

## Outils d'identification de sondes chimiques pour une meilleure compréhension du vivant

**Résumé** La perturbation d'un système biologique complexe par des molécules chimiques s'est avérée une approche gagnante pour une meilleure compréhension du vivant. Cette approche fait appel à des techniques à l'interface de la chimie et de la biologie qui seront détaillées dans cet article. Deux exemples concrets permettent d'illustrer comment le criblage de molécules a permis de résoudre une question biologique et de développer des composés à haute valeur ajoutée.

La chémogénomique, ou génétique chimique, a permis de faire des avancées considérables dans la caractérisation de processus biologiques complexes. Il est question ici d'utiliser des petites molécules pour venir perturber un système biologique d'intérêt au même titre qu'une mutation génétique ou qu'une inhibition de l'expression génétique par un oligonucléotide (siRNA). Comment identifier de telles molécules ? Cette identification s'appuie bien souvent initialement sur le criblage à haut débit (« high-throughput screening », HTS) ou à haut contenu (« high content screening », HCS), qui consiste à tester un très grand nombre de molécules et à sélectionner les molécules actives. Ces approches de criblage ont longtemps été exploitées par les industries pharmaceutiques comme étape préalable au développement de nouveaux médicaments.

Dès le début des années 2000, la communauté académique française s'est emparée de cette opportunité qu'offrait la chémogénomique pour comprendre ou moduler des processus biologiques impliqués dans des pathologies. Ainsi des chimistes, notamment M. Hibert et S. Rault, ont commencé à collecter les molécules synthétisées dans leur laboratoire dans le but de leur donner une seconde vie et de promouvoir les collaborations entre chimistes et biologistes. Cette initiative a conduit à la création de la Chimiothèque Nationale [1-2], qui regroupe maintenant plus de 70 000 composés et 15 000 extraits naturels issus de 45 laboratoires.

De la même façon, les avancées technologiques ont permis à des laboratoires académiques de se doter de plateformes robotisées pour cribler des collections de molécules afin de sélectionner des ligands spécifiques et affins. Le criblage peut être réalisé sur des cibles clairement identifiées ; on sélectionne les petites molécules capables de se lier à des protéines purifiées et de perturber leur fonction (ou leur interaction avec d'autres protéines). Il peut aussi être réalisé sur cellules entières ou sur des organismes modèles ; on l'appelle alors criblage phénotypique car il permet de sélectionner les molécules capables d'induire un phénotype d'intérêt. Dans tous les cas, l'étape préalable est la mise au point d'un test fiable et miniaturisé qui sera réalisé dans des plaques multipuits (de 96 à 1 536). Ce test sera bien évidemment développé en fonction de l'équipement disponible sur la plateforme et devra répondre à des critères stricts de reproductibilité et de robustesse. La communauté académique française autour du criblage et de la découverte de sondes chimiques bioactives s'est fédérée il y a une dizaine d'années en créant le réseau « ChemBioScreen », qui regroupe un ensemble d'équipes de recherche et de plateformes de criblage directement impliquées dans ces approches [3]. Ainsi, plusieurs plateformes

ont une expertise reconnue dans la mise au point de tests basés sur différents modes de détection optique (HTS) ou sur de la microscopie (HCS) et sont capables de réaliser des criblages de chimiothèques composées de quelques centaines jusqu'à des dizaines de milliers de molécules.

Le réseau de chémoinformaticiens organisé en GDR de chémoinformatique (maintenant GDR « BigDataChim ») a aussi joué un rôle essentiel dans ces approches de criblage. En effet, certaines campagnes de criblage ne peuvent pas être réalisées sur des milliers de composés pour des raisons de coût. Le criblage *in silico* permet alors de restreindre le nombre de composés à cribler expérimentalement. La constitution de sous-ensembles représentatifs de la diversité moléculaire (exemple de la Chimiothèque Nationale Essentielle, 1 040 molécules) ou consacrés à une classe de cibles biologiques (kinase, GPCR, inhibiteurs d'interactions protéine-protéine) a été déterminante pour augmenter le ratio de touches issues du criblage [4-5]. Enfin, l'utilisation des résultats de criblages expérimentaux comme sets d'apprentissage dans une approche d'intelligence artificielle peut permettre de comprendre/prédire des activités biologiques, et donc de sélectionner les molécules à cribler expérimentalement.

Plus de 385 criblages de la Chimiothèque Nationale ont été menés, conduisant à plus de 1 500 publications, 84 brevets et sept startups. Les activités de service sont dorénavant intégrées au sein d'une infrastructure de recherche nommée « ChemBioFrance » [6]. Parmi ses nombreux succès, le criblage a généré de nombreuses collaborations, a conduit à l'identification de 65 molécules bioactives qui ont permis de valider une cible *in cellula* ou *in vivo* [7]. Nous avons choisi d'illustrer ces progrès à travers deux exemples : le premier ayant conduit à une avancée en recherche fondamentale et le second ayant conduit à la création d'une startup.

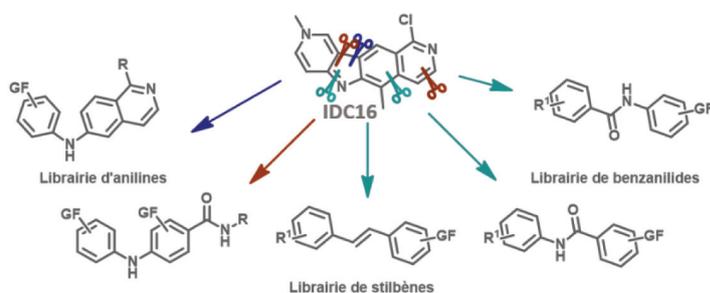
Le premier exemple s'appuie sur un projet visant à identifier des composés antiviraux à large spectre en cherchant non pas des molécules qui vont agir sur le virus, mais des composés qui vont cibler les cellules hôtes, en stimulant leur réponse antivirale innée. Il s'agit ici de rechercher des composés induisant l'expression des ISG (« interferon-stimulated genes »). Ces derniers constituent un vaste ensemble de gènes antiviraux et sont activés en particulier par les interférons de type I (IFN-I). Les IFN-I contrôlent l'expression de plusieurs centaines de gènes codant pour des facteurs capables d'inhiber l'entrée, la réplication et la propagation des virus à différentes étapes de leur cycle. Pour identifier des molécules pharmacologiques capables de mimer l'action des IFN-I, un test phénotypique [8] a été mis au point, basé sur des cellules humaines transfectées

avec un gène rapporteur luciférase sous contrôle d'un élément de réponse aux IFN-I. Cette séquence est classiquement présente dans le promoteur des gènes (ISG) induits par les IFN-I. Ces cellules rapportrices ont été utilisées pour cribler différentes chimiothèques comprenant plusieurs dizaines de milliers de molécules, dont la Chimiothèque Nationale. Cette approche a permis de montrer que des composés inhibant la biosynthèse des pyrimidines induisent un stress cellulaire qui potentialise la réponse aux IFN-I et confère à ces molécules des propriétés antivirales remarquables. Ainsi, un lien fonctionnel jusqu'alors inconnu a pu être établi entre l'homéostasie du pool de nucléosides intracellulaire et la réponse antivirale innée [8-10].

Le second exemple concerne la recherche de modulateurs d'épissage alternatif appliqués à l'ARN viral du VIH. En effet, le virus utilise la machinerie d'épissage alternatif de l'hôte pour préparer les ARN messagers nécessaires à sa réplication. Le criblage de la Chimiothèque Curie-CNRS (qui fait partie de la Chimiothèque Nationale) a été effectué grâce à un test phénotypique faisant aussi appel au gène rapporteur de la luciférase. Cependant, la molécule identifiée (nommée IDC16) possède une structure hétérocyclique plane responsable de propriétés d'intercalation dans l'ADN, toxiques pour la cellule [11-12]. Pour éviter cette toxicité, des molécules analogues plus flexibles ont été préparées par synthèse parallèle à l'Institut Curie (voir figure).

Ainsi, grâce à des efforts de synthèse, des molécules prometteuses ont été obtenues et ont permis de créer une startup en 2008. Cette société, nommée Abivax [13], a achevé en 2015 des essais cliniques en phase 2a dans le domaine du VIH [14] et mène actuellement des phases cliniques dans le domaine de l'inflammation (phase 3 en rectocolite hémorragique et phase 2b en maladie de Crohn et polyarthrite rhumatoïde prévues en 2022). L'identification de la cible, le complexe de liaison de la coiffe de l'ARN (« cap binding complex »), a été réalisée à l'aide d'un outil chimique. Celui-ci est un composé trifonctionnel comportant le principe actif, un groupement photoactivable générant un lien covalent avec la cible et une fonction reportrice permettant la capture (« pull-down ») du complexe covalent drogue-cible [14].

Les molécules identifiées, aussi appelées touches, sont des outils pour comprendre et perturber un phénotype ou peuvent à travers un long processus conduire à des candidats-médicaments. Ces succès n'auraient pas été possibles sans les liens étroits entre les communautés de chimistes et de biologistes. Bien sûr, il reste des défis importants. Le premier concerne la diversification des composés testés pour une plus grande couverture de l'espace chimique. Dans ce contexte, les composés originaux émanant de laboratoires de synthèse organique ne travaillant pas à l'interface chimie biologie représentent une diversité moléculaire nouvelle et riche et devraient intégrer la collection nationale. La Chimiothèque Nationale a aussi pour projet de synthétiser des châssis moléculaires innovants proposés par intelligence artificielle pour combler des espaces chimiques vacants. Le second défi porte sur l'annotation des résultats du criblage afin de permettre dans le futur d'affiner des modèles de prédiction de chémoinformatique et de mieux choisir les composés à cribler ainsi que les touches à développer. Pour le criblage, notamment dans le domaine des antimicrobiens, le problème est de trouver des touches capables de passer à travers la ou les membranes en fonction de l'organisme étudié. Ainsi, des criblages combinant un test biochimique sur une cible bien



Préparation de bibliothèques d'analogues de IDC16 par synthèse parallèle robotisée.

identifiée dans un contexte cellulaire (appelé « target-based whole cell screening ») sont de plus en plus mis en place ; ils permettent de résoudre le problème de la perméabilité tout en s'affranchissant de la recherche de la cible, qui reste dans le cas d'un criblage phénotypique une étape difficile à franchir. L'utilisation de sondes multifonctionnelles issues de la touche, comme celle présentée dans le second exemple, est une stratégie d'intérêt. Cette approche dite « activity-based protein profiling » [15] a déjà fait ses preuves, mais la découverte et l'amélioration de ces outils de chémobiologie restent nécessaires et permettront l'accélération de cette validation de nouvelles cibles. Dans le domaine des anticancéreux (mais pas uniquement), le développement de tests sur des sphéroïdes et des organoïdes est très prometteur car ils permettent de mieux approcher la complexité de la tumeur ou du tissu considéré.

- [1] M.F. Hibert, French/European academic compound library initiative, *Drug Discov.*, **2009**, *14*, p. 723-725.
- [2] F. Mahuteau-Betzer, The French National Compound Library: advances and future prospects, *Med. Sci.*, **2015**, *31*, p. 417-422.
- [3] [www.chembioscreen.fr](http://www.chembioscreen.fr)
- [4] N. Bosc *et al.*, Fr-PPIChem: an academic compound library dedicated to protein-protein interactions, *ACS Chem. Biol.*, **2020**, *15*, p. 1566-574.
- [5] S. Millhas *et al.*, Protein-protein interaction inhibition (2P2I)-oriented chemical library accelerates hit discovery, *ACS Chem. Biol.*, **2016**, *11*, p. 2140-148.
- [6] <https://chembiofrance.cn.cnr.fr/fr>
- [7] <https://chembiofrance.cn.cnr.fr/fr/success-stories>
- [8] M. Lucas-Hourani *et al.*, Inhibition of pyrimidine biosynthesis pathway suppresses viral growth through innate immunity, *PLoS Pathog.*, **2013**, *9*, e1003678.
- [9] P.O. Vidalain *et al.*, Stimulation of the antiviral innate immune response by pyrimidine biosynthesis inhibitors: a surprise of phenotypic screening, *Med. Sci.*, **2015**, *31*, p. 98-104.
- [10] M. Lucas-Hourani *et al.*, Original chemical series of pyrimidine biosynthesis inhibitors that boost the antiviral interferon response, *Antimicrob. Agents Chemother.*, **2017**, *61*, e00383-17.
- [11] J. Soret *et al.*, Selective modification of alternative splicing by indole derivatives that target serine-arginine-rich protein splicing factors, *Proc. Natl. Acad. Sci. USA*, **2005**, *102*, p. 8764-769.
- [12] N. Bakkour *et al.*, Small-molecule inhibition of HIV pre-mRNA splicing as a novel antiretroviral therapy to overcome drug resistance, *PLoS Pathog.*, **2007**, *3*, p. 1530-539.
- [13] [www.abivax.com/?lang=fr](http://www.abivax.com/?lang=fr)
- [14] N. Campos *et al.*, Long lasting control of viral rebound with a new drug ABX464 targeting Rev - mediated viral RNA biogenesis, *Retrovirology*, **2015**, *12*, 30.
- [15] H. Fang *et al.*, Recent advances in activity-based probes (ABPs) and affinity-based probes (AfBPs) for profiling of enzymes, *Chem. Sci.*, **2021**, *12*, p. 8288-310.

**Florence MAHUTEAU-BETZER,**

Directrice de recherche CNRS, Institut Curie, UMR 9187-U1196, Université Paris-Saclay.

**Hélène MUNIER-LEHMANN,**

Chargée de recherche Inserm, Institut Pasteur, Unité de Chimie et Biocatalyse, Paris.

\* [florence.mahuteau@curie.fr](mailto:florence.mahuteau@curie.fr) ; [hmunier@pasteur.fr](mailto:hmunier@pasteur.fr)