Assemblage et désassemblage des virus : mode d'emploi

Guillaume Tresset⁽¹⁾ (guillaume.tresset@u-psud.fr), Martin Castelnovo⁽²⁾ et Amélie Leforestier⁽¹⁾

(1) Laboratoire de Physique des Solides, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91405 Orsay

(2) Laboratoire de Physique, UMR5672 ENS Lyon – Univ. Claude Bernard – CNRS, École normale supérieure de Lyon, 69364 Lyon Cedex 07

Les virus biologiques sont des entités fascinantes qui opposent une relative simplicité dans leur structure à une surprenante sophistication dans leur fonction.

Bien qu'inertes, leurs centaines voire milliers de composants moléculaires s'assemblent et se désassemblent spontanément dans un milieu cellulaire hétérogène et encombré, avec une précision quasi-atomique et un taux d'erreur marginal.

Les concepts théoriques et expérimentaux de la physique statistique et des fluides complexes permettent aujourd'hui de décrire la thermodynamique et les phénomènes hors équilibre qui régissent l'assemblage et le désassemblage de virus naturels ou synthétiques.

Les virus sont des agents infectieux qui utilisent les cellules d'un organisme hôte pour se multiplier. Leur unique fonction est de transférer leurs gènes au cœur de cellules cibles qui, de ce fait, synthétisent les composants viraux à leur insu. Ils sont parfois capables de résister à des environnements très divers et souvent hostiles, et peuvent être transportés dans l'air ou bien dans l'eau, par des insectes ou dans les aliments. Après contact et internalisation dans une cellule cible, le virus se désassemble pour transférer son *génome** qui est ensuite traduit en nouveaux composants viraux. Ceux-ci s'assemblent spontanément en nouveaux virus qui sortent de la cellule pour en infecter d'autres, et le cycle viral recommence. Les mécanismes d'assemblage et de désassemblage sont de ce fait étroitement liés à la prolifération des virus.

Les virus sont constitués (fig. 1) : (i) d'une chaîne d'acide nucléique (ARN ou ADN), qui contient le code génétique ou génome ; (ii) d'une *capside**, qui est une coque protéique semi-rigide protégeant le génome de l'environnement extérieur ; et, parfois, (iii) d'une enveloppe souple faite de lipides enrobant la capside et contenant des molécules qui se lient avec les cellules cibles à infecter. Leur taille est généralement nanométrique mais peut atteindre le micromètre. Ils peuvent présenter des formes variées, souvent en icosaèdre ou en cylindre.

Le premier virus connu, le virus de la mosaïque du tabac (TMV en anglais pour *tobacco mosaic virus*), fut étudié en 1898 par le botaniste hollandais Martinus Beijerinck. La virologie a, par la suite, bénéficié des formidables avancées de la physique ondulatoire. Le TMV fut ainsi observé pour la première fois par microscopie électronique en 1939, et le premier diffractogramme aux rayons X de cristaux de TMV fut

obtenu en 1941. En 1955, Heinz Fraenkel-Conrat et Robley Williams parviennent à reconstituer un TMV infectieux par autoassemblage à partir du génome et des protéines de la capside. L'exploit sera réitéré dans les années 60 et 70 avec d'autres virus de plantes et en y incorporant des matériaux synthétiques. La voie était ainsi ouverte à l'assemblage contrôlé de virus et de particules virales synthétiques. Par ailleurs, la symétrie icosaédrique de nombreux virus a été anticipée dès 1956 par James Watson et Francis Crick, puis de manière plus élaborée par Donald Caspar et Aaron Klug en 1962 grâce aux progrès de la cristallographie.

Nous avons sélectionné dans cet article quelques avancées en virologie s'appuyant sur des approches physiques, afin d'illustrer les liens étroits qui existent entre ces deux sciences.

La prévalence de la symétrie icosaédrique

Dans leurs travaux de référence, Donald Caspar et Aaron Klug ont proposé que les capsides de forme globalement sphérique, qui représentent environ la moitié des virus sur Terre, adoptent en réalité des symétries icosaédriques [1]. Les protéines constituant la capside sont disposées sur les faces, les arêtes ou les sommets d'un icosaèdre (voir la figure E2 de l'encadré, p. 24). Ces travaux ont conduit à introduire la notion de quasi-équivalence : pour des capsides formées de protéines chimiquement identiques, l'environnement local d'une protéine est légèrement différent suivant sa localisation au sein de la capside. Ceci n'est possible que si cette protéine est capable d'avoir plusieurs conformations structurales proches (voir encadré).

Cet article est dédié à la mémoire de Vladimir Lorman.

Les termes en italiques suivis d'un astérisque sont définis dans le glossaire.

© Wiebke Drendel

Récemment, les concepts de quasi-équivalence ont été généralisés par la théorie de la cristallisation de Landau [2] pour décrire plusieurs virus incompatibles avec la théorie originale de Caspar et Klug (virus de la dengue, polyomavirus murin, etc.)

L'omniprésence de la symétrie icosaédrique peut aussi être comprise par des arguments géométriques simples (voir encadré) : les capsides virales présentent des motifs hexagonaux et pentagonaux clairement visibles sur les structures déterminées par cryomicroscopie électronique. Une structure purement faite d'hexagones aurait une densité surfacique maximale, mais elle ne pourrait pas être fermée car elle serait plate. Un résultat topologique attribué à Euler montre en effet qu'il est nécessaire d'avoir exactement douze pentagones pour introduire la courbure nécessaire à une

structure fermée, le reste de la structure étant composé d'hexagones. Si ces douze pentagones sont répartis de manière équidistante, la structure présente une symétrie icosaédrique. Lorsque les douze pentagones imposés par la topologie sont répartis de manière non uniforme, de nouvelles formes géométriques sont observées, et la symétrie icosaédrique est perdue. C'est le cas par exemple pour le virus de l'immunodéficience humaine (VIH-1) dont la capside est conique, ou encore pour certains *bactériophages** comme le T4 qui présente une forme allongée.

Les mécanismes à l'œuvre

Nous allons maintenant voir les principes physiques qui gouvernent l'assemblage spontané et réversible des centaines voire milliers de composants viraux en une structure complexe.

Thermodynamique de virus simples

La plupart des capsides virales sont formées grâce à un processus d'auto-assemblage qui peut être reproduit *in vitro* en mélangeant dans un milieu aqueux le génome avec les

>>>

a

b

1. Simulations par dynamique moléculaire de l'assemblage d'une capside virale dodécaédrique autour d'un polymère chargé (d'après J.D. Perlmutter *et al.*, *Annu. Rev. Phys. Chem.* **66** (2015) 217). Les sous-unités protéiques de la capside sont ici modélisées par des pentagones formés de billes bleues et le polymère chargé par une chaîne de billes rouges. Cette géométrie rappelle la structure d'une capside T = 1 (voir encadré). (a) Modèle coopératif résultant d'une forte interaction entre sous-unités et polymère, et d'une faible interaction entre sous-unités. (b) Modèle de nucléation-croissance résultant d'une faible interaction entre sous-unités et polymère, et d'une forte interaction entre sous-unités.

► La symétrie icosaédrique des virus

L'icosaèdre régulier est un solide platonicien comportant 20 faces triangulaires et 12 sommets (fig. E1a). En se basant sur l'observation que de nombreux virus de symétrie icosaédrique contiennent un nombre de protéines multiple de 60, Donald Caspar et Aaron Klug ont développé en 1962 un système de classification de ces virus fondé sur un nombre T appelé « nombre de triangulation », qui est égal au nombre de configurations distinctes des protéines : une capsid icosaédrique contient alors $60T$ protéines formant 12 motifs pentagonaux autour des sommets de l'icosaèdre et $10(T-1)$ motifs hexagonaux.

La figure E2 représente une capsid $T = 3$ à basse résolution. Les protéines sont ici de nature chimique identique, mais adoptent trois conformations différentes identifiées par trois couleurs. On y voit 12 motifs pentagonaux centrés sur les sommets de l'icosaèdre (protéines vertes) et 20 motifs hexagonaux (alternance de protéines orange et violettes). Notons qu'une capsid $T = 1$ ne contient que 12 motifs pentagonaux, tout comme le dodécaèdre (fig. E1b) qui est le dual de l'icosaèdre.

E1. (a) Icosaèdre régulier. (b) Dodécaèdre régulier.

E2. Capsid $T = 3$ à basse résolution.

»»»

sous-unités constitutives de la capsid. Celles-ci sont bien souvent des dimères de protéines, car plus stables que les protéines seules. Ce processus est favorable donc spontané, car le gain énergétique d'association des sous-unités est plus important que la perte d'entropie qui lui est aussi associée. Il existe toutefois une concentration seuil en sous-unités pour que l'autoassemblage soit initié. Le gain énergétique est associé à des interactions attractives à courte portée entre les sous-unités (forces de van der Waals et/ou hydrophobes). Après un régime transitoire, l'association entre les sous-unités finit par atteindre un état d'équilibre chimique régi par une loi de type action de masse. Il faut noter que la présence du génome permet de guider l'assemblage, mais que des capsides vides, donc sans génome, peuvent aussi être formées dans certaines conditions ioniques. Les interactions entre les sous-unités et le génome sont essentiellement d'origine électrostatique car les sous-unités comportent souvent des groupements chimiques chargés positivement, tandis que le génome porte invariablement des charges négatives. Lorsque le génome est présent, on note

une relation linéaire entre la charge électrostatique totale du génome et celle des sous-unités protéiques.

Suivre et modéliser la dynamique d'assemblage

L'autoassemblage d'une capsid vide est souvent interprété en termes de *nucléation**-croissance (fig. 1b). Les sous-unités commencent par former un germe sous l'action d'une fluctuation locale de concentration ; c'est la phase de nucléation. D'autres sous-unités restées libres viennent ensuite se lier séquentiellement à ce germe pour le faire croître rapidement jusqu'à la formation d'une capsid complète. Le germe varie selon le système viral considéré (ex : trimère de sous-unités pour la capsid du virus de l'hépatite B) et reste difficile à identifier du fait de sa nature transitoire. Plus l'énergie libre d'association est faible, plus le nombre de sous-unités constituant le germe augmentera. Par ailleurs, l'assemblage et le désassemblage présentent une importante hystérèse qui permet aux capsides assemblées de rester métastables même à forte dilution, ce qui est une propriété cruciale pour leur survie en dehors du milieu intracellulaire.

Le suivi expérimental des processus d'assemblage et de désassemblage des capsides est une tâche ardue, car elle nécessite de sonder à la fois des échelles spatiales nanométriques et des échelles temporelles s'étalant de la microseconde à l'heure. La diffusion statique de la lumière a permis de nombreuses études cinétiques et en particulier d'identifier la nature du germe de nucléation pour plusieurs systèmes viraux. Toutefois, cette technique ne donne pas d'information structurale, elle ne permet d'estimer que la masse moléculaire moyenne des espèces et elle nécessite des hypothèses parfois fortes. En revanche, la diffusion des rayons X résolue en temps fournit la structure des espèces transitoires les plus stables sans présupposer leur nature [3]. Seules des sources synchrotron délivrent un flux de photons suffisant pour acquérir les premières millisecondes de la cinétique. En outre, des techniques de molécule unique commencent à émerger au sein du domaine, notamment le comptage au travers de nanopores en nitrure de silicium : les espèces en cours d'assemblage sont continuellement forcées par pression à transiter dans un nanopore parcouru par un courant (fluide) imposé de l'extérieur. La chute de courant plus ou moins importante selon la taille des espèces obstruant le nanopore permet alors d'établir une statistique sur leur distribution de taille dans une fenêtre de temps donnée. Enfin, il faut noter que les techniques de calculs numériques ont beaucoup contribué à la compréhension des mécanismes à l'échelle moléculaire (fig. 1). Des simulations Monte Carlo et de dynamique moléculaire ont permis de construire des scénarios plausibles d'assemblage en fonction des énergies d'interaction entre les sous-unités et des conditions ioniques du milieu.

Nos connaissances sur l'empaquetage du génome dans une capsid sont encore embryonnaires. Les simulations par dynamique moléculaire font état de deux scénarios qui dépendent des énergies d'interaction entre les composants : dans le premier cas, les sous-unités se lient rapidement au génome puis se réorganisent de manière coopérative en capsid (fig. 1a) ; dans le second scénario, la capsid s'assemble selon le mécanisme de nucléation-croissance vu précédemment tout en empaquetant simultanément le génome (fig. 1b). Très peu de mesures expérimentales ont été effectuées à ce jour.

Des virus sous pression : le cas particulier des bactériophages

Les bactériophages (ou phages) désignent spécifiquement les virus infectant les bactéries. Parmi eux, les phages dotés d'une queue protéique sont les plus nombreux. Ils se caractérisent par des voies d'assemblage/désassemblage originales, où empaquetage et relargage du génome sont distincts de l'assemblage de la capside. Ces virus sont formés d'un ADN double brin, compressé à l'intérieur d'une capside icosaédrique dont un des sommets est occupé par un sas protéique auquel est fixée une queue (encart de la fig. 2a). Lors de l'infection, ils s'accrochent sur la surface de leur hôte et y injectent leur génome qui sera *exprimé** en nouveaux composants viraux par la bactérie ; des capsides vides sont ensuite assemblées, puis l'ADN est internalisé dans chaque capside à travers le sas protéique déjà présent, grâce à un puissant moteur moléculaire chargé d'enfiler la totalité de l'ADN. La queue est assemblée séparément et vient *in fine* se fixer sur le sas qui se referme ensuite.

L'étude de la physique des processus d'éjection et d'empaquetage de l'ADN s'est beaucoup développée ces dernières années grâce à la possibilité de les reproduire *in vitro*. Des mesures de force sur particules uniques (pinces optiques) ont permis d'analyser finement le fonctionnement du moteur

moléculaire lors de l'empaquetage (fig. 2b) ; la diffusion de la lumière et les microscopies optiques (fluorescence par exemple) ont ouvert l'accès à la cinétique de l'éjection, tandis que la cryomicroscopie électronique a révélé les structures adoptées par les moteurs moléculaires et l'ADN.

Les énergies mises en jeu sont considérables : la chaîne d'ADN, *polyélectrolyte** semi-flexible fortement chargé, de quelques dizaines de micromètres de long, est empaquetée à haute densité (50% en fraction volumique) dans un icosaèdre de quelques dizaines de nanomètres de diamètre. Dans de telles conditions, la pression atteint plusieurs dizaines d'atmosphères. La force développée par le moteur pour empaqueter la chaîne et surmonter la pression qu'elle génère peut atteindre 50 à 100 pN (fig. 2b), soit 10 à 50 fois plus que la force développée par les moteurs moléculaires cellulaires. Les expériences d'empaquetage *in vitro* ont aussi révélé des effets stochastiques, soulignant l'importance de la dynamique, et suggérant des conformations hors équilibre, le temps de relaxation de la chaîne étant supérieur au temps d'empaquetage.

L'éjection *in vitro* est un processus passif, beaucoup plus rapide que l'empaquetage (quelques secondes contre quelques minutes), et est pilotée par la différence de pression entre l'intérieur de la capside et le

milieu extérieur (fig. 2a). Mais elle ne résulte pas d'un simple équilibre entre milieux extérieur et intérieur, comme prédit par les modèles théoriques quasi-statiques : là encore, il s'agit d'un processus stochastique, la quantité d'ADN éjectée par chaque individu n'étant pas déterminée de manière univoque par son environnement.

Une des clés pour une meilleure compréhension de ce phénomène pourrait être l'organisation adoptée par l'ADN dans les conditions extrêmes régnant dans la capside. À cette densité, l'ADN est un cristal liquide, avec un ordre hexagonal mis en évidence dès les années 60 par diffraction des rayons X. Mais, au-delà de cet ordre local, le repliement de la chaîne, qui est unique à chaque phage, est mal compris, et beaucoup reste à faire pour comprendre l'intrication entre empaquetage et éjection, pression et organisation de l'ADN [4, 5].

Ennemis ou amis ?

La compréhension et la maîtrise de plus en plus fines des mécanismes d'assemblage et de désassemblage des virus produit une perception sans doute moins négative de ces agents infectieux. Du statut d'ennemis invisibles, les virus se révèlent avoir leurs failles, mais aussi leur utilité comme nous allons le voir par la suite. Les principes et les mécanismes physiques sont peu à peu mis à jour et laissent entrevoir de nouvelles stratégies thérapeutiques destinées à interrompre le cycle de vie viral. Le fait par exemple d'être capable de bloquer la réplication virale en inhibant soit l'assemblage soit le désassemblage constitue une approche très efficace. Des peptides, c'est-à-dire des petites protéines obtenues par synthèse, sont actuellement en cours de développement pour bloquer l'assemblage du virus de l'immunodéficience humaine (VIH). Des simulations atomistiques par dynamique moléculaire commencent à fournir des informations relatives aux sites des protéines virales sur lesquels ces molécules viennent se fixer, mais les temps de calculs restent prohibitifs pour en tirer les mécanismes complets.

Forts de leur capacité d'assemblage et de leurs exceptionnelles propriétés circulatoires dans un organisme, les virus ont suscité de nouvelles applications biomédicales. Si l'ingénierie génétique des virus se pratique déjà depuis plusieurs décennies, une voie alternative s'est ouverte et consiste à assembler *in vitro* des particules virales

2. Éjection et empaquetage *in vitro* de l'ADN des bactériophages.

(a) L'éjection de l'ADN hors de la capside des bactériophages est déclenchée par la fixation de l'extrémité de la queue sur un récepteur. Le rôle de la pression est mis en évidence en augmentant la pression du milieu extérieur (en présence d'un *osmolyte**). L'éjection est alors partiellement inhibée. Les bactériophages (ici T5 observés par cryomicroscopie électronique) éjectent des quantités d'ADN variables selon les individus. (b) Mesure de la force développée par le moteur moléculaire (tête de flèche noire) lors de l'empaquetage d'une molécule d'ADN. La capside vide (procapside) d'un côté, la molécule d'ADN de l'autre, sont attachées à des microsphères (en bleu) piégées et manipulées par des pinces optiques (d'après D.E. Smith, *Curr. Opin. Virol.* **1** (2011) 134).

>>>

inoffensives pour le patient. Ces particules virales synthétiques conservent la capside du virus natif et donc ses fonctionnalités d'assemblage et de désassemblage, mais elles sont dans l'incapacité de se multiplier car le génome est remplacé par un agent thérapeutique. On est ainsi capable d'assembler des virus contenant des médicaments, éventuellement solubilisés dans des nanoémulsions, à la place du génome (fig. 3a). Des recherches sont menées dans le même esprit pour transporter des gènes non viraux (fig. 3b) vers des cellules défectueuses et procéder à de la thérapie génique. Enfin, des nanoparticules métalliques peuvent être empaquetées aisément dans une capside virale (fig. 3c) et être utilisées pour « brûler » des cellules cancéreuses par application d'un champ magnétique ou par illumination dans l'infrarouge.

La régularité inégalée des capsides virales offre de nombreuses possibilités en science des matériaux. Des capsides icosahédriques vides ont par exemple servi de moules pour la synthèse de nanoparticules métalliques, ce qui a conféré à ces dernières une remarquable monodispersité en taille.

Pour terminer, des circuits de nanofils d'or ont été élaborés par croissance autour de capsides icosahédriques (fig. 3d). La symétrie des réseaux et les distances nanométriques entre les conducteurs laissent entrevoir une approche de fabrication novatrice pour des circuits électroniques avec une forte densité de composants.

Conclusion

Si dès ses origines, la virologie s'est adressée à la physique, il semble qu'aujourd'hui ce soit la physique qui s'est emparée d'un pan de la virologie. La discipline ainsi engendrée est dénommée *virologie physique*, et elle se propose d'appliquer les concepts expérimentaux et théoriques de la physique statistique et des fluides complexes aux particules virales au sens large. Par son approche parfois orthogonale aux sciences du vivant, elle permettra très certainement des progrès complémentaires dans la compréhension du cycle de vie viral ainsi que dans le développement de stratégies thérapeutiques et de matériaux bio-inspirés innovants. ■

a

b

c

d

3. Quelques représentations de particules virales synthétiques.

- (a) Nanogoutte d'huile stabilisée par des tensioactifs et empaquetée dans une capside virale (d'après Chang *et al.*, *ACS Nano* **2** (2008) 281).
 (b) Fragments d'ADN non viraux dans une capside (d'après Kwak *et al.*, *J. Am. Chem. Soc.* **132** (2010) 7834).
 (c) Nanoparticule d'or dans une capside (d'après Sun *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104** (2007) 1354).
 (d) Circuit de nanofils d'or construit autour d'un virus icosahédrique (d'après Blum *et al.*, *Small* **7** (2005) 702).

Glossaire

Génome. Chaîne d'acide nucléique (ADN ou ARN) contenant le code génétique du virus. Le génome viral est *exprimé* par la machinerie des cellules hôtes pour produire les protéines et les acides nucléiques qui s'assembleront en nouveaux virus.

Capside. Coque semi-rigide faite de protéines protégeant le génome du virus. Elle peut prendre diverses formes (icosahédrique, tubulaire, etc) selon les espèces virales.

Bactériophage ou phage. Virus n'infectant que des bactéries.

Nucléation. La nucléation est un phénomène par lequel des atomes ou des molécules s'agregent en un petit germe instable, à partir duquel va croître un assemblage plus conséquent. Originellement observée pour la formation des cristaux solides, la nucléation semble être un processus générique, y compris pour de nombreux assemblages moléculaires biologiques.

Exprimer. Au sens de la biologie, synthétiser des molécules fonctionnelles (ADN ou ARN, protéines...) à partir du code génétique.

Osmolyte. Petites molécules organiques solubles dans la solution intracellulaire, qui permettent de maintenir ou ajuster la pression osmotique des différents compartiments cellulaires et le volume de la cellule.

Polyélectrolyte. Polymère porteur de groupes ionisés, anioniques ou cationiques, et soluble dans l'eau.

En savoir plus

- 1• R.F. Bruinsma et W.S. Klug, *Annu. Rev. Condens. Matter Phys.* **6** (2015) 245.
- 2• V.L. Lorman et S.B. Rochal, *Phys. Rev. Lett.* **98** (2007) 185502.
- 3• G. Tresset *et al.*, *J. Am. Chem. Soc.* **135** (2013) 15373.
- 4• M. de Frutos, A. Leforestier et F. Livolant, *Biophys. Rev. Lett.* **09** (2014) 81.
- 5• M. Jeembaeva, M. Castelnovo, F. Larsson et A. Evilevitch, *J. Mol. Biol.* **381** (2008) 310.
- 6• W.H. Roos, R. Bruinsma et G.J.L. Wuite, *Nat. Phys.* **6** (2010) 733.