

# LA FORME DES VIRUS

Gilles HALBOUT

**Résumé.** Partant d'un exemple de TPE sur la forme des capsides, j'explique comment utiliser des mathématiques accessibles au niveau de la terminale pour décrire la géométrie des virus.

Pour cet exposé, je suis parti d'un TPE très intéressant réalisé pendant l'année 2002-2003 au LEGT Notre-Dame à Fontenay-le-Comte, Vendée (85) par Antoine Rousseau, Gêrôme Dieu, David Dupont sur « la géométrie des virus ». Je n'ai pas pour objectif de critiquer ce travail : bien au contraire, j'y ai beaucoup appris sur les virus. Je voudrais plutôt montrer comment ce thème, comme tant d'autres, peut donner lieu à des études ou simulations mathématiques très simples et j'ai essayé d'y rattacher de nombreux points du programme de mathématique de terminale.

Le plan du TPE dont je suis parti était le suivant :

- Présentation du TPE
- Introduction
- Étude biologique des virus (généralités)
  - Généralités sur les virus
  - La capsid
  - Mode d'action du virus
- Étude mathématique des polyèdres réguliers, et en particulier de l'icosaèdre
  - Les polyèdres
  - Les polyèdres de Platon
  - Démonstration de l'existence de cinq polyèdres réguliers
- Conclusion.

Et la problématique développée était :

*Les capsides des virus ont une forme géométrique. Quelle est-elle et peut-on la rapprocher des formes géométriques régulières en mathématiques ?*

Je reprendrai ce plan sans sa partie mathématique. Je compléterai mon étude en m'inspirant d'un très bon article [1] sur le virus de la grippe publié en mars 1999 dans *Pour la Science*. On pourra aussi consulter d'autres articles sur les virus ([2], [3]).

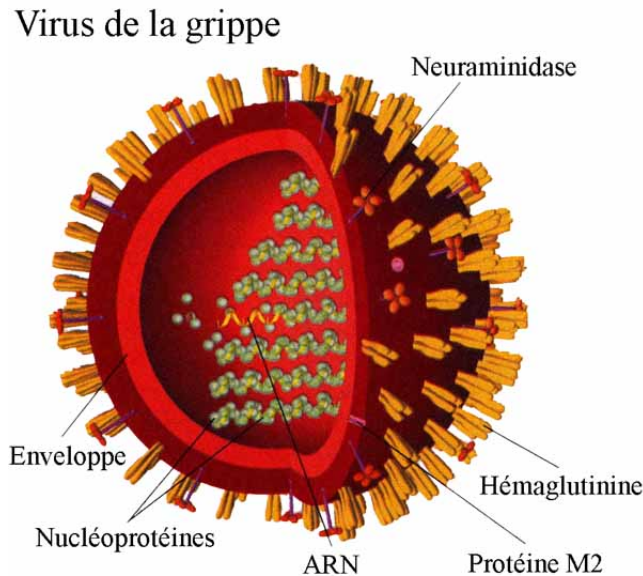
En préparant cet exposé, j'ai pris le même plaisir qu'ont eu, espérons-le, les élèves qui découvraient comme moi un sujet étranger à leurs préoccupations quotidiennes. Comme on n'apprend jamais tout seul je voudrais remercier Françoise Bringel et Mario Keller (de l'Institut de Biologie Moléculaire des Plantes à l'ULP) qui m'ont conseillé sur les virus et mon ami Nicolas Destainville du laboratoire de Physique Théorique de l'Université Paul Sabatier, grand connaisseur des quasi-cristaux. Je lui dois notamment la très bonne référence [4], d'un niveau assez difficile cependant.

## 1. La forme générale des virus

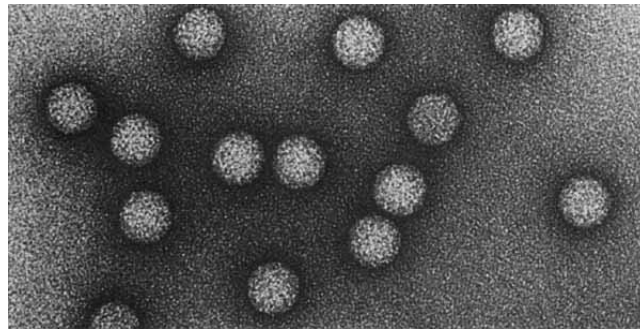
**1.1. Une description.** Un virus comporte toujours un génome (ADN ou ARN) emballé dans une structure protéique appelée capsid. La capsid protège le génome. Elle a une structure géométrique (soit tubulaire, soit polyédrale).

Par exemple, le virus de la grippe a une taille de  $100\text{ nm}$  et possède une enveloppe externe composée de molécules de lipides et tapissée de protéines (neuraminidase et hémagglutinine) : ce sont des chaînes repliées d'acides aminés.

Ces molécules ont presque toutes la même structure tridimensionnelle, toutefois les séquences des acides aminés varient notablement selon les protéines.



**1.2. Des contraintes physiques à une description géométrique.** Il est très difficile de voir les formes des virus même avec un grossissement de  $\times 300.000$  : on a l'impression qu'il s'agit de petites sphères, comme on le voit ici pour le virus de la poliomyélite.



La structure est en fait très souvent celle d'un polyèdre régulier. On peut expliquer cette configuration :

- les protéines ont tendance à se mettre dans une disposition « favorable énergétiquement » ;
- en première approximation, on peut supposer que toutes les protéines (ou plus précisément, les sous-unités protéiques) sont identiques et à symétrie sphérique.

Une étude mathématique (recherche du minimum d'une fonction) peut alors mettre en évidence la structure. L'énergie, que l'on cherche à minimiser est une somme, pour tous les couples de protéines  $(P_i, P_j)$  de termes  $c_i c_j / d_{ij}^2$  où  $c_i$  est une constante qui ne dépend que de la protéine  $P_i$  et  $d_{ij}$  est la distance entre  $P_i$  et  $P_j$ . Dans le cas où toutes les protéines

sont identiques (c'est-à-dire que les constantes  $c_i$  sont égales) on obtient une configuration comme celle décrite précédemment. Mais c'est très difficile à démontrer (voir [4]).

Avec trois protéines identiques, l'étude est plus simple. On commence par confiner nos protéines sur un cercle : on doit choisir une « échelle » (en faisant des homothéties de rapport  $\lambda$ , on ne change pas la nature du problème mais on divise l'énergie par  $\lambda^2$ ).

Fixons deux protéines et supposons la troisième entre les deux premières. La fonction que l'on doit minimiser est la fonction

$$f : \mathbb{R}/2\pi\mathbb{Z} \longrightarrow \mathbb{R}$$

$$\theta \longmapsto \frac{1}{|e^{i\theta} - 1|^2} + \frac{1}{|e^{i\theta} - e^{i\theta_0}|^2},$$

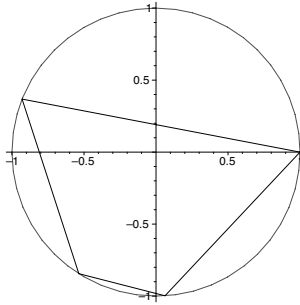
où  $\theta_0$  est une constante fixée. On vérifie que cette fonction est minimale en un seul point  $\theta = \theta_0/2$ , ce qui signifie que la protéine est à égale distance des deux autres. En faisant un raisonnement par l'absurde, on prouve ensuite que trois protéines sur le cercle se trouvent en position d'énergie minimale si elles forment un triangle équilatéral.

De manière générale, si on a plus de trois protéines sur un cercle, on fait le même raisonnement en supposant, en première approximation (pour être rigoureux, il faut des outils mathématiques plus sophistiqués), que la contribution énergétique d'une protéine provient uniquement des interactions avec ses deux voisines. D'après ce qui précède, chaque protéine est à égale distance de ses voisines. Toutes ensemble, elles forment donc un polygone régulier. Enfin, en raisonnant dans l'espace, on montre que celles-ci doivent former un polyèdre régulier, c'est-à-dire un polyèdre dont les faces sont des polygones réguliers isométriques qui s'assemblent en chaque sommet de la même façon.

On peut ensuite faire une modélisation (ici sur MAPLE) et retrouver ce résultat de distribution régulière de manière expérimentale. Attention, ceci n'est pas une preuve ! Mais si l'on suppose les protéines différentes (c'est-à-dire que les  $c_i$  sont différentes), on obtient une indication sur leur disposition mais surtout, réciproquement, si l'on rencontre dans la nature des protéines disposées de manière non régulière, on peut utiliser des moyens informatiques pour deviner les différentes constantes  $c_i$  et ainsi les différents types de protéines en jeu.

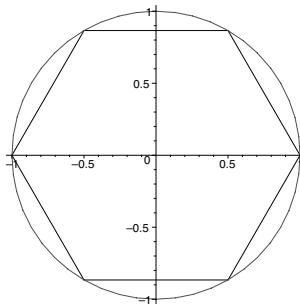
Voici un exemple avec quatre protéines : devinez lesquelles sont les mêmes !

```
> restart:with(plots):poly:=proc(p,a,b,c,d)
> local cercle, pol,j,bi,ci,di,n,t,u,uu,v,vv:n:=4:t[1]:=a:t[2]:=b:
> t[3]:=c:t[4]:=d:uu[2]:=1:uu[3]:=2:uu[4]:=3:vv:=100000000000:
> for bi from 1 to p do u[2]:=evalf(2*Pi*bi/p):
> for ci from bi+1 to p do u[3]:=evalf(2*Pi*ci/p):
> for di from ci+1 to p do u[4]:=evalf(2*Pi*di/p): v:=0:
> for j from 2 to n do v:=v+t[1]*t[j]/(sin((u[j])/2))^2:od:
> for j from 3 to n do v:=v+t[2]*t[j]/(sin((u[2]-u[j])/2))^2:od:
> for j from 4 to n do v:=v+t[3]*t[j]/(sin((u[3]-u[j])/2))^2:od:
> if v < vv then vv:=v:uu[2]:=u[2]:uu[3]:=u[3]:uu[4]:=u[4]:fi:od:od:od:
> pol:=plots[pointplot]([[cos(0),sin(0)],[cos(uu[2]),sin(uu[2])],[cos(uu[3]),sin(uu[3])],[cos(uu[4]),sin(uu[4])],[cos(0),sin(0)]]),style=line):
> cercle:=plot([sin(t),cos(t),t=-Pi..Pi):
> plots[display]({cercle,pol},scaling=constrained)
> end:
> poly(50,20,20,1,1);
```



Voici maintenant ce que l'on obtient avec six protéines identiques : c'est rassurant<sup>(1)</sup> !

```
> restart:with(plots):poly:=proc(p,a,b,c,d,e,f)
> local cercle, pol,j,bi,ci,di,ei,ffi,n,t,u,uu,v,vv:
> n:=6:t[1]:=a:t[2]:=b:t[3]:=c:t[4]:=d:t[5]:=e:t[6]:=f:
> uu[2]:=1:uu[3]:=2:uu[4]:=3:uu[5]:=4:uu[6]:=5:vv:=100000000000:
> for bi from 1 to p do u[2]:=evalf(2*Pi*bi/p):
> for ci from bi+1 to p do u[3]:=evalf(2*Pi*ci/p):
> for di from ci+1 to p do u[4]:=evalf(2*Pi*di/p):
> for ei from di+1 to p do u[5]:=evalf(2*Pi*ei/p):
> for ffi from ei+1 to p do u[6]:=evalf(2*Pi*ffi/p):
> v:=0:for j from 2 to n do v:=v+t[1]*t[j]/(sin((u[j])/2))^2:od:
> for j from 3 to n do v:=v+t[2]*t[j]/(sin((u[2]-u[j])/2))^2:od:
> for j from 4 to n do v:=v+t[3]*t[j]/(sin((u[3]-u[j])/2))^2:od:
> for j from 5 to n do v:=v+t[4]*t[j]/(sin((u[4]-u[j])/2))^2:od:
> for j from 6 to n do v:=v+t[5]*t[j]/(sin((u[5]-u[j])/2))^2:od:
> if v<vv then vv:=v: for j from 1 to n do
> uu[j]:=u[j]:od:fi:od:od:od:od:od:od:
> pol:=plots[pointplot]([[cos(0),sin(0)],[cos(uu[2]),sin(uu[2])],[cos(uu[
> 3]),sin(uu[3])],[cos(uu[4]),sin(uu[4])],[cos(uu[5]),sin(uu[5])],[cos(
> uu[6]),sin(uu[6])],[cos(0),sin(0)]],style=line):
> cercle:=plot([sin(t),cos(t),t=-Pi..Pi):
> plots[display]({cercle,pol},scaling=constrained)
> end:
> poly(18,1,1,1,1,1,1);
```



**1.3. Une classification.** L'étude mathématique a mis en évidence le fait que les contraintes physiques se traduisent par des contraintes géométriques, ici l'invariance par certaines rotations d'ordre fini. Cette condition était déjà énoncée par Crick et Watson<sup>(2)</sup> en 1956 qui parlaient de « symétrie cubique » : chaque sous-unité protéique doit être

<sup>(1)</sup>Ce n'est pas le moyen le plus simple pour dessiner un hexagone...

<sup>(2)</sup>Les découvreurs de la structure en double hélice des molécules d'ADN.

entourée par le même environnement, et celui-ci doit être invariant par une rotation d'ordre fini.

Contrairement aux polygones réguliers, il n'existe pas une infinité de polyèdres convexes réguliers. On peut démontrer que :

- il en existe au plus cinq ;
- il en existe effectivement cinq qui sont les cinq polyèdres de Platon : le tétraèdre régulier, le cube, l'octaèdre régulier, le dodécaèdre régulier et l'icosaèdre régulier
- Pour montrer le premier point, on note  $p$  le nombre d'arêtes de chaque polygone et  $q$  le nombre d'arêtes partant de chaque point et on démontre que

$$(1) \quad \frac{1}{p} + \frac{1}{q} > \frac{1}{2}.$$

Puisque  $p$  et  $q$  sont au moins égaux à 3, le couple  $(p, q)$  est nécessairement égal à l'un des couples suivants

$$(3, 3), (4, 3), (3, 4), (5, 3), (3, 5)$$

qui correspondent respectivement aux cinq polyèdres réguliers du deuxième point. Pour montrer l'inégalité stricte (1), on applique souvent la formule d'Euler  $F - A + S = 2$ , où  $F$ ,  $A$  et  $S$  comptent respectivement le nombre de faces, d'arêtes et de sommets d'un polyèdre. Cette formule se démontre en projetant le polyèdre sur une sphère et en mesurant de deux manières la surface de la sphère. Ceci repose sur la formule de Girard, calculant l'aire d'un triangle sphérique et est peut-être un peu difficile. On poursuit en remarquant que  $pF = 2A$  et que  $qS = 2A$  et donc que

$$\frac{1}{p} + \frac{1}{q} - \frac{1}{2} = \frac{1}{A}$$

d'où l'inégalité.

Mais cette inégalité peut aussi se démontrer directement, en remarquant que la projection « dilate » les angles : soit  $ABCD$  un tétraèdre tel que  $D$  soit la projection orthogonale de  $A$  sur  $(BCD)$ , en comparant les produits scalaires  $\overrightarrow{AB} \cdot \overrightarrow{AC}$  et  $\overrightarrow{DB} \cdot \overrightarrow{DC}$ , on démontre que

$$\cos(\overrightarrow{AB}, \overrightarrow{AC}) \geq \cos(\overrightarrow{DB}, \overrightarrow{DC}).$$

Ensuite, puisque l'angle entre deux arêtes d'un  $p$ -gone régulier est  $\frac{p-2}{p}\pi$ , en projetant les  $q$  angles issus d'un point, on obtient

$$q \frac{p-2}{p} \pi < 2\pi$$

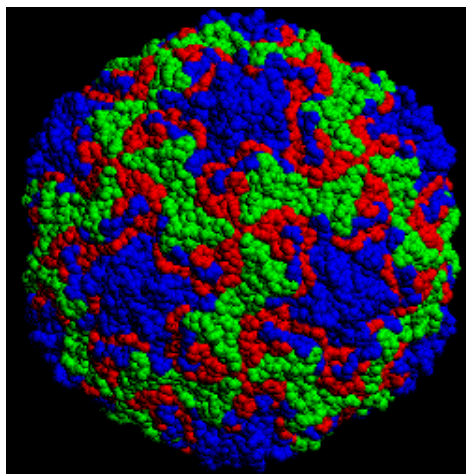
ce qui donne l'inégalité.

- On oublie souvent de vérifier le second point : il est aisé de construire un tétraèdre régulier, un cube ou un octaèdre régulier. On peut construire un icosaèdre régulier à partir d'un dodécaèdre régulier en prenant, pour sommets, les centres des faces du dodécaèdre. Enfin, on construit facilement un dodécaèdre régulier « à la main » : on construit un pentagone régulier auquel on attache cinq pentagones réguliers isométriques, un sur chaque arête, qu'on soulève d'un angle ayant pour sinus  $2 \sin(\frac{\pi}{10})$ . Ceci fournit quinze points dont on peut calculer les coordonnées dans l'espace. On répète l'opération et on vérifie que les deux figures obtenues s'assemblent correctement.

**1.4. Raffinement du modèle.** Nous allons discuter des hypothèses (symétrie sphérique et uniformité des protéines) faites jusque là :

*Les capsides peuvent comporter différentes protéines.* Le modèle du virus comme un polyèdre régulier comprenant une même protéine à chaque sommet (ou sur chaque face) est trop rudimentaire. La capside de la plupart des virus (comme le rhinovirus 14 ci-dessous ou celui de la poliomyélite) comporte en fait plusieurs types de protéines. Ces différentes protéines sont disposées de manière à conserver les symétries, soit au centre des faces, soit dans les coins des faces, soit sur les arêtes, soit sur les sommets. Une étude mathématique nous montre que c'est en étudiant la structure du « groupe » des isométries du polyèdre que l'on peut mieux comprendre comment peuvent s'organiser les différents groupes de protéines (sur ce sujet la littérature, même très accessible, ne manque pas sur les groupes de symétries  $A_4$ ,  $S_4$  et  $A_5$  des polyèdres réguliers).

Observons l'image ci-dessous. Elle montre la capside du rhinovirus 14 déterminée par cristallographie aux rayons x. Chaque acide aminé est représenté par une sphère de 4 angstroms de diamètre. La capside est constituée de quatre types de protéines : protéines VP1 (noir), VP (gris clair), VP3 (gris) et VP4 (pas visible). Sur cette image, on aperçoit bien les axes de rotation d'ordre 5 avec VP1 aux centres des pentagones (il doit y en avoir 12!).

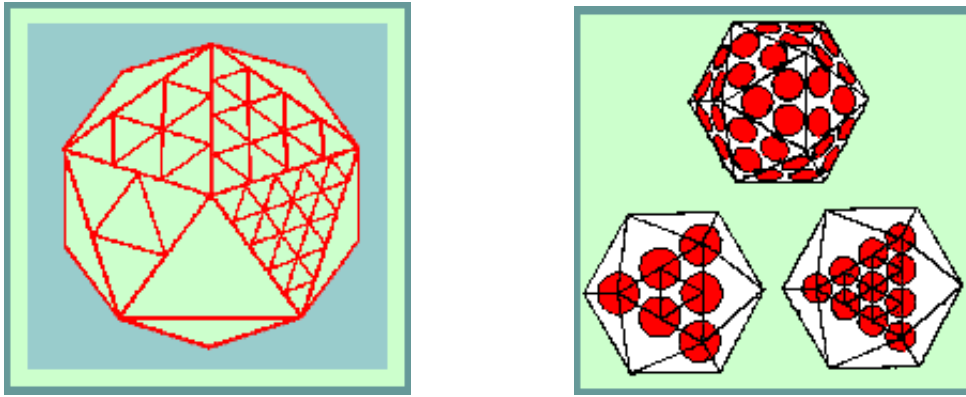


*La répartition des protéines ne respecte pas les faces.* De plus, la répartition des protéines ne respecte pas toujours les faces du polyèdre régulier : il faut alors faire un « pavage » plus fin du polyèdre. Prenons l'exemple de l'icosaèdre régulier. Les protéines (ou pasmers) ont presque toujours une symétrie (on les appelle alors dimères, trimères... selon l'ordre des rotations les préservant).

Caspar et Klug ont montré que, pour décrire précisément l'agencement des protéines, il fallait paver le dodécaèdre avec  $20T$  triangles où  $T = a^2 + ba + b^2$ . Si  $T = 1$ , on retrouve les vingt triangles formant les faces de l'icosaèdre régulier.

Pour obtenir une triangulation générale, on réalise un pavage du plan en triangles équilatéraux. On se déplace le long de  $a$  triangles dans le même sens, puis on tourne de  $2\pi/3$  et on se déplace le long de  $b$  triangles. Le chemin global détermine l'arête du triangle de base, face de notre icosaèdre régulier. On obtient ainsi une triangulation de l'icosaèdre régulier qui ne respecte pas nécessairement les faces : elle la respecte si et seulement si  $b = 0$  comme sur le dessin ci-après. On peut obtenir un nouveau polyèdre en « aplatissant » les triangles qui ne se trouvaient pas sur une même face. Ce polyèdre, même s'il est constitué de triangles équilatéraux, n'est en général pas régulier car d'un point peuvent partir cinq ou six arêtes. Cette nouvelle figure géométrique rend mieux compte de la structure de la capside et de la disposition des protéines.

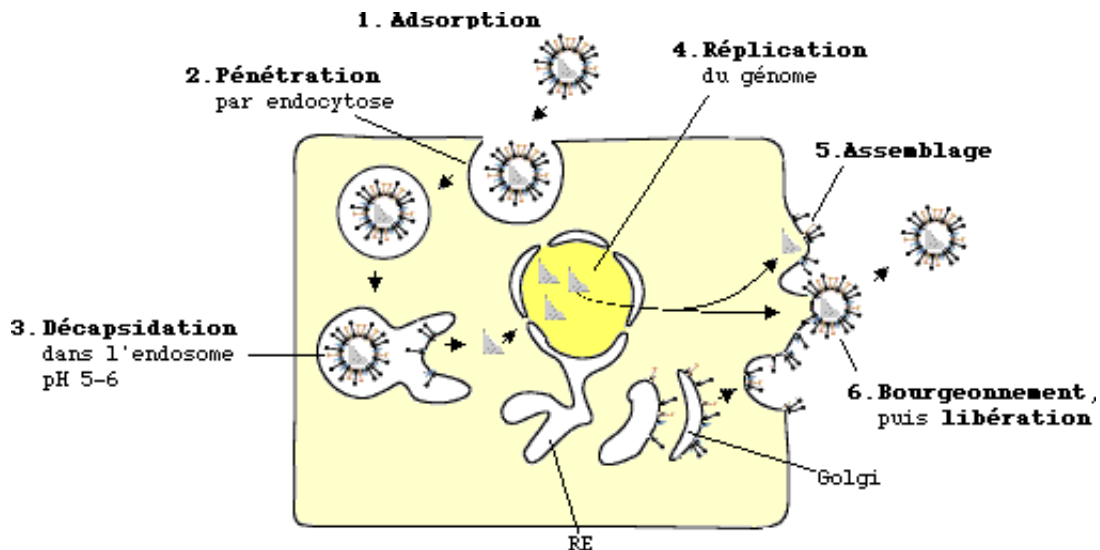
Voici quelques exemples de triangulations respectant les faces, c'est-à-dire pour  $b = 0$ .



Dans le dessin de droite, on a représenté plusieurs dispositions de protéines : dans les coins des faces (en haut), sur les sommets (en bas). On peut ensuite faire une étude combinatoire et montrer qu'il y a  $60T$  monomères,  $30T$  dimères,  $20T$  trimères, 12 pentamères et  $10(T - 1)$  hexamères.

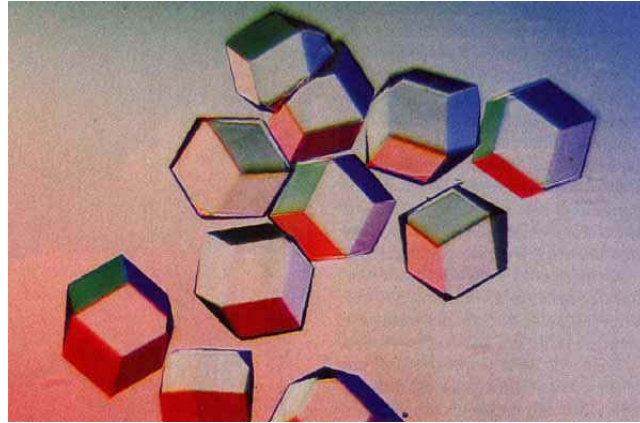
## 2. L'action des virus

**2.1. Description.** Tous les virus obéissent à une stratégie de multiplication fondamentale que l'on peut découper en cinq étapes illustrées par le cycle du virus de la grippe :



- L'adsorption : l'adsorption du virus consiste en la liaison d'une protéine virale à un récepteur de la surface cellulaire. Certains virus peuvent utiliser plusieurs molécules différentes pour s'attacher aux récepteurs. Les récepteurs peuvent être une protéine ou le sucre attaché à une glycoprotéine ou glycolipide. Certains virus portent à leur surface une neuraminidase qui leur permet de se détacher de leur récepteur en clivant l'acide neuraminique terminal porté par la chaîne polysaccharidique du récepteur.
- La pénétration : à la différence de l'adsorption, la pénétration est un processus qui requiert de l'énergie.
- La décapsidation : mise à nu du génome viral.
- La réplication du génome viral : l'élément clé dans la réplication virale est la synthèse des protéines virales par la machinerie cellulaire.
- L'assemblage et la libération.

**2.2. Les traitements.** Le virus de la grippe possède à sa surface une enzyme, la neuraminidase, qui sert à dissoudre la colle que constitue l'acide sialique adsorbé à la surface du virus. Grâce à cette enzyme, le virus ne reste pas collé à la cellule infectée, mais peut se libérer pour aller infecter d'autres cellules saines... C'est cette enzyme qui fait de la grippe une maladie infectieuse extrêmement contagieuse.



Nous allons étudier la structure de cette protéine. Elle est elle-même formée de particules qui, à leur tour, sont disposées de manière régulière et donc devraient former un polyèdre régulier. En fait, cette enzyme neuraminidase cristallise sous la forme de dodécaèdres rhombiques (c'est un polyèdre dont les douze faces sont des losanges isométriques).

Alors que l'on s'attendait à trouver encore un polyèdre, dans ce cas un cube, on trouve un dodécaèdre rhombique. C'est donc que la structure n'est pas régulière : en faisant pour le cube le même raisonnement que celui que l'on avait fait pour l'icosaèdre, on démontre que notre enzyme a une structure  $(1, 1)$  (dans le cas du cube, on fait un pavage en carrés et une structure  $(a, b)$  donne  $T = 6(a^2 + b^2)$  carrés, on a découpé chacune des six faces en  $a^2 + b^2$  carrés!), ce qui permet de la caractériser.

C'est grâce à l'analyse structurale de ces cristaux, que les chercheurs ont pu déterminer la structure tridimensionnelle de l'enzyme neuraminidase et élaborer de nouveaux médicaments qui, en bloquant le site actif de l'enzyme, l'empêchent de fonctionner normalement, inhibant ainsi la multiplication du virus.

**2.3. Autres exemples.** Je ne résiste pas au plaisir de donner aussi une liste de sujets liés à des problèmes géométriques similaires : c'est le cas, évidemment pour les quasi-cristaux (comme sur la figure suivante, à gauche).

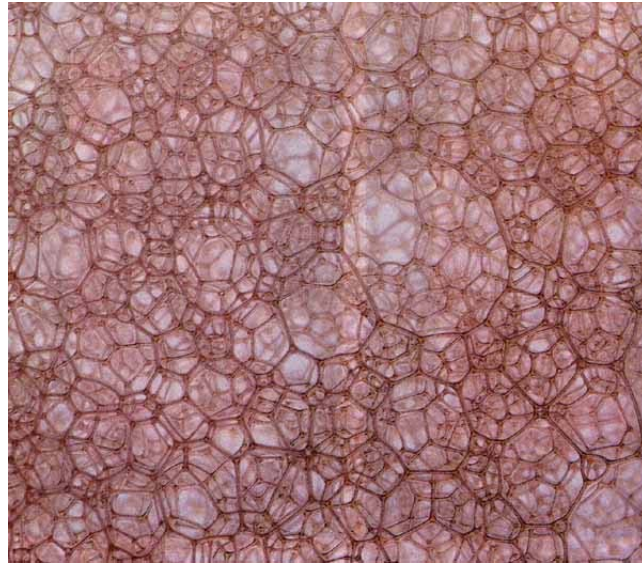
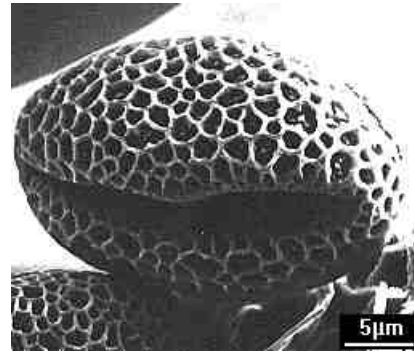
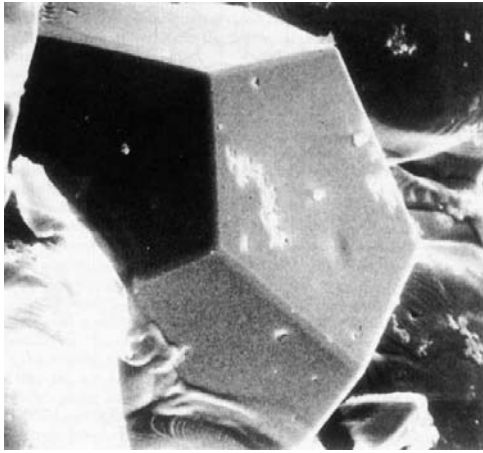
Mais on les retrouve aussi en

- chimie (chimie tétraédrique, octaédrique, cubique, polyédrales et prismatiques) ;
- botanique : les particules de pollen ont souvent des formes géométriques qui rappellent celles des polyèdres (figure ci-dessous, à droite) ;
- zoologie (les yeux d'insectes, les nids d'abeilles) ;
- physique : la mousse de savon devient polyédrale lorsque qu'elle ne contient plus de liquide. Les différents films ne se rejoignent que par trois, selon des angles de  $120^\circ$ .

### Bibliographie

- [1] Le virus de la grippe : *Pour la Science*, mars 1999. .
- [2] Le virus de la poliomyélite : *Pour la Science*, mai 1987.





[3] Le virus du sida : *Pour la Science*, février 1987.

[4] R. VANSELOW ET R. HOWE, *Chemistry and Physics of Solid Surfaces VII*, Springer 1988, pp. 367-405.

Gilles HALBOUT  
Institut de Recherche Mathématique Avancée  
Université Louis Pasteur  
Strasbourg  
halbout@math.u-strasbg.fr