## A study of the self-assembly of anisotropic particles mimicking the pathological agregation of proteins

Martin FLAMENT

Ecole Normale Supérieure, Biology Department, PSL University Paris

13/02/2023 - 30/06/2023

Some diseases are caused by the pathological aggregation of proteins into fibers. We can mimick the behavior of these proteins with anisotropic colloids and computer simulations which simulate the assembly of assymetric particles. Simulations have shown that the aggregation can be constrained because the shape of a particle prevents it from binding isotropically to the aggregate. This constraint, interpreted as frustration, fosters the elongation of the aggregate into a fiber. The role of this frustration has been well described for identical particles interacting, but the way this frustration appears for 2 different particles interacting is not yet understood. Therefore I simulated the self-assembly of a solution of 2 triangular particles with restricted directional interactions. From these simulations I detail the different types of aggregates the particles form, and how their interactions affect their self-assembly. I identify different forms of frustration that could constrain the assembly of the particles, and assess their importance in doing so.

In sec. 1 I show that we want to understand how proteins assemble into fibers when they are misfolded by simulating the assembly of particles with restricted interactions.

In sec. 2 I present a model which restricts the interactions between particles based on lock-key complementarity. I study the behaviour of a solution of 2 particles with these restricting rules by describing the aggregates formed at equilibrium.

In sec. 3 I detail the different aggregates formed by the particles at equilibrium. I measure different forms of frustration to assess whether the systems I simulated are frustrated, and evaluate how this frustration affects the self-assembly of the particles.

## 1 Introduction : frustrated self-assembly of protein-like particles

Some diseases are caused by the aggregation of misfolded proteins into fibers or plaques. Although the link between protein shape and function is well known, the reason why certain proteins aggregate into fibers is not well understood. The self-assembly of one shape of particle has been well described, especially the role frustration plays in constraining this self-assembly. The presence of a second particle type should add complexity to the way particles can assemble, though it is hard to guess how this complexity can affect the self-assembly of the particles. In subsection 1.1 I explain how the structure of a protein affects its self-assembly properties, and give examples of diseases caused by aggregation into fibers of misfolded proteins. I then explain in subsection 1.2 how colloids help understanding the self-assembly of anisotropic particles. Finally in subsection 1.3 I show that the study of the self-assembly of identical anisotropic particles has helped understand how incompatible interactions foster the formation of fibers, and how adding a second particle type creates more design space for systems with anisotropic particles to investigate.

### **1.1** Some diseases are characterised by the aggregation of mutated proteins into fibers

The function of proteins are tightly linked to their shape and their assembly into functional complexes. In particular diseases can be due to proteins being misfolded or mutated. In this section I discuss the example of diseases caused by the assembly of abnormal proteins into fibers. I show that this change of the protein's shape causes an abnormal aggregation by affecting the interactions of the protein with its environment.

Drepanocytosis is a genetic blood disease characterised by a change in the shape of the blood cells, which take the shape of a sickle. These blood cells are also less deformable, and can block blood vessels, causing anemia [1]. This disease is caused by a single mutation in the sequence of the gene coding for  $\beta$ -hemoglobin, replacing a glutamate amino-acid by a valine [2]. This modification causes hemoglobin tetramers to assemble into fibers when they unbind with dioxygen [3]. These fibers, tethered to the membrane, will then rigidify the blood cells, giving them their pathological sickle shape (**Figure 1**).

Other diseases, such as Parkinson or Alzheimer, are also caused by the misfolding of proteins which induces their self-assembly into pathogenic fibrous protein aggregates [4], which bodes for the existence of a mechanism explaining why some misfolded proteins spontaneously aggregate into fibers. Indeed, the self-assembly of proteins highly depends on the presence of specific sites with which proteins bind to one another and that are in most cases determined by the structure of the protein [5]. Misfolding can affect protein self-assembly either because the proteins gain an additional binding site, changing its interactions with other proteins(**Figure 2**,



Figure 1: Drepanocytosis is caused by the polymerisation of mutated hemoglobins (from Sundd *et al.*, 2019) HbA is the healthy version of a hemoglobin tetramer, which does not bind with other hemoglobins. A mutation on both genes coding for the  $\beta$ -hemoglobin creates a new form of hemoglobin tetramer, called HbS, with a new interaction site. This site auses these mutated tetramers to polymerize into fibers when the hemoglobins unbind with dioxygen. These fibers deform the blood cells into sickles.



Figure 2: Self-assembly of proteins is dictated by the way the proteins interact with each other(from Garcia-Seisdedos *et al.*, 2017)

in red), similarly to the way mutated hemoglobins polymerize in drepanocytosis, or because the proteins are unstable and stabilize by aggregating into fibers (**Figure 2**, top), like in the case of Parkinson disease [6].

#### **1.2** Colloids are great tools to study self-assembly of anisotropic particles

Self-assembly is the way particles spontaneously aggregate to form large aggregates. Proteins assemble by binding with specific sites, which cause inter-proteic interactions to be highly anisotropical. Therefore, the study of the self-assembly of proteins needs to simulate the self-assembly of anisotropic particles.

Colloids have for a long time made good models to study the aggregation of particles [7], as they can spontaneously aggregate together, for instance by the effect of a depletion force [8]. This entropic force appears when 2 particles are close with small solutes in water. The 2 particles will then bind together to minimize their excluded volume while chasing out the so-

lutes. In the particular example of protein self-assembly, micrometer-sized colloids can mimick protein behaviour to great effect, as the shape of the surface of the colloid restricts how it can bind with other particles. For example, it is possible to print complementary shapes on colloids to restrict their interactions, based on a lock-key complementarity. Colloids can also be printed into 2D polygons, such as hexagons or triangles, in which case they will pave a plane when binding. It is thus possible to restrict the interactions between the different particles by printing complementary shapes on the faces of the polygonal colloid.

However, colloid experiments can take a lot of time to run, therefore fine-tuned experimental designs can be hard to understand in a systematic way. Therefore, numerical simulations can help to simulate self-assembly of anisotropic particles by simulating finely-tuned particle designs in greater number, and the results obtained can then be confirmed with colloidal experiments.

# **1.3** Incompatible interactions in anisotropic particles induce frustration which favors aggregation into fibers

Numerical simulations can help describe in a more systematic way the self-assembly of anisotropic particles. These simulations have for example identified the role of frustration in the elongation of particle aggregates. Numerical simulations are a very interesting tool in unraveling the structure of particle self-assembly [9]. These simulations can for example simulate the interaction between particles which can deform upon binding with one another. These simulations show that when a free particle wants to bind to particles in an aggregate, it can either bind to a tip of the aggregate, where it will increase the surface of the aggregate, or to the side, where the free particle will have to deform itself as particles are more packed at this place of the aggregate (Figure 3). Thus, fibers form when the cost of the particle adding at the tip of the aggregate is lower than the deformation the particle should suffer to bind to the side [10]. This deformation cost is the result of an incompatibility between the shape of the surface of the aggregate and that of the particle, which needs to deform to fit perfectly.

Here frustration is associated to the shape of the particle, but it can also be associated, as it is with colloids, to complementary interactions between faces. Thus we can simulate the self-assembly of anisotropic particles by associating energies to the interactions between the different faces of the particle. Frustration then arises from the presence of less favorable interactions or even repulsive interactions. Such simulations have shown that many different aggregates can arise from the aggregation of such particles, including fibers [11]. These aggregates of lower dimensionality arise from the presence of these unfavorable interactions which constrain the binding of a particle to an aggregate.

However, proteins can have multiple binding sites with which they can interact with different species. Therefore, it is important to study the role of competition between the aggregation of different species together in the self-assembly of particles. Indeed, different particle types can



Figure 3: Fibers form when incompatible interactions favor the elongation of an aggregate rather than its broadening (from Lenz *et al.*, 2017)

segregate, if they cannot bind together and particles of each species will only bind with other particles of the same type, but if two particle types can bind with one another, the interaction between the two particle types can create new ways for particles to assemble. Therefore in this report I will study the self-assembly of particles with different shapes and compare it to the self-assembly of particles with themselves.

## 2 A model restricting the self-assembly of 2 particles in interaction

In this chapter I present a model suggested by our collaborators at ESPCI, O. du Roure, J. Heuvingh and Mayarani M. This model restricts directional interactions by introducing 3 different face types with specific interaction rules based on a lock-key complementarity. Using this model, I can design particles with different interaction rules, and pair these different particle types. I want to study how the different particle types interact by simulating the self-assembly of these particles on a lattice, and describing geometrically the aggregates these particles can assemble into.

In subsection 2.1 I present the model of restricted interactions I studied, which I called FLK model. Especially I present the diversity in the different designs I can build using this model. In subsection 2.2 I explain how I simulated the different pairs of particles on a lattice model developed by my advisor L. Koehler [11] and in section 2.3 how I ran the simulations to study the equilibrium states of each pair of particles. Then in subsection 2.4 I present the geometrical properties I measure on the aggregates formed to describe the shape of this aggregate. Finally in subsection 2.5 I present the different ways the systems I simulated can be frustrated, and how I measure them.

#### 2.1 Building particles with complementary faces

Colloidal particles interactions can be restricted by printing complementary shapes on different faces. The model I present here aims to mimick this by considering different face types which interact by complementarity rules. I then explore the diversity of particles and particle pairs we can achieve with this model.

The model suggested by our collaborators considers particles whose faces could be of one of 3 types, illustrated in **Fig 4 b**). 2 of these faces types have either a nodge or an indent, which restricts them to only bind with one another in by Lock-Key complementarity, which gives them their name respectively of Key (K) and Lock (L). The last face type has nothing at its surface, and is thus called Flat (F), and can only bind to itself, the other face types binding more rarely with it because of their shape. Moving forward, I call this model the FLK model, in reference to the 3 face types it considers. The different interaction rules between each face type are summarized on **Figure 4**, where I also show how I identified each face by a color.

To measure the different interactions, I associate to each distinct face pair an energy, denoted as  $E_{ab}$ , with a and b denoting the 2 faces in contact. By convention, favorable interactions are associated to a negative energy and unfavorable ones to a positive energy. The intensity of the interaction between colloidal surfaces depends on the size of the surface, as the bigger the surface is, the more their interaction reduces the surface tension of both colloids. Therefore, L and K faces should bind more strongly than F faces, as the nodge and indent increase the



Figure 4: The FLK model restricts interactions between faces with noncomplementary shapes. The 3 different face types are identified by a colour, F in yellow, K and L in dark and light blue. The interactions of each face with the other are represented by the arrows. An unfavorable interaction is associated to a positive energy, and a favorable one to a negative energy.

surface with which the faces bind. I thus chose  $E_{LK} = 1.5E_{FF} = -12k_bT$ . For the repulsive interactions, I chose  $E_{FL} = E_{FK} = E_{LL} = E_{KK} = E_{repul} = 10k_bT$ , which ensures that repulsion can destabilize the binding of 2 particles by being of the same magnitude as attraction.

I studied the self-assembly of triangular particles with these face types. I first enumerated the 11 different particles we can build using these 3 face types. To reference these particles, I denote them by the type of the 3 faces of the particle. The first letter corresponds to a face I chose as reference, and the next letters correspond to the face type of the other faces read clockwise. For example, FFF corresponds to a particle with only F faces, and FLK is the particle with the 3 different face types. The latter also has an enantiomer FKL which I also consider. I then designed the different pairs of particles by pairing every particle type with another particle type. However, since K and L have equivalent behaviours, some particle pairs were redundant with one another, so I only considered pairs of particle with distinct behaviours.

#### 2.2 Embedding the FLK model onto a lattice

To simulate the behaviour of a pair of particles, I simulate particles on a lattice. The way these particles interact is given by their identity and face composition.

I studied the self-assembly of each particle pair by simulating the self-assembly of triangular particles on a triangular lattice. This lattice is made up of  $N_{sites}$  sites, which can either be empty or occupied by a particle. Each particle *i* is described by its position on the lattice  $\mathbf{r}^{(i)} = (x^{(i)} \ y^{(i)})$ , its orientation  $\varphi^{(i)}$  and its identity. When 2 particles are on neighbouring sites, they interact each one of their faces. The faces engaged in this interaction are determined by the position, orientation and identity of the 2 particles, and are identified by a pair of faces. By denoting the faces of a particle of type 1 (a,b,c) and those of a particle of type 2 (a',b',c'), there are 21 distinct pair of faces particles can form, since we consider (a,b)=(b,a). Then, by knowing the composition the composition of each particle, I can associate an energy to every one of these 21 pair of faces. This returns a matrix which encodes the interaction map  $E_{ab}$  of

the particle pair (Figure 5).



Figure 5: The interaction matrix of the (FLK, FKL) (1 and 2, respectively) particle pair. This matrix associates to every face pair possible in the system an energy corresponding to the nature of the faces engaged in the pair. The dotted lines separate the interactions between 2 FLK particles (top left), thos between 2 FKL particles (bottom right), and those between FLK and FKL particles (bottom left). The pairs shaded in grey are redundant with other unshaded face pairs.

Therefore, at any given time, I can calculate the energy per particle of the system as  $E = \sum N_{kl} E_{kl} \frac{1}{N_{particles}}$  with k and l chosen within (a,b,c,a',b',c'). I calculate the energy per particle to have an intensive value of energy to describe the state of the system. When I simulate the particles, I want to minimize this energy per particle to describe the equilibrium state of the system.

#### 2.3 Equilibrating the systems with a Monte-Carlo algorithm

#### To study how particles aggregate I bring the system to an equilibrium with a Monte-Carlo algorithm which I detail here.

For each pair of particle (*part*, *part'*), I simulate a system composed of  $N_{particles} = N + N' = 140 + 140$ , which I fix to have a particle density  $\frac{N_{particles}}{N_{sites}}$  between 0.1 and 0.2. This ensures that self-assembly is driven solely by the interactions between the particles. I chose N and N' to be equal so the density of each particle is the same. Initially, the positions and orientations of the particles are chosen at random. For each pair of particle, I simulated 10 independent systems.

The system is then equilibrated by a Monte-Carlo algorithm: the system starts at a high energy  $k_bT_0$  and progressively decreases in temperature until reaching a temperature of  $1k_bT$  to reach an equilibrium state. The system goes through 100 temperature values evenly spaced between  $T_0$  and 1. For each temperature T, the code executes  $N_{steps} * N_{sites}$  steps. At each step,



Figure 6: To optimize the parameters of the simulation, I measure the energy of simulations run with different initial temperatures and steps. The simulations were run using a system with FFK and FLL particles.

one particle is chosen at random, and can do one of 2 actions:

- The particle can change orientation without moving. A new orientation  $\varphi'^{(i)}$  is chosen randomly, and the particle replaces its old orientation to this new one.
- The particle can move without rotating. A new position  $\mathbf{r}^{(i)}$  is chosen randomly between the remaining free sites, and the particle moves from  $\mathbf{r}^{(i)}$  to  $\mathbf{r}^{(i)}$ .

Each action changes the energy of the system by  $\Delta E$  based on the past and new interactions. If  $\Delta E < 0$ , the action is accepted and the particle keeps its new position and orientation. This ensures that the system always tries to minimize its energy. If  $\Delta E > 0$  it is accepted with a probability  $p = \exp(-\Delta E/T)$ . This allows that the simulation explores different equilibrium states for the system.

To choose the values for  $T_0$  and  $N_{steps}$ , I simulated one system with different values for  $T_0$  and  $N_{steps}$ , and measured their final energy levels, which I represent on **figure 6**. The goal was to find values for which I measured the lowest energy, so the system is equilibrated as most as possible, without making my simulations run for too long. On Figure 6, I observe that the more  $N_{steps}$  there were, the more the system was equilibrated, with a plateau near  $1000N_{steps}$ , therefore I chose  $N_{steps} = 1100$ . Initial temperature does not seem to have a strong effect on the equilibration of the system when the simulation is long, as long as they aren't too high. The choice of temperature is thus a tradeoff between spending time at high temperatures, to explore different stable configurations, and spending more time at low temperatures to reach a more accurate equilibrium state. Initial temperatures between 6K and 8K allow to better equilibrate the system as their plateau is clearer around  $900N_{steps}$ , therefore I chose  $T_0 = 7K$ .



#### Figure 7: A systematic view of the simulations using the FLK model.

The particles I enumerated are represented on the left part of the graph. Each square represents a simulation, with the particle types found in the simulations being at the left and top of the graph. A zoomed-in view of a simulation with FLK particles can be found at the top left.



Figure 8: The self-assembly of particles can be described by measuring geometrical properties of the aggregates formed. In the top simulation, particles aggregate into one dense aggregate. On the other side, in the bottom simulation particles form small aggregates. Some aggregates are spherical, and others are more elongated and form short fibers.

#### 2.4 Characterisation of the aggregates at equilibrium

The code I used to run the simulations returns the positions and orientations of each particles, from which it also gives information on the different aggregates formed. On Figure 7 I show an example of equilibrium state for each simulation, from which I can recognise different shapes of aggregates. In this chapter I explain how I use this information to describe the shape taken by these aggregates.

The goal of the simulations was to find pairs of particles for which I observed fibers. Fibers are generally elongated, have a high surface/volume ratio, and need a lot of particles to form. Therefore, I measure 3 geometrical properties to describe the shape of aggregates formed:

- The size of the aggregates, which I define as the number of particles in the aggregates  $n_p$ . This parameter serves mostly at separating small oligomers from larger aggregates. This value can range from 0 to  $N_part$ , here 280.
- The mean number of neighbours, or compactness of the system. This serves as a measure of the volume/surface ratio and is measured over the whole system, not individual aggregates. Since fibers tend to increase their surface compared to spherical aggregates of the same size, they should have smaller compactness compared to larger aggregates. This value ranges from 0 to the number of neighbours a particle can have at most, here with triangular particles it can go up to 3.
- The anisotropy of the aggregate, measured as  $1 \frac{\lambda_{-}}{\lambda_{+}}$ ,  $\lambda_{-}$  and  $\lambda_{+}$  being the smaller and greater eigenvalues of the inertia matrix of the aggregate. The inertia matrix is a 2x2 matrix measured as

 $I = \begin{pmatrix} \sum_{i=1}^{N_{part}} x_i^2 & -\sum_{i=1}^{N_{part}} x_i y_i \\ -\sum_{i=1}^{N_{part}} x_i y_i & \sum_{i=1}^{N_{part}} y_i^2 \end{pmatrix}$ . These eigenvalues always exist because the matrix is symmetrical, and they represent the principal axes of the ellipsoid formed by the cluster. Anisotropy is null if the aggregate is perfectly spherical, and is closer to 1 if the aggregate is elongated.



Figure 9: Stoechiometric frustration measures the balance between the number of L and K faces in the system

**Figure 8** shows the measurements of these values measured for 2 different pairs of particle, one for which particles mostly assemble into one large and dense aggregate (top), and the other for which particles aggregate either into small spheric aggregates or short elongated fibers (bottom). The assembly of the particles into the big aggregate seems reflected by the large size measured for the aggregates and the high compactness measured for these simulations, especially compared to the pair of particles shown below. The formation of elongated aggregates however seems more reflected by the high discrepancy in the anisotropy of the aggregates measured for the bottom simulation.

#### **2.5** Measuring frustration in the system

I have shown in section 1.3 that anisotropic particles can form unfavorable links which introduce frustration in the system, and that this frustration favors the formation of fibers. In this section I present how I measured frustration in the simulations I ran.

In the FLK model, L and K faces can only bind to one another. Therefore, if these 2 face types are not in the same quantity, there will be faces that cannot be paired in a favorable way. This creates a type of frustration which I call stoechiometric frustration, and measures the proportion of faces in the system that cannot pair favorably. On **Figure 9**, I show example of simulations for which stoechiometric frustration is null (left), medium (middle), and maximum (right). There I see that a system with no such frustration forms a dense aggregate, and that the more stoechiometrically frustrated the system is, the less particles seem to assemble into these dense aggregates.

Therefore, to measure the effect of stoechiometric frustration on the system, I compare for every pair of particle types the energy per particle measured in a system with a particle density of 1 and that measured in the dilute conditions explained in section 2.3. The dense simulations were equilibrated using the same annealing method as the dilute systems, only with more  $N_{steps}$  to obtain a fully equilibrated system. **Figure 10** shows an example of the simulation of a pair of particles, here (FFK, FKL), in a dense (left), and dilute (right) conditions. This system has too many K faces for all of them to be paired with L faces, therefore in the dilute condition



Figure 10: **Dense frustration measures how much energy the system has to lose by being diluted.** The left picture shows a solution of FFK and FKL particles in dense condition - the 2 vacancies are only here to allow particles to move since they cannot only move to empty sites- and the right picture shows the same particles but in the dilute conditions given in chapter 2.3. The energy per particle for each system was averaged over 10 simulations.





these K faces can be free and not destabilize the system, but in dense conditions this is not always possible, therefore the energy per particles measured is higher in the dense condition. Therefore, the lower the energy measured in a dilute system compared to the dense one, the more frustrated the system will be.

Stoechiometric frustration only defines the maximum number of faces a system has which cannot be paired favorably. Therefore, in my simulations these faces can either be found at the surface of the aggregates, or they can still interact with other faces, therefore creating an unstable interaction which frustrates the system. An example of such interaction can be found in **Figure 11**, where I identified 2 LLL particles interacting with L faces. These interactions are associated to a positive energy, and thus frustrate the system. Therefore, the count of unfavorable interactions represents the effective frustration of a simulation.

## **3** Triangular particles assemble into a low diversity of aggregates because they do not incorporate frustration.

The numerical simulations I ran return the position and orientation of particles which I illustrate in Figure 12 in the previous section. In this section I show how the geometrical measurements I introduced earlier describe the behaviour of the system. I then link these shapes to the amount of frustration measured in each system.

In subsection 3.1 I present the different types of aggregates I observe, discussing which geometrical properties introduced in the previous chapter are more relevant to differentiate each aggregate shape. In subsection 3.2 I show that effective frustration is not observed in my simulations, and that it does not foster the assembly into fibers. I then show in subsection 3.3 that is stoechiometric frustration affects particle self-assembly by inhibiting particles from assembling into dense aggregates.

#### 3.1 Most systems self-assemble into bulks, gas, and more rarely into fibers

# I observed different shapes particles aggregated into. In this chapter I differentiate the 3 most observed behaviours using geometrical properties of the aggregates.

I simulated the self-assembly of every pair of particle designed, and represented the results of these simulations on **Figure 12**. In most cases, particles either assemble into a small number of compact aggregates or they hardly aggregate at all and form oligomers of 1 to 3 particles. I label these cases respectively as bulks and gases. Bulks form when all the faces in the system can be paired, and can be more or less elongated. In a system composed of FFF and FFK particles (**Figure 12**, column 1 row 2), the K faces cannot be paired, and behave in the same way surfactants do, favoring the formation of smaller, but no less dense aggregates. Within a bulk, different particle types can be more or less segregated, depending on whether they have faces complementary to one another. For example, FKL and FLK particles (column 2 row 9) segregates within the bulk, KKL acting as a 'glue' between FFK aggregates to form the bulk.

In some rare cases, I identified some fibrous aggregates. These fibers can be either made up of 1 type of particles units, such as how FFK particles assemble -in which case the formation of fibers competes kinetically with the assembly into micelles-, or are a regular assembly of both particle types, FKK and LLL particles for example assemble in such a way. In this last case, I also observe that the fibers can be branched.

However, apart from these 3 types of aggregates, none other of the aggregates described by L. Koehler [11] were represented enough to be analysed. Therefore I concentrate on recognising fibers from the bulks and gases. To do this I illustrate the different aggregate shapes in **Figure 13**, discussing which geometrical properties describes better each shape.

Size and compactness are what describes the most bulks and gases. Indeed, gases are made



Figure 12: Most pair of particles either assemble into compact aggregates or form dispersed oligomers. I observe fiber-like aggregates for the (FFK,FFK), (FFK, KKL) and (FKK,LLL) pairs - respectively second column and row, second column and sixth row, third column and tenth row. Aggregates of intermediate sizes were hardto classify.

of at most 3 or 4 particles, and have therefore a very low compactness, whereas bulks are much larger and dense. Most bulks are made up of all the particles on the lattice, and are thus very



Figure 13: **Bulky, gaseous and fibrous aggregates can be recognised by their size, anisotropy, or compactness.** The pair of particles for which I show examples of simulation results are, from left to right: (FFF, FFF), (FFF, FFK), (FKK, LLL), (FFK, FFK), (FKK, KKK), (KKK, KKK). The 2 first form bulky aggregates, the next 2 form fibers, and the 2 last form gases. I measure for these systems compactnesses of 2.8, 2.4, 2.3, 1.9, 0.4 and 0.0. The size (in red) and anisotropy (in blue) distribution measure the proportion of particles engaged in aggregates with a given size , measured from 1 to 280, and anisotropy, measured from 0 to 1.

dense. However the presence of surfactant can limit the size of the aggregates, in this case they can be recognised by a rather high level of compactness. Anisotropy on the other hand is not fitting to describe these types of aggregates, as bulks can be more or less elongated, though the bigger they are the more spherical they get. Gases on the other hand ban either be perfectly spherical, if they are monomeric, or perfectly anisotropic if they are dimeric.

Fibers are quite the opposite: they can be of various sizes, but they are marked by a high anisotropy. However, anisotropy itself is not enough to recognise a fiber, since bulk aggregates can be as anisotropic as fibers. Therefore, fibers can be separated from bulks by their level compactness, as fibers tend to be less compact than bulks.

# **3.2** Effective frustration does not reflect the level of stoechiomettric frustration of the systems

When visualizing the results of the simulations, there was only 1 simulation in which I observed an unfavorable interaction, which I showed in section 2.5. Therefore, I lowered the intensity of  $E_{rep}$  to improve the odds for particles to form unfavorable interactions at equilibrium and observe an effect of effective repulsion on self-assembly.

On Figure 14 I represent for each pair of particles the evolution in the proportion of favorable (in blue) and unfavorable (in red) interactions when I lowered  $E_{rep}$ . The probability for



Figure 14: **Measuring the proportion of favorable and unfavorable links in the simulations.** A magnified version of the graph obtained for the system with FFK and KKK particles is shown on the top right as an example for the other graphs. Here I measure the proportion of favorable and unfavorable links observed at equilibrium in the different simulations for different repulsion intensities. The proportions measured were averaged over 10 simulations.



Figure 15: The proportion of unfavorable links decreases exponentially when repulsion energy increases. I fitted systems for which the proportion of unfavorable links changed when I reduced  $E_{repul}$ . Each point represents one pair of particles.

an unfavorable interaction to be maintained decreases exponentially when the strength of this interaction increases, therefore I expect the proportion of unfavorable junctions to be higher for low values of repulsion. I observed this for most pairs of particles: for high values of repulsion values, I observe close to no unfavorable junctions, and when repulsion is weak enough, the proportion of unfavorable junctions increases. Therefore, it is possible to study the effect of unfavorable junctions on the shape of the aggregates by studying the self-assembly of particles with a low repulsion energy.

I also observed that the proportion of favorable interactions remains constant even when I nullified repulsion. This suggests that the particles first aggregate by binding favorably, and satisfying as many favorable interactions as possible. Then, the surplus faces can either be integrated, if they do not destabilize the system too much, or they will stay in an unbound state. In the last case, the particles with these unbound faces can either be found at the surface of the aggregates, and these faces can then act as surfactants, or they will be isolated and behave like a gas.

To understand why I measured so few unfavorable interactions in the systems, I wanted to calculate the repulsion energy for which unfavorable junctions become too rare to be observed, which I label  $E_{tresh}$ . The probability of such interactions being fixed in the system decreases exponentially with the energy of repulsion, therefore I tried fitting my data with a function of the shape  $p_{E_{repul}} = (p_0 - p_{10})e^{-\frac{E_{repul}}{E_t hresh}} + p_{10}$ , where  $p_{E_{repul}}$  represents the proportion of unfavorable junctions at a given repulsion energy. In this function,  $E_{thresh}$  should measure the speed with which  $E_{repul}$  cuts unfavorable interactions out: the higher  $E_{thresh}$  is, the less unfavorable interactions I observe for high repulsion values.

I show on **Figure 15** that this equation fits well with my data, I measure a correlation coefficient of 0.99. Therefore, I could measure  $p_0$ ,  $p_{10}$  and  $E_{thresh}$  for every pair of particle. I measured no  $p_{10}$  higher than 0.1%, therefore I always considered it as null. However, I observe that  $p_0$  increases from 0 to 8% when stoechiometric frustration increases. This dependency to



Figure 16: Hexagonal particles form loops which allow the establishment of unfavorable interactions more easily than triangular particles. For an unfavorable interaction (in red) to happen, it needs to be balanced out by a favorable interaction (in green).

stoechiometric frustration was expected, because when repulsion is null, the probability for an unfavorable interaction to occur is linked to the odds with which 2 unpairable faces to meet, and thus to the density of unpairable faces in the system, which is measured by stoechiometric frustration.

Finally, the different pairs of particles had an  $E_{thresh}$  around  $1.25k_bT$ , which is low, and explains why I measure close to no unfavorable interactions with high repulsion intensity. This may be because of the triangular shape of the particles, which restricts loop formation and thus prevents unfavorable links to be stabilised. Indeed, unfavorable links can be stabilised within an aggregate if its energetic cost is balanced by another stabilising link, which happens when particles loop. However, triangular particles need to be 6 in order to form a loop, where hexagonal particles only need to be 3 to do so, as illustrated on **Figure 16**. Therefore, loops are slower to form with triangular particles than with square or hexagonal particles, which suggests that the incorporation of unfavourable interactions is almost impossible for triangle particles for kinetic reasons. This hypothesis could be tested by simulating square or hexagonal particles and measuring the proportion of unfavorable links. I also observe that the value for  $E_{thresh}$  does seem to not depend on the face composition of the system.

Decreasing repulsion energy increases the proportion of unfavorable links. Therefore I tested whether this increase had an effect on the particle self-assembly. To do this I measured the compactness of the different systems for  $E_{repul} = 0k_bT$  and  $E_{repul} = 10k_bT$  (Figure 17). I observed that for every pair of particle compactness increased when I lowered repulsion, which is consistent with the previous observation that the proportion of unpaired faces decreased with the repulsion energy, as the unpaired faces were more frequently engaged in unfavorable interactions when the repulsion got weak enough. However this change in compactness is very small, and I did not observe a significant change in the types of aggregates formed with and without repulsion.

Indeed, the compactness seems more closely linked to stoechiometric frustration than it is to effective frustration. For example, compactness increases more in systems with a high stoechiometric frustration when repulsion is nullified than what it increases in less stoechiometrically frustrated systems. Therefore, effective frustration does not significantly affect the self-assembly of particles, which seems much strongly affected by stoechiometric frustration.



Figure 17: Effective frustration does not significantly affect the compactness of aggregates. Each point represents one pair of particles.

# **3.3** Triangle particles accomodate stoechiometric frustration by forming less dense aggregates

In the previous section, I have shown that triangular particles do not self-assemble by incorporating unfavorable links. Therefore in this chapter I show that stoechiometric frustration expresses itself as dense frustration.

Dense frustration measures how much particles lose energy when they go from a dense to a diluted state. In a dense state, each face is engaged in an interaction, therefore the energy of a dense system should reflect its face composition. Indeed, a system in dense conditions can be forced to form unfavorable interactions if there are too much faces which cannot be paired favorably. The dense frustration of a system should thus be linked to its stoechiometric frustration. To test this link I compared the energy of a system in dense and dilute conditions for every pair of particle, which I sorted according to their stoechiometric frustration. The results are shown on Figure 18. I observe that the more a system is stoechiometrically frustrated, the higher the energy per particle is, be it in dense or dilute conditions. This shows that stoechiometric frustration is indeed a form of frustration as it defines how stable a system can get by measuring the proportion of favorable interactions particles can establish. I also observe that most systems had lower energy in dilute conditions than in dense conditions, and the more stoechiometrically frustrated the system is the lower the energy in dilute condition is compared to that in dense conditions. Indeed, the only systems which are more stable in dense conditions are those with no stoechiometric frustration. These systems can favorably pair all of their faces, therefore their dilution induces an energetic cost associated to forming an aggregate with unpaired faces at its surface. The more a system is stoechiometrically frustrated, the more faces there are in it which can only bind unfavorably with other faces. In a dense system, these faces are forced to interact, therefore destabilizing the system. However, when entering a dilute state



Figure 18: **Stoechiometric frustration increases a system's dense frustration.** Each point represents one pair of particles. The energy values were averaged over 10 simulations.



Figure 19: The compactness of a system reflects its face composition. Each point represents one aggregate.

these faces are mostly unbound, as shown in the previous section, which stabilizes the systems and explains why systems with higher stoechiometric frustration are more stable in dilute conditions than dense ones. This suggests that dense frustration is a good way to measure the effect of stoechiometric frustration, and that systems with high stoechiometric frustration tend to form smaller and less dense aggregates, and should therefore be less compact.

To test this, I measured for each pair of particle the compactness and size of the aggregates formed, and compared it to the stoechiometric frustration of each system (**Figure 19**). From this measurement I observe that the more stoechiometrically frustrated a system is, the smaller the aggregates particles in the system assemble into, forming only monomers when this frustration is maximum. Moreover, systems with comparable levels of stoechiometric frustration have comparable levels of compactness. This clearly shows that with this FLK model particle self-assembly is mostly determined by the proportion of each F, L and K face present in the system, and an unbalance in the number of L and K faces prevents particles from self-assembling into large aggregates, as they prefer to be vaporized into oligomers.



Figure 20: Hexagonal particles can have designs which force particles to assemble into fibers.

### 4 Perspectives

In this section I explore how I could improve my simulations to solve the problems I faced. In particular I will focus on how to improve the model and its implementation to observe more diversity in aggregate shapes and better understand frustration. I have shown previously that triangular particles scarcely bind in an unfavorable way. One of the main reasons for which they do not is because of their shape. Indeed, I already explained that since they need 6 particles to form a loop, it was harder for them to balance the cost of an unfavorable link. Moreover, since they only have 3 faces, it is harder for a particle engaged in an unfavorable interaction to mitigate the cost of this interaction. Therefore, simulating similar systems with square or hexagonal particles could help understand how satisfying unfavorable links can produce new aggregate shapes. These shapes also have more complex interaction rules, for example by having non-adjacent faces which can favor a linear assembly in some precise conditions (Figure 20), which triangles do not have.

I also noted a low diversity in particle aggregates, which could come from the model itself, as it introduces only face types which have a complementary face type. I think that this could explain why I observed so many dense aggregates. Indeed, all particles pairs with perfect proportions of F, L and K faces will assemble into dense aggregates because they will be able to satisfy all their links. I have already explained how this can be a problem of the triangular shape, but I suggest adding another face type to the model, which would not have a complementary face type to bind to, which could be made into colloids by printing a shape which fits with no other shape. This face type could be called Neutral, and would severely restrict the orientation in which particles would assemble, especially in pairs of particles for which L and K faces are not found on the same particle types.

## **5** Conclusion

All in all, I have shown that the FLK model suggested by our collaborators does present the advantage to study in a systematic way the way directional interactions can affect particle self-assembly. However, using this model on triangular particles does not produce diverse aggregates. This is mostly due to the triangular shape which is kinetically too slow to incorporate unfavorable links. Therefore, self-assembly of triangular particles is mostly dictated by whether particles can satisfy all their links or not. Thus the model could be improved either by adding new faces or by being tested on square or hexagonal particles.

### References

- [1] P. Sundd, M. T. Gladwin, and E. M. Novelli, "Pathophysiology of sickle cell disease," *Annu Rev Pathol.*, pp. 263–92, 14 2019. DOI: 10.1146.
- [2] G. J. Katol *et al.*, "Sickle cell disease," *Nat Rev Dis Primers*, vol. 4, no. 18010, 2018.
  DOI: 10.1038/nrdp.2018.10.
- [3] C. T. Noguchi and A. N. Schechter, "Sickle hemoglobin polymerization in solution and in cells," *Annu Rev Biophys Biophys Chem*, vol. 14, pp. 239–63, 1985. DOI: 10.1146/annurev.bb.14.060185.001323.
- [4] K. P. Tuomas, M. Vendruscolo, and C. M. Dobson, "The amyloid state and its association with protein misfolding diseases," *Nature reviews Molecular cell biology*, vol. 15, pp. 384–96, DOI: 10.1038/nrm3810.
- [5] H. Garcia-Seisdedos, C. Empereur-Mot, N. Elad, and E. D. Levy, "Proteins evolve on the edge of supramolecular self-assembly.," *Nature*, vol. 548, pp. 244–247, 2017. DOI: 10.1038/nature23320.
- [6] S. e. a. Ray, "α-synuclein aggregation nucleates through liquid–liquid phase separation," *Nature Chemistry*, vol. 12, pp. 705–16, 2020. DOI: 10.1038/s41557-020-0465-9.
- H. Fanga, M. F. Hagana, and W. B. Rogersa, "Two-step crystallization and solid-solid transitions in binary colloidal mixtures," *PNAS*, vol. 117, no. 45, pp. 27927–33, 2020. DOI: 10.1073/pnas.2008561117.
- [8] Y. Mao, M. Cates, and H. D. Lekkerkerker, "Depletion force in colloidal systems," *Physics 1: Statistical Mechanics and its Applications*, vol. 222, pp. 10–24, 1995.
- [9] M. Li, A. Kloczkowski, M. Cieplak, and M. Kouza, "Computational models for the study of protein aggregation," *Methods in Molecular Biology*, vol. 2340, 51<sub>78</sub>, 2022. DOI: 10. 1007/978-1-0716-1546-1\_4.
- [10] M. Lenz and T. A. Witten, "Geometrical frustration yields fibre formation in self-assembly," *Nature physics*, vol. 12, pp. 1100–04, 2017. DOI: 10.1038/NPHYS4184.
- [11] L. Koehler, P. Ronceray, and M. Lenz, "Principles of self-assembly for particles with simple geometries and complex interactions," *In preparation*,