

# Self-Assembly of $\alpha$ -Tocopherol Transfer Protein Nanoparticles – a Patchy-Protein Model

Raphael Peltzer,<sup>†</sup> Hima Bindu Kolli,<sup>†</sup> Achim Stocker,<sup>\*,‡</sup> and Michele  
Cascella<sup>\*,†</sup>

<sup>†</sup>*Department of Chemistry, and Hylleraas Centre for Quantum Molecular Sciences, University of  
Oslo, PO Box 1033 Blindern, 0315 Oslo, Norway*

<sup>‡</sup>*Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, 3012 Bern,  
Switzerland*

E-mail: [achim.stocker@dcb.unibe.ch](mailto:achim.stocker@dcb.unibe.ch); [michele.cascella@kjemi.uio.no](mailto:michele.cascella@kjemi.uio.no)

## Abstract

We describe the mechanism of self-aggregation of  $\alpha$ -tocopherol transfer protein into a spherical nano-cage employing by Monte Carlo simulations. The protein is modelled by a patchy coarse-grained representation, where the protein-protein interfaces, determined in the past by x-ray diffraction, are represented by simplified two-body interaction potentials. Our results show that the oligomerization kinetics proceeds in two steps, with the formation of meta-stable trimeric units, and the subsequent assembly into the spherical aggregates. Data are in agreement with experimental observations regarding the prevalence of different aggregation states at specific ambient conditions. Finally, our results indicate a route for the experimental stabilization of the trimer, crucial for the understanding of the physiological role of such aggregates in vitamin E body trafficking.

# 1 Introduction

2  $\alpha$ -tocopherol transfer protein ( $\alpha$ -TTP hereafter) is the liver factor responsible for the retention of  
3 RRR- $\alpha$ -tocopherol ( $\alpha$ -tol), the active isoform of vitamin E, in the human body.<sup>1-3</sup>  $\alpha$ -TTP solubi-  
4 lizes  $\alpha$ -tol from the external leaflet of maturing endosomal compartments, promoting its release  
5 into the blood. Structural studies over the years<sup>4,5</sup> indicate that  $\alpha$ -TTP is active as a monomer,  
6 similarly to other transporters of its family.<sup>6-8</sup> Unlike for retention, the mechanism(s) by which  
7  $\alpha$ -tol is secreted into the blood, and then absorbed into the target tissues, is at present not well  
8 understood. Possible pathways for the secretion and blood transport of  $\alpha$ -tol include enrichment  
9 into the leaflets of the plasma membrane by a lipid-exchange mechanism,<sup>9,10</sup> and transport into  
10 the blood by aggregating to very-low density lipoprotein vesicles.<sup>2</sup> The absorption from the blood  
11 into the target tissues is even less understood, but it must imply some mechanism of recognition in  
12 order to bypass endothelial barriers, like the blood-brain barrier or the placenta.

13 The involvement of  $\alpha$ -TTP into  $\alpha$ -tol trafficking has not been clearly defined. In a recent  
14 work, Arai and coworkers have suggested that transfer of  $\alpha$ -tol to the plasma membrane is coupled  
15 to the extraction of phosphatidylinositolphosphates (PIPs) from the same membrane by  $\alpha$ -TTP.<sup>3</sup>  
16 Interestingly, they also suggested that lipid-exchange at the plasma membrane may involve higher  
17 order aggregates of  $\alpha$ -TTP than the monomers.<sup>3</sup>

18 Very recently, we provided structural evidence that upon binding to  $\alpha$ -tol,  $\alpha$ -TTP acquires the  
19 tendency to oligomerize.<sup>11</sup> The oligomerised proteins form stable, regular spherical nanoparticles  
20 composed of 24  $\alpha$ -TTP units ( $\alpha$ -TTP<sub>S</sub>), which could be characterized by a series of methods,  
21 including cryo-EM and X-ray diffraction (PDB: 5MUE and 5MUG).<sup>11</sup> Thermal analysis demon-  
22 strated that  $\alpha$ -TTP<sub>S</sub> is thermodynamically stable; furthermore, oxidative conditions enhance its  
23 stability by promoting the formation of twelve disulfide-bonds cross-linking different  $\alpha$ -TTP units.  
24 The X-ray crystallography data of  $\alpha$ -TTP<sub>S</sub> revealed a regular assembly of 24 monomers organized  
25 in a cubic symmetry. Each  $\alpha$ -TTP unit is located on one vertex of a cantellated cube, and it is  
26 involved in two kinds of molecular contacts with four neighboring proteins (Figure 1). The first  
27 interface builds around the  $C_4$  symmetry axis of the assembly, and it is constituted by a patch of

28 surface amino acids that are exposed to the solvent in the native monomeric folding of  $\alpha$ -TTP . The  
29 second interface is responsible for the assembly of  $\alpha$ -TTP around the trimeric  $C_3$  axis. This inter-  
30 face is located on the surface of the proper SEC-14-like binding domain; in  $\alpha$ -TTP it is screened  
31 from the solvent by the N-terminal helical domain. In  $\alpha$ -TTP<sub>S</sub>, the interfaces are accessible to the  
32 partner proteins thanks to the unfolding of the first N-terminal helix, which is not detectable in  
33 the corresponding X-ray structure. The partial unfolding of the N-terminal helix is triggered by  
34 external conditions, including binding to  $\alpha$ -tol or interaction with negatively-charged lipids.<sup>11</sup>

35  $\alpha$ -TTP<sub>S</sub> shows selective and efficient transport properties through *in vitro* models of endothelial  
36 barriers,<sup>11</sup> making it a potential candidate as one of the physiological route for the delivery of  
37 vitamin E into the brain.<sup>12,13</sup> As much as oligomerization of  $\alpha$ -TTP is crucial for its transfecting  
38 properties, its mechanistic aspects remain obscure. In particular, chromatographic data showed that  
39 when the monomeric form is the most stable aggregation state for  $\alpha$ -TTP , this is at equilibrium  
40 with a small presence of low-weight dimeric or tetrameric aggregates. On the contrary, when the  
41 aggregation into the high weight is triggered, the only species present in the solution are monomeric  
42  $\alpha$ -TTP or regular  $\alpha$ -TTP<sub>S</sub> constructs, while no other low-, middle-weight assemblies coexist at  
43 detectable concentration.<sup>11</sup>

44 In the present study, we investigated the formation of  $\alpha$ -TTP<sub>S</sub> using Monte Carlo simulations  
45 of a toy patchy-model of  $\alpha$ -TTP. Molecular simulations using patchy models are proven to be very  
46 effective in understanding the nature of self assembly in systems like patchy colloids, soft func-  
47 tionalized nanoparticles and biomolecules.<sup>14-17</sup> In particular, models with anisotropic and highly  
48 directional interacting patches are particularly suited to describe protein assembly.<sup>18-23</sup>

## 49 **Computational Methods**

50 **Computational model** The patchy model for one  $\alpha$ -TTP consisted of one hard sphere, with four  
51 interaction sites (IS) located on its surface (Figure 1). The relative orientation of the IS was chosen  
52 to mimic the geometric organization of  $\alpha$ -TTP<sub>S</sub>. Interactions were imposed between IS<sub>A</sub> and IS<sub>B</sub>,

53 and between  $IS_C$  and  $IS_D$  types, consistently with the experimental structure of  $\alpha$ -TTP<sub>S</sub>.

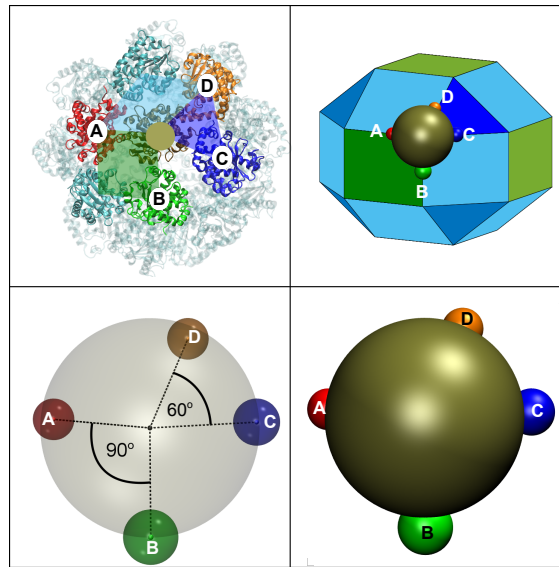


Figure 1: Coarse Grained model of  $\alpha$ -TTP<sub>S</sub>. *Top*: (left) In the native  $\alpha$ -TTP<sub>S</sub>, any monomeric  $\alpha$ -TTP (gold spot) is in contact with four other proteins along the edges of a cantellated cube (top right). *Bottom*:  $\alpha$ -TTP is described as a sphere with four interaction sites corresponding to the protein-protein contacts in  $\alpha$ -TTP<sub>S</sub>.

54 The interaction potentials for the two IS pairs were described by toy potential wells of depth  
55  $E_{A/B} = u$ ,  $E_{C/D} = 3u$  ( $u$  being an arbitrary unit of energy), dependent on both the distance  
56 between the IS, and the relative orientation of the proteins. The initial 1:3 ratio between  $E_{A/B}$  and  
57  $E_{C/D}$  was calibrated on an estimate of the dimerization free-energies from atomistic models using  
58 a standard thermodynamic cycle,<sup>24</sup> computing the solvation free energy of individual and dimeric  
59 structures solving the linearized Poisson-Boltzmann equation using the APBS software,<sup>24</sup> and the  
60 binding energy in vacuo using the Amber force field.<sup>25</sup> Protein dimers were extrapolated from the  
61 X-ray structure of  $\alpha$ -TTP<sub>S</sub> (PDB:5MUE).<sup>11</sup>

62 The four IS are identified by four vectors with origin in the center of the hard sphere, and ends

63 in:

$$\begin{aligned} \text{IS}_A &= (0, R, 0) \\ \text{IS}_B &= (-R, 0, 0) \\ \text{IS}_C &= \left(0, \frac{R}{\sqrt{2}}, -\frac{R}{\sqrt{2}}\right) \\ \text{IS}_D &= \left(\frac{R}{\sqrt{2}}, 0, -\frac{R}{\sqrt{2}}\right) \end{aligned} \quad (1)$$

64 where  $R = 2$  nm is the radius of a hard sphere with its center in  $O(0, 0, 0)$ .

65 The potential energies between the  $\text{IS}_A$ - $\text{IS}_B$  and  $\text{IS}_C$ - $\text{IS}_D$  couples are described by the following  
66 well potentials:

$$E_{A/B}(r, \psi, \phi) = \begin{cases} 0 & r > r_{cut} \\ E_{A/B} \cos \psi \cos \phi & r \leq r_{cut} \end{cases} \quad (2)$$

67

$$E_{C/D}(r, \psi, \theta) = \begin{cases} 0 & r > r_{cut} \\ E_{C/D} \cos \psi \cos 2\theta & r \leq r_{cut} \end{cases} \quad (3)$$

68 where  $r$  is the IS-IS distance,  $r_{cut} = 0.2 R$  is the maximum range of the interaction.  $\cos \psi$ ,  $\cos \phi$ ,  $\cos \theta$   
69 are defined from the scalar multiplication of the normalized vectors as in figure 2. The angular de-  
70 pendency is necessary to model both the chirality of the protein, and that protein binding occurs  
71 over an extended surface that requires a well-defined orientation of the two partners.

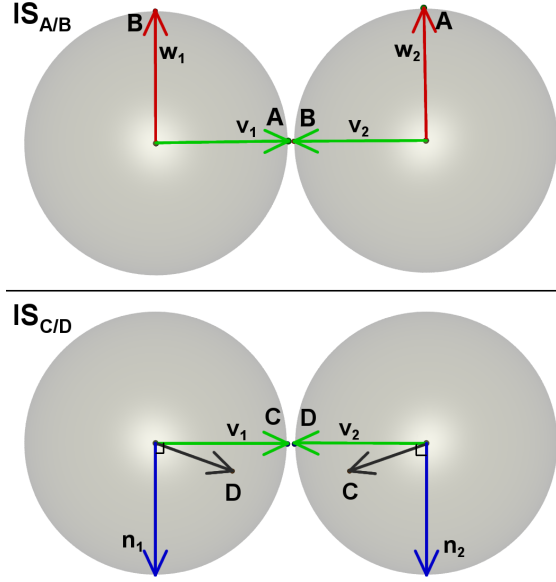


Figure 2: Normalized vectors used to define the angular dependency of the interaction energy.  $\cos \psi = \mathbf{v}_1 \cdot \mathbf{v}_2$ , for any A/B or C/D interaction (green arrows);  $\cos \phi = \mathbf{w}_1 \cdot \mathbf{w}_2$  for any A/B interaction (top panel, red arrows);  $\cos \theta = \mathbf{n}_1 \cdot \mathbf{n}_2$ , for any C/D interaction (bottom panel, blue arrows)

72 **System setup** We simulated a system having  $N = 216$  particles at thermal equilibrium. The  
 73 protein particles were initially distributed uniformly in a periodic cubic simulation box of edge  
 74  $24.625 R$  (where  $R$  is the radius of the protein), corresponding to roughly the experimental con-  
 75 centration at which  $\alpha$ -TTP aggregation is observed.<sup>11</sup> The accessible conformational space in the  
 76 canonical NVT ensemble was explored using a Metropolis Monte Carlo (MC) algorithm.<sup>26</sup>

77 Random moves included the rotation or translation of the single particles or whole clusters of  
 78 bound particles.<sup>27-29</sup> A bond between two particles was assumed to exist if the distance between  
 79 the corresponding interaction sites was less than  $0.2 R$ . Two particles were considered to belong  
 80 to same cluster if they were connected by a chain of bonds.<sup>30</sup> Rotational moves made use of  
 81 quaternion representation of the particle's orientation, which was modified by a smaller random  
 82 orientation and then renormalized.<sup>27,28</sup>

83 The canonical ensemble was sampled at different values of the temperature, in order to deter-  
 84 mine thermodynamic regimes at which different  $\alpha$ -TTP aggregates exist. All the simulations at  
 85 different temperatures started from a random configuration of the proteins in the box. Simulations

86 were organized in cycles, each cycle consisting of a number of attempted particle moves. Typical  
 87 equilibration runs consisted of  $6 \times 10^6$  -  $9 \times 10^6$  MC cycles and were followed by a production  
 88 run of additional  $3 \times 10^5$  MC cycles, during which averages of energy and cluster abundance were  
 89 calculated. Convergence of the results was tested by evaluating both the convergence of the expect-  
 90 tation value of the energy, and comparing the variance to the typical short time (1000 steps) energy  
 91 fluctuations at a given temperature over the last  $6 \times 10^5$  MC cycles.

92 The acceptance probability for the single particle moves like translation or rotation was evalu-  
 93 ated according to:

$$\left\{ \begin{array}{l} \text{if } E_t \leq E_i, \quad P(acc) = 1 \\ \text{if } E_t > E_i, \quad P(acc) = \exp(-\beta(E_t - E_i)) \end{array} \right\} \quad (4)$$

94 Here P is the acceptance probability,  $E_i$  the initial Energy and  $E_t$  the energy after the test step.  $\beta$   
 95 is the reciprocal thermodynamic temperature of the system. The cluster moves were implemented  
 96 following the early rejection scheme.<sup>27</sup>

97 All results here are presented in reduced units:  $\mathbf{U}^* = U/u$  and  $\mathbf{T}^* = k_B T/u$  for the inner  
 98 energy and the temperature of the system.

99 Trajectory analysis was performed using the tools available in the VMD 1.9.2 package.<sup>31</sup>

## 100 **Results and Discussion**

101 In a first set of MC runs, we investigated the existence of  $\alpha$ -TTP aggregates when only the  $IS_{A/B}$   
 102 interface is active. This setup mimics the experimental conditions at which the N-terminal region  
 103 is folded, and the  $IS_{C/D}$  is not exposed to the solvent. The top panel of figure 3 shows the relative  
 104 abundance of  $\alpha$ -TTP aggregates as a function of the thermal energy. At high temperatures, only  
 105 monomeric species are present. At around  $\mathbf{T}^* u = 0.09 E_{AB}$ , we observed the appearance of low  
 106 weight aggregates, mostly tetramers, with also a non-negligible presence of trimers and dimers.  
 107 Lowering the temperature stabilises the tetrameric packing, which corresponds to the aggregation  
 108 state of four  $\alpha$ -TTP proteins around the  $C_4$  Symmetry axis in  $\alpha$ -TTP<sub>S</sub> (Figure S1). Higher molec-



109 ular weight structures, for example, linear chain-like structures along sequences of  $IS_{A/B}$  contacts  
 110 were not observed, as they are forbidden by the chirality condition on the interaction energy, which  
 111 only allows the formation of ring-like tetramers.

112 In a second set of MC simulations, both the  $IS_{A/B}$ ,  $IS_{C/D}$  interactions were active. This setup  
 113 mimics  $\alpha$ -TTP with an unfolded N-terminus. In this case, we observed three temperature regimes  
 114 at which distinct aggregation states appear (Figure 3 (Bottom)). In the high temperature range  
 115 ( $T^* > 0.25$ ), only  $\alpha$ -TTP monomers were present. In the narrow ( $0.10 < T^* < 0.25$ ) region, we  
 116 detected the formation of trimeric species ( $\alpha$ -TTP<sub>3</sub>), while for ( $T^* < 0.12$ ) the systems rapidly  
 117 evolved in high-weight aggregates ( $\alpha$ -TTP<sub>S</sub>) (Figure 6).

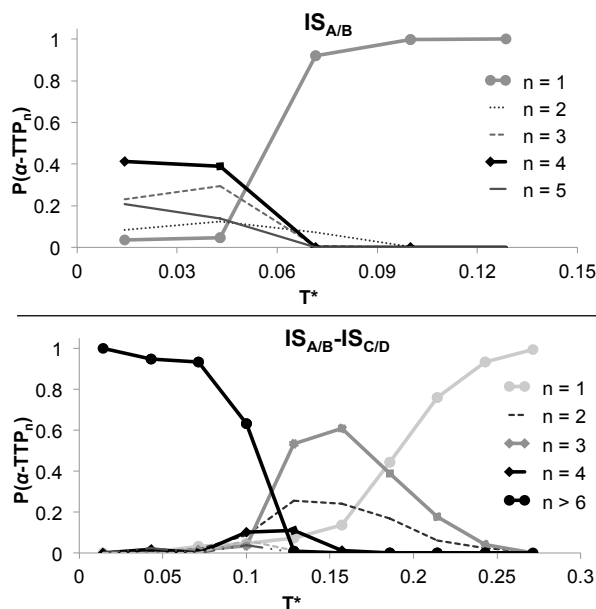


Figure 3: Aggregation states of  $\alpha$ -TTP ( $\alpha$ -TTP<sub>n</sub>) as a function of the reduced temperature  $T^*$ , (*top panel*) when only the  $IS_{A/B}$  interface is active, or (*bottom panel*) when both the  $IS_{A/B}$  and  $IS_{C/D}$  interfaces are active.

118 The trimeric phase includes aggregates built along the  $IS_{C/D}$  interface. In fact,  $\alpha$ -TTP<sub>3</sub>s are  
 119 stable in a temperature range at which the thermal energy is too high to allow the formation of A/B  
 120 contacts. The structure of  $\alpha$ -TTP<sub>3</sub> corresponds to the assembly of three  $\alpha$ -TTP proteins around  
 121 the  $C_3$  axis of  $\alpha$ -TTP<sub>S</sub>. The high-weight aggregates appearing for ( $T^* < 0.12$ ) are constituted by  
 122 oligomerization of  $\alpha$ -TTP<sub>3</sub>, up to  $\alpha$ -TTP<sub>S</sub> by formation of  $IS_{A/B}$  contacts.

123 During the MC runs, we observed a marginal degree of poly-dispersion especially near the  
 124 transition temperatures (Figure 3). Nonetheless, the profile of the energy as a function of the  
 125 temperature shows two clear sigmoidal jumps, indicating the presence of two distinct first-order  
 126 phase transitions from  $\alpha$ -TTP to  $\alpha$ -TTP<sub>3</sub>, to  $\alpha$ -TTP<sub>S</sub> (Figure 4).

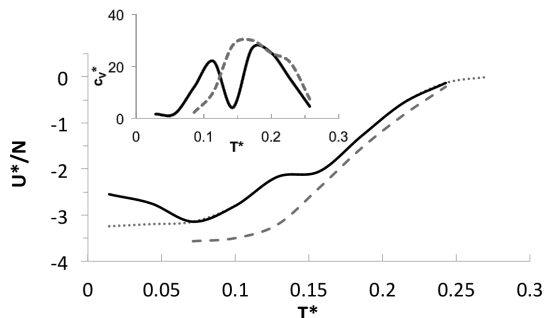


Figure 4: Inner energy per protein  $U^*/N$ , and specific heat  $c_v^*$  (inset) as a function of the reduced temperature for the system with both active  $IS_{A/B}$  and  $IS_{C/D}$  (continuous line). The dotted line was obtained by simulated annealing starting from converged data at  $T^*=0.07$ . The dashed lines report the same data for the system with  $E_{A/B} = 1.5 u$ .

127 Oligomerization to  $\alpha$ -TTP<sub>S</sub> begins at a higher temperature ( $T^* = 0.12$ ) than the one charac-  
 128 terizing  $\alpha$ -TTP aggregation when only  $IS_{A/B}$  are active ( $T^* = 0.07$ ). In fact,  $\alpha$ -TTP<sub>3</sub> dimerization  
 129 involves binding over two  $IS_{A/B}$  contacts, producing a hexameric structure centered around a  $C_2$   
 130 symmetry axis corresponding to one of the  $C_2$  axes of  $\alpha$ -TTP<sub>S</sub>. Simultaneous formation of two  
 131  $IS_{A/B}$  interactions is facilitated by the pre-organization of the interaction sites along the edges of  
 132 the rigid  $\alpha$ -TTP<sub>3</sub>.

133 The cooperative effect of the  $IS_{A/B}$  onto the binding of  $\alpha$ -TTP<sub>3</sub> is responsible for the absence  
 134 of intermediate weight aggregates between  $\alpha$ -TTP<sub>3</sub> and  $\alpha$ -TTP<sub>S</sub>. Practically, the assembly of  
 135  $\alpha$ -TTP<sub>S</sub> may be schematically seen as the progressive dimerization of  $\alpha$ -TTP<sub>3</sub>,  $\alpha$ -TTP<sub>6</sub>, and  $\alpha$ -  
 136 TTP<sub>12</sub> over two, four, and eight  $IS_{A/B}$  contacts (Figure 5). Below the critical temperature that  
 137 allows the first dimerization of  $\alpha$ -TTP<sub>3</sub>, further assemblies involve increasingly larger number of  
 138  $IS_{A/B}$  interactions, yielding  $\alpha$ -TTP<sub>S</sub>.

139 The existence of a region of thermodynamic stability for  $\alpha$ -TTP<sub>3</sub> depends on the relative mag-  
 140 nitude of  $E_{A/B}$  versus  $E_{C/D}$ . To verify that, we ran one additional set of MC simulations on a

141 system where  $E_{A/B} = 1.5 u$ . In this case, we expected the critical temperature for the formation  
 142 of  $\alpha\text{-TTP}_3$  to be very similar to that of  $\alpha\text{-TTP}_3$  dimerization. In fact, we observed only one sig-  
 143 moidal profile of the  $U^*$  vs.  $T^*$  plot, indicating the coalescence of the two phase transitions into  
 144 one (Figure 4), and a direct aggregation from  $\alpha\text{-TTP}$  to  $\alpha\text{-TTP}_S$ .

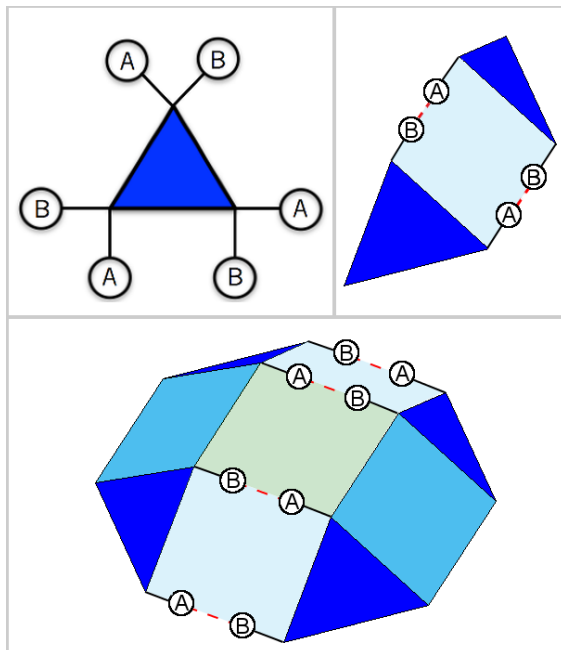


Figure 5: Oligomerization of  $\alpha\text{-TTP}_3$  (Blue triangle). Each oligomerization step involves the formation of at least two  $IS_{A/B}$  contacts. Newly formed contacts are represented by dashed lines.

145 Although our MC runs depict the clear tendency of the system to form  $\alpha\text{-TTP}_S$ , statistically,  
 146 we obtained the formation of imperfect spherical complexes with an average aggregation number  
 147 of 20.2. The presence of defects is visible from the average energy per protein reported in Figure  
 148 4, which is always larger than the ideal value of  $-4u$  even for low values of the thermal energy. The  
 149 formation of defectuous  $\alpha\text{-TTP}_S$  assemblies is due convergence issues related to the appearance of  
 150 kinetically trapped states at lower temperatures. These states are dominant in MC runs at values of  
 151  $T^*u \ll E_{A/B}$ , which yielded poorly aggregated structures, with energies consistently higher than  
 152 the best organized  $\alpha\text{-TTP}_S$ -like structures found at higher values of the temperature. Improvement  
 153 of the sampled structures at  $T^* \leq 0.06$  was obtained by applying 30 cycles of simulated annealing,  
 154 between  $T^* = 0.07$  and the target temperature. In this case, we could observe the formation more

155 regular  $\alpha$ -TTP<sub>S</sub> structures with aggregation number  $\approx 23$ .

156 Overall, our data are in optimal agreement with the native gel electrophoresis experiment re-  
157 ported in ref.<sup>11</sup> In particular, natively folded  $\alpha$ -TTP, which can oligomerize only through the A/B  
158 interface, showed the predominance of a monomeric form, with residual presence of low-weight  
159 aggregates (dimer, tetramer, figure 3). On the contrary, after triggering aggregation by unfolding  
160 of the N-terminus, the proteins assembled into stable  $\alpha$ -TTP<sub>S</sub>, which showed no tendency to dis-  
161 aggregate back into lighter oligomers in further incubation tests over a time-window of 24 hours.<sup>11</sup>  
162 Thus, the experimental condition of the real system would correspond to the region of  $\mathbf{T}^* \approx 0.06$ -  
163 0.08 in our toy system, where the folded state is mostly monomeric, while the partially unfolded  
164 state yields almost pure  $\alpha$ -TTP<sub>S</sub> (figure 6).

165 Apart from  $\alpha$ -TTP and  $\alpha$ -TTP<sub>S</sub>,  $\alpha$ -TTP<sub>3</sub> is another oligomerization state for which, when  
166 IS<sub>C/D</sub> is active, there exists a region of thermodynamical stability. The nature of  $\alpha$ -TTP<sub>3</sub> as a true  
167 thermodynamic stable aggregate is confirmed by a diverging specific heat in correspondence of the  
168 boundary transition temperatures  $\mathbf{T}^* \approx 0.11, 0.18$  (figure 4, inset), which is a clear indication of the  
169 presence of two separate phase transitions. Experimentally, native  $\alpha$ -TTP rapidly evolves into  $\alpha$ -  
170 TTP<sub>S</sub>, indicating that ambient thermodynamic conditions fall in the region of stability of the phase  
171 diagram for  $\alpha$ -TTP<sub>S</sub>, nonetheless  $\alpha$ -TTP<sub>3</sub> should be the dominating species in an intermediate  
172 region at higher temperature. The narrowness of such a region depends on the relative strength of  
173 IS<sub>A/B</sub> and IS<sub>C/D</sub>.

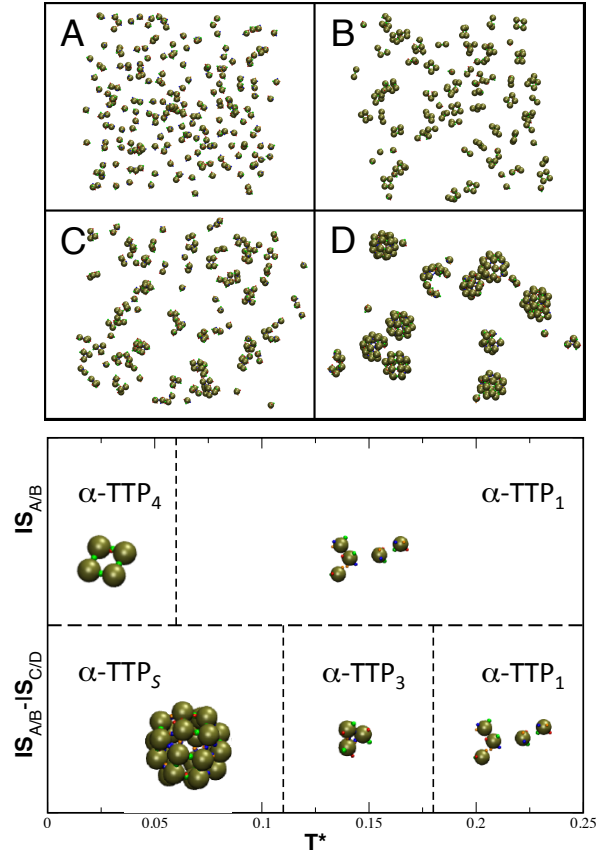


Figure 6: *Top*: Aggregation states of  $\alpha$ -TTP at different conditions. *Panel A*: monomeric dispersion at  $\mathbf{T}^* = 0.27$ ; *Panel B*: aggregation of low-weight oligomers at  $\mathbf{T}^* = 0.04$ , for the system with only active  $IS_{A/B}$ ; *Panel C*: system with all active  $IS$ 's at intermediate  $\mathbf{T}^* = 0.13$ , where trimers begin to form; *Panel D*: same system at  $\mathbf{T}^* = 0.07$ , characterized by formation of  $\alpha$ -TTP<sub>S</sub>. *Bottom*: Phase diagram with dominating species at different conditions of temperature and folded state. The region between 0.06 and 0.11  $\mathbf{T}^*$  corresponds to the experimentally observed behavior, with either properly folded monomers, or assembled  $\alpha$ -TTP<sub>S</sub>.

174 According to our results, point mutations at the surface of the protein that either weaken the  
 175  $IS_{A/B}$  interaction or, alternatively, strengthen the  $IS_{C/D}$  interaction would both yield to an en-  
 176 largement of the region of stability for  $\alpha$ -TTP<sub>3</sub>. In our previous study, the analysis of the  $\alpha$ -  
 177 TTP<sub>S</sub> protein-protein interfaces at the four-fold symmetry revealed hydrophobic contact areas that  
 178 are mostly responsible for binding (figure 7).<sup>11</sup> Introducing specific point-mutations of these key  
 179 residues can have stark impact on particle assembly. Any disruptive mutation like, for example  
 180 F165R, where a positively charged residue is introduced into the hydrophobic patch through site-  
 181 directed mutagenesis, should weaken the cooperative effects of the  $IS_{A/B}$ . In this way, it should be

182 possible to inactivate progressive dimerization, favoring instead  $\alpha$ -TTP<sub>3</sub> as the dominating species.

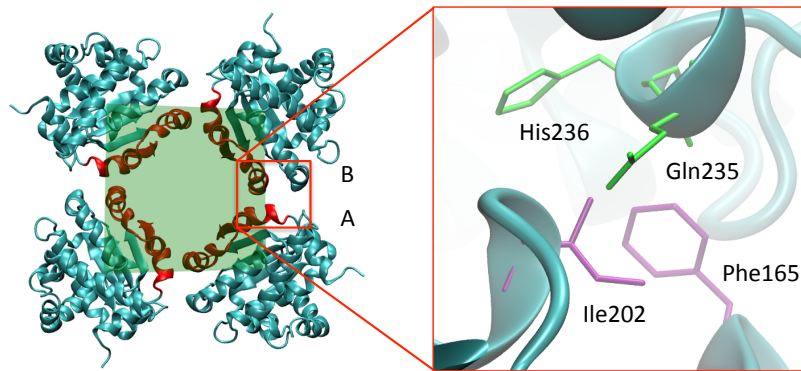


Figure 7:  $\alpha$ -TTP interaction at the *Left*: Structure of the assembly of  $\alpha$ -TTP around the  $C_4$  symmetry axis of  $\alpha$ -TTP<sub>S</sub>. The structure interlocks via multiple  $IS_{AB}$ , topologically located at the ligand binding site region of  $\alpha$ -TTP. *Right*: the most relevant residues responsible for the binding at this interface are shown in licorice. The green and purple colors refer to amino acids belonging to two different  $\alpha$ -TTP units.

183 Interestingly, partial  $\alpha$ -TTP aggregation is not strictly bound to the unfolding of the N-terminus.  
184 Rather, low-weight aggregates of folded  $\alpha$ -TTP can be formed by binding through the natively  
185 solvent-exposed A/B interface. The transition from monomers to A/B dimers or tetramers is de-  
186 termined by the balance between the A/B binding energy and the dimerization entropy loss. In  
187 solution, this balance is in favor of the monomeric species. Nonetheless, external factors like  
188 pre-organization of the monomers on a surface, may favor the formation of such oligomers. Data  
189 by Arai and co-workers<sup>3</sup> reported that a mixture of  $\alpha$ -TTP,  $\alpha$ -tol, and lipid fractions containing  
190 different PIPs such as PI(3,4)P2 or PI(4,5)P2 induced the formation of  $\alpha$ -TTP tetramers. In  
191 this study, also crystals of such ternary mixtures were analyzed by x-ray crystallography at 2.6 or  
192 2.0 Å resolution, respectively. Superposition of the open (PDB: 1OIZ), closed (PDB: 1OIP), and  
193 PI(4,5)P2-bound (PDB: 3W68) structures revealed a semi opened mobile gate conformation in the  
194 ternary structure of  $\alpha$ -TTP.<sup>3</sup> It was also shown that such ternary complexes possess inter-membrane  
195 transfer activity *in vitro* when using donor or acceptor liposomes doped with PIPs. Neither  $\alpha$ -TTP<sub>3</sub>  
196 nor  $\alpha$ -TTP<sub>S</sub> aggregates were reported in this study. This is in accordance with our previous ob-  
197 servations that aggregation into spherical particles occur only when  $\alpha$ -tol is bound to  $\alpha$ -TTP with  
198 the mobile gate being in its fully closed state, and subsequent unfolding of the N-terminus has

199 unmasked the trimeric interaction interface of  $\alpha$ -TTP.

## 200 **Concluding remarks**

201 Our model provides a description of the thermodynamically stable aggregation states of  $\alpha$ -TTP  
202 that is consistent with experimental data. We report the existence of a metastable low-weight  
203 oligomerization state ( $\alpha$ -TTP<sub>3</sub>) that is key to the fast and regular assembly of  $\alpha$ -TTP<sub>S</sub>.

204 Expression of functional mutants with different assembling properties should be feasible by  
205 minimal modifications of the native sequence. Studies by Kortemme et al.<sup>32</sup> have shown in other  
206 systems that in general a single mutation is sufficient to redesign functional protein-protein inter-  
207 faces and thus alter specificity. Interestingly, self-assembly into a similar spherical homo-multimer  
208 structure composed by 24 monomers has been reported in ferritin, an evolutionarily unrelated pro-  
209 tein than  $\alpha$ -TTP.<sup>33</sup> In a very recent study, Dmochowsky and coworkers showed that that a homo  
210 dimeric state represents a common intermediate during protein cage assembly of the 24-meric fer-  
211 ritin, and that the dimer/24mer balance can be experimentally altered by introducing single positive  
212 charges at sites along the dimer-dimer interface.<sup>34</sup> It is important to notice that even small varia-  
213 tions in the binding affinity can have a very large impact in the assembling process, due to the  
214 cooperativity effects taking place during the assembling, as evidenced in the present study.

215 Functional low-weight oligomers such as  $\alpha$ -TTP<sub>3</sub> may play a crucial role for transcytosis  
216 through endothelial membranes. Our former transfection studies<sup>11</sup> showed that the transcytotic  
217 flux does not follow a diffusive regime, with larger  $\alpha$ -TTP<sub>S</sub> transfecting at a faster rate than smaller  
218 monomeric  $\alpha$ -TTP. In fact, the delay observed in the  $\alpha$ -TTP flux may imply that aggregation of a  
219 minimal unit larger than the monomer is required to activate the transport. Further studies in-  
220 troducing disruptive mutations at the trimeric or tetrameric interfaces could help the understanding  
221 of the assembly kinetics and thermodynamics at different physiological conditions. Most impor-  
222 tantly, they may lead to the identification of the minimal biologically active units that are active for  
223 the transfection through the endothelium, a crucial step toward the engineering of these protein for

224 targeted drug delivery.

## 225 **Associated Content**

226 The authors declare no competing financial interest.

## 227 **Acknowledgement**

228 This work was supported by the Research Council of Norway (RCN) through the CoE Hylleraas  
229 Centre for Quantum Molecular Sciences Grant No. 262695, and by the Norwegian Supercomput-  
230 ing Program (NOTUR) (Grant No. NN4654K). HBK received funding from the European Union  
231 Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agree-  
232 ment No. 704491.

## 233 **References**

- 234 (1) Traber, M. G.; Sokol, R. J.; Burton, G. W.; Ingold, K. U.; Papas, A. M.; Huffaker, J. E.; Kay-  
235 den, H. J. Impaired ability of patients with familial isolated vitamin E deficiency to incorpo-  
236 rate  $\alpha$ -tocopherol into lipoproteins secreted by the liver. *J. Clin. Invest.* **1990**, *85*, 397–407.
- 237 (2) Traber, M. G.; Burton, G. W.; Hamilton, R. L. Vitamin E trafficking. *Ann. N. Y. Acad. Sci.*  
238 **2004**, *1031*, 1–12.
- 239 (3) Kono, N.; Ohto, U.; Hiramatsu, T.; Urabe, M.; Uchida, Y.; Satow, Y.; Arai, H. Impaired  $\alpha$ -  
240 TTP-PIPs interaction underlies familial vitamin E deficiency. *Science* **2013**, *340*, 1106–1110.
- 241 (4) Meier, R.; Tomizaki, T.; Schulze-Briese, C.; Baumann, U.; Stocker, A. The molecular basis  
242 of vitamin E retention: Structure of human  $\alpha$ -tocopherol transfer protein. *J. Mol. Biol.* **2003**,  
243 *331*, 725–734.

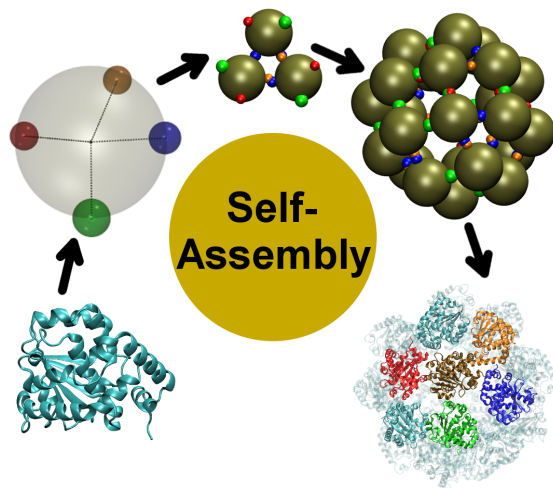


- 244 (5) Min, K. C.; Kovall, R. a.; Hendrickson, W. a. Crystal structure of human alpha-tocopherol  
245 transfer protein bound to its ligand: implications for ataxia with vitamin E deficiency. *Proc.*  
246 *Natl. Acad. Sci. U. S. A.* **2003**, *100*, 14713–14718.
- 247 (6) Sha, B.; Phillips, S. E.; Bankaitis, V. A.; Luo, M. Crystal structure of the *Saccharomyces*  
248 *cerevisiae* phosphatidylinositol-transfer protein. *Nature* **1998**, *391*, 506–510.
- 249 (7) Stocker, A.; Tomizaki, T.; Schulze-Briese, C.; Baumann, U. Crystal structure of the human  
250 supernatant protein factor. *Structure* **2002**, *10*, 1533–1540.
- 251 (8) He, X.; Lobsiger, J.; Stocker, A. Bothnia dystrophy is caused by domino-like rearrangements  
252 in cellular retinaldehyde-binding protein mutant R234W. *Proc. Natl. Acad. Sci. U. S. A.* **2009**,  
253 *106*, 18545–18550.
- 254 (9) Stocker, A. Molecular mechanisms of vitamin E transport. *Ann. N. Y. Acad. Sci.* **2004**, *1031*,  
255 44–59.
- 256 (10) Chung, S.; Ghelfi, M.; Atkinson, J.; Parker, R.; Qian, J.; Carlin, C.; Manor, D. Vitamin E and  
257 phosphoinositides regulate the intracellular localization of the hepatic  $\alpha$ -tocopherol transfer  
258 protein. *J. Biol. Chem.* **2016**, *291*, 17028–17039.
- 259 (11) Aeschimann, W.; Staats, S.; Kammer, S.; Olieric, N.; Jeckelmann, J. M.; Fotiadis, D.;  
260 Netscher, T.; Rimbach, G.; Cascella, M.; Stocker, A. Self-assembled  $\alpha$ -Tocopherol Trans-  
261 fer Protein Nanoparticles Promote Vitamin E Delivery Across an Endothelial Barrier. *Sci.*  
262 *Rep.* **2017**, *7*, 1–13.
- 263 (12) Saraiva, C.; Praça, C.; Ferreira, R.; Santos, T.; Ferreira, L.; Bernardino, L. Nanoparticle-  
264 mediated brain drug delivery: Overcoming blood-brain barrier to treat neurodegenerative  
265 diseases. *J. Control. Release* **2016**, *235*, 34–47.
- 266 (13) Deli, M. A.; Ábrahám, C. S.; Kataoka, Y.; Niwa, M. Permeability studies on in vitro blood-

- 267 brain barrier models: Physiology, pathology, and pharmacology. *Cell. Mol. Neurobiol.* **2005**,  
268 25, 59–127.
- 269 (14) Zhang, Z.; Glotzer, S. C. Self-assembly of patchy particles. *Nano Lett.* **2004**, 4, 1407–1413.
- 270 (15) Sciortino, F.; Giacometti, A.; Pastore, G. Phase Diagram of Janus Particles. *Phys. Rev. Lett.*  
271 **2009**, 103, 237801.
- 272 (16) Kraft, D. J.; Ni, R.; Smallenburg, F.; Hermes, M.; Yoon, K.; Weitz, D. A.; van Blaaderen, A.;  
273 Groenewold, J.; Dijkstra, M.; Kegel, W. K. Surface roughness directed self-assembly of  
274 patchy particles into colloidal micelles. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, 109, 10787–  
275 10792.
- 276 (17) Wilber, A. W.; Doye, J. P. K.; Louis, A. A.; Noya, E. G.; Miller, M. A.; Wong, P. Re-  
277 versible self-assembly of patchy particles into monodisperse icosahedral clusters. *J. Chem.*  
278 *Phys.* **2007**, 127, 085106.
- 279 (18) Carlsson, F.; Linse, P.; Malmsten, M. Monte Carlo simulations of polyelectrolyte-protein  
280 complexation. *J. Phys. Chem. B* **2001**, 105, 9040–9049.
- 281 (19) Fantoni, R.; Gazzillo, D.; Giacometti, A.; Miller, M. A.; Pastore, G. Patchy sticky hard  
282 spheres: Analytical study and Monte Carlo simulations. *J. Chem. Phys.* **2007**, 127, 234507.
- 283 (20) Villar, G.; Wilber, A. W.; Williamson, A. J.; Thiara, P.; Doye, J. P.; Louis, A. A.;  
284 Jochum, M. N.; Lewis, A. C.; Levy, E. D. Self-assembly and evolution of homomeric protein  
285 complexes. *Phys. Rev. Lett.* **2009**, 102, 118106.
- 286 (21) Liu, H.; Kumar, S. K.; Douglas, J. F. Self-Assembly-Induced Protein Crystallization. *Phys.*  
287 *Rev. Lett.* **2009**, 103, 018101.
- 288 (22) Li, Y.; Shi, T.; An, L.; Huang, Q. Monte Carlo Simulation on Complex Formation of Proteins  
289 and Polysaccharides. *J. Phys. Chem. B* **2012**, 116, 3045–3053.

- 290 (23) Li, W.; Persson, B. A.; Morin, M.; Behrens, M. A.; Lund, M.; Zackrisson Oskolkova, M.  
291 Charge-induced patchy attractions between proteins. *J. Phys. Chem. B* **2015**, *119*, 503–508.
- 292 (24) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosys-  
293 tems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*,  
294 10037–10041.
- 295 (25) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C.; Brook, S.;  
296 Brook, S.; Brook, S. Comparison of multiple AMBER force fields and development of im-  
297 proved protein backbone parameters. *Proteins* **2006**, *65*, 712–725.
- 298 (26) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of  
299 state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092.
- 300 (27) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*, 2nd ed.; Academic Press, Inc.:  
301 Orlando, FL, USA, 2001.
- 302 (28) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: New York,  
303 NY, USA, 1987.
- 304 (29) Wu, D.; Chandler, D.; Smit, B. Electrostatic analogy for surfactant assemblies. *J. Phys. Chem.*  
305 **1992**, *96*, 4077–4083.
- 306 (30) Miller, M. A.; Amon, L. M.; Reinhardt, W. P. Should one adjust the maximum step size in a  
307 Metropolis Monte Carlo simulation? *Chem. Phys. Lett.* **2000**, *331*, 278–284.
- 308 (31) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.*  
309 **1996**, *14*, 33–38.
- 310 (32) Kortemme, T.; Joachimiak, L. A.; Bullock, A. N.; Schuler, A. D.; Stoddard, B. L.; Baker, D.  
311 Computational redesign of protein-protein interaction specificity. *Nat. Struct. Mol. Biol.*  
312 **2004**, *11*, 371–379.

- 313 (33) Lawson, D. M.; Artymiuk, P. J.; Yewdall, S. J.; Smith, J. M. A.; Livingstone, J. C.; Treffry, A.;  
314 Luzzago, A.; Levi, S.; Arosio, P.; Cesareni, G.; Thomas, C. D.; Shaw, W. V.; Harrison, P. M.  
315 Solving the structure of human H ferritin by genetically engineering intermolecular crystal  
316 contacts. *Nature* **1991**, *349*, 541–544.
- 317 (34) Pulsipher, K. W.; Villegas, J. A.; Roose, B. W.; Hicks, T. L.; Yoon, J.; Saven, J. G.; Dmo-  
318 chowski, I. J. Thermophilic ferritin 24mer assembly and nanoparticle encapsulation modu-  
319 lated by interdimer electrostatic repulsion. *Biochemistry* **2017**, *56*, 3596–3606.



## Table of Contents