

> La Chimiothèque Nationale (CN) est une fédération de chimiothèques académiques locales, une chimiothèque étant une collection de petites molécules de synthèse ou isolées de produits naturels. Une « extractothèque » contenant des extraits naturels est aussi proposée. La force de la Chimiothèque Nationale réside dans la diversité de sa collection et la possibilité de proposer des interlocuteurs chimistes aux biologistes cribleurs. Les collaborations post-criblage permettent, à travers la synthèse d'analogues et d'outils chimiques, d'identifier de nouvelles sondes chimiques pour comprendre le vivant et, dans certains cas, d'envisager des applications dans le domaine thérapeutique. Dans cet article, nous montrerons les avancées et les perspectives que cette fédération a permis d'engendrer. >

Achévé en 2003, le séquençage du génome humain a permis l'identification d'un grand nombre de cibles biologiques. Un énorme besoin en sondes chimiques utiles pour étudier ces nouvelles cibles a donc émergé. Parallèlement, les avancées en termes de robotisation et de miniaturisation ont permis d'envisager le criblage d'un grand nombre de molécules pour identifier des ligands à ces nouvelles cibles. Afin d'accéder à de grandes collections de petites molécules, deux stratégies complémentaires ont été adoptées : la préparation de chimiothèques par synthèse parallèle<sup>1</sup> (cas des chimiothèques commerciales et d'autres initiatives académiques comme le *NIH molecular libraries program*) [1], l'exhumation des composés synthétisés au cours du temps. Les « big pharma » ont souvent eu recours aux deux stratégies. Le criblage de chimiothèques est largement utilisé dans l'industrie pharmaceutique dans le but d'accélérer la découverte de

Dans le monde académique, le criblage n'a pas la même voca-

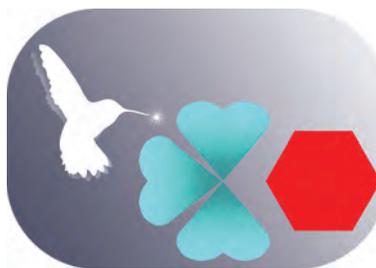
(→) voir la Synthèse de D. Rognan et P. Bonnet, *m/s* n° 12, décembre 2014, p. 1152

Cet article fait partie de la série « Chémobiologie » dont la parution a débuté dans le n° 12, vol. 30, décembre 2014 ([www.medecinesciences.org](http://www.medecinesciences.org)).

## Chémobiologie (9) Chimiothèque Nationale

### Avancées et perspectives

Florence Mahuteau-Betzer



Institut Curie,  
UMR 9187-U1196,  
bâtiment 110,  
Centre Universitaire,  
Chimie, modélisation  
moléculaire et imagerie  
pour la biologie,  
91405 Orsay, France.  
[florence.mahuteau@curie.fr](mailto:florence.mahuteau@curie.fr)

tion. S'il peut faciliter dans certains cas la découverte de molécules susceptibles d'avoir des applications thérapeutiques, il vise surtout à l'identification de sondes précieuses pour étudier le vivant en modulant ou contrôlant une activité biologique. En effet, la petite molécule identifiée par criblage, nommée « touche », facilite l'étude de la nouvelle cible. Lorsque le criblage est enzymatique, la petite molécule permet de valider la fonction biologique de cette enzyme. Lorsque le criblage est phénotypique, elle peut permettre d'identifier la cible [3].

### Naissance des chimiothèques académiques françaises

Ce formidable potentiel d'innovations et de collaborations académiques entre biologistes et chimistes a conduit, dès la fin des années 1990, quelques laboratoires de pharmacochimie français à organiser les composés de synthèse obtenus au fil du temps en chimiothèques, pour faciliter le criblage virtuel et/ou expérimental. Rapidement, en 2003, ces initiatives locales ont été fédérées en Chimiothèque nationale (CN) [4]. Le premier succès de la Chimiothèque nationale est venu du recensement du patrimoine des laboratoires. En effet, une fois que les composés synthétisés dans le cadre de programme de recherche étaient publiés dans des articles scientifiques, des brevets ou des thèses, ils étaient voués à l'oubli. Heureusement, la plupart de ces « trésors » ont été conservés dans les laboratoires, et la constitution de chimiothèques académiques a permis leur exhumation. Ce patrimoine recensé par la Chimiothèque Nationale est très important puisqu'il comprend plus de 56 000 composés en octobre 2014, et

<sup>1</sup> La synthèse parallèle consiste à préparer dans des réacteurs différents des composés analogues en changeant uniquement une des entités chimiques engagées. Elle est souvent automatisée et conduit à la préparation de chimiothèques focalisées offrant peu de diversités structurales.

peut être évalué à 1 000 € le produit, soit 55 millions d'euros pour la Chimiothèque Nationale. Enfin, cette constitution de chimiothèque offre « une seconde vie » à ces molécules oubliées qui peuvent être testées sur de nouvelles cibles ou de nouvelles fonctions biologiques. Cette initiative française n'est pas isolée. À partir de 2004, les *National institutes of health* aux États-Unis ont identifié neuf laboratoires constituant le « *Molecular libraries probe production center* » capables de produire de nouvelles molécules par synthèse parallèle. Cette initiative a permis de préparer 350 000 molécules criblées à travers un réseau de plates-formes de criblage académique. Des initiatives de constitution de collections pour le criblage ont été menées dans d'autres pays, mais les collections sont restées modestes et peu organisées.

Devant l'essor du criblage en milieu industriel et académique, un certain nombre de sociétés se sont dédiées à la préparation et la commercialisation de chimiothèques. Celles-ci contiennent des molécules difficilement brevetables. En effet, ces molécules peuvent être brevetées pour leur application sur une cible biologique, mais leurs structures ne sont pas brevetables en tant que telles. De plus, leurs propriétés ADMETox<sup>2</sup> ne sont pas optimisées, ce qui empêche leur développement comme candidat médicament ou sonde chimique. Il est donc admis que les touches identifiées par criblage ne sont pour la plupart pas utilisables en l'état, et qu'un travail d'optimisation de structure est nécessaire.

Les chimiothèques académiques présentent les mêmes limites (brevetabilité, propriétés ADMETox). En revanche, elles permettent d'établir des collaborations entre chimistes et biologistes, et donc d'envisager l'optimisation des touches. En effet, les chimistes qui déposent leurs produits dans la Chimiothèque Nationale sont intéressés par les résultats de criblage et sont capables de produire des analogues des touches. Cette possibilité d'optimisation des touches identifiées par criblage représente un atout très fort pour le biologiste. Enfin, la Chimiothèque Nationale, parce qu'elle regroupe une collection de molécules issues de laboratoires qui développent des familles chimiques très différentes, comporte donc une grande diversité de structures [5]. Bien que plus petites que les collections commerciales ou industrielles, les collections académiques possèdent des structures originales et très diverses, encore peu criblées au regard des autres collections.

## La Chimiothèque Nationale : missions, fonctionnement, utilisateurs

### Rôle du groupement d'intérêt scientifique – Chimiothèque Nationale (GIS CN)

Depuis 2007, la Chimiothèque Nationale repose sur deux structures : un groupement d'intérêt scientifique (GIS CN : fédération des établissements universitaires et organismes de recherche adhérents du projet) et l'unité de gestion de la Chimiothèque Nationale (UGCN : unité support du GIS CN). Le GIS compte à l'heure actuelle 42 établissements publics de recherche. Un comité de pilotage est chargé des décisions stratégiques. Des ateliers thématiques sont créés selon les

besoins du GIS CN. Dans les laboratoires, des personnels nommés « chimiothécaires » organisent la collection localement, en gérant les collectes de composés et les conditionnements en plaques (voir ci-dessous). Ils sont les interlocuteurs de l'UGCN et font le lien entre les chimistes de synthèse, pourvoyeurs des collections, et leurs utilisateurs. La Chimiothèque Nationale forme tout nouveau chimiothécaire aux bonnes pratiques grâce au réseau déjà existant.

La collection de molécules est vivante. Si certaines molécules peuvent disparaître, la Chimiothèque nationale a aussi vocation à s'agrandir. Pour ce faire, le GIS CN accompagne les laboratoires de chimie intéressés par cette initiative en proposant des aides logistiques et matérielles. En 2013, certaines plates-formes chimiothèques équipées d'automates ont été sollicitées pour mettre en plaques les collections de laboratoires non équipés et ainsi permettre d'accroître la taille de la collection nationale en plaques. De plus, le GIS s'occupe de promouvoir la constitution de chimiothèques par des actions de communication (présence à des congrès, prise de contacts ciblés avec des laboratoires) et de formation. Une autre voie encore peu explorée consiste à créer des chimiothèques ciblées autour de structures originales synthétisées dans certains laboratoires.

### La collection

La Chimiothèque Nationale regroupe à l'heure actuelle plus de 56 000 molécules de synthèse et produits naturels issus des laboratoires académiques français, dont 48 000 sont accessibles sous forme de plaques 96 puits. Cette collection est répertoriée dans une base de données et possède plusieurs modes de stockage physique : des solutions dans le DMSO (diméthylsulfoxyde) au format microplaque et des poudres dans des piluliers (Figure 1). Ces poudres servent à une nouvelle fourniture du produit pour confirmer l'activité des touches. Ceci permet de réévaluer le produit en s'affranchissant d'une éventuelle dégradation en solution. Il est souvent réalisé un criblage secondaire sur ces touches, ce criblage, différent du premier, permet de valider l'activité biologique par un autre test. La communauté des chimiothécaires, organisée en ateliers, a permis d'adopter des protocoles communs de gestion et de préparation des microplaques destinées au criblage. Les plaques livrées sont des plaques 96 puits comportant 80 substances à des concentrations massiques de 10 mg/mL dans le DMSO. L'atelier des chimiothécaires permet une réflexion constante sur tous les aspects de la gestion physique et informatique des collections de substances, la préparation des plaques, mais aussi sur les procédures d'envois de poudres ou de plaques. L'atelier chémo-informatique a notamment permis de créer un nouveau logiciel de base de données chimiques adapté

<sup>2</sup> Ce sont les critères de criblage *in vitro* de cinétique et de métabolisme : absorption, distribution, métabolisme, excrétion, toxicité.

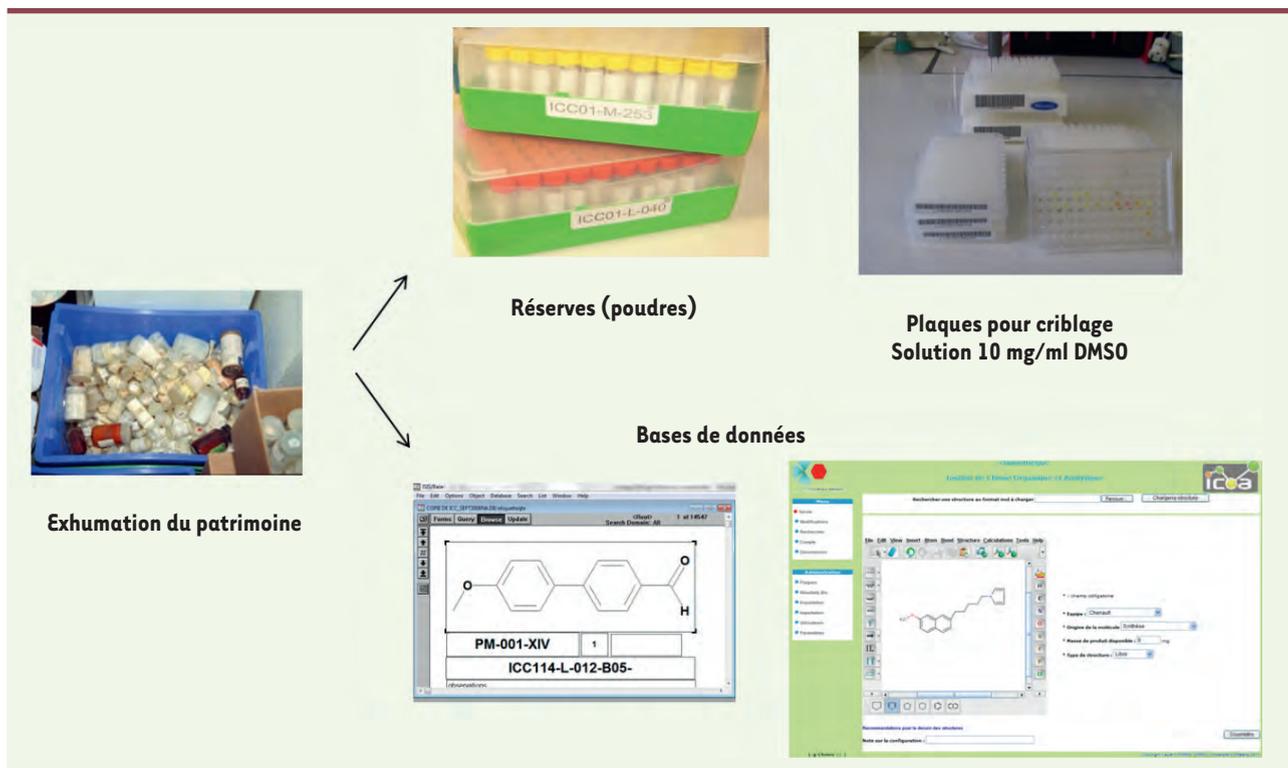


Figure 1. Constitution des chimiothèques locales.

à la gestion des chimiothèques de laboratoires multi-équipes : LgChimio. Enfin, l'atelier « descripteurs moléculaires » a défini les standards internes pour la représentation des structures et l'échange de données chimiques et biologiques. Cet atelier a également permis d'enrichir le système d'information par de nouveaux champs, donnant ainsi la possibilité d'explorer de manière plus fine les collections de substances.

### Accès à la collection

Depuis 2005, le fichier SDF<sup>3</sup> des structures présentes dans la collection est accessible sur internet (<http://chimiotheque-nationale.enscm.fr/>) après un enregistrement préalable de l'utilisateur (Figure 2). Une requête peut ensuite être faite, soit pour l'accès à quelques composés (en poudre) sélectionnés par sous-structure ou par criblage virtuel, soit pour des plaques multipuits. L'UGCN transmet cette demande aux laboratoires détenteurs des molécules demandées. Après réponse de ces derniers, l'UGCN engage les procédures juridiques nécessaires à l'accord de transfert de matériel (MTA en anglais). Si les tutelles du laboratoire demandeur sont signataires du GIS, alors une simple lettre d'intention signée par l'UGCN et le laboratoire demandeur est requise. Si une des tutelles du laboratoire n'est pas membre du GIS, alors un accord de transfert de matériel est nécessaire et requiert la signature des tutelles du laboratoire concerné par l'accord. Cette procédure mise en place depuis 2009 a considérablement facilité la mise à disposition de la collection.

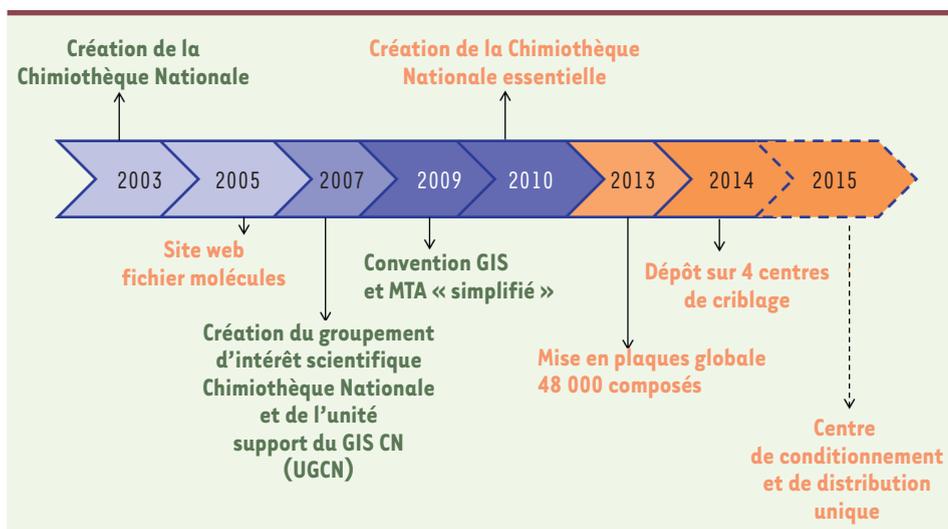
<sup>3</sup> Fichier extension (standard database format) optimisé pour des ensembles de données d'informations très larges.

L'accord régit aussi les modalités de la collaboration et notamment :

- la copropriété des résultats entre l'utilisateur et les membres du GIS CN concernés, uniquement pour les produits dont ils sont propriétaires ;
- des retours financiers aux membres du GIS CN concernés en cas d'exploitation commerciale desdits résultats ;
- la transmission des résultats par le tiers utilisateur à l'UGCN qui se chargera de les transmettre aux laboratoires du GIS CN concernés, uniquement pour les produits dont ils sont propriétaires ;
- un droit de priorité pour le(s) laboratoire(s) du GIS CN concerné(s) de participer à un programme de recherche entrepris par l'utilisateur sur les éventuelles touches ;
- un droit d'utilisation libre et gratuit des résultats pour les copropriétaires, uniquement dans le cadre de projet de recherche interne.

Une participation financière est demandée aux utilisateurs de la Chimiothèque Nationale, elle ne couvre évidemment pas les frais de synthèse (ce qui la différencie des chimiothèques commerciales), mais permet de partager le coût des envois et des consommables.

Il est à noter que le criblage peut être réel ou virtuel. Dans le cas d'un criblage virtuel initial, seuls quelques composés sont testés expérimentalement. Ces deux approches sont souvent complémentaires et parfois utilisées dans un même programme.



**Figure 2. Avancées de la Chimiothèque Nationale.** CN : Chimiothèque Nationale ; UGCN : unité support du GIS CN ; GIS CN : groupement d'intérêt scientifique Chimiothèque Nationale.

### Développements post-criblage

Une fois le criblage effectué, le cribleur et le chimiothécaire doivent valider la touche en confirmant la pureté et l'activité biologique à partir de poudres (Figure 3). Une fois la touche confirmée, les chimistes et biologistes concernés par les molécules touchées et par la cible ou la fonction biologique testée décident ou non de continuer l'aventure à travers une collaboration de recherche. Celle-ci peut contribuer à identifier la cible biologique (la touche est une sonde permettant la compréhension de fonctions biologiques), et/ou à améliorer l'activité biologique (dans un contexte thérapeutique). Ces développements post-criblage sont pluridisciplinaires, et le criblage virtuel peut apporter une aide importante à la constitution de chimiothèques ciblées d'analogues de touches (Figure 4). Si la cible biologique est bien identifiée et caractérisée, des approches basées sur le récepteur comme le *docking*<sup>4</sup> sont utilisées. Sinon, des approches basées sur le ligand peuvent être adoptées. Dans tous les cas, des chimiothèques ciblées d'analogues sont générées. Quand la cible n'est pas connue, son identification peut être entreprise avec notamment l'aide d'outils chimiques (colonne d'affinité, photocrosslinker, etc.). Enfin, la sonde ou le *lead* obtenu doivent avoir des propriétés de sélectivité et d'activité améliorées.

### Promouvoir le criblage de la Chimiothèque Nationale

Un des rôles du GIS CN est de promouvoir l'outil qui constitue les chimiothèques et donc le criblage. Depuis sa création, la Chimiothèque Nationale a toujours travaillé en étroite collaboration avec d'une part la communauté des chimioinformaticiens à travers le groupement de

recherche (GDR) chémoinformatique, et, d'autre part, les responsables de plates-formes de criblage regroupés dans le GDR ChemBioScreen. Un certain nombre d'actions ont été menées : la constitution d'une chimiothèque essentielle, le dépôt de copies de la Chimiothèque Nationale sur des plates-formes de service identifiées, et la mise en place d'un centre de gestion et de distribution unique.

### Chimiothèque essentielle

Le criblage à haut ou moyen débit d'une grande collection de composés requiert la miniaturisation des tests et peut entraîner des coûts importants. Le criblage est évalué à 1 € le puits pour les tests les moins onéreux. Pour faciliter le criblage manuel de la collection et rendre la chimiothèque accessible à un plus grand nombre d'utilisateurs, une chimiothèque essentielle (CNE) a été constituée en 2009. Elle constitue un sous-ensemble représentant la diversité moléculaire de la Chimiothèque Nationale (640 substances sur les 40 000 présentes dans la base en 2009). Cette sélection a été réalisée grâce à l'aide de chimioinformaticiens. Outre l'annotation biologique de ce sous-ensemble, des propriétés physicochimiques comme la solubilité, le pKa, ont été évaluées expérimentalement et confrontées à des modèles prédictifs [6]. Cette première expérience de constitution de chimiothèque essentielle a révélé aussi quelques difficultés. En effet, un certain nombre de composés sélectionnés n'ont pas pu faire partie de ce sous-ensemble pour des raisons de disponibilité de poudre ou pour des raisons de pureté insuffisante. Ceci n'a pas pour autant entaché le bilan positif de cette initiative. Tout d'abord, un programme interdisciplinaire de recherche (PIR) innovation thérapeutique a été lancé pour financer le criblage de cette collection. Douze projets ont été soutenus. Depuis, la chimiothèque essentielle a été criblée à treize autres reprises. Les résultats connus à ce jour montrent que 90 % des criblages ont conduit à l'identification de touches et ce malgré la taille relativement petite de la collection. Une seconde version de cette chimiothèque essentielle est envisagée en élargissant cette chimiothèque existante à 640 autres composés.

<sup>4</sup> Le *docking* (ou ancrage) à haut débit consiste à prédire à la fois la conformation active et l'orientation relative de chacune des molécules de la chimiothèque sélectionnée par rapport à la cible d'intérêt. (Didier Rognan).

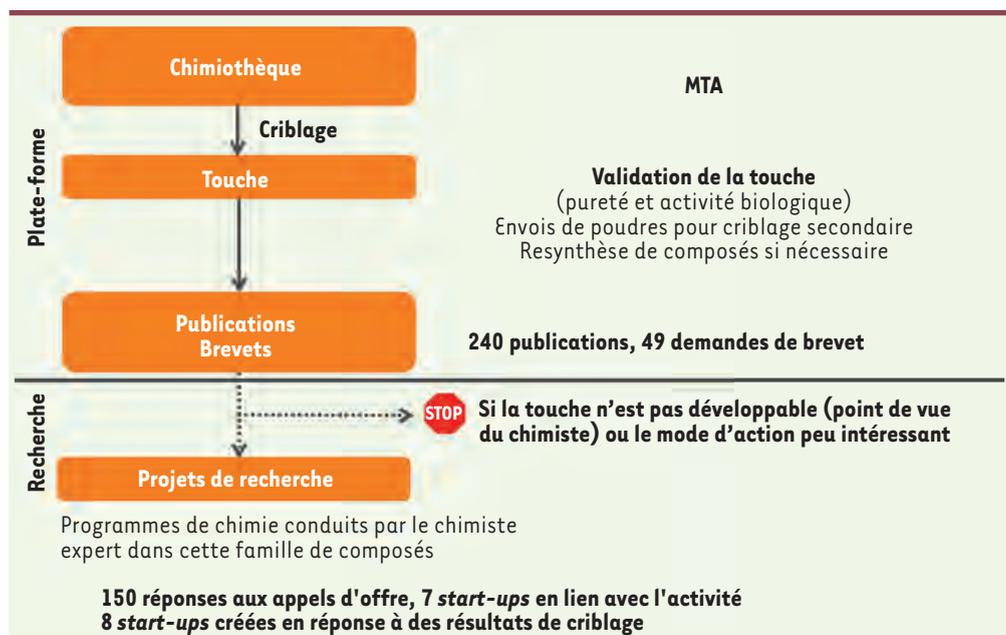


Figure 3. Fonctionnement et bilan de la Chimiothèque Nationale.

### Copies mises à disposition sur quelques plates-formes de criblage

Récemment, des copies de la Chimiothèque Nationale ont été mises à disposition sur quatre plates-formes de criblage de service de manière à faciliter leur accès aux utilisateurs potentiels. Ces plates-formes ont été identifiées et sélectionnées comme plates-formes de service pouvant répondre à des sollicitations extérieures et réaliser des criblages de l'ordre de 50 000 composés. Les centres concernés sont : la plate-forme PCBIS (*platform of integrative chemical biology* de Strasbourg), la plate-forme de l'Institut Pasteur à Paris, la plate-forme de l'USR3388 Pierre Fabre à Toulouse et la plate-forme du CMBA (criblage de molécules bioactives) à Grenoble (→). Cette initiative est née de plusieurs constats. Tout d'abord, ce dépôt permet de favoriser le criblage de la Chimiothèque Nationale dans son ensemble. Il réduit les coûts en termes d'envoi de plaques puisque les plates-formes gèrent les jeux de plaques. Enfin, le fait que ces plates-formes disposent de copies intégrales des collections Chimiothèque Nationale ouvre aux laboratoires qui possèdent de petites collections la possibilité de voir leurs substances participer à de grandes campagnes de criblage, ce qui, à terme, devrait permettre d'attirer de nouveaux contributeurs à la Chimiothèque Nationale (cercle vertueux).

(→) voir la Synthèse de F. Chevalier et E. Maréchal, *m/s* n° 3, mars 2015, page 320

### Centre unique de gestion et de distribution de la Chimiothèque Nationale

Il apparaît déjà depuis un certain temps que la Chimiothèque Nationale serait plus efficace si elle était regroupée et distribuée de manière unique et homogène. Si la distribution sur certains centres

de criblage représente déjà une avancée remarquable dans ce sens, un centre unique de gestion et de distribution de plaques permettrait de la distribuer à d'autres cribleurs intéressés et de pouvoir aussi distribuer la chimiothèque essentielle. Des demandes de financement ayant échoué à plusieurs reprises, de nouvelles options sont envisagées comme le recours à un prestataire de service pour l'hébergement et la distribution de la Chimiothèque Nationale.

### Actions de fédération des communautés

Dans un souci de promouvoir le criblage à l'échelle européenne, les communautés de chimiothécaires, chémoinformaticiens et cribleurs ont participé activement à l'ESFRI (*European strategy forum on research infrastructures*) EU-Openscreen [7]. Cette ESFRI arrive bientôt au terme de sa phase préparatoire et a vocation à proposer le criblage à travers des plates-formes à très haut débit d'une chimiothèque européenne (composée de collections académiques et de composés commerciaux sélectionnés par les chémoinformaticiens) [8]. Si ce projet aboutit, la Chimiothèque Nationale deviendra le point d'entrée national pour cette infrastructure européenne. Il faut noter que la Chimiothèque Nationale est la seule chimiothèque européenne composée de molécules collectées dans

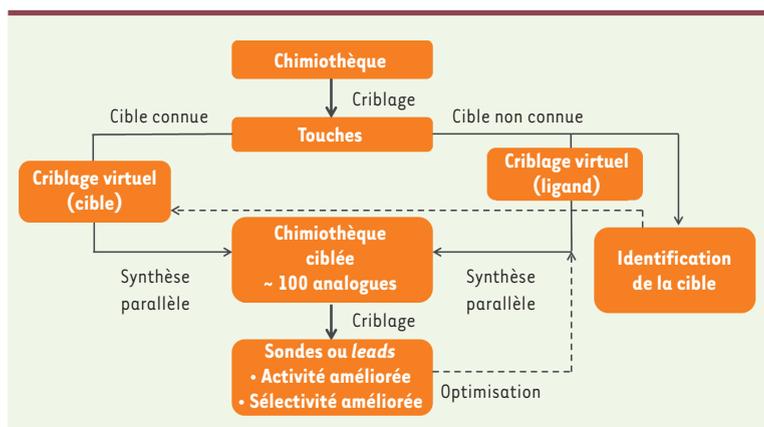


Figure 4. Stratégies post-criblage.

des laboratoires académiques de taille importante et organisée. En effet, d'autres pays possèdent des collections organisées commerciales et quelques chimiothèques collectées en milieu académique mais de tailles beaucoup plus faibles.

### Bilan et perspectives

La Chimiothèque Nationale a permis de fédérer les initiatives locales et présente un bilan très positif concernant son criblage en milieu académique. Les nombreuses collaborations générées par le criblage ont débouché sur des projets très fructueux. Plus de 240 publications ont été recensées ainsi que 49 demandes de brevets. Enfin, 14 *start-ups* ont été créées à partir de résultats de criblage ou d'activités associées aux chimiothèques.

Ces succès reposent sur des criblages partiels de la Chimiothèque Nationale. On peut donc s'attendre à l'augmentation de ces succès. En effet, la mise en plaques des petites collections réalisée en 2013 a permis d'accroître la taille de la collection. Enfin, les actions récentes de mise à disposition de la Chimiothèque Nationale sur des plates-formes de criblage, ainsi que la constitution d'un centre unique de gestion de la Chimiothèque Nationale, devraient accentuer le criblage total de la collection et augmenter le nombre de touches, et donc, à terme, de succès.

Les plates-formes de criblage peuvent sans difficulté cribler la totalité de la collection (environ 50 000 molécules) puisque celle-ci est dorénavant disponible sur un certain nombre de plates-formes et que la constitution d'un centre unique de gestion et de distribution semble imminente. Sachant que le taux moyen de touches sur un criblage avoisine les 1 %, il est nécessaire d'avoir une collection importante et très diverse pour trouver des touches intéressantes. Le criblage total et régulier de la Chimiothèque Nationale permettrait aussi de pouvoir annoter un large panel de résultats biologiques de manière à mieux prédire ou comprendre l'activité des petites molécules.

Le seul frein à ces criblages totaux reste leur coût, même si ce coût n'est finalement pas exorbitant (de 1 à 10 € par composé criblé, selon la complexité du test). Le problème réside notamment dans la difficulté à trouver des appels d'offres qui permettent de financer le criblage ainsi que les premiers travaux de post-criblage. Des financements sont nécessaires pour transformer le potentiel que représente le criblage de la Chimiothèque Nationale en termes à la fois de compréhension des mécanismes du vivant et de valorisation. ♦

### SUMMARY

#### The French National Compound Library: advances and future prospects

The French National Compound Library (*Chimiothèque Nationale*) has been created in 2003 and is the federation of local collections. It contains more than 56 000 small molecules and natural compounds synthesised or isolated in different laboratories over the past years. This explains the diversity of the collection. The strength of this initiative is the ability to connect chemists and biologists for the development of hits. This development involves the synthesis of analogues or/and chemical tools to find new targets. These collaborations lead to the identification of new chemical probes. These probes able to modulate a biological function are essential to study biological pathways. They can also be useful for therapeutic applications. This article will describe the major achievements and perspectives of the French Chemical Library. ♦

### LIENS D'INTÉRÊT

L'auteur déclare n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

### RÉFÉRENCES

1. McCarthy A. The NIH molecular libraries program: identifying chemical probes for new medicines. *Chem Biol* 2010 ; 17 : 549-50.
2. Macarron R, Banks MN, Bojanic D, et al. Impact of high-throughput screening in biomedical research. *Nat Rev Drug Discov* 2011 ; 10 : 188-95.
3. Prudent R, Soleilhac E, Barette C, et al. Les criblages phénotypiques ou comment faire d'une pierre deux coups : découvrir la cible et la molécule pharmacologique capable de la réguler. *Med Sci (Paris)* 2013 ; 29 : 897-905.
4. Hibert MF. French/European academic compound library initiative. *Drug Discov Today* 2009 ; 14 : 723-5.
5. Krier M, Bret G, Rognan D. Assessing the scaffold diversity of screening libraries. *J Chem Inf Model* 2006 ; 46 : 512-24.
6. Ruggiu F, Gizzi P, Galzi JL, et al. Quantitative structure-property relationship modeling: a valuable support in high-throughput screening quality control. *Anal Chem* 2014 ; 86 : 2510-20.
7. Frank R. EU-OPENSREEN : a European infrastructure of open screening platforms for chemical biology. *ACS Chem Biol* 2014 ; 9 : 853-4.
8. Horvath D, Lisurek M, Rupp B, et al. Design of a general-purpose European compound screening library for EU-OPENSREEN. *ChemMedChem* 2014 ; 9 : 2309-26.
9. Rognan D, Bonnet P. Les chimiothèques et le criblage virtuel. *Med Sci (Paris)* 2014 ; 30 : 1152-60.

### TIRÉS À PART

F. Mahuteau-Betzer



Tarifs d'abonnement m/s - 2015

**Abonnez-vous**

**à médecine/sciences**

> Grâce à m/s, vivez en direct les progrès des sciences biologiques et médicales

Bulletin d'abonnement page 453 dans ce numéro de m/s

