential application in the production of ethanol from hemicelluloses, embedded in the cell wall of plants. Bioconversion of renewable biomass to fermentable sugar and ethanol is important in view of the rapid depletion of fossil fuels.

Moreover due to the remarkable tolerance of this enzyme for epimers, as well as chemically modified substrate analogues bearing non-natural features, such as azido or fluoro substituents, a large variety of interesting and novel synthetic applications has become feasible for this remarkably useful and easy-to-handle biocatalyst [5]. As such much of the information on producer organisms and on developed processes is in the form of patents.

Architecture. Gluci is a homotetramer, Fig. S2. Each chain with a length of 388 residues consists of 2 domains; the N-terminal domain is the core domain and has an alpha-beta barrel structure; the C-terminal domain has a long extended loop structure with 5 helices. The latter domain "embraces" the core domain of an adjacent subunit forming a tight dimer. The active site includes residues from 2 subunits and is located in a deep cleft near the C-terminal ends of the strands of the barrel domain [6].

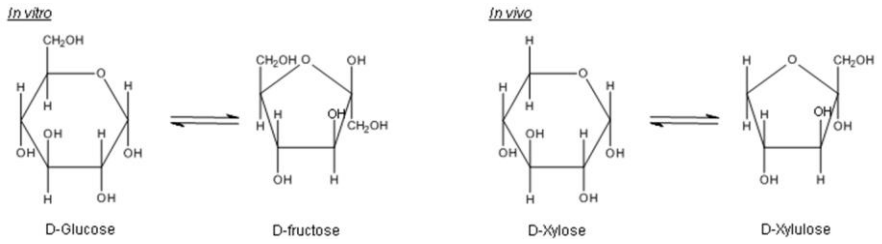

Fig. S1. Reactions catalyzed by glucose isomerase *in vivo* and *in vitro*.

Identification of the crystal packing contacts. The solvent-accessible surface area (ASA) of a given residue within a macromolecule is given by the trace of the center of a spherical probe (r = 1.4 Å) moving around the macromolecule. The ASA values were calculated using NACCESS [7] with the default Van der Waals radii [8]. Each residue with an ASA>0 within the glucose isomerase tetramer is part of the tetramer surface and capable of mediating crystal contacts. For each residue the buried solvent-accessible surface area (ΔASA) upon crystallization was calculated by subtracting the ASA of the residue within the context of the crystal from the ASA within the isolated tetramer. The residues involved in crystal contacts were identified as those having a positive ΔASA and are highlighted in Fig. S3.

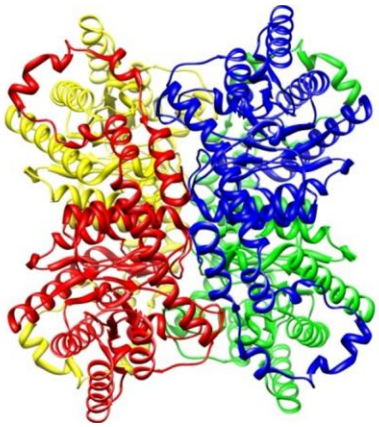

Fig. S2. The glucose isomerase tetramer. Each chain is shown in a different color. Created from Protein Data Base entry 2GLK (6) using UCSF Chimera [3].

Each monomer has two patches involved in crystal packing contacts: patch A (residues Arg76, Asp80, Asp81, Thr82, Gly83 and Lys85) located at the C-terminal end of a helix in the core domain and patch B (residues Arg331, Asp336, Arg340 and Pro341) located in the loop-domain. In the body centered unit cell the central tetramer makes one crystal contact with the 8 tetramers located at the cell corners. Every contact is a contact between an A and a B patch, making all the crystal packing contacts identical. The residues partaking in the crystal contacts are mostly, with the exception of Gly and Pro, acidic (Asp), basic (Arg and Lys), and polar (Thr). This indicates that the interactions in the crystal are not hydrophobic but rather ionic and hydrogen-bond.

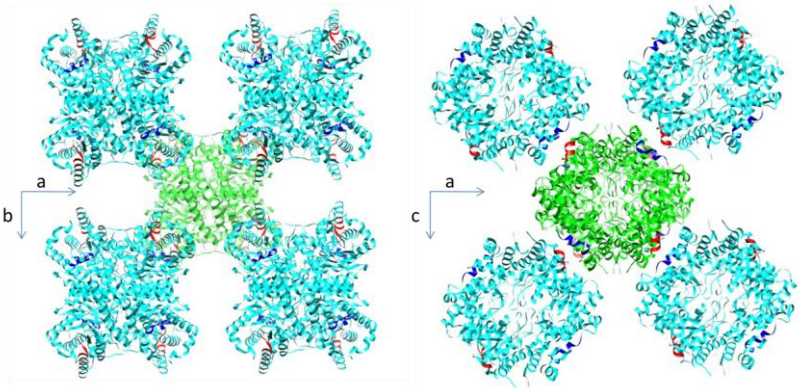

Fig. S3. Two views of the central tetramer in the unit cell (green) with the 4 tetramers involved in packing contacts on top. The four tetramer below make similar contacts and are not shown. Left: view along c-axis, and right: view along b-axis. Created from Protein Data Base entry 2GLK (6) using UCSF Chimera [3]. The sites involved in crystal packing contacts are highlighted in red (patch A) and blue (patch B): each crystal contact is a contact between a patch A and a patch B.

Materials and methods

Crystallization of glucose isomerase. Glucose isomerase from *Streptomyces rubiginosus* was purchased from Microcrystal Oy (Helsinki, Finland). The protein solution was dialyzed against 10mM Hepes (N-2-hydroxyethylpiperazine-N′-2-ethanesulfonic) buffer pH= 7.0 containing 1mM $MgCl_2$. Protein concentrations were determined by UV absorbance at 280 nm using extinction coefficient of 1.042 ml $mg^{-1}cm^{-1}$. Crystals were nucleated and grown in batch within an optical quartz cuvette compatible with the laser confocal microscope. The crystallization conditions were 30 mg ml^{-1} protein in 100mM Hepes

buffer at pH = 7.0, 200 mM $MgCl_2$ and 5 weight % polyethylene glycol with molecular weight 1000 g mol^{-1} [1] (PEG 1000) at 20°C. At these conditions, the protein crystallizes exclusively in the orthorhombic form (I222, a=94 Å, b=99 Å, c=103 Å) [9].

Glycerol was purchased from Sigma-Aldrich. Pure glycerol was added to the crystallization solvent to achieve the desired volume concentration.

Surface imaging using Laser Confocal Microscopy (LCM) combined with Differential Interference Contrast Microscopy (DIC). We monitored *in situ* the growth of the (011) faces of the orthorhombic glucose isomerase crystals by LCM-DIM. For this, a D-Eclipse C1 confocal system was attached to an inverted Eclipse TE-2000 U optical microscope (Nikon, Brussels, Belgium). An air objective lens PlanFluor 20x with a numerical aperture of 0.45 and a working distance of 7.6 mm was used (theoretical lateral resolution 660 nm). The diameter of the confocal aperture used for the observation was 103 µm. A Nomarski prism was inserted into the optical path to utilize differential interference contrast. The solutions and crystals were held in a 45 mm x 12.5 mm x 0.5 mm quartz cuvette with two detachable windows (Hellma, Müllheim, Germany). The temperature of the sample stage was controlled using a curve-matched thermistor, two Peltier elements and a PR-59 (Supercool ®) temperature controller PC-interfaced through a serial connection by means of a Labview ® driver (precision of the temperature control is ± 0.1°C). Crystals were illuminated with an Argon-ion laser with a wavelength of 488 nm. Photomicrographs of 512 x 512 pixels were acquired over a 4 s scan time with a pixel-dwell of 11 µs [10]. These experiments were carried out in Brussels.

The dependence of step velocity on solution flow velocity. *In situ* observation of the gluci surface during forced flow experiments was done by a laser confocal differential interference contrast microscope at Laboratorio de Estudios Crystalograficos, Granada. This setup consists of a confocal system (FV300, Olympus) attached to an inverted optical microscope (IX70, Olympus) equipped with a 10x objective lens (LUCplan FLN 10x, Olympus) and a Nomarski prism is introduced into the optical path. A partially coherent superluminiscent diode (Amonics Ltd., model ASLD68-050-B-FA: 680 nm) is used as illumination source to avoid diffraction noise. The observation cell, made out of Teflon, was mounted inside a copper sample stage which completely surrounds the observation cell, except for the observation area with a diameter of 7 mm at the bottom side. The temperature of the sample stage was controlled using a setup identical to the one used in the confocal imaging experiments in Brussels. Controlled solution flow was produced by a peristaltic pump (Minipuls 3, Gilson) using Tygon tubing (1.65 mm internal diameter). The rotation rate of the peristaltic pump can be varied between 0.5 and 8 RPM. A calibration curve was set up to convert the pump rotation rate to flow rates by measuring the time required to flow 10ml through the system as a function of the pump rotation rate, Fig. S4. Using the cross section of the cell 1.5 × 1.75 mm^2 at the location where the studied crystals rest, the flow rates were converted to linear flow velocities.

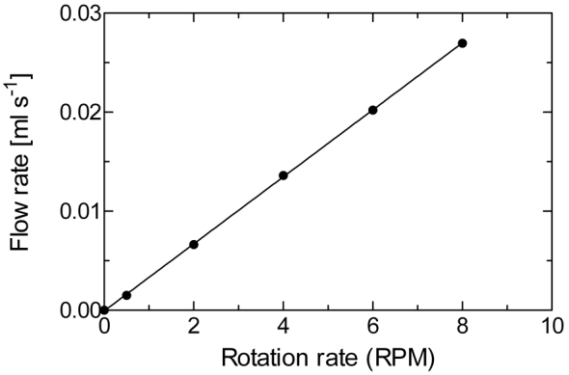

Fig. S4. Calibration curve of the peristaltic pump.

Atomic force microscopy imaging of the surface of growing crystals. Glass discs containing adsorbed crystals were glued to metal pucks with two component epoxy glue and were placed onto the AFM scanner. Approximately 50 µl of growth solution was added to the sample, the O-ring was installed to prevent evaporation and the fluid cell was mounted into the scanner head. The system was left to

stabilize for 60 minutes to avoid drift of the set point. AFM imaging of the crystallization processes was conducted in tapping mode using Nanoscope IIIa multimode AFM (Veeco, Santa Barbara, USA). Images were acquired in buffer solutions of identical composition to the mother liquor, but at varying glucose isomerase concentrations. Standard silicon nitride and oxide sharpened Nanoprobe SPM tips (Veeco, Santa Barbara, USA) were used. Cantilevers with nominal force constants of 0.01 N/m were typically utilized. The tapping drive frequency was adjusted in the range 8 - 10 kHz to the resonance value of the employed tips. In order to minimize the force applied to the crystalline surface during scanning, the set point voltage was continually adjusted to the lowest level for which tip-crystal contact was maintained.

Dynamic light scattering characterization of glucose isomerase solution. Dynamic light scattering (DLS) data were collected at 20°C in 10 mm cylindrical cuvettes at an angle of 90° employing an ALV-5000/EPP static and dynamic light scattering device with a 22 mW He-Ne laser with a wavelength λ = 632.8 nm (Houston) or in 96 well flat bottom plates at an angle of 158° employing a Wyatt DynaPro Plate Reader with a 60 mW Ga-As laser with a wavelength of λ = 823.9 nm (Brussels). Both devices output the correlation function of the intensity variations g_2 as a function of the delay time τ for τ values in the range 10^{-7}-10^{2} s. From the correlation function $g_2(\tau)$, the characteristic delay times τ were computed using ALV-Correlator Software 3.0 and Dynamics 6.9.2.9. Depending on the solution viscosity, the delay times τ were in the range 0.01 – 0.1 ms.

Static light scattering. Static light scattering (SLS) data were collected at 20°C on the same ALV device used for DLS in Houston. The range of concentrations was 0.1 – 5 mg/ml. The method outputs the ratio KC/R_θ, where *C* is the protein concentration, $R_\theta = I_\theta / I_0$ is the Rayleigh ratio of the intensity of the light scattered at angle θ to the incident intensity, K is a system constant defined as $K = N_A^{-1}\left(2\pi n_0 / \lambda^2\right)^2 (dn/dC)^2$ Ref. [11], N_A is the Avogadro number, *n* is the solution refractive index and n_0 = 1.334 is the refractive index of the solvent at the wavelength of the laser beam λ Ref. [12]. The refractive index increment *dn*/*dC* = 0.213 $cm^3 g^{-1}$ was measured using an Abbe 60-ED refractometer.

The KC/R_θ ratio is proportional to the osmotic compressibility

$$\frac{KC}{R_\theta} = \frac{1}{RT}\left(\frac{\partial \Pi}{\partial C}\right), \qquad (1.)$$

where Π is the contribution of the scattering species to the osmotic pressure of the solution, *R* is the universal gas constant and *T* is the absolute temperature.

The osmotic compressibility allows evaluation of the protein chemical potential μ [13]

$$\mu = \int\left(\frac{\partial \Pi}{\partial C}\right)\frac{M_w}{C}dC, \qquad (2.)$$

where M_w = 173,000 g mol^{-1} is the protein molecular mass.

The relation between the KC/R_θ ratio and the protein chemical potential μ in Eqs. (1) and (2) indicates that an increase in KC/R_θ at fixed protein concentration, solution composition and temperature reflects an increase in μ and *vice versa*—decrease in KC/R_θ reflects a decrease in μ.

Determination of solvent viscosity. Two methods were used to determine the viscosity of the solvent through which gluci molecules diffuse. In the dynamic light scattering method, we use the correlation function $g_2(\tau)$ of the suspension of probe particles of radius R_0 = 200 nm, whose volume fraction was 2×10^{-6} and the CONTIN algorithm [14] to evaluate the corresponding distribution function of relaxation times $G(\tau)$. Since the peaks in $G(\tau)$ corresponding to scattering from the probe particles are typically sharp, they directly indicate the characteristic diffusion time τ_p of the particles. The viscosity η of the solution is obtained from τ_p via the Einstein-Stokes equation

$$\eta = \frac{k_B T}{6\pi R_0} q^2 \tau_p \quad . \tag{3.}$$

In Eq. (3) q is the wave vector $q = 4\pi n_0/\lambda \sin(\theta/2)$ [15, p. 83]; with the above value of the laser wavelength λ, the solvent refractive index $n_0 = 1.334$, and scattering angle θ, $q = 1.87 \times 10^7$ m^{-1}. To determine the size of the probe particles R_0, we use their characteristic diffusion time in pure water with viscosity 1.002 mPa s. We get $R_0 = 200$ nm. These tests were repeated for all glycerol concentrations and for solutions without glycerol.

The rheological measurements were performed on a stress-controlled rheometer (Anton Paar MCR501) with a double-wall Couette geometry requiring a sample volume of 3.5 ml. The sample temperature was controlled with a Peltier element with accuracy better than 0.1 K. For steady-shear flow, viscosity at different shear rates was probed. The sample was analyzed in the linear viscoelastic regime. The linear limit was determined from a strain sweep experiment at a fixed frequency of 10 rad.s^{-1}. The frequency sweep experiments were conducted at a fixed strain of 1 (within the linear viscoelastic regime) covering an angular frequency range from 1 to 100 rad.s^{-1}. The RheoPlus software was used for data acquisition and analysis.

Determination of solution hydrodynamic and transport properties

Tests for possible solution viscoelasticity. Viscoelasticity is a behavior in which the kinetic energy of a diffusing molecule is not completely dissipated, but part of it is stored in the medium and returned to the molecule [16]. Its effects on the kinetics of reactions in solution have not been studied and possible viscoelasticity could obscure the interpretation of the effects of increased viscosity on the kinetics of aggregation, as suggested in [17,18]. To determine if the hydrodynamic response in the tested solutions is purely viscous or partially elastic, we record the correlation between the mean squared displacement $\langle r^2 \rangle$ of probe particles and the observation time τ. This $\langle r^2 \rangle(\tau)$ was determined by dynamic light scattering from the probe particles of radius $R_0 = 200$ nm. The volume fraction of the probe particles was 2×10^{-6}.

We record the correlation function $g_2(\tau)$ of light scattered by probe particles as in Ref. [19]. The mean squared displacement $\langle r^2 \rangle(\tau)$ is calculated from $g_1(\tau) = \exp[-q^2\langle r^2 \rangle(\tau)/6]$ and $g_2(\tau) - 1 = [g_1(\tau)]^2$. From $\langle r^2 \rangle(\tau)$, we calculate the logarithmic increment α

$$\alpha = \frac{d \ln\left(\langle r^2 \rangle(\tau)\right)}{d \ln \tau} \quad . \tag{4.}$$

Values of $\alpha < 1$ are used as evidence for partially elastic behavior of the solution; purely viscous

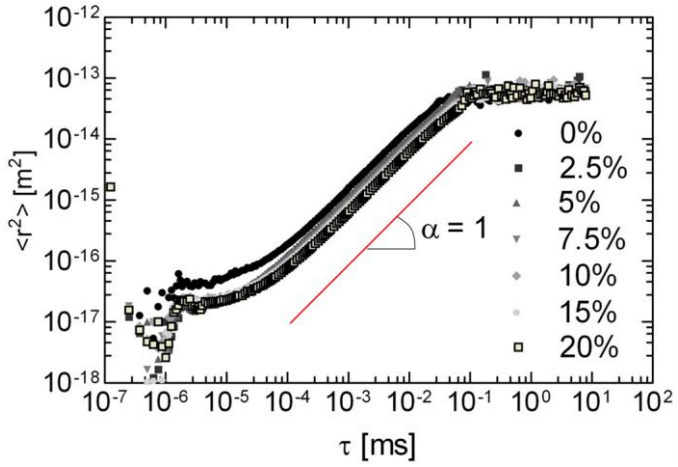

Fig. S5. Mean squared displacement of R_0 = 200 nm latex spheres as a function of lag time τ for different glycerol concentrations (in 100mM Hepes pH = 7.0, 200mM $MgCl_2$, 5 mass % PEG 1000, 0 vol. % - 20 vol. % glycerol) monitored by DLS. Red line has slope $\alpha = 1$.

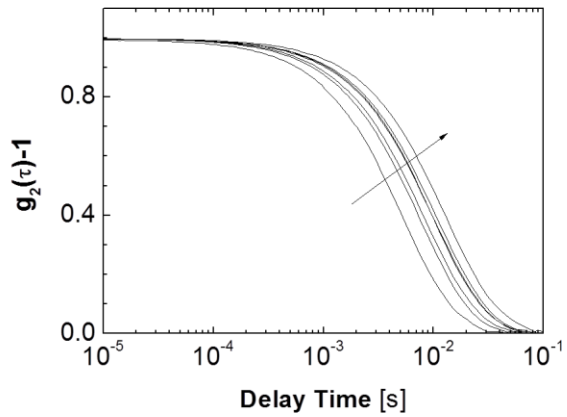

Fig. S6. The intensity correlation functions of light scattered form latex spheres of radius 200 nm and occupying 2×10^{-6} of the solution volume, suspended in the crystallization solvent with concentrations of glycerol 0, 2.5, 5, 7.5, 10, 15 and 20 vol. % arranged in increasing sequence as indicated by the arrow.

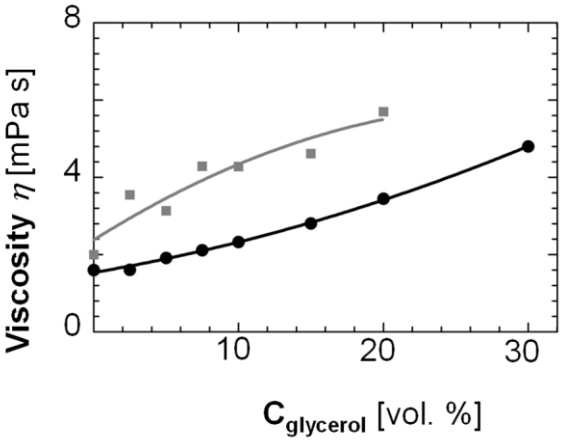

Fig. S7. The effect of glycerol concentration on the solution viscosity. Gray points: determination by dynamic light scattering using 400 nm diameter latex probe particles; the characteristic time scale of this determination is 10 – 30 ms. Black points: determination using rheometry; these results are shear-rate dependent and data at the minimal tested shear rate 20 s^{-1} are provided, higher shear rates lead to even lower values of η. Curves are polynomial interpolations.

Brownian transport leads to $\alpha = 1$; $\alpha > 1$ indicates active transport by a carrier [20].
Figure S5 shows that for the six tested concentrations of glycerol and in solutions without glycerol $\alpha \cong 1$ at time scales up to 0.1 ms. This includes the range of characteristic times of diffusion of glucose isomerase: 0.01 – 0.1 ms. Hence, we conclude that the hydrodynamic response in the growth solutions is purely viscous.

The solvent viscosity. Fig. S7 displays the solvent viscosity, determined by the two methods discussed above. Both sets of data show that the viscosity increases upon increase associated with the addition of glycerol. The viscosity values determined by a Couette rheometer are lower than those determined using probe particles. This discrepancy may be due to: (i) solution slip due to the high shear stress of the rheometry technique, or (ii) viscoelasticity of the solution, similar to the one demonstrated in Ref. [21], leading to different apparent viscosities at the two characteristic timescales of the two determinations: 10-30 ms for the probe particle method and ~1 s for the rheometer.

Since the results in Fig. S5 above indicate that the hydrodynamic behavior in the crystallization solutions is purely viscous at the relevant timescales, we conclude that the higher values of the rheometer-determined viscosity are due to solution slip. Thus, we choose the viscosity determined by the probe particle method as representative of the one which governs gluci diffusion. More importantly, since the solution is purely viscous, it has a unique viscosity which can be tested for its effects on the aggregation kinetics.

Determination of the protein diffusion coefficient. Using gluci concentration 5 mg ml^{-1}, for each glycerol concentration 30 correlation functions $g_2(\tau)$ of light scattered by the protein molecules were acquired over 20 seconds, representative $g_2(\tau)$ are shown in Fig. S8. The g_2's were individually analyzed with the Dynamics 6.9.2.9 software package yielding characteristic decay times τ. From these

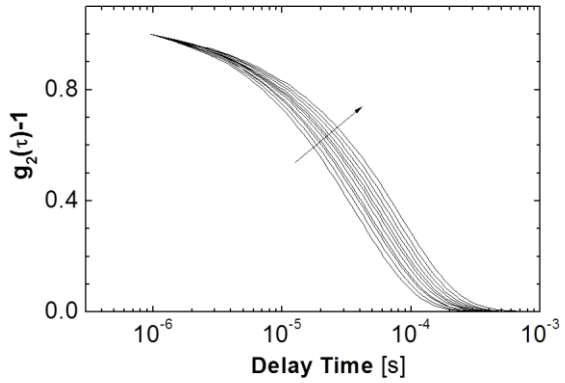

Fig. S8. The intensity correlation functions of gluci solutions of concentration 5 mg ml^{-1} at glycerol concentrations 0, 2, 4, 6, 8, 10, 12, 14, 16, 18 and 20 vol. % arranged in increasing sequence as indicated by the arrow.

τ, the protein diffusion coefficient was determined as $D = (q^2\tau)^{-1}$, where the wave vector $q = 2.09 \times 10^7$ m^{-1}. The average of 30 diffusion coefficients was taken to be the final value.

Tests of protein stability

Differential Scanning Calorimetry (DSC) Measurements. Calorimetric scans were performed using a Nano DSC Series III System (TA Instruments) with a capillary cell (0.3 ml volume) at a single scan rate of 1.0 ^{o}C min^{-1}. All samples were filtered and degassed for 10 min at 283 K prior to the experiment. Sample conditions were 10mM Hepes pH = 7.0, 1mM $MgCl_2$, 0.6mg ml^{-1} and 0 or 20 vol. % glycerol. In Fig. S9, the DSC thermograms are presented as molar heat capacity Δc_p of gluci obtained from the raw signal corrected for the buffer contribution and normalized per mole of gluci tetramer contained in the measuring cell. Fig. S9 shows that, as expected, the addition of glycerol stabilizes the conformation of gluci and increases the melting temperature from 85 to 92 ^{o}C.

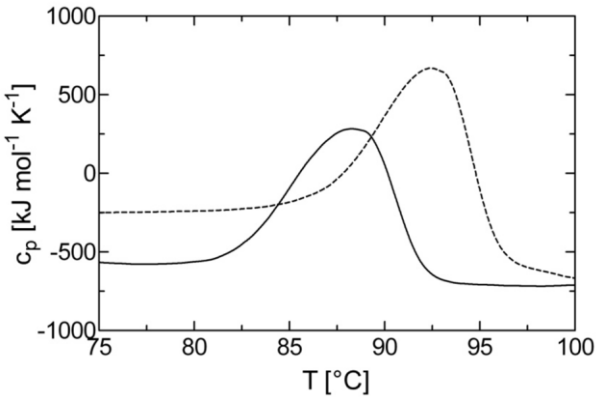

Fig. S9. DSC thermograms: Thermal denaturation of gluci homotetramers at a scan rate of 1°C min^{-1}, at 0 vol. % glycerol (solid line) and 20 vol. % glycerol (dashed line).

Does glycerol affect the activation barrier for attachment?

Static light scattering characterization of the interactions between gluci molecules. Glycerol might enhance repulsion between the molecules in a kink and an incoming molecule and in this way slow down the incorporation of a molecule into a kink. In a kinetic analysis, this would appear as increase of the activation barrier for incorporation [22]. As a first test, we searched for a correlation between the lower step velocity in Fig. 1F of the main text and lower dielectric permeability ε of the solution [23] and we found that ε decreases linearly from 80 to 75 by the addition of up to 20 vol. % glycerol; this effect is clearly insufficient to explain the complete growth cessation in Fig. 1F of the main text.

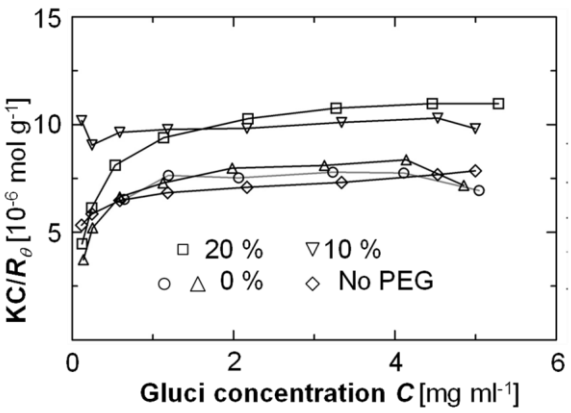

Fig. S10. Towards the tests of the potential effects of glycerol on the activation barrier for attachement of molecules. Debye plots of the ratio KC/R_θ, where C is the protein concentration, $R_\theta = I_\theta/I_0$ is the Rayleigh ratio of the intensity of the light scattered at angle θ to the incident intensity, and K is a system constant, determined by static light scattering from solutions containing 0, 10 and 20 vol. % glycerol and, for comparison, in a solution similar to the growth solution with 0 vol. % glycerol, but in the absence of PEG.

For further tests, as a model for the interactions in the kink we use the interactions between the gluci molecules in solution, which can be studied by static light scattering. Figure S10 contains the Debye plot for gluci solutions in the presence and absence of glycerol defined above. Figure S10 shows that the addition of glycerol increases the values of KC/R_θ in the whole range of tested C's. This indicates that glycerol leads to greater chemical potential of gluci in solution, i.e., glycerol enhances the repulsion between the gluci molecules. However, the enhancement is relatively mild and does not depend on the concentration of glycerol: the values of KC/R_θ for 10 and 20 vol. % glycerol nearly coincide (the deviation at $C < 0.2$ mg ml^{-1} at $C_{glycerol}$ = 10 vol. % is likely an error due to weak scattering at this low protein concentration).

Thermodynamic analyses of the effects of glycerol on intermolecular interactions in solutions of other proteins have concluded that glycerol is excluded from the immediate environment of the protein molecules [24,25]. Enhanced intermolecular repulsion was attributed to strengthening of the hydration layer around the protein molecules [25]. It is likely that the increased repulsion indicated by the data in Fig. S10 is due to a similar mechanism.

The initial slope of the concentration dependence of the ratio KC/R_θ has been used to evaluate the second osmotic virial coefficient B_{22} (or, in the case of multi-component systems, such as the one here, its equivalent $B_{2,eff}$ [26,27]). Thus, the positive slopes of the $KC/R_\theta(C)$ dependencies in Fig. S10 at low concentrations indicate that the values of $B_{2,eff}$ are also positive [28]. Positive values of both B_{22} and $B_{2,eff}$ indicate repulsion between the protein molecules and this is the basis of the foundation of the protein crystallizability rule, according to which only moderately negative values of B_2 are conducive of crystallization [29,30]. The observation in Fig. S10 is in clear violation of this rule. Similar exceptions from the crystallizability rule have been recorded for other proteins, e.g., lumazine synthase [19]. The obvious reason for them is that while $B_{2,eff}$ is an azimuthally and radially averaged characteristic of the intermolecular interaction in solution, the crystals are formed by strongly directional and often short-ranged bonds. Hence, $B_{2,eff}$ may be a poor indicator of local interactions capable of supporting crystal packing.

Does glycerol affect the crystal packing contacts? To further elucidate the role of glycerol for the molecular interactions important for crystallization, we analyze the contacts between the molecules in the crystal, see above for details. The results conclude that the bonds between the molecules in the crystal are ionic. Such bonds would not be affected by a stronger hydration layer around the molecules, induced by glycerol. The highest resolution gluci crystal structure in the protein database (2GLK), which is the only one exhibiting the presence of glycerol in the structure, reveals glycerol bound to four locations: two at the interface between the gluci subunits, one location in the active center, and one on the protein surface; all four locations are away from the crystal contact sites [6]. We conclude that while

glycerol mildly affects the azimuthally averaged interactions between gluci molecules in solution (as evidenced by the static light scattering results in Fig. S10), its effects on the crystal packing sites are minimal or non-existent and no effects on the activation barrier for attachment of molecules to kinks are expected.

References

(1) Bhosale, S. H.; Rao, M. B.; Deshpande, V. V.Molecular and industrial aspects of glucose isomerase. *Microbiol Rev* **1996**, *60*, 280.

(2) Srivastava, P.; Shukla, S.; Choubey, S. K.; V.S., G.Isolation, Purification & Characterization of Glucose Isomerase Enzyme form Streptomyces species isolated from Parbhani Region. *Journal of Enzyme Research,* **2010**, *1*, 1.

(3) Pettersen, E.; Goddard, T.; Huang, C.; Couch, G.; Greenblatt, D.; Meng, E.; Ferrin, T.UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* **2004** *25*, 1605.

(4) Ding, M.; Teng, Y.; Yin, Q.; Chen, W.; Zhao, F.Identification, expression, and characterization of the highly conserved d-xylose isomerase in animals. *Acta Biochimica et Biophysica Sinica* **2009**, *41*, 116.

(5) Hausler, H.; Stutz, A. E.D-xylose (D-glucose) isomerase and related enzymes in carbohydrate synthesis. *Top. Curr. Chem.* **2001**, *215*, 77.

(6) Katz, A. K.; Li, X.; Carrell, H. L.; Hanson, B. L.; Langan, P.; Coates, L.; Schoenborn, B. P.; Glusker, J. P.; Bunick, G. J.Locating active-site hydrogen atoms in d-xylose isomerase: Time-of-flight neutron diffraction. *Proceedings of the National Academy of Sciences* **2006**, *103*, 8342.

(7) Hubbard, S. J.; Thornton, J. M. NACCESS; Department of Biochemistry and Molecular Biology, University College London, 1993.

(8) Chothia, C.The nature of the accessible and buried surfaces in proteins. *Journal of Molecular Biology* **1976**, *105*, 1.

(9) Carrell, H. L.; Glusker, J. P.; Burger, V.; Manfre, F.; Tritsch, D.; Biellmann, J. F.X-ray analysis of D-xylose isomerase at 1.9 A: native enzyme in complex with substrate and with a mechanism-designed inactivator. *Proc Natl Acad Sci U S A* **1989**, *86*, 4440.

(10) Sleutel, M.; Maes, D.; Wyns, L.; Willaert, R.Kinetic Roughening of Glucose Isomerase Crystals. *Crystal Growth & Design* **2008**, *8*, 4409.

(11) Eisenberg, D.; Crothers, D. *Physical Chemistry with Applications to Life Sciences*; The Benjamin/Cummins: Menlo Park, 1979.

(12) *Handbook of Chemistry and Physics, 81-th Edition*; CRC Press: Boca Raton, 2000-2001.

(13) Petsev, D. N.; Wu, X.; Galkin, O.; Vekilov, P. G.Thermodynamic functions of concentrated protein solutions from phase equilibria. *J. Phys. Chem. B* **2003**, *107*, 3921.

(14) Provencher, S. W.CONTIN: a general purpose constrained regularization program for inverting noisy linear algebraic and integral equations. *Comp. Phys. Communications* **1982**, *27*, 229.

(15) Berne, B.; Pecora, R. *Dynamic Light Scattering with Applications to Chemistry, Biology, and Physics*; Dover: Mineola, NY, 2000.

(16) Mason, T. G.; Weitz, D. A.Optical Measurements of Frequency-Dependent Linear Viscoelastic Moduli of Complex Fluids. *Physical Review Letters* **1995**, *74*, 1250

(17) Fleming, G. R.; Wolynes, P. G.Chemical Dynamics in Solution *Phys. Today* **1990**, *43*, 36.

(18) Schroeder, J.; Troe, J.Elementary Reactions in the Gas-Liquid Transition Range. *Annual Review of Physical Chemistry* **1987**, *38*, 163

(19) Gliko, O.; Pan, W.; Katsonis, P.; Neumaier, N.; Galkin, O.; Weinkauf, S.; Vekilov, P. G.Metastable liquid clusters in super- and undersaturated protein solutions. *J. Phys. Chem. B* **2007**, *111*, 3106.

(20) Mason, T. G.Estimating the viscoelastic moduli of complex fluids using the generalized Stokes-Einstein equasion. *Rheol. Acta.* **2000**, *39*, 371.

(21) Pan, W.; Filobelo, L.; Pham, N. D. Q.; Galkin, O.; Uzunova, V. V.; Vekilov, P. G.Viscoelasticity in Homogeneous Protein Solutions. *Physical Review Letters* **2009**, *102*, 058101.

(22) Petsev, D. N.; Chen, K.; Gliko, O.; Vekilov, P. G.Diffusion-limited kinetics of the solution-solid phase transition of molecular substances. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 792.

(23) Sedgwick, H.; Cameron, J. E.; Poon, W. C. K.; Egelhaaf, S. U.Protein phase behavior and crystallization: Effect of glycerol. *Journal of Chemical Physics* **2007**, *127*, 6.

(24) Gekko, K.; Timasheff, S. N.Mechanism of protein stabilization by glycerol: proferential hydration in glycerol-water mixtures. *Biochemistry* **1981**, *20*, 4667.

(25) Farnum, M.; Zukoski, C.Effect of glycerol on the interactions and solubility of bovine pancreatic trypsin inhibitor. *Biophys J* **1999**, *76*, 2716.

(26) Kirkwood, J. G.; Goldberg, R. J.Light Scattering Arising from Composition Fluctuations in Multi-Component Systems. *The Journal of Chemical Physics* **1950**, *18*, 54.

(27) Pan, W.; Uzunova, V. V.; Vekilov, P. G.Free heme in micromolar amounts enhances the attraction between sickle cell hemoglobin molecules. *Biopolymers* **2009**, *91*, 1108.

(28) Petsev, D. N.; Thomas, B. R.; Yau, S.-T.; Vekilov, P. G.Interactions and Aggregation of Apoferritin Molecules in Solution: Effects of Added Electrolytes. *Biophysical J.* **2000**, *78*, 2060.

(29) George, A.; Wilson, W. W.Predicting protein crystallization form a dilute solution property. *Acta Crystallogr. Section D* **1994**, *50*, 361.

(30) Haas, C.; Drenth, J.; Wilson, W. W.Relation between the solubility of proteins in aqueous solutions and the second virial coefficient of the solution. *J. Phys. Chem. B* **1999**, *103*, 2808.

Legends for Supplementary movies

Movie S1. Generation of steps by multiple dislocations and their propagation on a (011) face of a glucose isomerase crystal. Crystal grown over weekend from a glucose isomerase solution with initial concentration 30mg ml^{-1} (actual protein concentration at the time of recording unknown), in the presence of 100mM HEPES buffer at pH = 7.0, 200mM $MgCl_2$ and 5% PEG 1000. Time-lapse imaging using LCM-DIM; time delay between successive frames is 150 seconds, which are shown at a rate 15 frames per second.

Movie S2. Generation of steps by two-dimensional nucleation on a (011) face of a glucose isomerase crystal. Crystal grown over weekend from a glucose isomerase solution with initial concentration 30mg ml^{-1} (actual protein concentration at the time of recording unknown), in the presence of 100mM HEPES buffer at pH = 7.0, 200mM $MgCl_2$ and 5% PEG 1000. Time-lapse imaging using LCM-DIM; time delay between successive frames is 30 seconds, which are shown at a rate 15 frames per second.

Movie S3. A 3D view of a model of the incorporation site on a step on a (011) face of a glucose isomerase crystal created from Protein Data Base entry 2GLK with package QuteMol 0.4.1, see text for references. Underlying layer is shown in ochre, growing upper layer—in purple, a molecule entering the kink—in teal. The contact patches on the molecular surfaces, see the Supporting Text for details, are highlighted in navy (patch A) and silver (patch B).