
RAPPORT DE CONJONCTURE

DU COMITÉ NATIONAL DE LA RECHERCHE SCIENTIFIQUE

ÉDITION 2014

Extrait



CNRS ÉDITIONS

15, rue Malebranche – 75005 Paris

SECTION 20

BIOLOGIE MOLÉCULAIRE ET STRUCTURALE, BIOCHIMIE

Extrait de la déclaration adoptée par le Comité national de la recherche scientifique réuni en session plénière extraordinaire le 11 juin 2014

La recherche est indispensable au développement des connaissances, au dynamisme économique ainsi qu'à l'entretien de l'esprit critique et démocratique. La pérennité des emplois scientifiques est indispensable à la liberté et la fécondité de la recherche. Le Comité national de la recherche scientifique rassemble tous les personnels de la recherche publique (chercheurs, enseignants-chercheurs, ingénieurs et techniciens). Ses membres, réunis en session plénière extraordinaire, demandent de toute urgence un plan pluriannuel ambitieux pour l'emploi scientifique. Ils affirment que la réduction continue de l'emploi scientifique est le résultat de choix politiques et non une conséquence de la conjoncture économique.

L'emploi scientifique est l'investissement d'avenir par excellence

Conserver en l'état le budget de l'enseignement supérieur et de la recherche revient à prolonger son déclin. Stabiliser les effectifs ne suffirait pas non plus à redynamiser la recherche : il faut envoyer un signe fort aux jeunes qui intègrent aujourd'hui l'enseignement supérieur en leur donnant les moyens et l'envie de faire de la recherche. On ne peut pas sacrifier les milliers de jeunes sans statut qui font la recherche d'aujourd'hui. Il faut de toute urgence résorber la précarité. Cela suppose la création, sur plusieurs années, de plusieurs milliers de postes supplémentaires dans le service public ainsi qu'une vraie politique d'incitation à l'emploi des docteurs dans le secteur privé, notamment industriel.

Composition de la section

Marie-Thérèse GIUDICI-ORTICONI (présidente de section); Valérie GEOFFROY (secrétaire scientifique); Rodolphe AUGER; Sandrine BOSCHI; Benoît CHENAIS; Alexandre DE BREVERN; Aude ÉCHALIER; Patrice GOUET; Anne HARDUIN-LEPERS; Jean-Michel JAULT; Fabrice LECLERC; Hugues LORTAT-JACOB; Roland MARQUET; Stéphane MASSOU; Gérard PEHAU-ARNAUDET; Bruno ROBERT; Isabelle SCHALK; Patrick TAUC; Emmanuel TETAUD; Carine TISNE-VICROBECK; Véronique TREZEGUET BUSQUET.

Résumé

Les travaux réalisés par les chercheurs de la section 20 ont pour objectif d'élucider les processus moléculaires qui régissent le vivant. Le caractère général de nos préoccupations et questionnements conceptuels et expérimentaux fait que la section 20 constitue un dénominateur commun à la plupart des domaines de la biologie et se situe au carrefour de nombreuses disciplines. Ainsi, développant et utilisant de nouveaux outils, mettant en place de nouveaux modèles d'étude, elle est interdisciplinaire par nature et joue un rôle majeur à l'interface de la biologie, chimie, physique, informatique... Présents dans tous les grands centres du CNRS sur le territoire, nos laboratoires couvrent de vastes champs d'étude : de l'homme aux micro-organismes, des applications environnementales aux développements thérapeutiques.

Introduction

Le développement des connaissances sur les systèmes biologiques passe nécessairement par le décryptage des processus moléculaires qui les régissent. Il s'agit là du cœur des activités de la section 20 qui a pour objectifs l'élucidation, l'analyse et la compréhension des aspects moléculaires, métaboliques et structuraux du fonctionnement du vivant. Les priorités se doivent d'évoluer avec les progrès réalisés en séquençage à haut débit ou encore dans le développement de nouvelles technologies qui permettent la résolution de structures de plus en plus complexes. Ainsi, la génomique structurale et intégrative se doit d'investir les aspects fonctionnels et métaboliques afin d'appréhender les défis de la biologie des systèmes, et d'ouvrir ainsi la porte à la biologie synthétique.

Les thèmes de recherche de la section 20 sont par essence interdisciplinaires. Ils trouvent écho aussi bien dans d'autres sections de l'INSB

(sections 21, 23 et 27 dans les domaines respectifs de la génomique, du modèle végétal et des virus) qu'à l'interface chimie biologie avec l'INC et dans l'environnement et les énergies renouvelables avec l'INEE. Notre ancrage moléculaire nous permet de développer ces différentes activités d'interface sans perdre notre identité. Parmi les grands axes de notre recherche nous pouvons citer : les bases moléculaires et structurales des protéines et des acides nucléiques, la catalyse, la transduction du signal et les modifications post-traductionnelles, la dynamique et la réactivité des mécanismes macromoléculaires et de leurs complexes, les approches biophysiques et en molécule unique ou encore le métabolisme bactérien. La restructuration des sections a permis de regrouper en section 20 toute la bioénergétique eucaryote et microbienne ainsi que les protéines membranaires. L'ensemble de ces thématiques, majeures pour appréhender le fonctionnement du vivant et développées par les chercheurs dans les laboratoires de la section 20, se retrouvent dans les grands centres du CNRS : des campus Paris, Strasbourg, Lyon-Grenoble, Marseille, Montpellier, Lille et dans une moindre mesure à Toulouse et Bordeaux. Il convient de souligner que le caractère intrinsèquement interdisciplinaire de nos approches et questionnements fait que nombre de chercheurs de la section 20 font partie de laboratoires pilotés par l'INC ou encore de laboratoires dont la section principale de rattachement est la section 21 ou la section 23.

Enfin, en parallèle à notre cœur de métier, la recherche fondamentale, les travaux réalisés dans les laboratoires de la section 20 peuvent trouver un développement dans le domaine de la bioéconomie tant par la compréhension à l'échelle moléculaire des pathologies que par des applications dans le domaine environnemental.

I. Dynamique, flexibilité, et interactions des protéines

A. Dynamique, flexibilité

La nature dynamique des protéines est une propriété intrinsèque ayant des conséquences majeures sur leur fonction. De nos jours, les protéines sont considérées comme des acteurs de réseaux fluctuant, avec une flexibilité locale et une propension à la plasticité structurale globale. De manière générale, les interactions protéine-ligand (protéines, acides nucléiques, lipides...) impliquées dans la transduction du signal, l'assemblage de machines macromoléculaires, la régulation allostérique et l'adaptation thermique des enzymes requièrent une flexibilité structurale. La compréhension détaillée de la dynamique des protéines, qui nécessite des informations sur une large gamme d'échelles de temps, a été rendue possible grâce au développement d'approches interdisciplinaires associant biologistes, chimistes, physiciens et bioinformaticiens. Le prochain défi sera le développement d'une vision intégrée de la structure des protéines, de leur dynamique et de leur fonction. L'impact des fluctuations thermodynamiques dues au solvant sera à prendre en compte. Cette intégration est indispensable à la compréhension des mécanismes biologiques à l'échelle atomique. Dans ce contexte, les enzymes constituent une classe attractive de protéines et elles sont utilisées comme des systèmes modèles pour comprendre les incidences des fluctuations des structures sur les fonctions biologiques. La flexibilité du site actif est actuellement considérée comme essentielle pour la réduction des barrières d'énergie libre et l'accélération de la réaction enzymatique bien que le lien entre la flexibilité et le «turn-over» n'ait pas été clairement établi. En revanche, le rôle de la flexibilité conformationnelle a été clairement établi pour l'accessibilité du site actif, la liaison et l'orientation des ligands et la libération de produits, la

stabilisation des intermédiaires et l'adaptation de l'environnement du site actif pour la catalyse. Une vision intégrée de la structure des enzymes, de leur dynamique et de leur fonction sera nécessaire au décryptage des effets allostériques et coopératifs mais aussi à la conception de catalyseurs plus efficaces. Avec l'importance croissante de la biotechnologie industrielle et les besoins en enzymes de transformation pour la production de produits chimiques ou de nouveaux médicaments, la conception rationnelle de nouveaux catalyseurs par ingénierie enzymatique constitue un défi majeur à relever. Toutefois, à l'heure actuelle, la flexibilité et les mouvements moléculaires ne sont pas encore suffisamment pris en compte. L'intégration de la composante dynamique dans les modèles moléculaires permettra d'augmenter les capacités de construction par ingénierie de nouvelles protéines dotées de nouvelles fonctions.

B. Interactions, complexes et machineries protéiques

Les interactions entre protéines régulent la quasi-totalité des processus essentiels à la vie des cellules, incluant la transduction du signal, les métabolismes, les catalyses et l'expression des gènes. Ces interactions de physico-chimie, temporalité et fonction extrêmement diverses, forment un réseau dynamique appelé «interactome». Dans ce domaine, les mécanismes mis en jeu au cours de l'assemblage des complexes protéiques restent encore peu décrits. Il sera donc important d'identifier les différents acteurs responsables des genèse, fonction, plasticité, dissociation de ces assemblages et de caractériser la dynamique de leur relation. La compréhension de l'organisation structurale de ce type de machinerie passe par la modélisation en intégrant des données de diverses approches expérimentales (microscopie électronique, RMN, cristallographie aux rayons X, protéomique, SAXS, etc.) afin de construire un ensemble de modèles structuraux répondant aux contraintes expérimentales générées. Plus

globalement, il est important de connaître qui interagit avec qui au niveau de la cellule et d'analyser les mécanismes moléculaires impliqués dans ces interactions. La mise en évidence récente d'un nombre élevé de protéines multifonctionnelles souligne l'importance de décrypter ces réseaux d'interactions et leurs bases moléculaires afin de mieux comprendre le fonctionnement des cellules et d'envisager un contrôle des voies métaboliques à des fins thérapeutiques ou environnementales.

La caractérisation du génome entier, du protéome ou du transcriptome ouvre la voie à des projets beaucoup plus ambitieux visant à comprendre au niveau moléculaire le fonctionnement global d'une cellule. Cela implique de pouvoir combiner différents paramètres cinétiques et/ou thermodynamiques avec des mesures à l'échelle du génome par modélisation mathématique dynamique afin de favoriser une compréhension approfondie des mécanismes complexes sous-jacents. De toute évidence, obtenir une compréhension complète du comportement de toutes les interactions qui se produisent au sein d'une cellule, par exemple, est extrêmement ambitieux et nécessite non seulement une approche descriptive, mais aussi quantitative de ces nœuds. Il est donc évident que de nouvelles approches expérimentales et informatiques devront être développées pour faire face à ce degré de complexité.

C. Protéines membranaires

Les protéines membranaires constituent entre 25 et 33 % du protéome de chaque cellule et jouent un rôle primordial dans différents processus biologiques (transport, énergie, reconnaissance). Leur étude fonctionnelle et structurale reste un enjeu majeur pour la compréhension des mécanismes moléculaires et cellulaires au niveau des membranes. Le développement de nouveaux surfactants (e.g. maltose néopentyl glycols, amphipols, peptidogènes, tensioactifs fluorés...) plus efficaces pour solubiliser et cristalliser les protéines

membranaires et de nouvelles méthodes pour appréhender leur structure tridimensionnelle (phase cubique lipidique et phase éponge, RPE de type «DEER», RMN du solide) ou leur fonction (nanodisques) a ouvert la voie à une meilleure connaissance de ces protéines. Il reste donc indispensable de développer cet axe stratégique et interdisciplinaire source de connaissance et d'innovations technologiques tant en santé qu'en biotechnologies blanches (environnement, énergie).

Un axe majeur de la recherche porte sur le rôle que jouent les lipides associés à ces protéines. Longtemps considérés comme un simple « solvant » des protéines membranaires, ils sont maintenant reconnus comme des composants à part entière de ces protéines où ils assurent non seulement un rôle structural mais également biologique comme co-facteurs pour des activités enzymatiques ou de messagers intracellulaires. Cela a été démontré récemment dans des systèmes d'études aussi différents que les GPCR ou encore le complexe photosynthétique. Leur hétérogénéité de distribution, tant transverse que latérale, est responsable de la courbure et de la dynamique membranaires et par conséquent impliquée dans la dynamique cellulaire (signalisation, transport, architecture, etc.). Mais ces interactions ne sont pas encore bien comprises : sont-elles spécifiques et impliquent-elles une interaction directe avec la protéine ou sont-elles plutôt liées à un changement physicochimique de la membrane entourant la protéine ? Une meilleure connaissance de ces interactions permettra de mieux comprendre au niveau moléculaire le fonctionnement et la régulation des protéines membranaires.

II. Les ARN

L'étude des ARNs et de leurs fonctions a été modifiée de façon profonde par le développement et la démocratisation du séquençage de l'ADN à très haut débit et il est vraisemblable

que cette (r)évolution se poursuivra par la combinaison d'approches à différentes échelles.

A. Les ARN régulateurs

Les ARN régulateurs ont pris une place centrale dans l'étude des ARN et de l'expression génique et il est vraisemblable qu'il reste plusieurs classes d'ARN régulateurs à découvrir en raison de la diversité de leurs voies de biogenèse. Il est donc important de développer de nouvelles stratégies expérimentales et bio-informatiques d'identification (RNomics) et d'élucider leur rôle biologique, leur cible primaire et les mécanismes d'action.

Longtemps focalisé sur l'étude des petits ARN, ce champ d'investigation donne de plus en plus de place aux longs ARN non-codants (lncRNA), linéaires ou circulaires. Les ARN régulateurs forment des réseaux complexes impliquant des régulateurs transcriptionnels. De plus, ils sont soumis à divers niveaux de régulation, aussi bien *via* de longs ARN non codants circulaires ou circRNA, que *via* leur dégradation, qui peut être modulée par des modifications post-transcriptionnelles. Il est donc primordial de considérer les systèmes de manière plus globale et intégrée en définissant les liens entre les régulateurs, le métabolisme cellulaire et les facteurs de signalisation.

Les ARN régulateurs ne se limitent pas aux seuls eucaryotes : ils jouent également un rôle important dans la virulence des bactéries pathogènes et se retrouvent chez certains virus ; ils commencent aussi à être identifiés chez les archaea où le niveau d'expression d'ARN antisens est relativement élevé. De nombreuses études des effets des ARN régulateurs cellulaires sur la réplication virale et *vice-versa* ont été entreprises et devront se poursuivre. En comparaison, l'étude du rôle des ARN dans les relations entre hôtes et bactéries pathogènes ou parasites n'en est qu'à ses débuts et pourrait déboucher sur de nouvelles stratégies de lutte contre ces infections.

Par ailleurs, de nombreux ARN sont multifonctionnels. Par exemple, des ARN non codants fonctionnels sont produits à partir des introns ou d'intermédiaires de dégradation des ARNm et il a été démontré récemment qu'un court ARN produit à partir d'un ARNm se fixe sur les ribosomes et module leur activité. Il a également été suggéré que des produits de dégradation d'ARNnc (ARNt) pourraient se lier au ribosome. À l'inverse, il devient de mieux en mieux établi que de nombreux longs ARN « non codants » codent en fait pour des peptides qui pourraient jouer des rôles clés, par exemple dans le développement embryonnaire. Jusqu'à présent les diverses activités des ARN multifonctionnels n'ont été étudiées que de façon individuelle et l'étude de l'influence mutuelle de ces multiples fonctions constitue un enjeu important.

B. ARN, épigénétique et organisation chromatinienne

Un nombre croissant d'études commence à mettre en évidence le rôle des ARN dans l'épigénétique et l'organisation spatiale de la chromatine en domaines fonctionnels dynamiques qui permettent de réguler à distance tout un ensemble de gènes. Il apparaît donc important d'intensifier les efforts pour identifier les ARN qui induisent des changements durables de l'expression génétique et de caractériser les facteurs et les mécanismes qui permettent cette « reprogrammation » génétique.

Actuellement, il est possible d'analyser la transcription non seulement dans une large population cellulaire, mais aussi sur un petit groupe de cellules, voire une cellule unique. La faisabilité d'une cartographie tridimensionnelle de l'expression génétique dans un tissu a ainsi été démontrée. Une cartographie spatio-temporelle des ARN codants et non-codants dans les cellules, les tissus, et les organes, chez l'homme et les organismes modèles, dans des conditions normales et pathologiques, ou encore lors du développement

devient donc envisageable. Toutefois, ces projets nécessiteront des capacités de séquençage considérables et les moyens informatiques permettant de stocker, manipuler et analyser les données correspondantes. Parallèlement, les techniques de marquage des ARN *in vivo* nécessaires à leur imagerie, qui ont connu récemment des avancées considérables, devront encore être améliorées.

C. Maturation des ARN et complexes RNP impliqués

Le séquençage à haut débit a non seulement permis d'identifier de nouvelles classes d'ARN et d'analyser leur expression, mais aussi d'aborder la présence de nucléotides modifiés, la polyadénylation, l'épissage alternatif, la traduction, la stabilité et de manière générale le métabolisme des ARN. On peut envisager à terme de cataloguer les événements d'épissage alternatif, de recodage traductionnel, de modification post-transcriptionnelle, ou d'altération de la stabilité des ARN dans des conditions normales ou pathologiques, ou chez divers micro-organismes ou virus.

La plupart des fonctions biologiques impliquant l'ARN font intervenir des complexes ribonucléoprotéiques de taille importante (transcription, épissage, transport, stockage dans les granules de stress ou les P-bodies...). Si la cristallographie aux rayons X a bouleversé notre compréhension de la traduction, la cryo-EM a maintenant largement pris le relais, permettant d'obtenir des structures de complexes impossibles à purifier en quantité et homogénéité suffisantes pour l'obtention de cristaux. Des étapes « non-classiques » de la traduction deviennent ainsi progressivement accessibles à l'analyse structurale, de même que des sous-populations spécifiques de ribosomes (ribosomes spécialisés) notamment avec l'avènement de nouveaux algorithmes pour la classification des images couplés à des détecteurs extrêmement performants et permettant ainsi d'approcher des résolutions quasi atomiques. Ces études devraient permettre

d'appréhender la traduction dans toute sa subtilité et sa variabilité, avec de possibles applications thérapeutiques. Il en est de même pour les complexes impliqués dans l'épissage, le transport, le stockage et la dégradation des ARN. Le nombre de structures à (très) haute résolution de complexes ARN-protéines restant relativement limité, il est crucial de poursuivre nos efforts dans ce domaine, si l'on souhaite comprendre les règles de reconnaissance mises en jeu et la dynamique des interactions.

D. Structure secondaire et tertiaire des ARN

L'analyse de la structure secondaire et tertiaire des ARN *in vitro* (et dans de rares cas dans des particules virales) a fait des progrès notables grâce au développement de la technique SHAPE alliée à l'analyse par électrophorèse capillaire. Elle permet une analyse relativement fine de la structure des ARN en solution ainsi que l'identification des sites de fixation des protéines. Cette technique qui nécessite une quantité importante de matériel continue à se développer par la combinaison de plusieurs sondes chimiques mais reste peu répandue en France.

Parallèlement, ces dernières années ont vu la publication de nouvelles méthodologies permettant de cartographier la structure secondaire des ARN au niveau d'un transcriptome entier, soit après extraction, soit directement dans la cellule. À ce stade, ces études permettent de répondre à des questions générales sur le degré de structuration des ARN, mais les données obtenues pour un ARN particulier restent fragmentaires, permettant de confirmer l'existence de structures déjà connues, mais rarement de proposer des structures sur la base de ces seules données. Il est vraisemblable que ces approches démontreront leur potentiel dans l'étude structurale des ARN viraux, dont la petite taille permettra l'obtention d'informations structurales à haute densité, aussi bien dans les cellules que dans les particules virales. Une comparaison et une évaluation systéma-

tique à large échelle des différentes stratégies disponibles seront donc nécessaires.

À côté de ces approches expérimentales, plusieurs approches de modélisation de la structure tertiaire des ARN ont vu le jour ces dernières années. Toutefois, elles pèchent par leur manque de précision en particulier pour de grosses molécules d'ARN. Des approches alternatives utilisant des algorithmes originaux semblent prometteuses pour la modélisation de grosses structures d'ARN. Les méthodes de simulation « gros-grain » utilisent une représentation des macromolécules où les résidus sont décrits de façon simplifiée (généralement par une sphère). Ce type de méthode fournit un outil intéressant pour la modélisation des gros complexes ARN-ARN et ARN-protéines. Des développements supplémentaires sont toutefois nécessaires pour adapter ces méthodes qui ont été utilisées essentiellement sur les protéines. L'utilisation de méthodes hybrides combinant une représentation « gros-grain » et une représentation atomique voire quantique de différentes parties de complexes macromoléculaires ribonucléoprotéiques est sans doute prometteuse pour des complexes ARN-protéines sujets à des changements de conformation et/ou porteurs d'activités catalytiques.

E. Nanotechnologies et ARN

La modularité de l'ARN et la connaissance de la structure tridimensionnelle de modules d'interaction ARN-ARN permettent de concevoir des nanostructures d'ARN s'assemblant spontanément. Le champ d'application de telles nanoparticules est très large et inclut les applications biomédicales, d'autant plus qu'il est possible d'envisager l'assemblage intracellulaire de ces nanostructures. Ce champ de recherche prometteur est peu représenté en France. Son développement requiera des avancées dans la modélisation de l'ARN, la visualisation de molécules uniques d'ARN et la chimie de l'ARN nécessaire pour fonctionner les nanostructures d'ARN.

III. Modifications post-traductionnelles et transduction du signal

À la variété importante des protéines synthétisées par une cellule ou un être vivant s'ajoute une combinatoire immense résultant de leur modification post-traductionnelle (MPT). Cela élargit considérablement la capacité de codage de l'information (épigénétique) et augmente d'autant la diversité des protéomes et des régulations métaboliques. Des champs disciplinaires très diversifiés se sont fédérés pour déchiffrer ces sous-ensembles protéomiques et les dix dernières années ont vu l'émergence de la caractérisation des phosphoprotéomes, des glycoprotéomes, et des ubiquitomes, pour n'en mentionner que quelques-unes.

Les MPT sont des maturations protéolytiques ou des modifications chimiques d'acides aminés par adduits covalents plus ou moins grands. Les MPT permettent une adaptation rapide de la cellule à des changements de son environnement et englobent des processus allant de la modification de l'activité et de la stabilité d'une protéine jusqu'à sa localisation et la régulation de sa dégradation. Elles ont des répercussions également sur la nature des interactions protéine-protéine, sur la transduction du signal, le phénotype et la multiplication cellulaire, l'apoptose ou l'autophagie. Il est donc indéniable que les MPT contrôlent la quasi-totalité des processus cellulaires.

Les MPT sont des événements dynamiques, parfois interdépendants et transitoires, ce qui rend leur étude complexe et restreinte à un type d'altération. Pourtant la caractérisation des réseaux de régulation par les MPT est essentielle à la compréhension du vivant ainsi que des nombreuses dérégulations pathologiques associées. Elle offre ainsi de multiples cibles d'intervention thérapeutique. Cela dit, cette étude nécessite le développement d'outils chimiques, biologiques et technologiques spécialisés.

A. Glycosylation

La glycosylation est une modification covalente majeure qui concerne environ 80 % des protéines et 30 % des lipides dont les propriétés biochimiques et biologiques dépendent très souvent de la structure glycanique qu'elles portent. Les glycanes, jouent un rôle très général dans l'expression, le repliement, le trafic, la localisation et la durée de vie des protéines auxquelles ils sont liés. Ils sont aussi impliqués dans de nombreux processus de reconnaissance et interviennent dans la formation de complexes multimoléculaires. Ces assemblages et leur dynamique contribuent à donner à la cellule et à l'organisme les bases de sa robustesse, de son adaptabilité et de sa complexité. Pourtant, à cause de leur complexité, les glycanes demeurent encore mal connus. La glycobiologie représente aujourd'hui un domaine d'étude en pleine expansion visant à connaître la structure et la fonction des glycanes.

L'assemblage des unités saccharidiques de base des glycanes n'est pas codé génétiquement mais résulte de l'activité d'une machinerie biosynthétique complexe comprenant des centaines d'enzymes (CAZymes qui dégradent, modifient ou créent des liaisons glucidiques). L'identification et la caractérisation des CAZymes des organismes vivants et le développement d'essais à haut débit pour l'identification de nouvelles activités enzymatiques et de spécificités de substrats constituent un véritable défi à relever. Par ailleurs, un enjeu important visera à déchiffrer les règles de biosynthèse de ces molécules. Il est donc nécessaire d'explorer la diversité structurale des glycoconjugués (10 000 à 20 000 épitopes glycaniques dans le glycome humain), leur dynamique et la variabilité de leur expression.

Cela soulève le problème de l'accès à la structure 3D des glycanes et de la reproductibilité de leur production. Il existe peu de méthodes analytiques et la méthode de choix reste la spectrométrie de masse bioanalytique pour établir un profilage du glycome, grâce à de nouvelles approches très sensibles

(nanoLC-ESI-MS, tandem MS). Les notions de clustering de glycanes et d'impact de la fraction aglycone ajoutent un niveau de complexité supplémentaire. Un enjeu majeur est donc de développer de nouveaux outils analytiques des glycanes et glycoconjugués et d'améliorer les méthodes de séparation et de purification. Ainsi, la production de glycanes pour la fabrication de « glycan array » et la constitution de banques de structures glycaniques variées et exhaustives sont en pleine expansion.

Les glycanes naturels sont très hétérogènes, en faible proportion et peu accessibles. Leur synthèse par voie chimique soulève le problème de la stéréosélectivité et de la régiosélectivité des réactions. Développer de nouvelles méthodes de synthèse chimique et/ou chimio-enzymatique simples et rentables demeure un défi majeur et repose sur des approches disciplinaires variées à l'interface entre la biologie et la chimie. Cela passe par l'élargissement de la communauté des glycobiologistes à celle des chimistes pour la synthèse chimique automatisée de glycanes et l'élaboration d'analogues glycaniques et de glycovecteurs pour la thérapie ciblée. Des stratégies d'usines cellulaires pour la production de molécules recombinantes humanisées et d'ingénierie combinatoire par évolution dirigée d'enzymes spécifiques pour la synthèse de nouveaux biopolymères à partir d'agro-ressources abondantes et peu coûteuses sont également en cours de développement. Pour comprendre l'origine de la complexité structurale des glycanes et leur fonction, de nouvelles approches de glycomique comparative et d'imagerie des glycanes voient le jour avec l'utilisation de groupements chimiques greffés sur les molécules de sucre pour quantifier des changements dynamiques de glycosylation. Un nouveau défi sera le développement d'outils bio-informatique pour la prédiction de la structure des glycanes et leur modelage basé sur les résultats d'analyse de glycomes. D'ores et déjà la communauté de bioinformaticiens doit être engagée dans la constitution de bases de données et la conception d'outils d'analyse des glycanes.

La phosphorylation intracellulaire a été largement reconnue dans l'efficacité de la trans-

duction du signal. Mais la surface cellulaire elle-même, de par sa composition, joue également un rôle essentiel dans la signalisation. Elle est caractérisée par une couche dense pouvant atteindre plusieurs μm d'épaisseur, le « glycoalyx », constituée de glycolipides et glycoprotéines (principalement glycosaminoglycannes). Elle est un « passage obligé » pour de nombreuses protéines et occupe une position unique pour participer à l'ensemble des processus qui interviennent au niveau des membranes cellulaires. De fait, la plupart des protéines de signalisation interagissent avec les glycosaminoglycannes en amont de la reconnaissance des récepteurs eux-mêmes. Ces interactions contrôlent ainsi l'activation et le déclenchement des cascades de signalisation spécifiques. Là encore, la caractérisation des interfaces protéine-glycane, l'étude des mécanismes d'interaction, l'ingénierie d'oligosaccharides interférant avec ces mécanismes, offrent de multiples opportunités d'intervention sur de nombreux systèmes, tant physiologiques que pathologiques.

B. MPT, transduction du signal et cycle cellulaire

La phosphorylation, au niveau intracellulaire, est largement reconnue dans la propagation de la transduction du signal chez les eucaryotes. Chez les bactéries, ces modifications ont été longtemps ignorées ou considérées comme marginales. Cependant, un renouveau récent de l'étude de la phosphorylation, en particulier des résidus sérine, thréonine et tyrosine, révèle maintenant leur implication dans le contrôle du cycle cellulaire, de la forme et de la pathogénie des bactéries. Des phosphoprotéomes sont maintenant connus chez certaines bactéries mais cela concerne généralement une seule condition de croissance et ne représente sans doute que le sommet de l'iceberg (les phosphorylations les plus abondantes). De par leur nature transitoire, il faudra étendre ces études, pour avoir une vision cinétique de ces phosphorylations

dans le futur, et intensifier l'identification de ces modifications. Ces études sont certainement le commencement d'une ère nouvelle pour la bactériologie de demain.

C. Nouvelles MPT

Bien que certaines des MPTs soient connues depuis plusieurs décennies (glycosylations, phosphorylation), on entrevoit seulement maintenant leur amplitude et leur diversité. Ainsi, des analyses récentes montrent qu'au moins 5 à 10 % des protéines mitochondriales sont phosphorylées sur des tyrosines, sachant que la phosphorylation des résidus sérine/thréonine est plus fréquente. L'amélioration de nos connaissances des MPT a bénéficié des avancées technologiques récentes des spectromètres de masse qui abaissent considérablement les seuils de détection. Outre l'élargissement du spectre des protéines modifiées par une MPT donnée, ces progrès ont permis d'identifier de nombreuses nouvelles MPT. Ainsi, on dénombre actuellement plus de 12 types de modifications différentes et il est fort probable que bien d'autres MPT seront identifiées dans les années à venir. Il faudra appréhender leur distribution et leur occurrence et comprendre leur rôle pour chacune des protéines modifiées. Un travail considérable reste à entreprendre pour intégrer ces nouvelles connaissances à l'échelle de la cellule, de l'organe, de l'organisme et de la relation entre hôte et pathogène.

D. Modifications post-traductionnelles multiples et associées : des réseaux de réseaux

À cette vision quasi unidimensionnelle, s'ajoute celle – plus réaliste – des modifications post-traductionnelles multiples, associées et

interdépendantes qui définissent des réseaux de réseaux de régulation dont la complexité échappe encore à l'exploration et à la compréhension. Il est pourtant nécessaire d'évaluer cette complexité qui conduira à une image certes plus complexe mais plus fiable de la transduction du signal. En effet, les évidences de modifications multiples d'une même protéine s'accumulent et pourraient étendre les possibilités d'adaptation spatiale et temporelle de l'activité des facteurs de transcription. L'exploration à grande échelle de MPT multiples devrait se poursuivre et s'intensifier grâce au développement d'outils analytiques de plus en plus sensibles et sophistiqués. Un des grands défis à relever est l'intégration de ces réseaux de modifications à l'ensemble des réseaux de signalisation cellulaire par une approche de biologie systémique moléculaire.

En conclusion, les dix dernières années ont vu des progrès spectaculaires dans l'identification des MPT, de leurs occurrences biologiques et de leur caractérisation moléculaire et atomique. Pour intégrer ces développements avec des avancées cellulaires et biomédicales, il est nécessaire de renforcer les études atomiques et moléculaires.

IV. La production et l'utilisation d'énergie

La bioénergétique regroupe l'ensemble des processus permettant aux cellules de produire, véhiculer, transmettre, et utiliser l'énergie chimique ainsi que les mécanismes de régulation de l'homéostasie énergétique permettant à un organisme d'adapter son fonctionnement à son environnement. C'est une discipline en perpétuelle évolution associant la biologie cellulaire, la biochimie, la biophysique et la chimie. Les dérèglements de ces processus souvent vitaux engendrent de nombreuses pathologies.

La connaissance précise d'un nombre croissant de structures tridimensionnelles de

complexes macromoléculaires transducteurs d'énergie des voies photosynthétiques et d'oxydation phosphorylante (procaryotes et eucaryotes) devrait conduire à la compréhension de leur fonctionnement au niveau moléculaire mais également au niveau des voies dans lesquelles ils sont impliqués. De nouvelles questions se posent telles que la localisation des transducteurs, la présence de certains pools de transfert d'électrons (ubiquinones et cytochrome c), la canalisation des substrats ou encore les interactions entre les complexes macromoléculaires dans la membrane. Les progrès de diverses techniques (microscopie électronique, utilisation de sondes optiques, études spectroscopiques, pontages chimiques et électrophorèse en conditions natives) concourent à rendre possibles ces études. Ainsi, un point actuellement très débattu concerne la dynamique des supercomplexes localisés dans les membranes photosynthétiques, mitochondriales et procaryotes. Ces complexes constituent des unités fonctionnelles qui pourraient également contrôler la structure membranaire à grande échelle. Les super-complexes contrôlent-ils les flux d'électrons? Comment s'y effectue le transfert d'électrons? Ont-ils réellement un rôle physiologique? Quelle est la dynamique de formation et de dissociation de ces édifices? Est-elle en lien avec le cycle cellulaire? Ce sont autant de questions qui attendent encore des réponses.

Dans le cas de la photosynthèse, l'étude des mécanismes moléculaires impliqués évolue extrêmement rapidement, en particulier grâce à l'approche « photosynthèse artificielle » qui fait naître des espoirs pour relever les défis énergétiques qui attendent la planète au cours des prochaines décennies. Des mécanismes comme l'oxydation de l'eau par le photosystème II ou la réduction des protons par les hydrogénases, font l'objet de recherches intenses. Par ailleurs le mécanisme de séparation des charges primaires (et donc de la transformation de l'énergie lumineuse en énergie chimique) est toujours mal compris en détail. La récente mise en évidence de phénomènes oscillatoires, pouvant s'apparenter à des cohérences quantiques qui optimiseraient les processus de collecte de la lumière et de sa transformation, a

entraîné l'émergence d'un nouveau domaine de recherche dynamique et productif.

D'un point de vue général, les interconnexions entre ces processus énergétiques et le métabolisme cellulaire sont encore mal caractérisées. Des observations assez anciennes, comme l'inhibition de la respiration par l'activation de la glycolyse (l'effet Warburg) qui a été mise en évidence dans de nombreux types cellulaires, n'ont toujours pas trouvé d'explication satisfaisante même si certains facteurs de régulation ont pu être identifiés. Au cours de ces dernières années, de nombreuses études portant sur des mécanismes aussi variés que la réponse immunitaire, la dynamique du réticulum endoplasmique, les maladies neuro-dégénératives (protéine Parkin), semblent impliquer les mitochondries et leur capacité énergétique dans la régulation de ces mécanismes. Par ailleurs, les nouvelles techniques de séquençage ont révélé l'existence de nombreuses voies métaboliques jusqu'ici insoupçonnées montrant ainsi que les limites de nos connaissances actuelles sont sans cesse repoussées. L'ensemble de ces études souligne la nécessité de développer une approche plus globale du métabolisme qui mette en œuvre modélisation informatique, quantification rigoureuse des différents flux métaboliques et analyse des métabolites par des méthodes très sensibles (*e.g.* spectrométrie de masse, RMN), avec pour objectif d'obtenir une description précise de la dynamique de ces systèmes complexes.

À l'autre bout de cette chaîne énergétique, on retrouve l'ensemble des protéines « consommatrices » qui nécessitent un apport en énergie pour changer de conformation et ainsi accomplir leur fonction physiologique (*e.g.* GTPases, ATPases, transporteurs primaires et secondaires, GPCRs, etc.). Parmi celles-ci, il existe de nombreuses protéines membranaires qui permettent notamment les échanges avec le milieu extérieur ou avec l'hôte colonisé (informations, nutriments, exportation de protéines extracellulaires, injection de toxines par les systèmes de sécrétion bactériens, rejets de déchets...) et qui jouent également un rôle dans les communications intercellulaires. Elles parti-

cipent aussi à la cohésion intracellulaire en régulant des flux de signaux entre les différents organites cellulaires.

La découverte de nouveaux micro-organismes vivant dans des habitats extrêmement divers et mettant en œuvre des stratégies très variées de production d'énergie a donné naissance à une bioénergétique, différente de celle qui s'intéressait quasi exclusivement à la respiration mitochondriale et à la photosynthèse des chloroplastes. Ces travaux ont conduit à la conviction qu'outre son rôle de carburant pour toute réaction cellulaire, l'énergie est un des facteurs, sinon le, déterminants de l'évolution de la vie sur terre depuis ses origines jusqu'à ce jour. L'émergence de capacités à utiliser tout déséquilibre électrochimique présent dans l'environnement à des fins de conversion d'énergie paraît donc avoir été cruciale pour que les procaryotes puissent coloniser la Terre entière. La source de nouveaux systèmes moléculaires et métaboliques doit permettre de progresser dans la recherche de corrélations et similitudes entre des voies bioénergétiques apparemment différentes et permettre de mieux comprendre entre autres, la transition évolutive procaryotes/eucaryotes.

La bioénergétique moderne qui est en plein essor analyse donc toute la diversité de ces mécanismes et de ces modèles. Cet axe majeur de la recherche qui étudie la synthèse d'énergie et donc la base de la vie, est de plus source de valorisation potentielle tant dans le domaine de la santé que dans le domaine de l'environnement et des biocarburants avancés.

V. Nouveaux défis pour la biologie structurale intégrative

L'évolution rapide, au cours des dernières décennies, des différentes approches de biologie structurale permet aujourd'hui d'espérer

faire le lien entre les événements ayant lieu aux différentes échelles du vivant (atomique/moléculaire/multimoléculaire/cellulaire/tissulaire/organismes). Cette démarche implique de déterminer la structure des biomolécules seules et en complexe avec leurs partenaires, d'étudier la dynamique et l'énergétique de leurs interactions et de leurs réactions ainsi que leur organisation spatio-temporelle dans la cellule. Elle nécessite la combinaison de méthodes structurales (cristallographie aux rayons X, RMN, SAXS, microscopie électronique), d'études dynamiques (RMN, diffusion de neutrons) avec des approches mécanistiques qui impliquent souvent des spectroscopies optiques et magnétiques avancées, des approches de molécules uniques et de microscopies optiques (FCS et méthodes associées) et à super-résolution (PALM, STORM, STED).

La biologie structurale s'est en général déjà largement tournée vers une approche intégrative. La détermination des structures 3D des macromolécules biologiques par RMN ou cristallographie est de plus en plus souvent associée à des méthodes à plus basse résolution (SAXS, microscopie électronique) qui permettent l'intégration des structures 3D obtenues dans les complexes supramoléculaires auxquelles elles participent *in vivo*. La détermination de la structure des protéines membranaires reste un processus difficile, et seul un nombre restreint de laboratoires s'est aujourd'hui engagé dans cette voie. Les études structurales concernent par ailleurs de plus en plus des systèmes macromoléculaires complexes, certains de très grande taille (ribosome, polymérase, protéasome, virus, etc.). Elles sont rendues possibles grâce aux nouvelles synergies entre les méthodes structurales classiques et les développements récents d'approches en microscopie électronique à très haute résolution, reposant en général sur l'analyse d'un très grand nombre de particules. Il est à noter que les études les plus récentes utilisant ce type d'approche en microscopie électronique tendent à intégrer la dynamique de ces grands complexes dans leur analyse structurale. Les approches à partir de RMN liquide et solide ont aussi connu de nombreux développements pour l'analyse de la dynamique des structures

ainsi que pour l'analyse des structures macromoléculaires *in vivo* et la détermination des conformations macromoléculaires peu peuplées.

L'ensemble de ces approches fournit un cadre structural qui permet d'utiliser avec beaucoup plus de pertinence toute une palette d'analyses des processus d'interaction entre macromolécules biologiques. Parallèlement à la généralisation de l'utilisation des méthodes bien établies (ultracentrifugation analytique, DSC/ITC, SPR), l'explosion des développements méthodologiques (par exemple SPRI, «backscattering interferometry») devrait permettre une analyse de plus en plus facile et précise de la nature et des aspects énergétiques des interactions entre macromolécules biologiques. La détermination de la structure des macromolécules biologiques permet aussi, en utilisant soit des méthodes spectroscopiques optiques et/ou magnétiques, soit des méthodes de molécules uniques, de caractériser les mécanismes très fins du fonctionnement des biomolécules. Il est à noter cependant qu'aujourd'hui encore, il n'existe que très peu de laboratoires au niveau national allant jusqu'à la modélisation complète des résultats obtenus par ces approches. Parallèlement aux développements de la microscopie électronique, ces dernières années ont vu une évolution rapide et remarquable des approches en microscopie optique qui permettent aujourd'hui de suivre la mobilité de molécules *in vitro* et *in vivo*, de suivre quelques molécules marquées dans des cellules intactes ou d'analyser par des techniques de FRET les interactions entre macromolécules. L'un des défis dans ce domaine consiste à intégrer les informations obtenues par ces différentes techniques et les intégrer sur des cellules intactes. Par ailleurs, le franchissement de la limite d'Abe au tournant du siècle, permet aujourd'hui à de plus en plus de laboratoires d'effectuer des mesures d'imagerie avec une résolution de l'ordre de 50 nm (en utilisant la microscopie STED) ou une précision de localisation d'une vingtaine de nanomètres (méthodes PALM/STORM). Ces méthodes sont encore assez exigeantes (très hauts flux dans le STED et nécessité de protéines marquées avec des sondes fluorescentes

photoactivables pour les microscopies de localisation) mais elles devraient connaître des développements spectaculaires dans les prochaines années. Parallèlement à ces approches purement optiques, la microscopie à force atomique permet de visualiser des protéines (souvent membranaires) avec des résolutions de l'ordre du nanomètre, et a connu des avancées technologiques remarquables permettant aujourd'hui une résolution temporelle très inférieure à la seconde.

Il y a eu au cours des dix dernières années, une évolution nette de la conception même de la structure des macromolécules biologiques, qui a émergé aussi bien de la découverte de conformations macromoléculaires peu peuplées par RMN, d'ensembles de conformations différentes révélées par microscopie électronique et des mesures utilisant des molécules uniques qui ont parfois révélé de très grandes variations structurales d'une macromolécule à l'autre ou d'une même macromolécule au cours du temps ou dans différentes conditions. Par ailleurs, les mesures de forces sur molécule unique ont permis un abord original des mécanismes d'interaction ADN-protéine et récemment de la structure des protéines.

Notons pour finir que les méthodes théoriques d'approche des macromolécules biologiques et de leurs composants ont connu des développements spectaculaires, soit liés à la puissance des ordinateurs qui permettent aujourd'hui la simulation de très grands systèmes sur des échelles de temps avoisinant la microseconde, soit liés aux développements d'approches gros grains, d'approches de type QM/MM, et pour les plus petites molécules le développement de la DFT et td-DFT. Ces méthodes sont aujourd'hui susceptibles de fournir les éléments d'un cadre théorique pour le fonctionnement de systèmes très complexes dans des situations proches de *in vivo*.

VI. Micro-organismes unicellulaires et virus, biodiversité, relations hôtes-pathogènes et évolution

Les microbes sont des modèles de choix pour décrypter tant les modules de bases du fonctionnement cellulaire (protéines, complexe, ARN, lipides...) que la structure de machineries cellulaires universelles ou les interactions et les réseaux. Ces modèles ont longtemps été étudiés pour leur simplicité par comparaison au modèle eucaryote. Toutefois, la diversité des organismes et des métabolismes ou l'idée que leur comportement au sein d'une communauté constituerait le prélude aux interactions cellulaires existant chez les eucaryotes, font de ces systèmes des outils de choix pour comprendre l'organisation du vivant, son fonctionnement et son évolution.

Enfin, ces organismes issus de la biodiversité sont une source d'innovations technologiques, qu'il s'agisse des molécules extraites d'organismes se développant dans les milieux extrêmes ou encore du fait de leur métabolisme, qui ont des applications dans le domaine de la remédiation, de la lixiviation ou encore de l'énergie.

A. Micro-organismes unicellulaires et biodiversité

En quelques décennies, toutes les nouvelles approches « omiques » (génomique, protéomique, métabolomique, etc.) et la résolution de structures de macromolécules isolées ou en complexes ont permis une explosion de nos connaissances du monde des micro-organismes. L'essentiel de nos connaissances fondamentales en microbiologie a pour origine des études réalisées en mode de vie plancto-

nique. Or les micro-organismes « vivent » en communautés parfois attachées aux surfaces, appelées biofilms, ou encore pour coloniser un habitat, ou par exemple pour effectuer la dégradation de la biomasse. Toutefois, à l'heure actuelle, l'étude d'une communauté microbienne dans son ensemble reste très embryonnaire et la compréhension parfaite de son fonctionnement est limitée. La métagénomique a permis l'accès aux génomes entiers de bactéries non cultivées et d'avoir une vision plus globale de la diversité génétique présente au sein d'une communauté entière. Néanmoins, le séquençage de plusieurs centaines de génomes bactériens et l'annotation des gènes a surtout permis la prédiction fonctionnelle des gènes. Il reste un travail considérable en études fonctionnelles, structurales et métaboliques pour confirmer ces prédictions et comprendre le rôle de ces gènes dans les processus biologiques de ces organismes, l'impact environnemental sur leur expression, mais également pour comprendre les interactions entre bactéries et leurs conséquences sur leur métabolisme. La structuration et le comportement bactérien dans des biofilms sont probablement le comportement communautaire qui est le plus étudié. Ces associations denses et structurées de micro-organismes entourés d'une matrice extra-cellulaire (exopolysaccharides, ADN, protéines) permettent aux organismes de résister à de nombreux stress environnementaux et ont souvent un rôle néfaste dans des environnements industriels ou médicaux. L'intérêt grandissant porté aux biofilms cette dernière décennie est lié à l'inefficacité des méthodes conventionnelles de lutte antimicrobienne sur ces structures. Une meilleure compréhension de l'écologie des biofilms et des mécanismes d'adhésion est probablement un des grands défis de la microbiologie.

B. Micro-organismes unicellulaires et relations hôtes-pathogènes

Il est important également de s'intéresser aux relations hôtes/pathogènes/environnement. Aucun organisme n'échappe à l'interaction avec son environnement qui est une source d'adaptation et d'évolution. Dans le cadre médical, avec l'accélération alarmante de l'apparition de souches multirésistantes aux antibiotiques qui laisse présager un manque frappant d'antibiotiques efficaces dans un futur proche, les relations hôtes-pathogènes restent un enjeu majeur pour le développement de stratégies thérapeutiques et vaccinales. La compréhension au niveau moléculaire des mécanismes infectieux constitue un outil puissant pour la mise en place d'approches thérapeutiques et préventives ciblées. Mais l'étude d'un agent pathogène et de son hôte a souvent été réduite aux seules relations entre ces deux partenaires, alors que cet hôte est en interaction avec plusieurs autres micro-organismes qui peuvent être pathogènes ou non. Ainsi, le multi-parasitisme devrait être considéré dans les interactions hôte/pathogène.

Un autre domaine en très rapide évolution est l'étude du microbiote humain défini comme l'ensemble des espèces microbiennes (plus de 500) colonisant durablement un individu dans les premières années de sa vie et représentant plus de 10 fois le nombre de cellules de l'organisme humain. Le microbiote, spécifique à chaque individu, varie en fonction de facteurs environnementaux (nutrition, antibiotiques, hygiène) et est dynamique au cours de la vie. Ces relations hôte/microbiote commencent à être explorées et révèlent une complexité considérable. Ces interactions sont très largement supérieures à celles des interactions hôte-pathogène et pourront, très probablement, également apporter de nouvelles approches thérapeutiques.

C. Virus et évolution

Même s'ils ne sont pas considérés comme des micro-organismes vivants, les virus sont étroitement associés aux processus d'évolution car ils doivent utiliser un hôte pour se répliquer. La balance des interactions entre le virus et son hôte détermine les capacités de survie de chaque entité.

L'évolution des virus à ARN et des rétrovirus est particulièrement rapide et ces virus existent sous forme de quasi-espèces. Cette propriété a des conséquences importantes pour le contournement des défenses immunitaires et l'émergence de virus résistants aux traitements antiviraux. L'étude de l'évolution des virus et des conséquences de l'existence des quasi-espèces doit permettre de mieux appréhender les menaces qu'ils représentent pour la santé publique. Les virus capables de zoonoses fréquentes (par exemple les virus *influenza A*, qui passent des oiseaux aquatiques à certains mammifères) et ceux transmis par l'intermédiaire d'un vecteur (par exemple les arbovirus) représentent des cas particulièrement importants, car ils évoluent à des vitesses différentes dans leurs différents hôtes.

Les récentes découvertes sur les virus géants nucléocytoplasmiques à ADN (NCLDV) montrent qu'un génome viral comme celui des Pandoravirus peut comporter plus de gènes qu'un petit génome eucaryote. Elles posent la question d'un mode d'évolution réductrice chez ces virus. La connaissance du génome viral doit être complétée par celle du protéome et de l'interactome et l'observation du cycle viral peut mener à des découvertes surprenantes. L'étude d'un NCLDV a récemment fait apparaître le concept de virophage, un virus infectant un autre virus. Il est donc primordial de poursuivre l'étude de la virosphère, tant pour prévenir les risques d'épidémie virale que pour mieux comprendre les origines de la vie.

D. Virus et technologie

Des progrès technologiques majeurs sont souvent liés à des avancées significatives de nos connaissances sur les virus. Les premiers essais de laser X à électrons libres ont permis d'obtenir des images inédites d'un NCLDV (le mimivirus) à la résolution du nanomètre. La cryo-microscopie électronique permet maintenant de déterminer la structure atomique d'un bactériophage sans nécessité de cristallisation. La spectroscopie de fluorescence permet d'apprécier la compacité de l'ADN génomique dans ces mêmes virus. La RMN du solide s'avère pleinement appropriée pour étudier la dynamique d'assemblage d'une capsid virale, comme celle du VIH-1. Enfin, l'amélioration des techniques de tomographie permet d'observer en détails les virus *in situ* dans les cellules infectées. La poursuite des avancées technologiques permettant d'explorer les cycles viraux est donc indispensable pour améliorer les connaissances en virologie fondamentale et appliquée.

VII. Biologie des systèmes et biologie synthétique

La biologie des systèmes a pour objectif d'intégrer les différents niveaux de description du vivant (de la molécule à la cellule et à l'organisme) afin de mieux en comprendre son fonctionnement et dont la description théorique deviendrait alors possible. Elle se concentre sur les interactions complexes existant dans les systèmes biologiques et offre une perspective plus globale (ou moins réductionniste) que les approches classiques de la recherche biologique et biomédicale. Émergeant dans les sciences de la vie avec le nouveau millénaire, ce concept a depuis été largement utilisé dans une variété de contextes (signalisation intracellulaire, voies métaboliques, interactions ADN-

protéines et ARN-protéines, etc.). Dans tous les cas, les données actuelles ont été acquises essentiellement par des méthodes permettant un criblage à haut débit. Les interactions mises en évidence sont généralement binaires et statiques, et un bon nombre d'entre elles restent à valider *in vivo*. Dans l'avenir, il faudra établir la hiérarchie des interactions cataloguées et leur interdépendance. Celle-ci dépend de plusieurs facteurs : l'affinité entre les partenaires, leur localisation (espace), la séquence des événements qui conduit à leur interaction (temps), l'état physiologique de la cellule (présence ou non d'autres acteurs moléculaires à un instant donné) et la synergie/compétition entre différents partenaires. Il faudra aussi combiner les résultats avec les données régulièrement obtenues par les autres domaines de la biologie intégrative (génome, transcriptome, protéome, glycome, métabolome). De tels développements n'iront pas sans la mise en œuvre d'outils informatiques et mathématiques spécialement dédiés et adaptés.

S'il reste du chemin à parcourir avant de décrire d'un point de vue qualitatif et quantitatif l'organisation d'une cellule et de fournir un modèle rendant compte globalement de son fonctionnement et permettant la modélisation de la dysrégulation d'éléments spécifiques rencontrée dans des pathologies, il n'en demeure pas moins que d'ores et déjà apparaissent d'intéressantes applications à l'interactome. Par exemple, la fusion d'un interactome viral à l'interactome humain a permis de simuler une infection virale ; la comparaison des réseaux d'interactions dans six maladies neurodégénératives a mis en évidence des protéines-clés communes. En étendant l'interactome à l'étude des interactions inter-cellulaires, il est possible d'imaginer d'autres applications en médecine (développement, renouvellement et régénération des tissus) et en microbiologie (réseau qui s'établit dans un biofilm ou dans un consortium de micro-organismes dans un biotope reconstitué). Enfin, la synthèse de modulateurs de réponses conçus à partir de réseaux d'interactions devrait constituer des outils permettant de mieux comprendre, voire de maîtriser, la réponse cellulaire globale.

L'analyse du métabolome permet maintenant d'accéder à l'ensemble des petites molécules (métabolites, coenzymes, cofacteurs, etc.) produit par un système biologique reflétant les capacités du métabolisme cellulaire. Il sera néanmoins indispensable d'y inclure un paramétrage de la localisation sub-cellulaire des enzymes impliquées, ce qui n'est en général pas fait actuellement, bien que les voies métaboliques puissent être partiellement redondantes entre organites de la même cellule. Combinées aux analyses bio-informatique des génomes séquencés, ces approches devraient conduire à la caractérisation de nouvelles fonctions/voies métaboliques et à l'identification des gènes cibles, qui pourront avoir un intérêt fondamental, biomédical ou biotechnologique.

La biologie synthétique ne peut être dissociée de la biologie des systèmes. Cette discipline qui s'impose au niveau européen et mondial comme une approche majeure associant sciences dures et sciences humaines et sociales devrait porter ses fruits dans les domaines de la bioéconomie : environnement, énergie, nutrition, santé. Elle s'est matérialisée avec la parution de l'article de Graig Venter sur une bactérie synthétique sur la base du génome de *Mycoplasma*.

Il convient toutefois de reconnaître qu'en dépit de quelques projets fortement soutenus par le PIA (TWB, BIOCAR et SYNTHACS) peu de laboratoires de recherche fondamentale s'attachent à relever ce défi. Pourtant les compétences, tant au niveau moléculaire que métabolique et génomique, sont présentes dans les laboratoires. La biologie synthétique, à tort, n'est pas perçue comme un réel axe scientifique pouvant répondre à des questionnements fondamentaux, comme les stratégies mises en place par l'évolution pour optimiser la stabilité des protéines, les éléments constitutifs nécessaires à la vie d'une cellule ou encore les aspects thermodynamiques de la régulation des voies métaboliques, mais comme un outil de recherche et développement. Il convient donc de prendre conscience de cet état de fait et de promouvoir des interactions fortes entre enzymologistes, biochimistes, chimistes,

génétiens, écologues et SHS si l'on ne veut pas que la France s'exclue d'elle-même d'un des enjeux majeurs de la recherche en biologie du XXI^e siècle.

Ainsi, alors que les objectifs de la biologie synthétique apparaissent souvent lointains, ils sont à portée de main dans le domaine de la virologie. Il a déjà été démontré qu'une approche totalement synthétique permet de gagner un temps précieux dans la production de vaccins antigrippaux. D'une façon plus générale, la taille réduite des génomes des virus à ARN permet d'envisager l'utilisation des outils de la biologie synthétique pour produire de manière rationnelle des virus aux propriétés modifiées, par exemple pour la production de vaccins vivants atténués. L'établissement des réseaux de régulation impliquant l'ARN et l'analyse quantitative des propriétés particulières de ces réseaux sera aussi un atout pour le développement de la biologie synthétique.

Conclusion

Maintenant que les outils et cadres conceptuels ont été développés, les nouveaux défis adressés aux chercheurs de la section 20 couvrent un vaste domaine d'investigation. Le décryptage du code post-traductionnel et l'intégration des réseaux de modifications post-traductionnelles à l'ensemble des réseaux de signalisation cellulaire ainsi que la compréhension de la multifonctionnalité de certains ARN et protéines qui révolutionne notre façon de voir le fonctionnement cellulaire constituent des enjeux majeurs. Le développement d'approches multi-échelles et le développement d'une réelle approche RNomics devraient nous permettre d'aborder cette problématique avec l'aide de la biologie structurale intégrative qui associe les techniques de haute et basse résolution aux approches dynamiques.

Le développement de la biologie synthétique s'inscrit dans une démarche interdiscipli-

naire de grande envergure conceptuelle et technologique qui dépasse de loin les frontières scientifiques actuelles et doit nous mobiliser, afin de rester, malgré la concurrence globale, une stratégie affichée claire et ambitieuse qui doit être associée à des moyens humains et financiers.

L'apport des techniques de biologie des systèmes, en particulier les séquençages à haut débit de banques d'ARN issues d'échantillons et de conditions biologiques multiples, ne doit pas faire oublier la nécessité d'une étude approfondie des mécanismes biologiques, dont ceux impliquant l'ARN et leurs complexes macromoléculaires, et des relations souvent transitoires et hiérarchiques entre ces molécules. Ces études très pointues incluent des approches expérimentales et théoriques et ne sont souvent possibles qu'au cas par cas. Mais elles sont souvent les seules à pouvoir *in fine* apporter une compréhension des mécanismes biologiques aux niveaux moléculaires et atomiques. Elles restent cruciales pour comprendre les mécanismes pathologiques et le détournement des mécanismes biologiques par les agents infectieux et pour le développement de stratégies thérapeutiques pertinentes.

Dans la situation économique actuelle, la tentation est grande de se limiter à l'étude d'un seul organisme supérieur (l'homme) et d'un petit nombre de micro-organismes pathogènes. Mais ce serait là la façon la plus sûre de passer à côté de découvertes majeures. Rappelons simplement que la GFP, qui a illuminé notre vision de la biologie, provient d'une méduse, que le système CRISPR, qui promet de révolutionner la génétique, a été découvert en étudiant la résistance de bactéries non pathogènes à l'infection par des phages et que l'avancée spectaculaire de nos connaissances de l'immunité innée chez l'homme doit beaucoup à l'étude des insectes.

L'évolution actuelle va clairement dans le sens d'une diminution des moyens d'intervention du CNRS tant au niveau financier qu'en personnel. En parallèle, du fait, entre autres, du Programme d'Investissement d'Avenir, il y a un accroissement déraisonnable des écarts de financements et personnels alloués à certaines

équipes, sans qu'un accroissement significatif de la performance ne semble le justifier, du moins pour notre instance.

Cette politique associée à une orientation de la recherche étouffe de nombreuses équipes

performantes laissées pour compte et porte en elle le déclin de la recherche française.

« Parce que la tortue a le pied sûr, est-ce une raison pour couper les ailes de l'aigle ? » Edgar Allan Poe.

Comité national de la recherche scientifique. « Section 20- Biologie moléculaire et structurale, biochimie ». *Rapport de conjoncture 2014*, [édition PDF en ligne]. ISBN : 978-2-271-08746-1. Disponible sur : <http://rapports-du-comite-national.cnrs.fr/>