

► Les interactions protéine-protéine sont impliquées dans de nombreux processus cellulaires, ainsi que dans leur dysfonctionnement, ce qui en font des cibles thérapeutiques de choix. Toutefois, la conception de composés capables de moduler ce type d'interactions reste difficile et requiert la mise en place d'outils spécifiques, permettant d'accélérer les campagnes de développement de molécules bioactives et de diminuer leur coût. Les succès récents ont permis de caractériser certaines propriétés structurales et physicochimiques des interfaces protéine-protéine, ce qui a abouti à une possibilité d'inhibition de ces interactions par des petites molécules chimiques non peptidiques, ainsi qu'à la définition d'un profil caractéristique des composés chimiques associés. Dans cette revue, nous présentons le développement de collections de composés dédiées à ces cibles innovantes. ►

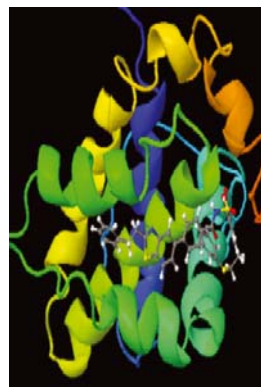
Le besoin en chimiothèques focalisées

Le nombre de molécules obtenant une autorisation de mise sur le marché stagne depuis de nombreuses années, en dépit de l'augmentation considérable des sommes investies par les sociétés pharmaceutiques en recherche et développement. Ainsi, seulement 27 nouveaux médicaments ont été approuvés en 2013, dix de moins que l'année précédente. Plus inquiétant, au cours des cinq dernières années, seuls une dizaine de nouveaux composés ont été approuvés en moyenne chaque année par les plus grandes sociétés pharmaceutiques. Les causes sont multiples et sont notamment liées à des raisons économiques et au renforcement des règles de précautions sanitaires, mais également aux stratégies de découvertes de nouveaux médicaments traditionnellement utilisées depuis

Cet article fait partie de la série « Chémobiologie » qui a débuté dans le n° 12, vol. 30, décembre 2014 (www.medecinesciences.org).

Chémobiologie (7) Les chimiothèques ciblant les interactions protéine-protéine

Olivier Sperandio¹, Bruno O. Villoutreix¹,
Xavier Morelli², Philippe Roche²



¹ Molécules thérapeutiques *in silico* (MTi), université Paris Diderot, Inserm UMR-S973, 35, rue Hélène Brion, 75205 Paris Cedex 13, France ;

² Centre de recherche en cancérologie de Marseille (CRCM), CNRS UMR7258 ; Inserm U1068 ; institut Paoli-Calmettes ; université d'Aix-Marseille UM105, 27, boulevard Lei Roure, 13009, Marseille, France.
philippe.roche@inserm.fr

plusieurs années. Il apparaît nécessaire de reconsidérer certaines étapes du processus de découverte. En particulier, deux sources d'amélioration peuvent être envisagées dans les premières étapes : d'une part, la sélection de cibles innovantes et, d'autre part, la qualité et la pertinence des collections de composés utilisées lors des campagnes de criblage pour identifier des molécules bioactives. La définition de la chimiothèque « idéale » a évolué au fil des ans. Historiquement, la stratégie de la plupart des sociétés pharmaceutiques et de plusieurs groupes académiques était de concevoir une collection de composés très vaste et, dans certains cas, très diversifiée, pouvant être testée contre toutes les cibles protéiques dans les différentes campagnes de criblage à haut débit [43] (→). Pour tenter de réduire le fort taux d'attrition, ainsi que les coûts lors des phases précoces de découverte de nouveaux médicaments, la qualité de ces chimiothèques a été améliorée par l'application de filtres simples se basant sur les propriétés physicochimiques et pharmacocinétiques de médicaments déjà approuvés ou en phase clinique. Les filtres les plus couramment acceptés correspondent aux règles de Lipinski (Ro5) [1], Veber [2], Oprea [3], Walters [4], Rhiston [5] ou la règle de toxicité de Pfizer [6]. Ces filtres sont appliqués afin d'éliminer des chimiothèques les composés qui ont peu de chances de conduire à des médicaments administrables par voie orale [6]. Ces règles s'appuient sur l'application de seuils simples pour certains descripteurs moléculaires standard, tels que la masse moléculaire (MW), le coefficient de partage octanol/eau

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

(log P) ou la surface topologique polaire (TPSA)¹ [44] (→). Plus récemment, des estimations plus quantitatives ont été développées afin d'éviter l'utilisation de seuils binaires [7]. Par ailleurs, les composés connus pour interagir avec les essais biologiques ou engendrer un fort taux de faux positifs sont, soit retirés des collections, soit, au minimum, annotés [8].

Les stratégies de criblage basées sur l'utilisation de collections de plusieurs centaines de milliers de composés impliquent des coûts énormes et des ressources importantes, qui limitent leur application à une poignée de centres de criblage dans le monde. Par ailleurs, les taux de succès de cette approche globale sont plus faibles que prévus, surtout pour les cibles non traditionnelles [9]. Une alternative de plus en plus acceptée consiste à concevoir des chimiothèques focalisées [10, 11]. Une chimiothèque focalisée est une collection de molécules chimiques dédiées à une cible ou un ensemble de cibles (généralement protéiques) appartenant au même espace chimique ou biologique [44, 45]. Historiquement, des chimiothèques focalisées ont été conçues pour cibler les grandes classes de cibles thérapeutiques, telles que les récepteurs couplés aux protéines G (GPCR), les kinases, les récepteurs nucléaires, ou encore les protéases. Plus récemment, des chimiothèques focalisées ont été développées contre des classes émergentes de cibles thérapeutiques, comme les interactions protéine-protéine (PPI). Les principales stratégies permettant de développer des chimiothèques focalisées reposent sur l'utilisation des propriétés de composés actifs déjà connus pour leur action sur la cible en question (approche dite *ligand-based*), ou sur la connaissance de la structure 3D de la cible (approche dite *structure-based*) grâce, notamment, aux progrès des méthodes de *docking* [12-15, 45] (→).

(→) Voir la Synthèse d'I. Krimm, m/s n°2, février 2015, page 197

(→) Voir la Synthèse de Y.S Wong, m/s n° 1, janvier 2015, page 93

Les chimiothèques ciblant les interactions protéine-protéine

Il existe actuellement moins de 500 cibles thérapeutiques recouvrant environ une dizaine de familles de protéines [16, 17]. Par ailleurs, l'ensemble des médicaments (en ne considérant ici que les petites molécules chimiques) disponibles sur le marché ne représentent qu'un nombre limité de molécules différentes (environ 1 200 composés). L'identification de nouvelles cibles thérapeutiques, ainsi que celle de molécules associées permettant de moduler leur activité, constitue donc un enjeu majeur. Parmi celles-ci, la modulation des réseaux d'interactions protéine-protéine représente une voie thérapeutique nouvelle très prometteuse. En effet, certaines protéines sont au cœur de réseaux d'interactions qui jouent un rôle fondamental dans de nombreuses fonctions cellulaires, telles que le développement, la prolifération ou la mort cellulaires. Le dysfonctionnement de ces réseaux est impliqué dans le développement de certaines maladies et,

notamment, dans différents types de cancer. Le nombre de protéines intervenant dans le réseau d'interactions protéine-protéine, chez l'homme, est estimé à plusieurs centaines de milliers, ce qui représente ainsi un important réservoir de cibles thérapeutiques potentielles [18, 19]. La modulation des interactions protéine-protéine suscite donc un engouement croissant, bien que celles-ci soient considérées comme des cibles particulièrement difficiles, en raison des caractéristiques de leur interface et de leur diversité structurale [20]. Plusieurs centaines d'inhibiteurs ont été développés à ce jour pour une cinquantaine de cibles, et plusieurs bases de données leur sont consacrées [21-25]. L'analyse des propriétés des interfaces protéiques a conduit au développement de fonctions de score spécifiques pour évaluer leur « druggabilité » (capacité à lier une petite molécule chimique) et, ainsi, faciliter la sélection de nouvelles cibles candidates [26, 27]. Par ailleurs, l'analyse des caractéristiques physicochimiques des inhibiteurs a permis de définir un profil caractéristique d'un inhibiteur « type » d'interactions protéine-protéine [28-31].

Malgré le nombre croissant de modulateurs développés, le taux de succès des campagnes de criblage contre des interactions protéine-protéine reste généralement faible, principalement en raison de l'inadéquation des chimiothèques utilisées. Un effort général a donc été entrepris en milieu académique, dans l'industrie pharmaceutique et chez les fournisseurs de molécules, afin de développer des collections de molécules appartenant à d'autres régions de l'espace chimique et dédiées à ce type de cibles. Les premiers efforts pour développer des molécules capables de perturber les interactions protéine-protéine reposaient sur la conception de mimés d'éléments de structures secondaires présents à l'interface, tels que les hélices α , les brins β , les coudes ou les poly-prolines [32-34]. Ces peptidomimétiques constituent une excellente source de composés pour la construction de chimiothèques focalisées. Toutefois, cette classe de composés se limite essentiellement aux interactions protéine-peptide et ne peut être utilisée pour des interfaces plus globulaires. Au cours de la dernière décennie, un grand nombre de composés non peptidomimétiques capables de perturber des complexes protéine-protéine ont été développés. Plusieurs études ont été entreprises pour caractériser les propriétés physicochimiques de ces inhibiteurs, en les comparant à d'autres classes de molécules, telles que les médicaments [28-30, 35].

En moyenne, les modulateurs d'interactions protéine-protéine sont des composés de relativement haute masse moléculaire, hydrophobes, rigides, plus rare-

¹ « La surface topologique polaire correspond à la surface des atomes polaires (azote, oxygène, etc.) ; le coefficient de partage octanol/eau correspond au logarithme du rapport des concentrations de la molécule étudiée dans l'octanol et dans l'eau, $\log P = \log(C_{\text{oct}}/C_{\text{eau}})$ » (voir [44]).

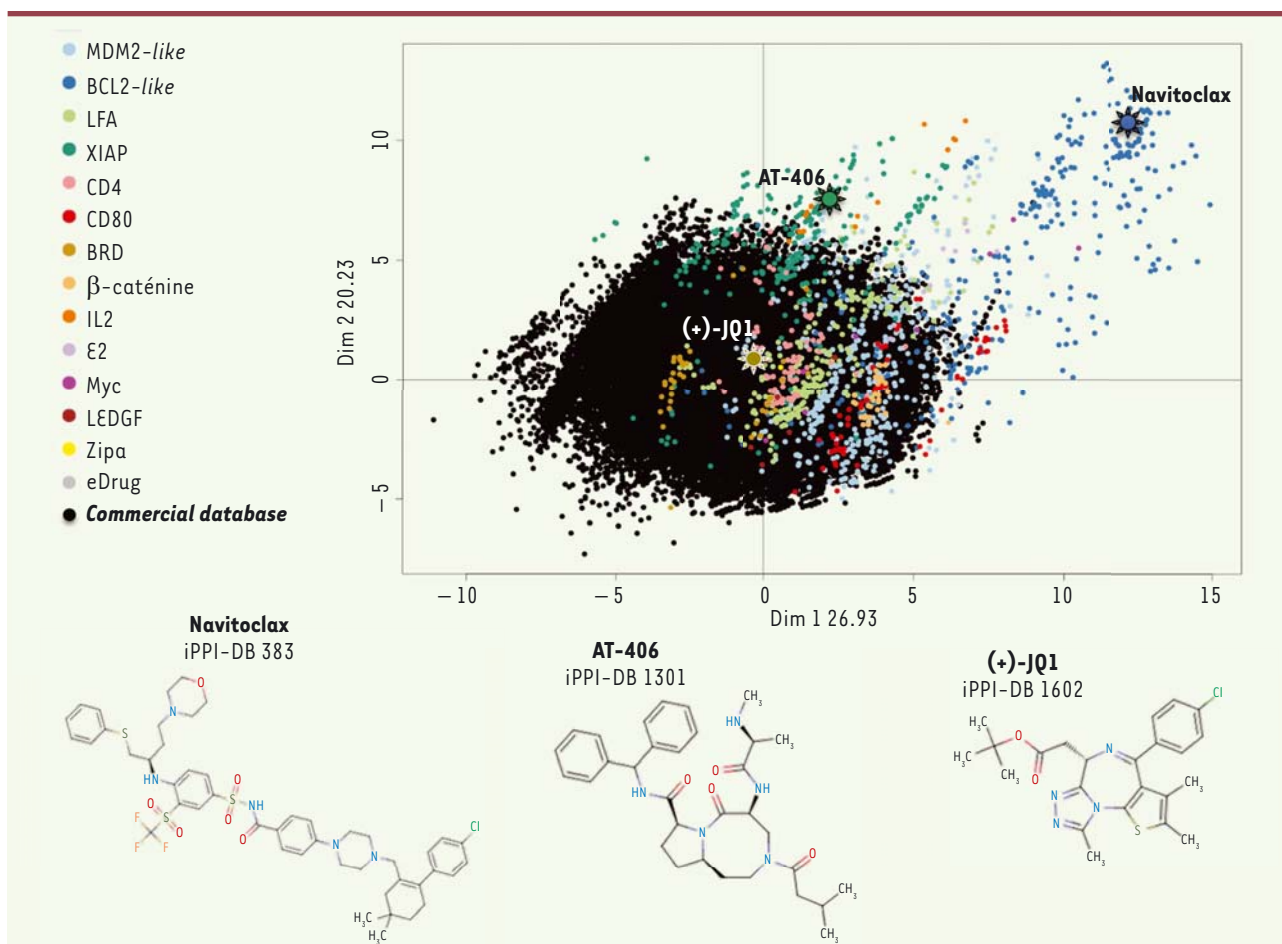


Figure 1. Visualisation de l'espace chimique des inhibiteurs de cibles PPI. Cette visualisation est faite par analyse en composante principale des propriétés physicochimiques des composés de l'iPPI-DB (<http://www.ippidb.cdithem.fr/>). Le décalage (en haut à droite) dans la position des composés de iPPI-DB (points de couleur) par rapport aux composés de la chimiothèque commerciale témoin (points noirs) illustre clairement l'inadaptation de ce type de chimiothèque pour identifier des inhibiteurs de cibles PPI. BCL2 : *B-cell lymphoma 2* ; BRD : *bromodomain-containing protein* ; IL2 : *interleukine 2* ; LEDGF : *lens epithelium-derived growth factor* ; LFA : *lymphocyte function-associated antigen* ; MDM2 : *mouse double minute 2* ; XIAP : *X-linked inhibitor of apoptosis*.

ment linéaires, et contenant plusieurs noyaux aromatiques, qui les différencient des médicaments standard ou des inhibiteurs d'enzyme [23, 28-31, 34]. Ces propriétés caractéristiques ont été utilisées pour assister la conception de chimiothèques dédiées aux interactions protéine-protéine, en extrayant des inhibiteurs potentiels des collections de criblage à l'aide de méthodes d'apprentissage, telles que les arbres de décision [29, 30, 36] ou les machines à vecteur support² [37]. Deux exemples de développements académiques sont présentés ci-après.

² « La classification consiste à classer des individus en fonction de certaines de leurs caractéristiques. Il existe différents types de classification, mais un des plus intuitifs et des plus utilisés est la classification supervisée. L'idée de la classification supervisée est d'apprendre une règle de classement à partir d'un ensemble de données dont le classement est déjà connu. Une fois la règle apprise, il est possible de l'appliquer pour catégoriser de nouvelles données, dont le classement est inconnu. Les machines à vecteurs de support, ou SVM (Support Vector Machines), sont une technique relativement récente (elles ont été introduites en 1992) de classification supervisée qui suscite beaucoup d'intérêt pour ses bonnes performances dans un large éventail d'applications pratiques ». (Dominik Françoer, Université de Sheerbroke).

PPI (protein-protein interaction)-HitProfiler

En 2009, une étude originale portant sur des composés modulateurs d'interactions protéine-protéine et sur leur comparaison à des médicaments agissant sur des cibles conventionnelles, a permis de déterminer, quantitativement chez les premiers, un profil physicochimique caractéristique [28, 29]. En effet, il est ressorti de cette analyse que ces composés se distinguent, entre autres, par des masses molaires moyennes sensiblement plus élevées (400-500 g \cdot mol⁻¹), une hydrophobicité plus importante (log P d'environ 3,5-4,5), ainsi qu'un plus grand nombre de cycles aromatiques (environ 3-4). Autant de critères quantitatifs pouvant être utilisés pour sélectionner les composés d'une chimiothèque. En outre, l'utilisation de ces données a aussi permis de concevoir

Fournisseur	Chimiothèque	Nombre de composés	Lien
Asinex	PPI	11 177	http://www.asinex.com/PPI_Library.html
ChemDiv	PPI	125 418	http://us.chemdiv.com/index.php?option=com_content&view=article&id=85&Itemid=184
ChemDiv	Eccentric	6 684	http://us.chemdiv.com/index.php?option=com_content&view=article&id=85&Itemid=184
Life chemicals	Similarity	23 532	http://www.lifechemicals.com/downloads/9356/13062/13074
Life chemicals	Machine learning	869	http://www.lifechemicals.com/downloads/9356/13062/13074
Life chemicals	Ro4	4 364	http://www.lifechemicals.com/downloads/9356/13062/13074
Otava	TreeTM	1 332	http://www.otavachemicals.com/products/target-focused-libraries/protein-protein-interaction
Otava	AnalogsTM	1 027	http://www.otavachemicals.com/products/target-focused-libraries/protein-protein-interaction
Otava	BRD4	373	http://www.otavachemicals.com/targets/bromodomain-containing-protein-4-brd4-focused-library

Tableau I. Chimiothèques commerciales dédiées aux interactions protéine-protéine.

des modèles statistiques, basés sur des arbres de décision, faisant ressortir deux propriétés supplémentaires caractéristiques de 80 % de ces modulateurs. La première décrit leur forme tridimensionnelle (descripteur RDF070m) : elle est caractérisée par des structures préférentiellement en étoile, en forme de T, ou de L. La seconde implique la nécessité, pour la plupart de ces composés, de posséder un nombre critique de 15 liaisons multiples, et plus spécifiquement aromatiques. La construction d'un tel modèle a permis le développement d'un logiciel librement accessible, PPI-HitProfiler (<http://www.cdithem.fr/getPPIHitProfiler.php>), dont la fonction consiste à sélectionner, parmi les composés d'une chimiothèque, ceux ayant le plus fort potentiel pour moduler une interaction protéine-protéine. PPI-HitProfiler peut, en outre, être utilisé directement au sein de l'outil en ligne FAF-drugs2 (<http://fafdrugs2.mti.univ-paris-diderot.fr>) [38, 39], qui permet plus globalement de filtrer toute chimiothèque selon de nombreux critères, notamment d'ADME-Tox (absorption, distribution, métabolisme, excrétion, toxicité). La validation de PPI-HitProfiler a été réalisée sur les données expérimentales de criblage de 11 cibles PPI (*protein-protein interaction*) différentes et un nombre cumulé de 500 000 molécules chimiques. Ce logiciel démontre, en outre, des sensibilités tout à fait comparables sur les composés d'iPPI-DB (*non-peptidic inhibitor-data base*) (<http://www.ippidb.cdithem.fr>) [21], une base de données de modulateurs d'interactions protéine-protéine récemment développée et qui contient 1 650 composés pour 13 familles de cibles PPI avec leur données pharmacologiques. En effet, près de 92 % des composés d'iPPI-DB sont correctement prédits comme modulateurs d'interactions protéine-protéine par PPI-HitProfiler.

Approche 2P2I

2P2I-database (2P2I-DB) est une base de données structurales annotée manuellement, consacrée aux inhibiteurs orthostériques (qui se lient au niveau de l'interface), ainsi qu'aux complexes protéine-protéine et protéine-inhibiteur dont la structure 3D a été résolue expérimentalement [24, 25]. Une telle base de données, disponible en ligne (<http://2p2idb.cnrs-mrs.fr/>), a l'avantage de ne proposer que des inhibiteurs dont la structure, en présence de la cible, est connue, permettant d'analyser les propriétés des interfaces et des inhibiteurs. La base de données comporte actuellement plus de 250 modulateurs, correspondant à 17 familles de cibles biologiques et 29 complexes protéine-protéine (version 2015). Les interfaces des complexes protéine-protéine et protéine-ligand ont été analysées afin de définir les paramètres structuraux qui gouvernent l'inhibition des complexes protéine-protéine par des petites molécules [25, 26]. Par ailleurs, les propriétés des modulateurs orthostériques ont été analysées et comparées à celles d'autres classes de molécules [31, 37]. Sur la base de cette comparaison, nous avons développé 2P2I_{HUNTER}, un outil permettant de filtrer des modulateurs d'interaction protéine-protéine orthostériques à partir de larges collections de molécules [37]. L'algorithme, basé sur une méthode d'apprentissage de type machine à vecteur support (SVM), a été appliqué à un ensemble de

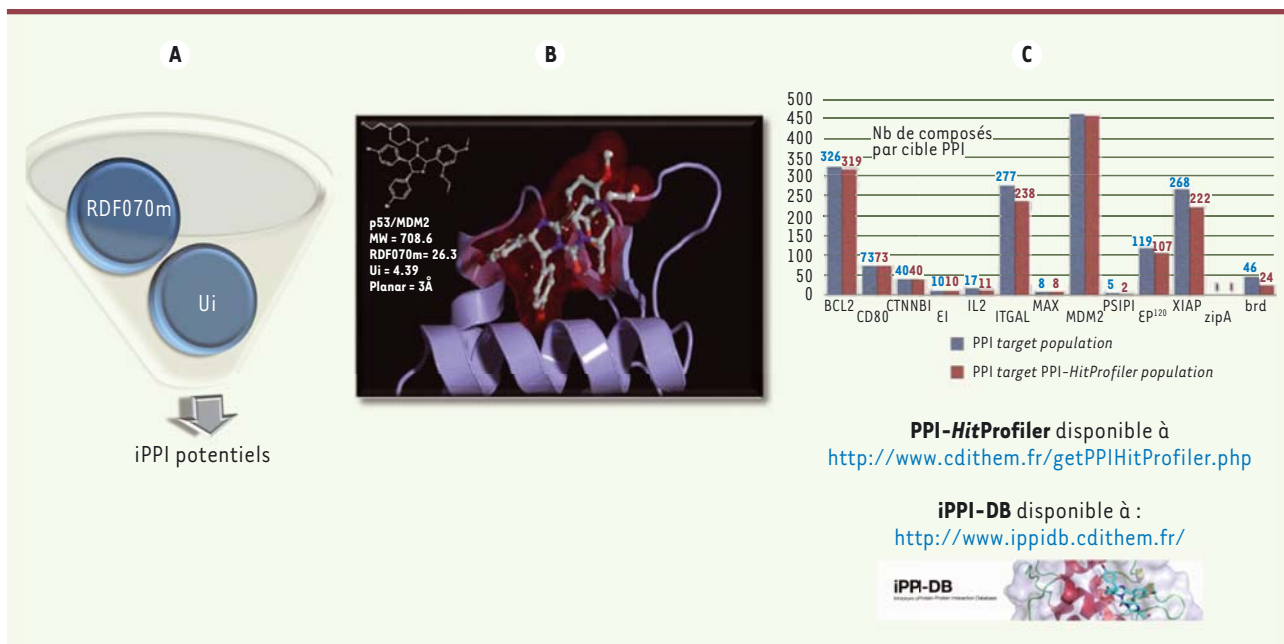


Figure 2. PPI-Hit Profiler. **A.** PPI-HitProfiler est un logiciel gratuit pour prédire le potentiel de toute molécule à être un inhibiteur de cible PPI. Il est basé sur un modèle statistique (arbre de décision) reposant sur deux propriétés physicochimiques. La première décrit la forme 3D spécifique des inhibiteurs de cibles PPI (iPPI), illustrée par des valeurs plus hautes du descripteur RDF070m qui favorisent les structures chimiques ramifiées (en forme de T, de L ou surtout d'étoile). La deuxième décrit le nombre de liaisons chimiques insaturées (exemple, aromatiques) illustrée par le descripteur Ui. **B.** Exemple d'inhibiteur de cible PPI : nutline-2 cocrystallisée avec MDM2 (Code PDB 1RV1). La forme en étoile de cette molécule, ainsi que la présence dans sa structure de 18 liaisons insaturées, illustrent les propriétés privilégiées de ce type d'inhibiteur pour moduler leur cible PPI respective. **C.** Validation de PPI-HitProfiler sur les données de l'iPPI-DB. PPI-HitProfiler prédit correctement 92 % des composés de l'iPPI-DB avec une bonne sensibilité par cible distincte PPI pour la quasi-totalité des cibles.

8,3 millions de composés provenant des principaux fournisseurs de molécules chimiques [40]. La chimiothèque résultante, composée de 140 000 composés, a été filtrée avec plusieurs outils permettant la sélection de structures privilégiées retrouvées dans les ligands de nombreuses cibles biologiques, ainsi que les composés les plus « tridimensionnels » [41, 46]. Enfin, la chimiothèque de diversité 2P21_{3D}, constituée de 1 664 composés, a été élaborée. Les composés ont été achetés, mis en plaque et validés *in vitro* sur plusieurs types d'interactions protéine-protéine structurellement diverses, incluant des domaines PDZ (*postsynaptic density protein*, *Drosophila disc large tumor suppressor*, and *zonula occludens-1 protein*), des bromodomaines, des interfaces contenant des hélices α (P53) conduisant à des pourcentages de succès relativement élevés en comparaison de criblages avec des chimiothèques commerciales standard (Milhas *et al.*, en cours de rédaction).

Chimiothèques commerciales

Plusieurs chimiothèques commerciales, dédiées aux interactions protéine-protéine et constituées de quelques centaines à plus de 100 000 composés, sont disponibles chez différents fournisseurs (Tableau 1). Ces chimiothèques ont été assemblées en utilisant des méthodes d'apprentissage telles que définies dans les paragraphes

précédents, des règles simples dérivées de la caractérisation des inhibiteurs d'interactions protéine-protéine, la recherche de composés similaires dans les collections de criblage (*scaffold hopping*) ou par synthèse orientée (Tableau 1). Les principales chimiothèques sont présentées brièvement :

- Asinex a conçu un ensemble de 111 177 composés dédiés aux interactions protéine-protéine en portant une attention particulière aux propriétés pharmacocinétiques et à la solubilité des composés. En se basant sur les propriétés des interfaces protéine-protéine, les chercheurs ont élaboré une stratégie reposant sur la synthèse de composés possédant un cœur hydrophile riche en accepteurs et donneurs de liaisons hydrogène, et des groupements lipophiles spécifiques des interactions protéine-protéine à la périphérie. Cette librairie focalisée a été en partie créée en utilisant les résultats de PPI-HitProfiler et, notamment, le descripteur RDF070m pour mettre l'accent sur la sélection finale des composés.
- Chemdiv propose un ensemble de chimiothèques dédiées à différents sous-espaces chimiques ou biologiques en fonction de la nature des cibles (MDM2 [*mouse double minute 2*], PDZ et CD16A), de la

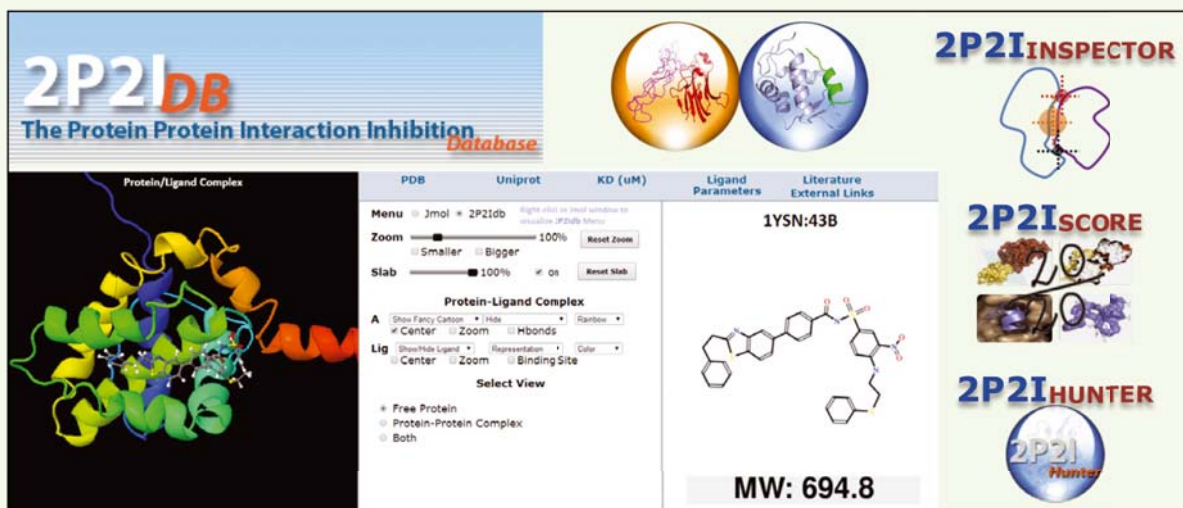


Figure 3. La base de données 2P2I-DB et ses outils associés. La base de données structurale 2P2I-DB (<http://2p2idb.cnrs-mrs.fr/index.html>) recense toutes les familles pour lesquelles la structure 3D des complexes protéine-protéine et protéine-inhibiteur a été caractérisée expérimentalement. Pour chaque famille, de nombreuses informations sont disponibles, telles que les caractéristiques des interfaces ou des inhibiteurs, ainsi que des liens vers d'autres sites d'intérêt (UniProt, PubMed, PDBsum, PDB, ChemSpider, etc.). Les structures peuvent être visualisées et manipulées de façon interactive. Différents outils sont disponibles afin d'analyser les propriétés des interfaces protéine-protéine (2P2I_{INSPECTOR}) ou de prédire leur « druggabilité » (2P2I_{SCORE}). Les propriétés des inhibiteurs, ainsi que les jeux de données utilisés pour développer l'algorithme d'apprentissage qui a permis de construire la chimiothèque focalisée 2P2I_{3D}, sont également disponibles (2P2I_{HUNTER}).

structure des composés (composés cycliques provenant de réactions d'Ugi³, dérivés spiro ou hétérocycles non conventionnels) ou du type de structure secondaire à inhiber (hélices α , brins β , β -turns, ou boucles). La chimiothèque globale est composée de 125 418 molécules qui ont, entre autres, été sélectionnées en prenant en compte leur structure tridimensionnelle [41, 42].

- Life chemicals propose trois chimiothèques dédiées aux interactions protéine-protéine. Une chimiothèque de similitude (23 532 composés) a été construite en recherchant dans la collection complète de molécules, des composés dont la structure est proche (Tanimoto 0,85) de celle d'inhibiteurs présents dans la base de données TIMBAL [22, 23]. Une chimiothèque de 4 364 composés a été obtenue par application sur la collection totale du filtre « Ro4 », tel que défini dans l'approche 2P2I précédemment décrite [31]. La 3^e librairie (869 composés) a été construite en utilisant l'approche *HitProfiler* [29, 30]. Deux paramètres (RDF070m et le degré d'insaturation U_i) ont été utilisés pour sélectionner les composés. De plus, seuls les composés dont la masse moléculaire est inférieure à 400 $\text{g}\cdot\text{mol}^{-1}$ ont été conservés dans la chimiothèque finale.

- Otava propose trois collections de molécules dédiées aux interactions protéine-protéine. La chimiothèque « Tree™ » (1 332 composés) a été obtenue par l'application d'un arbre de décision utilisant des

descripteurs de forme et la présence de fonctions ester [36]. La chimiothèque « Analogs™ » (1 027 composés) a été construite en recherchant dans la collection totale d'Otava, des composés similaires (ECFP4 [extended connectivity fingerprint with diameter of four bonds] fingerprints et Tanimoto 0,40) aux inhibiteurs présents dans la base de données TIMBAL [22, 23]. Ces deux chimiothèques ont ensuite été filtrées en utilisant une version modifiée de la règle de Lipinski dans laquelle la masse moléculaire autorisée est entre 300 and 700 $\text{g}\cdot\text{mol}^{-1}$, et log P entre 1 et 6. De plus, les composés possédant des fonctions réactives, ainsi que ceux connus comme induisant un fort taux de faux positifs, ont été retirés. La 3^e chimiothèque composée de 373 molécules, est dédiée aux inhibiteurs de bromodomaines (famille BRD4-1) et a été construite par *docking* de différents ligands dans le site de liaison du domaine 1 de la protéine BRD4 humaine, en se basant sur la structure cristallographique en présence de l'inhibiteur 3,5-diméthylisoxazol (PDB: 4J0S).

Perspectives

Les interactions protéine-protéine sont responsables d'un grand nombre de fonctions vitales mais elles sont aussi

³ La réaction d'Ugi fait intervenir 4 composants : un aldéhyde, une amine, un acide carboxylique et un isocyanide et permet la préparation rapide de dérivés α -aminoacyl amides.

impliquées dans de multiples pathologies. Certaines de ces interactions importantes pour la santé humaine ont déjà été ciblées avec succès par des produits biologiques. Toutefois, ces molécules ne sont pas adaptées à toutes les maladies et leur administration se fait généralement par voie parentérale. En revanche, cibler les interactions protéine-protéine avec des petites molécules chimiques permet d'envisager un large éventail d'indications. Actuellement, 15 à 20 inhibiteurs d'interactions protéine-protéine sont en développement clinique et le produit des ventes mondiales associées est estimé autour de 800 millions de dollars par an à partir de 2018 (*EvaluatePharma database : consensus analyst forecast of drug sales for 2018*). À plus long terme, les inhibiteurs d'interactions protéine-protéine pourraient prendre une part substantielle de marché dans des indications médicales (maladies auto-immunes ou en oncologie) actuellement traitées avec des produits biologiques. Ces innovations seront possibles rapidement et à moindre coût, si collectivement nous enrichissons nos connaissances des interactions protéine-protéine et si nous arrivons à faire de la conception rationnelle de chimiothèques dédiées à ce type de cibles et de mécanismes moléculaires. Des bases de données annotées ont été développées dans ce but, et des premiers filtres biostatistiques prometteurs permettant de préparer des chimiothèques dédiées ont été publiés. Il est nécessaire à présent d'augmenter l'efficacité de ces filtres, de concevoir des filtres spécifiques de grandes familles d'interactions, et de proposer des filtres optimisés opérant sous contraintes ADME-Tox (absorption, distribution, métabolisme, excretion, toxicité).

Plus spécifiquement, le développement de bases de données manuellement annotées comme celles de la 2P2I-DB ou de l'iPPI-DB, avec des stratégies différentes, la première se focalisant sur les données cristallographiques, l'autre sur les données pharmacologiques, offre une réelle opportunité de faciliter l'identification de nouvelles molécules « touches » de qualité. En effet, la disponibilité de telles données facilite la mise au point de nouveaux modèles prédictifs et, plus largement, la rationalisation des propriétés qui caractérisent l'efficacité avec laquelle certaines molécules peuvent lier des interfaces aussi complexes que celles des interactions protéine-protéine. Notamment, l'analyse des données combinées de la 2P2I-DB et celles de l'iPPI-DB dans une étude récente (en cours de publication) a permis de décrire les caractéristiques tridimensionnelles des modulateurs d'interactions protéine-protéine et, plus particulièrement, de décrire leur capacité à lier efficacement les interfaces souvent hydrophobes et plates des interactions protéine-protéine.

Ces études devraient permettre d'identifier, dans un futur proche, les propriétés majeures des modulateurs d'interaction protéine-protéine et, ainsi, de comprendre ce qui les caractérise. La prise en compte de spécificités pour quelques grands groupes d'interactions sera, en outre, un atout important pour gagner en spécificité. Ce gain de connaissance sera la clé dans la génération de chimiothèques focalisées enrichies en modulateurs potentiels d'interaction protéine-protéine. ♦

SUMMARY

Chemical libraries dedicated to protein-protein interactions

The identification of complete networks of protein-protein interactions (PPI) within a cell has contributed to major break-

throughs in understanding biological pathways, host-pathogen interactions and cancer development. As a consequence, PPI have emerged as a new class of promising therapeutic targets. However, they are still considered as a challenging class of targets for drug discovery programs. Recent successes have allowed the characterization of structural and physicochemical properties of protein-protein interfaces leading to a better understanding of how they can be disrupted with small molecule compounds. In addition, characterization of the profiles of PPI inhibitors has allowed the development of PPI-focused libraries. In this review, we present the current efforts at developing chemical libraries dedicated to these innovative targets. ♦

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Lipinski C, Lombardo F, Dominy B, Feeney P. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 2001 ; 46 : 3-26.
2. Veber DF, Johnson SR, Cheng HY, et al. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem* 2002 ; 45 : 2615-23.
3. Oprea TI. Property distribution of drug-related chemical databases. *J Comput Aided Mol Des* 2000 ; 14 : 251-64.
4. Walters WP, Murcko MA. Prediction of drug-likeness. *Adv Drug Deliv Rev* 2002 ; 54 : 255-71.
5. Rishton GM. Nonleadlikeness and leadlikeness in biochemical screening. *Drug Discov Today* 2003 ; 8 : 86-96.
6. Hughes JD, Blagg J, Price DA, et al. Physicochemical drug properties associated with *in vivo* toxicological outcomes. *Bioorg Med Chem Lett* 2008 ; 18 : 4872-5.
7. Bickerton GR, Paolini GV, Besnard J, et al. Quantifying the chemical beauty of drugs. *Nat Chem* 2012 ; 4 : 90-8.
8. Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem* 2010 ; 53 : 2719-40.
9. Gupta S. New drug development. In : *Drug discovery and clinical research*. Delhi, India : JayPee Brothers Medical Publishers Ltd, 2011 : 1-135.
10. Sheppard DW, Lipkin MJ, Harris CJ, et al. Strategies for small molecule library design. *Curr Pharm Des* 2013 ; 19 : 1-9.
11. Lipkin MJ, Stevens AP, Livingstone DJ, Harris CJ. How large does a compound screening collection need to be? *Comb Chem High Throughput Screen* 2008 ; 11 : 482-93.
12. Pirhadi S, Shiri F, Ghasemi JB. Methods and applications of structure based pharmacophores in drug discovery. *Curr Top Med Chem* 2013 ; 13 : 1036-47.
13. AbdulHameed MD, Chaudhury S, Singh N, et al. Exploring polypharmacology using a ROCS-based target fishing approach. *J Chem Inf Model* 2012 ; 52 : 492-505.
14. Cheng T, Li Q, Zhou Z, et al. Structure-based virtual screening for drug discovery: a problem-centric review. *AAPS J* 2012 ; 14 : 133-41.
15. Ripphausen P, Nisius B, Bajorath J. State-of-the-art in ligand-based virtual screening. *Drug Discov Today* 2011 ; 16 : 372-6.
16. Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? *Nat Rev Drug Discov* 2006 ; 5 : 993-6.
17. Rask-Andersen M, Almén MS, Schiöth HB. Trends in the exploitation of novel drug targets. *Nat Rev Drug Discov* 2011 ; 10 : 579-90.
18. Venkatesan K, Rual J, Vazquez A, et al. An empirical framework for binary interactome mapping. *Nat Methods* 2009 ; 6 : 83-90.
19. Stumpf M, Thorne T, de Silva E, et al. Estimating the size of the human interactome. *Proc Natl Acad Sci USA* 2008 ; 105 : 6959-64.

RÉFÉRENCES

20. Mullard A. Protein-protein interaction inhibitors get into the groove. *Nat Rev Drug Discov* 2012 ; 11 : 173-5.
21. Labbé CM, Laconde G, Kuenemann MA, et al. iPPI-DB: a manually curated and interactive database of small non-peptide inhibitors of protein-protein interactions. *Drug Discov Today* 2013 ; 18 : 958-68.
22. Higuero AP, Jubb H, Blundell TL. TIMBAL v2: update of a database holding small molecules modulating protein-protein interactions. *Database (Oxford)* 2013 ; 2013 : bat039.
23. Higuero AP, Schreyer A, Bickerton GRJ, et al. Atomic interactions and profile of small molecules disrupting protein-protein interfaces: the TIMBAL database. *Chem Biol Drug Design* 2009 ; 74 : 457-67.
24. Basse MJ, Betzi S, Bourgeois R, et al. 2P21db: a structural database dedicated to orthosteric modulation of protein-protein interactions. *Nucleic Acids Res* 2013 ; 41 : D824-7.
25. Bourgeois R, Basse MJ, Morelli X, Roche P. Atomic analysis of protein-protein interfaces with known inhibitors: the 2P21 database. *PLoS One* 2010 ; 5 : e9598.
26. Hamon V, Morelli X. Druggability of protein-protein interactions. In : *Understanding and exploiting protein-protein interactions as drug target*. Future Science Ltd, 2013 : 18-31.
27. Pérot S, Regad L, Reynès C, et al. Insights into an original pocket-ligand pair classification: a promising tool for ligand profile prediction. *PLoS One* 2013 ; 8 : e63730.
28. Pagliaro L, Felding J, Audouze K, et al. Emerging classes of protein-protein interaction inhibitors and new tools for their development. *Curr Opin Chem Biol* 2004 ; 8 : 442-9.
29. Reynes C, Host H, Camproux AC, et al. Designing focused chemical libraries enriched in protein-protein interaction inhibitors using machine-learning methods. *PLoS Comput Biol* 2010 ; 6 : e1000695.
30. Sperandio O, Reynès C, Camproux A, Villoutreix B. Rationalizing the chemical space of protein-protein interaction inhibitors. *Drug Discov Today* 2010 ; 15 : 220-9.
31. Morelli X, Bourgeois R, Roche P. Chemical and structural lessons from recent successes in protein-protein interaction inhibition (2P21). *Curr Opin Chem Biol* 2011 ; 15 : 475-81.
32. Akram ON, Degraff DJ, Sheehan JH, et al. Tailoring peptidomimetics for targeting protein-protein interactions. *Mol Cancer Res* 2014 ; 12 : 967-78.
33. Jayatunga MK, Thompson S, Hamilton AD. α -Helix mimetics: outwards and upwards. *Bioorg Med Chem Lett* 2014 ; 24 : 717-24.
34. Isvoran A, Craciun D, Martiny V, et al. Computational analysis of protein-protein interfaces involving an alpha helix: insights for terphenyl-like molecules binding. *BMC Pharmacol Toxicol* 2013 ; 14 : 31.
35. Villoutreix BO, Labbé CM, Lagorce D, et al. A leap into the chemical space of protein-protein interaction inhibitors. *Curr Pharm Des* 2012 ; 18 : 4648-67.
36. Neugebauer A, Hartmann RW, Klein CD. Prediction of protein-protein interaction inhibitors by chemoinformatics and machine learning methods. *J Med Chem* 2007 ; 50 : 4665-8.
37. Hamon V, Bourgeois R, Ducrot P, et al. 2P21 HUNTER: a tool for filtering orthosteric protein-protein interaction modulators via a dedicated support vector machine. *J R Soc Interface* 2014 ; 11 : 20130860.
38. Lagorce D, Sperandio O, Galons H, et al. FAF-Drugs2: free ADME/tox filtering tool to assist drug discovery and chemical biology projects. *BMC Bioinformatics* 2008 ; 9 : 396.
39. Lagorce D, Maupetit J, Baell J, et al. The FAF-Drugs2 server: a multistep engine to prepare electronic chemical compound collections. *Bioinformatics* 2011 ; 27 : 2018-20.
40. Hamon V, Brunel JM, Combes S, et al. 2P21chem: focused chemical libraries dedicated to orthosteric modulation of protein-protein interactions. *Med Chem Comm* 2013 ; 4 : 797-809.
41. Fry DC, So SS. Modulators of protein-protein interactions: importance of three-dimensionality. In : *Protein-protein interactions in drug discovery*. Wiley-VCH Verlag GmbH and Co KGaA, 2013 : 55-62.
42. Lovering F, Bikker J, Humblet C. Escape from flatland: increasing saturation as an approach to improving clinical success. *J Med Chem* 2009 ; 52 : 6752-6.
43. Rognan D, Bonnet P. Les chimiothèques et le criblage virtuel. *Med Sci (Paris)* 2014 ; 30 : 1152-1160.
44. Krimm I. le criblage de fragments : une voie prometteuse pour la conception de médicaments. *Med Sci (Paris)* 2015 ; 31 : 197-202.
45. Wong YS. Synthèse orientée vers la diversité structurale pour explorer le vivant. *Med Sci (Paris)* 2015 ; 31 : 93-7.
46. Kuenemann MA, Bourbon LM, Labbé CM, et al. Which three-dimensional characteristics make efficient inhibitors of protein-protein interactions? *J Chem Inf Model* 2014 ; 54 : 3067-79.

TIRÉS À PART
P. Roche

**3 JOURNÉES
QUI FONT GAGNER
DES ANNÉES !**

31 mars > 2 avril 2015
PARIS EXPO / PORTE DE VERSAILLES / HALL 4

Demandez votre BADGE GRATUIT :
www.forumlabo.com

Forum
LABO
BIOTECH

Le salon des fournisseurs
de matériels et services
pour le Laboratoire

Made by
GVL
events

UNE MANIFESTATION DU
GVL

