POLYPUBLIE
Polytechnique Montréal

POLYTECHNIQUE MONTRÉAL
UNIVERSITÉ D'INGÉNIERIE

Titre: Les protéines : des molécules de génie : rapport sur la recherche en bio-informatique, en protéomique et en génie protéique
Title:

Auteurs: Simon Hardy, & Pierre N. Robillard
Authors:

Date: 2001

Type: Rapport / Report

Référence: Hardy, S., & Robillard, P. N. (2001). Les protéines : des molécules de génie : rapport sur la recherche en bio-informatique, en protéomique et en génie protéique. (Rapport technique n° EPM-RT-2001-04). https://publications.polymtl.ca/9709/
Citation:

Document en libre accès dans PolyPublie
Open Access document in PolyPublie

URL de PolyPublie: https://publications.polymtl.ca/9709/
PolyPublie URL:

Version: Version officielle de l'éditeur / Published version

Conditions d'utilisation: Tous droits réservés / All rights reserved
Terms of Use:

Document publié chez l'éditeur officiel
Document issued by the official publisher

Institution: École Polytechnique de Montréal

Numéro de rapport: EPM-RT-2001-04
Report number:

URL officiel:
Official URL:

Mention légale:
Legal notice:

Ce fichier a été téléchargé à partir de PolyPublie, le dépôt institutionnel de Polytechnique Montréal
This file has been downloaded from PolyPublie, the institutional repository of Polytechnique Montréal

https://publications.polymtl.ca

LES PROTÉINES :
DES MOLÉCULES DE GÉNIE

Rapport sur la recherche en bio-informatique, en protéomique et en génie protéique

Réalisé par
Simon Hardy
Étudiant M.Sc.A.

Sous la supervision de
Pierre N. Robillard, Ph.D., ing.
Professeur titulaire

Département de génie informatique
Laboratoire de Recherche en Génie Logiciel
15 août 2001

Avant Propos

Quel rôle une École de génie peut-elle jouer dans le domaine très vaste de la nouvelle biologie?

La recherche en bioinformatique est surtout dominée actuellement par des chercheurs qui sont issus du domaine de la biologie, de la médecine et autres domaines connexes des sciences de la santé. Certains informaticiens commencent à s'intéresser à ce domaine surtout à cause de l'outil informatique qui devient un élément essentiel de la recherche.

On découvre maintenant que la clé de l'énigme du vivant repose sur autre chose que le décodage du génome. Les protéines qui sont produites à partir des gènes sont les prochaines candidates à l'exploration systématique.

Les protéines sont des structures moléculaires dont les propriétés dépendent de leurs structures mécaniques, des leurs liens chimiques, de leurs caractéristiques physiques, et de leurs environnement biologiques.

Le défi de cette décennie sera de concevoir un modèle opérationnel de la protéine. Cette entreprise nécessitera des équipes grandement multidisciplinaires dont la pierre angulaire sera l'informatique.

Et si le génie se lançait dans l'aventure?

Nous apporterons des méthodes de travail éprouvées et nouvelles pour ce domaine. Nous sommes de par notre formation des concepteurs et des bâtisseurs de modèles.

Les retombés de ces recherches sont considérables. Une meilleure compréhension du vivant, et des applications dans tous les domaines des sciences de la vie, de la médecine et de la pharmacologie.

Ce document est une introduction pour l'ingénieur à ce domaine fascinant. Il présente également toutes les activités en cours de par le monde. On note qu'une approche orientée génie serait une première mondiale....

Pierre N. Robillard, Ph.D., Ing.
Directeur
Département de génie informatique.

Table des matières

AVANT PROPOS **2**

TABLE DES MATIÈRES **3**

INTRODUCTION **4**

BIO-INFORMATIQUE : POUR SAISIR LA LOGIQUE DU VIVANT **6**

GÉNOMIQUE 7
PROTÉOMIQUE 9
LA BIO-INFORMATIQUE, UN OUTIL INDISPENSABLE 13

LES PROTÉINES DANS LE MONDE **19**

RECHERCHES UNIVERSITAIRES ET INDUSTRIELLES 19
Laboratoires universitaires *23*
Compagnies *27*
CONFÉRENCES 33

LES PROTÉINES AU CANADA **48**

CHAIRES DE RECHERCHE 49
LES CHERCHEURS CANADIENS EN BIO-INFORMATIQUE 65
PROJETS DE RECHERCHE SUBVENTIONNÉS PAR LE CRSNG 70
Projets en génomique 1999-2000 ..... *74*
Projets en génétique moléculaire et développementale 2000-2001 ..... *79*
Projets en biologie cellulaire 2000-2001 ..... *83*
RÉSEAU CANADIEN DES CENTRES D'EXCELLENCE EN GÉNIE PROTÉIQUE 89
PROJETS DE GÉNOME CANADA 93

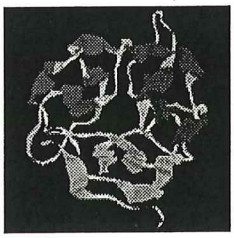
LA BIO-INFORMATIQUE DANS LES UNIVERSITÉS CANADIENNES ... **96**

CONCLUSION **97**

ANNEXE 1 ..... **103**

ANNEXE 2 ..... **106**

RÉFÉRENCES **108**



Introduction

Depuis quelques années, le monde de la biologie vit une révolution qui bouleverse la recherche. Un des accomplissements de cette révolution s'est réalisé par la complétion d'une tâche monumentale : le séquençage du génome humain qui s'est terminé en février 2001. Et maintenant que nous connaissons notre code génétique, une molécule autre que l'ADN retient l'intérêt des biologistes, la protéine.

En effet, après avoir décodé la chaîne de nucléotides qui est à l'intérieur de chacune de nos cellules, les scientifiques ont réalisé qu'ils en savaient toujours très peu sur le fonctionnement biochimique des organismes vivants et de leurs pathologies. Pour y arriver, ils doivent à présent étudier les protéines. Ces macromolécules sont le produit des gènes et assument plusieurs fonctions : ce sont des hormones, des enzymes, des transporteurs, des éléments structurants, etc.

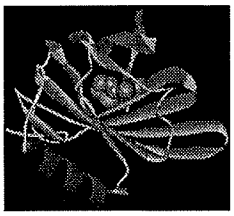
L'étude des protéines prend plusieurs formes, mais deux champs ont pris davantage d'importance : la protéomique et le génie protéique (ou ingénierie des protéines). La protéomique se définit comme l'analyse systématique de l'expression protéique des cellules. En d'autres mots, on cherche à identifier et caractériser les centaines de milliers de protéines qui sont présentes autant dans les bactéries que dans le corps humain. De son coté, le génie protéique vise la compréhension des principes reliant la structure des protéines à leur fonction dans le but de pouvoir concevoir et produire de nouvelles protéines avec de nouvelles fonctions et activités. Le potentiel des découvertes reliées aux protéines rend très optimistes les scientifiques et les investisseurs. Les protéines sont remplies de promesses.

L'analyse des protéines nécessite l'utilisation et le développement de technologies de pointe et de logiciels particuliers. L'utilisation d'applications informatiques aux sciences de la vie a donné naissance au terme bio-informatique. Cette association de la biologie et de l'informatique peut surprendre le néophyte, mais les biologistes ont depuis plusieurs années appris à utiliser la puissance des ordinateurs en plus de

manipuler les éprouvettes dans leur laboratoire. Il devient même de plus en plus acquis que la recherche en génomique et en protéomique est impossible sans des installations informatiques performantes. Il y a actuellement une pénurie de main-d'œuvre qualifiée en bio-informatique dans le secteur de la biotechnologie. Le développement de ce besoin fut si soudain que le milieu académique commence tout juste à offrir une formation adéquate. La bio-informatique fait même les manchettes. Deux articles de journaux sont disponibles en annexe de ce document pour consultation. Le titre du deuxième article, publié le 3 août 2001, est éloquent : *Biologie et informatique, les nouveaux inséparables.*

Les ingénieurs ont-ils un rôle à jouer dans ce bouillonnement scientifique de la biologie? Pour répondre à cette question, il fallait explorer la recherche actuelle sur les protéines. Les différentes sections de ce rapport ont pour but d'éclairer le lecteur, d'abord en l'informant grâce à une synthèse théorique et ensuite en lui présentant différentes problématiques de la recherche sur les protéines au Canada et dans le monde.

Note : Ce rapport contient un très grand nombre d'hyperliens. Une disquette se trouve à la fin de ce document afin de pouvoir le consulter sur un ordinateur et ainsi accéder à beaucoup plus d'informations et de pages personnelles de scientifiques via Internet.



Bio-informatique : pour saisir la logique du vivant

La bio-informatique est une discipline reliant les sciences de la vie aux technologies de l'information. Cette science interdisciplinaire est née de l'intégration des technologies de l'information à la biologie et se définit à la base par l'utilisation de l'informatique pour résoudre des problèmes complexes posés par la biologie. Les efforts de la dernière décennie pour décoder le génome humain et celui d'autres espèces n'auraient pu porter fruit sans un apport grandissant de la bio-informatique à l'analyse et à l'entreposage des données biologiques.

La bio-informatique est utile, voire même indispensable, aux différents domaines de la biologie, mais c'est réellement en génomique que cette science interdisciplinaire s'est avérée cruciale pour la recherche. Alors qu'au départ du projet du génome humain en 1988, on croyait que son séquençage complet se terminerait vers 2025, l'échéance a été maintes fois réévaluée et devancée jusqu'au 15 février 2001, où la carte du génome complétée à 97% a été publiée. Cette tâche monumentale, qui a été accomplie en coopération internationale grâce à Internet comme aucun autre projet scientifique auparavant, n'aurait pu être terminée si rapidement sans le développement de la bio-informatique et la robotisation des laboratoires.

On s'accorde maintenant pour dire que la bio-informatique sera à la source des découvertes futures de la biotechnologie. La recherche en biologie s'effectue de plus en plus devant un ordinateur alors qu'un nombre croissant de projets informatiques s'attaquent à des problèmes biologiques. Le couple biologie-informatique devrait être aussi important au vingt-et-unième siècle que l'a été le couple mathématique-physique au siècle dernier.

Génomique

Le défi de la génomique est de lire le code de la vie que constitue notre ADN. Langage universel du vivant, les lettres de l'ADN ne sont que 4 molécules appelées nucléotides, formées d'une base azotée (adénine, cytosine, guanine ou thymine), d'un sucre et d'un groupement phosphate. Ces nucléotides s'alternent répétitivement pour former la structure spiralée de l'ADN. Dans le cas du génome humain, ce sont 3 milliards de nucléotides qui sont alignés en 23 paires de chromosomes et qui constituent le plan d'architecture de notre organisme. En effet, ce plan situé au centre des noyaux de nos 60 000 milliards cellules est le maître d'œuvre des activités cellulaires. Afin de le lire, il faut décomposer cette gigantesque série de nucléotides associés aux lettres A, C, G et T en mots, c'est-à-dire en gènes.

Repérer les gènes n'est pas une tâche aisée. Les régions codantes de notre ADN sont séparées par une grande quantité de nucléotides non codants, que l'on a longtemps considérés comme n'ayant aucune fonction biologique, mais dont on commence à étudier les propriétés de régulation des gènes. En fait, les gènes fonctionnels (qui expriment un phénotype particulier) ne constitueraient que 1 à 1,5% de la totalité des 3 milliards de paires de base. Dans le noyau des cellules, les gènes fonctionnels activés sont transcrits en ARN messager, des molécules chimiquement semblables à l'ADN. Ces brins d'ARNm quittent ensuite le noyau pour participer à la fabrication de protéines constituées d'assemblage d'acides aminés. À chaque triplet de nucléotides de l'ARNm est associé un des 20 différents acides aminés. Alors qu'est décodée l'information génétique de l'ARNm par le ribosome, ce dernier joint les acides aminés appropriés en une longue chaîne peptidique qui formera une protéine précise. C'est ainsi que les gènes sont responsables de la synthèse des protéines produites dans la cellule.

Pour arriver à identifier les gènes, les scientifiques peuvent avoir recours à 2 méthodes bio-informatiques. Premièrement, ils comparent une séquence d'ADN avec des bases de données contenant des gènes antérieurement séquencés qui ont

été découverts chez l'humain ou d'autres espèces. Deuxièmement, ils utilisent des algorithmes informatiques pour prédire l'emplacement de gènes inconnus. Ces gènes doivent ensuite être validés expérimentalement puisque les algorithmes ne sont pas fiables : ils peuvent à l'occasion omettre un gène existant ou en localiser un alors qu'il n'y en a pas. Par ces techniques, on a réduit le nombre de gènes humains, initialement estimé à 100 000, à 30 000.

Pour étudier le génome de différentes espèces, la recherche en génomique se répartit en quelques activités[1]. Comme il a été précédemment mentionné, on doit d'abord identifier les gènes à partir des séquences d'ADN. On compare aussi ces séquences entières à d'autres qui ont été recensées dans des bases de données afin d'y repérer des séquences homologues. Ces séquences homologues ont déjà été expérimentalement caractérisées et annotées de façon appropriée. Ceci permet de regrouper les gènes et les protéines qui en découlent en familles. Cette association de gènes permet une validation des observations faites sur le gène à l'étude et donne des pistes sur sa fonction. Cette analyse complétée, il faut correctement annoter le gène étudié pour des références futures. Cette étape peut être décisive pour d'éventuelles recherches puisque les données annotées servent de base aux prochaines découvertes.

Déterminer correctement les fonctions d'un gène peut s'avérer l'exercice le plus périlleux de l'annotation. En effet, cet aspect de l'analyse d'un génome n'est pas une analyse classique de séquence. Il faut être en mesure d'intégrer les produits exprimés par les gènes, les protéines, dans des réseaux de processus cellulaires dont font partie les voies métaboliques, les mécanismes de transcription et les cascades de contrôle intracellulaire.

Au cours des dernières années, le rythme de la recherche en génomique s'est grandement accéléré. L'automatisation du séquençage des génomes a mis à la disposition des chercheurs une très grande quantité de données qui leur reste maintenant à analyser. La génomique n'est plus l'apanage des centres universitaires

[1] Tsoka S et Ouzounis A C, *Recent developments and future directions in computational genomics*, FEBS Letters, 480, 2000, p 42-48

et le séquençage est maintenant une activité privatisée effectuée par des laboratoires commerciaux qui ont à leur disposition une quantité impressionnante d'appareils sophistiqués. Avec la fin du séquençage du génome humain qui est presque atteinte, la course ne sera plus de décoder les nucléotides de l'ADN, mais plutôt d'en comprendre le fonctionnement en analysant les données tirées du séquençage. Et c'est justement à cette tâche complexe que s'attaque la protéomique.

Protéomique

En étudiant l'ADN, les chercheurs ont vite compris que le séquençage du génome ne livrerait pas les secrets de la vie. La génomique ne fait que lire le génome, elle n'en donne pas la compréhension. C'est pourquoi les spécialistes s'entendent pour dire que c'est la « post-génomique » qui apportera ces réponses. La principale discipline de la « post-génomique », la protéomique, s'intéresse aux protéines exprimées par les gènes. Elle pourrait bien constituer cette clé de la compréhension recherchée par les biologistes et alors répondre à des questions fondamentales de la biologie. Les protéines ont une importance singulière puisqu'elles sont ultimement responsables de presque tous les processus se déroulant dans la cellule. Les applications de la protéomique sont nombreuses tant en médecine, en agrochimie, qu'en environnement et leur commercialisation est intéressante. Pour le moment toutefois, c'est en pharmaceutique que la plupart des recherches en protéomique sont effectuées et d'énormes ressources y sont consacrées. Par le fait même, on croit que la protéomique redéfinit actuellement les méthodes de recherche de médicaments.

Si on compare les techniques utilisées, la quantité de données produites et la complexité des connaissances qui en découlent, la protéomique constitue un défi scientifique beaucoup plus grand que la génomique. Les biologistes travaillant en génomique étaient déjà submergés d'informations, mais ce n'était rien par rapport au flot de données qui sera généré par la protéomique. Alors que la génomique n'a que 30 000 gènes à déterminer chez l'humain, la protéomique s'attaque à plus de 600 000 protéines humaines différentes. De plus, l'analyse des protéines ne peut

pas être aussi facilement automatisée que le séquençage du génome. La technique actuellement utilisée est basée sur les gels 2D d'électrophorèse qui permettent la différenciation des protéines en fonction de leur poids moléculaire et de leur charge électrique. Les protéines qui sont isolées sur les gels et qui n'ont pas déjà été répertoriées peuvent alors être l'objet d'analyses plus poussées telles la résonance magnétique nucléaire, la cristallographie à rayons X ou la spectrométrie de masse, ce qui permet de connaître leur composition chimique.

Les chercheurs qui oeuvrent en protéomique ont présentement plusieurs avenues. La première consiste à prédire ou déterminer la protéine qui est issue d'un gène. Ensuite, on cherche à connaître les conditions particulières qui mènent à l'expression des gènes et à la production de protéines spécifiques. D'autres tentent de caractériser les protéines, c'est-à-dire d'en connaître la structure et les fonctions. Finalement, une intégration de l'ensemble des connaissances sur les protéines permet de reconstituer les mécanismes biologiques et pathologiques complexes se déroulant dans les organismes.

Déterminer l'expression d'un gène en protéine est un des premiers défis. Malheureusement, il ne suffit pas de connaître la séquence d'un gène pour déduire la protéine qui en est le résultat. L'épissage des gènes, un processus pendant lequel les régions non codantes de l'ADN sont éliminées, peut produire un réarrangement alternatif des fragments restants d'ARNm (l'ensemble des ARNm générés est aussi appelé transcriptomes). Ainsi, une même séquence d'ADN peut produire différentes séquences d'ARNm, donc différentes protéines. Un autre facteur embrouille le problème : les protéines subissent des modifications post-traductionnelles. Lorsque le ribosome a terminé la production de la protéine, celle-ci peut toujours se greffer à un sucre, à un lipide ou même à une autre protéine. Ces changements de structure et de composition peuvent complètement changer la fonction de la macromolécule.

Jusqu'à il y a une trentaine d'années, les scientifiques pensaient qu'à un gène correspondait une seule protéine qui à son tour n'avait qu'une seule fonction. On

sait maintenant que ce raisonnement est erroné. Désormais, on associe à un gène, *n* protéines qui ont *n* fonctions. Les rôles des protéines sont variés et indispensables au bon fonctionnement de la cellule. Ce sont des enzymes accélérant les réactions chimiques, des anticorps qui reconnaissent les molécules étrangères à l'organisme, des récepteurs qui informent la cellule et aussi des produits structurants qui solidifient les parois.

Il faut savoir que ce n'est pas à chaque cellule de notre corps qu'il revient d'exprimer la totalité des gènes. Des cellules de phénotypes différents, par exemple un neurone ou un leucocyte, ne sollicitent pas les mêmes gènes et produisent donc des protéines distinctes. En conséquence, chaque cellule contient un ensemble protéinique caractéristique et c'est ce qu'on appelle le protéome. La technologie des puces à ADN s'avèrera utile pour déterminer les conditions d'expression d'un gène. Ces puces permettent justement de vérifier quel gène est exprimé par une cellule d'un type particulier soumise à des conditions spécifiques. Face à cette réalité, les compagnies pharmaceutiques sont intéressées à comparer les protéomes de cellules saines et de cellules atteintes d'une pathologie à différents stades d'évolution pour repérer des différences au niveau des protéines. Certaines pourraient être en surplus ou en déficit. Ces protéines constituent d'éventuelles cibles thérapeutiques.

Lorsque la composition d'une protéine est connue, l'étape suivante consiste à découvrir ses fonctions. Les meilleurs indices menant à l'identification des fonctions d'une protéine sont donnés par sa structure. Celle-ci permet de déduire les macromolécules avec lesquelles la protéine peut interagir et donc les réactions biochimiques auxquelles elle prend part. Néanmoins, il ne suffit pas de connaître la séquence des acides aminés pour déduire son repliement tridimensionnel. La séquence ne constitue que le premier niveau structurel d'une protéine et la biologie moléculaire en a identifié quatre. Modéliser l'ensemble des interactions dans une protéine pour en connaître l'organisation spatiale est une tâche extrêmement complexe nécessitant une très grande puissance de calcul. La compagnie IBM a annoncé en 1999 la construction de Blue Gene, un ordinateur pouvant effectuer

10^{15} opérations par seconde et dont la première tâche sera d'analyser la structure d'une protéine complexe de 324 acides aminés.

Pour l'instant, des techniques de corrélation sont employées pour identifier les fonctions des protéines[2]. La méthode des arbres phylogéniques associe les fonctions de certaines protéines en vérifiant leur présence dans différentes espèces. Dans un exemple simplifié, si deux protéines sont présentes ou absentes simultanément dans le génome de deux espèces, on peut émettre l'hypothèse que ces protéines sont impliquées dans le même complexe fonctionnel ou la même voie métabolique. Une autre méthode, la Pierre de Rosette, permet d'identifier des réactions entre paire de protéines. Si dans un organisme donné, on retrouve des gènes fusionnés, alors que dans un autre organisme, des gènes homologues sont séparés, alors il est fort probable que les protéines provenant des deux gènes séparés interagissent.

La protéomique ne se limite pas à l'étude des protéines individuellement. Bien plus important encore est l'analyse des mécanismes de régulation, des interactions gènes-protéines, des cascades de réactions biochimiques et des circuits biologiques. Voilà où se situent toutes les riches possibilités de la protéomique. Élaborer ce modèle dynamique de la biochimie du vivant est d'une complexité ahurissante. Imaginer 600 000 protéines interagissant entre elles! Ces connaissances nous permettront d'influencer le développement d'une pathologie et nous mèneront à la compréhension du vivant. Certaines institutions se spécialisent dans cette reconstruction métabolique : le KEGG (Kyoto Encyclopedia of Genes and Genomes) élabore des cartes fonctionnelles d'interactions de protéines pour comprendre certains mécanismes ; EcoCyc (base de données métaboliques sur l'*E. coli*) intègre le savoir sur les voies biochimiques connues chez cette bactérie et l'information sur son génome.

La protéomique devrait connaître un important développement au cours des prochaines années. Les revenus mondiaux de la protéomique ont atteint le milliard

[2] Eisenberg D, Marcotte E M, Xenarios I, Yeates T O, *Protein function in the post-genomic era*, Nature, vol 405, juin 2000, p 823

de dollars américains en 2000 et les projections prédisent des revenus plus de 5 fois supérieure en 2005.[3]

La bio-informatique, un outil indispensable

La recherche en génomique et en protéomique n'aurait pu atteindre ce niveau sans la bio-informatique. L'impressionnante quantité de données générées nécessitait une grande capacité de stockage et une bonne organisation de l'information. L'archivage de la séquence complète des nucléotides du génome d'une seule personne nécessite 50 teraoctets[4]. Ainsi, on a vu se mettre sur pied plusieurs grandes bases de données publiques, qui servent de référence à l'ensemble de la communauté scientifique : GenBank pour le continent américain, la banque du EMBL (European Molecular Biology Laboratory) à Heidelberg en Allemagne pour couvrir l'Europe et la DDBJ (DNA Data bank of Japan) pour l'Asie-Pacifique. Ces banques regroupent principalement des données sur des séquences génomiques. D'autres bases de données se concentrent sur les protéines : par exemple SWISS-PROT et TrEMBL qui font partie du projet du protéome humain. Accessibles par Internet, ces banques riches en informations sont parcourues quotidiennement par différentes équipes de recherche à travers le monde.

Comme mentionné précédemment, une grande partie des analyses des nouvelles séquences d'ADN se fait en les comparant avec d'autres séquences. La plupart d'entre elles proviennent de ces banques de données. Ceci permet de trouver des séquences similaires menant à l'identification de gènes. Si des similarités sont trouvées entre protéines en utilisant le même procédé de comparaison, c'est déjà un indice menant à la découverte de leur structure et éventuellement de leurs fonctions.

Pour effectuer ces comparaisons à grande échelle, les biologistes doivent avoir à leur disposition des ordinateurs ayant des capacités de traitement à haute puissance.

---

[3] Frost & Sullivan, *World Proteomics Market*, Rapport # 5822-55, octobre 2000, reproduit dans Conseil de la science et de la technologie, *La bio-informatique au Québec : un levier essentiel du développement des bio-industries*, Gouvernement du Québec, Québec, janvier 2001, p. 19

[4] Butte A J, *Challenges in bioinformatics : infrastructure, models and analytics*, TRENDS in biotechnology, Vol 19, May 2001, p159-160

Les applications en biologie sont même en voie de détrôner la météorologie et l'astrophysique au niveau des besoins en traitement de données. Certaines compagnies offrent des logiciels gratuits en ligne permettant d'accomplir ces analyses. Le plus populaire d'entre eux, *BLAST* (Basic Local Alignment Search Tool), date de 1990 et utilise la méthode des mots pour chercher une séquence de nucléotides à travers différents génomes[5]. Ce logiciel ne convient pas lorsqu'il s'agit de comparer des génomes entiers entre eux. D'autres logiciels ont été développés pour ce genre d'analyse, par exemple *RAPID* (Rapid Analysis of Pre-Indexed structures) qui fait une estimation de la similarité entre deux séquences. Différents algorithmes ont été utilisés afin de détecter les répétitions et leurs occurrences dans les séquences ainsi que les motifs originaux ou dégénérés. De même, le modèle mathématique de Markov caché fut employé dans la prédiction de la localisation des gènes.

Les découvertes doivent être cataloguées afin d'être réutilisées et les séquences contenues dans les bases de données doivent être annotées. Par annotation, on signifie rattacher à une séquence les diverses informations reliées à l'état des connaissances sur cette séquence. On peut vouloir connaître l'organisme dont elle provient, les fonctions qui lui sont attribuées, les SNP (*single nucleotide polymorphism*) qui y ont été répertoriés, les pathologies qui y sont associées, etc. Le méta-logiciel de recherche *Entrez* du National Computational Biology Institute des États-Unis couvre la plupart des bases de données de l'Institut, incluant celles des structures tri-dimensionnelles des protéines, des génomes complets de plusieurs organismes et des références bibliographiques de publications scientifiques. Pour y arriver, les bases de données doivent être inter-opérationnelles, ce qui serait très difficile sans le Web. Un groupe du centre français de ressources Infobiogen a suggéré d'utiliser le langage XML pour structurer et véhiculer l'information biologique entre bases de données.[6] L'encapsulation des données biologiques par

[5] Bottomley S, *Rapid searching of sequences databases*, Drug discovery today, vol 4, octobre 1999, p 482-484

[6] Achard F, Vaysseix G et Barillot E, *XML, bioinformatics and data integration*, Bioinformatics, vol 79, février 2001, p 115-125

secteur d'activités qui se fait même à l'intérieur d'une compagnie ou entre les départements d'une institution académique est condamnée à disparaître.

De plus, afin de pouvoir utiliser des techniques automatisées de recherche et d'analyse, les bio-informaticiens ont senti le besoin de développer des modèles de données plus stricts. Certains ont utilisé UML (*Unified Modeling Language*) dans leur représentation.[7] [8] L'intégration des données constitue une des activités de la bio-informatique où il est important de systématiser et de standardiser le processus d'annotation à mesure que les données sont collectées. Ceci peut devenir crucial pour les recherches scientifiques où on a noté que certaines maladies sont déjà trop complexes pour que l'on puisse en conserver tous les détails dans un modèle mental.[9] Ces derniers ont aussi le défaut d'être inaccessibles à d'autres groupes de travail. Aussi, il ne suffit pas de développer de bons modèles pour correctement représenter des données, il doit aussi y avoir consensus sur la sémantique.[10] Le Gene Ontology Consortium travaille à la conception d'un vocabulaire pour le savoir en génomique et en protéomique. Finalement, un autre aspect sur lequel les bio-informaticiens travaillent en relation avec l'intégration des données est leur représentation visuelle pour en faciliter la consultation.

Pour trouver de nouvelles pistes de recherche, il faut être en mesure de parcourir l'immensité des informations amassées provenant des différentes disciplines de la biologie : microbiologie, toxicologie, biologie cellulaire, immunologie, biologie moléculaire, essais cliniques, etc. Cette exploitation intelligente des données brutes pour obtenir une vue globale est appelées *data mining*. Contemporain à l'analyse de données, le *data mining* a à sa disposition l'intelligence artificielle, les techniques avancées de statistiques et les réseaux neuronaux pour détecter des schémas relationnels inapparents à l'aide des méthodes traditionnelles. Les compagnies qui mineront et exploiteront le mieux l'information biologique pourront se tailler la part

[7] Paton N W et al., *Conceptual modeling of genomic information*, Bioinformatics, vol 16, mai 2000, p 548-557
[8] Bader G D et Hogue C, *BIND – A data specification for storing and describing biomolecular interactions, molecular complexes and pathways*, Bioinformatics, vol 16, mai 2000, p 465-477
[9] Butte A J, op cit.
[10] Rechenmann F, *From data o knowledge*, Bioinformatics, vol 16, mai 2000, p 411

du lion dans le marché compétitif des biotechnologies. Une grande quantité des données scientifiques se trouvent toujours dans les articles scientifiques, dans un format difficilement accessible pour les applications informatiques. Afin d'être en mesure d'exploiter cette ressource riche en connaissances, des logiciels ont été développés pour analyser les textes et en retirer l'information biologique recherchée.[11] [12]

Afin d'automatiser l'analyse des données de façon à tirer réellement profit du déluge d'informations provenant des résultats de recherche, Düx et Moille[13] indiquent qu'il ne faudra non seulement intégrer les bases de données, mais aussi les logiciels d'analyse (outils de recherche dans les bases de données de similarités, d'alignement multiple de séquences biologiques et d'alignement de structures de protéines, outils de prédiction de structures des protéines et modélisation de structures).

Au départ, les recherches ne s'orientaient que vers les pathologies causées par un seul gène. Cependant, celles-ci ne constituent qu'une faible partie de tous les troubles biologiques. Des dizaines ou même des centaines de gènes peuvent être à l'origine d'une pathologie, ce qui complexifie davantage la cascade des réactions à démystifier. Résoudre de tels problèmes sera un travail fastidieux qui, sans l'assistance d'outils bio-informatiques, pourra s'avérer presque impossible.

La bio-informatique a déjà commencé à modifier l'approche réductionniste de la biologie. Auparavant, les études en biologie se faisaient en observant et en analysant les parties d'un tout : on a décortiqué les organes pour comprendre le fonctionnement du corps humain comme on a subdivisé la cellule en organites. Avec la compréhension des mécanismes de l'ADN, il est impossible de réduire davantage, du moins du point de vue des sciences biologiques. À présent, ce sont

[11] Ono T et al, *Automated extraction of information on protein-protein interactions form the biological literature*, Bioinformatics, vol 17, février 2001, p 155-161
[12] Marcotte E M, Xenarios I, Eisenberg D, *Mining literature for protein-protein interactions*, Bioinformatics, vol 17, avril 2001, p 359-363
[13] Düx P et Moille F, *Bioinformatics : An inventory and analysis of recent developments in bioinformatics and related areas of research and developement*, p 16

des relations entre les découvertes, les mécanismes et les interactions qu'il faudra faire. On cherche à comprendre exactement comment, à partir de l'ADN, un organisme se bâtit, évolue et réagit.

De la même façon, les techniques de recherche de médicaments pourront délaisser la méthode empirique d'essais-erreurs pour adopter une méthode plus scientifique, basée sur une meilleure compréhension des interactions moléculaires. La démarche des compagnies pharmaceutiques sera plus prédictive, les cibles thérapeutiques clairement identifiées et les coûts de développement moindres. De plus, les effets secondaires d'une thérapie pourront être grandement diminués. Présentement, pour un médicament mis en marché, 10 000 substances chimiques ont été testées en laboratoire. Ce taux de succès est appelé à grandement s'améliorer. Ce passage de la recherche traditionnelle en laboratoire à une utilisation grandissante des ordinateurs a fait naître le terme biologie *in Silico* (en opposition à biologie *in Vivo* ou *in Vitro*).

Un autre domaine de recherche découlant des récentes découvertes attire l'attention : la pharmacogénomique. Ce domaine veut miser sur l'avancement de la génomique et de la protéomique pour développer des traitements individualisés. Le protéome de chacun est unique puisqu'il est à la fois influencé par le génome et l'environnement de l'individu, dont entre autre son alimentation. Pour le moment, on développe des médicaments généraux pour l'ensemble de la population. Toutefois, d'autres traitements pourraient s'avérer beaucoup plus efficaces pour certains individus en fonction de leur protéome. Ces mêmes médicaments pourraient s'avérer toxiques pour d'autres personnes ayant un protéome différent.

Actuellement, les médecins disposent de plusieurs médicaments pour une même maladie. Dans certains cas, les patients doivent essayer différents remèdes avant de sentir les effets escomptés. Et pourquoi ça? Entre individus, certaines protéines de l'organisme impliquées dans une maladie peuvent être légèrement différentes. En conséquence, certaines thérapies seront inopérantes (et parfois nocives) alors que d'autres mèneront à la guérison à cause de ces petites différences. La

pharmacogénomique permettra justement de prescrire avec précision le traitement le plus prometteur pour une personne selon ce qui la distingue des autres, son protéome.

Les protéines dans le monde

L'univers de la protéomique et du génie protéique est en constante évolution. Maintenant, suite aux explications théoriques, il serait intéressant de faire une revue de la recherche qui se fait actuellement dans le monde. Ainsi, il sera possible de dégager les problématiques et les enjeux qui se vivent dans l'étude des protéines.

Pour y arriver, deux relevés sont proposés. Tout d'abord, une liste des laboratoires universitaires et des compagnies faisant de la recherche en protéomique donnera un aperçu des intérêts des chercheurs travaillant dans des institutions ou dans l'industrie. Ensuite, une série de conférences donnera des indices sur les aspects de la protéomique de l'heure.

Recherches universitaires et industrielles

La recherche en protéomique est déjà relativement répandue dans le monde. Plusieurs laboratoires et compagnies se sont lancés dans la course aux protéines. Toutefois, un clivage important sépare les objectifs des organisations académiques de ceux des industries. Presque la totalité des industries n'ont qu'un seul but : accélérer la découverte de médicaments. Ce sont les applications de la protéomique reliées à la pharmaceutique qui indiquent les avenues de la recherche industrielle. Dans les universités, on s'attaque à des problématiques différentes : compréhension de mécanismes cellulaires, découverte de structures protéiques, analyse protéomique de différentes espèces et étude du développement de pathologies. Il y a aussi un autre problème d'une incroyable complexité qui attise la curiosité des chercheurs : la repliement de protéines.

Dans les pages qui suivent se trouve un recensement des laboratoires universitaires et des compagnies du globe qui font de la recherche en protéomique et dans des sujets connexes. Leur nom provient d'un portail Internet sur la protéomique (www.proteinscience.com). Des informations sur leur localisation et leurs sujets de

recherche sont disponibles. Chacun des laboratoires a aussi une cote entre 1 et 4 qui indique l'intérêt potentiel (1 : bas, 4 : élevé) d'un groupe de l'École Polytechnique pour leurs travaux.

Deux pays semblent faire davantage de recherche en protéomique. Premièrement les États-Unis, dont un grand nombre d'universités sont actives en protéomique et ensuite le Royaume-Uni, qui dépasse tous les pays européens. Ceci est très apparent pour les groupes universitaires, mais l'est relativement moins pour les industries dont plusieurs ont des bureaux dans quelques pays.

Au Canada, l'Université de Toronto attire l'attention puisqu'elle semble s'être dotée de plusieurs groupes s'intéressant aux protéines. Plus spécifiquement le groupe de protéomique et de bio-informatique dont fait partie le laboratoire du professeur de biochimie Christopher Hogue. Ce dernier, qui a aussi de l'expérience en programmation, s'intéresse à trois champs de recherche complémentaires : le repliement des protéines, la stabilité des protéines et l'assemblage macromoléculaire. Il est aussi intéressant de constater que l'Université de Toronto offre un programme d'études doctorales de recherche en bio-informatique et protéomique impliquant plusieurs départements de l'établissement.

Dans l'industrie de la protéomique, on peut regrouper l'ensemble des activités en différents types. Il y a d'abord les grandes compagnies à mission pharmaceutique ou biotechnologique qui disposent d'une division particulière qui effectue de la recherche en protéomique afin de trouver de nouvelles cibles thérapeutiques et de nouveaux médicaments. C'est le cas chez Merck Frosst. D'autres compagnies se spécialisent uniquement dans la recherche en protéomique et leur principal financement provient souvent de la vente d'informations aux industries de la pharmaceutique. Celles-ci font du criblage à haut-débit de protéines (analyse systématique à grande échelle), identifient des cibles thérapeutiques et des marqueurs de maladie et bâtissent des bases de données sur les protéines et les interactions protéine-protéine. Par exemple, la compagnie Proteom, basée à Cambridge. Bien sûr, ces compagnies ont besoin de fournisseurs particuliers qui

développent les technologies qui permettent l'analyse des protéines : spectrométrie de masse, cristallographie à rayons X et gels d'électrophorèse bidimensionnelle. Certaines de ces entreprises tentent aussi d'adapter la technologie des puces à ADN aux protéines. Finalement, il reste les concepteurs des inévitables outils bio-informatiques.

Comme il a été mentionné plus haut, le principal objectif des industries en protéomique est d'accélérer l'analyse de protéines afin d'obtenir plus rapidement un médicament qui peut être commercialisé. Les employés en marketing en ont fait leur adage. Plusieurs entreprises rivalisent entre elles en offrant des plates-formes complètes, intégrant différentes technologies, spectrométrie de masse et logiciels avancés d'analyse. Même les biopuces sont sensées remplacer plusieurs processus faits normalement en laboratoire en plus de consommer beaucoup moins de ressources et de temps. L'industrie ne concentre pas ses efforts sur la compréhension de cette nano-machine qu'est la protéine. Pour le moment, il s'agit davantage d'identifier et de caractériser un très grand nombre de ces macromolécules. On fait de la collecte d'information et on catalogue les protéines afin de trouver de nouveaux remèdes.

Les universités ne peuvent pas concourir là où les entreprises investissent à coups de millions. C'est pourquoi elles s'attaquent à d'autres problèmes. Plusieurs chercheurs s'intéressent à la prédiction de la structure moléculaire, aux mouvements macromoléculaires et au repliement de la chaîne peptidique. D'autres se concentrent sur des mécanismes moléculaires bien précis : la neurotransmission, la transcription, la régulation génique et les signaux cellulaires pour ne donner que quelques exemples. Ces champs de recherche ressemblent davantage à de la recherche fondamentale, mais les découvertes qui y sont faites auront aussi des applications bien concrètes. La compréhension du fonctionnement de la cellule et des protéines permet de comprendre les déséquilibres qui se produisent lors de maladies et donne la possibilité de développement de plans de thérapie plus ciblés.

Les scientifiques font aussi de la modélisation. Modélisation de molécules et de repliement, modélisation de réseaux biochimiques, modélisation d'activités cellulaires, modélisation de systèmes. Synthèse des connaissances actuelles, la modélisation simule un comportement. La modélisation n'est toutefois pas arrivée à maturité, trop de morceaux sont manquants afin d'être en mesure de reconstituer le casse-tête.

Laboratoires universitaires

Nom du laboratoire	Localisation	Sujets de recherche	Cote
Aberdeen Proteome Facility (University of Aberdeen)	Aberdeen, Écosse	Identification de protéines; Caractérisation de modifications post-traductionnelles; Mécanismes de résistance; Réponses cellulaires au stress environnemental	2
Andrew Miller Proteomics Group (Imperial College)	Londres, Royaume-Uni	Étude de structure, fonction et chimie des protéines de stress; Mécanique de repliement des protéines; Interactions protéine-protéine	3
Andrews Lab (Univ. of Michigan)	Ann Arbor, Michigan	Technologies de spectrométrie de masse; Automatisation d'analyse de protéomes; Protéomes de cellules membranaires; Gels 2D virtuels	1
Argonne National Lab (Proteomics Project) Laboratoires de protéomique, de bio-informatique et de génie protéique	Argonne, Illinois	Analyse de protéomes d'archeabactéries extrêmophiles (gels 2D, spectrométrie de masse);	2
Arrowsmith Lab (U of Toronto)	Toronto, Ontario	Aspects structuraux d'interaction protéine-ADN des facteurs de transcription	4
Baker Lab (U of Washington)	Seattle, Washington	Repliement tridimensionnel de protéines	3
BioMolecular Engineering Research Center (Boston U)	Boston, Massachussets	Analyse de génome complet; Prédiction de la structure de protéines; Développement d'outils d'analyse	2
Biomolecular Structure and Modelling Group (U College London)	Londres, Royaume-Uni	Analyse de structure et de séquence de protéines; Classification de structures protéiques; Cristallographie; Étude des glycoprotéines; Résonance nucléaire magnétique(RNM)	3
Biomolecular NMR Group (Ontario Cancer Institute; U of Toronto)	Toronto, Ontario	Étude par spectroscopie RNM de protéines impliquées dans le cancer; Signalisation du calcium; Régulation des gènes; Communication intercellulaire	2
Bloomsbury Centre for Structure Biology (U College London)	Londres, Royaume-Uni	Évolution moléculaire et phylogénétique; Analyse de séquences protéiques; Prédiction de structure protéique; Design de protéines	2

Brookhaven National Lab Proteome Project	Long Island, New York	Génomique structurelle; Structure et fonction de protéines réparatrices d'ADN	1
Brunel Bioinformatics Group (Brunel U)	Middlesex, Royaume-Uni	Prédiction de structure protéique; Analyse de séquence protéique; Analyse de génome; Classification de structure de protéines; Modélisation de protéines membranaires	2
Brutlag Bioinformatics Group (Stanford)	Stanford, Californie	Prédiction de fonctions biologiques et de structures de protéines; Outils de comparaison de données sur protéines	2
Bystroff Lab (Rensselaer Polytechnic Institute)	Troy, New York	Modélisation et simulation de repliement de protéines	3
Cambridge Crystallographic Data Centre	Cambridge, Royaume-Uni	Analyse de données provenant de cristallographie à rayons X	1
Center for Biological Sequence Analysis (Technical University of Denmark)	Lyngby, Danemark	Structure et fonction de ADN/ARN et protéines; Analyse comparative de séquences	3
Center for Genomics Research (Karolinska Institutet)	Stockholm, Suède	Transcription; Localisation intracellulaire de protéines; Cibles thérapeutiques	3
Center for Proteome Analysis (U of Southern Denmark)	Odense, Danemark	Site en construction	N/A
Center for Proteome Studies (U. of Michigan)	Ann Arbor, Michigan	Site en construction	N/A
Center for Structural Biology (Yale)	New Haven, Connecticut	Regroupement de 10 laboratoires étudiant des sujets variés : mécanisme moléculaire de neurotransmission, mouvements macromoléculaires, structure des ARN, repliement de protéines membranaires, etc.	4
Cohen Lab (U California San Francisco)	San Fransisco, Californie	Relations structure-séquence; Évolution des protéines; Design de bases de données sur protéines	3
Columbia University Bioinformatics Center	New York, New York	Analyse et prédiction de structure de protéines	2
Computational Biology Group (U California Santa Cruz)	Santa Cruz, Californie	Prédiction de structure de protéines	2
Computational Protein Structure Group (Oak Ridge National Labs)	Oak Ridge, Tennessee	Énergétique du repliement des protéines, Prédiction de structure	3
Darst Lab (Rockefeller U)	New York, New York	Enzymes de transcription	2
E. coli Structural Proteomics Group (U of Toronto)	Toronto, Ontario	Analyse du protéome de E. Coli : Identification de protéines, Détermination de structure et d'interaction	1
Edwards Lab (U of Toronto)	Toronto, Ontario	Transcription; Réplication d'ADN; Protéomique structurelle;	3
Ellenburger Lab (Harvard)	Boston, Massachusetts	Étude des interactions entre protéines et facteurs de transcription impliquées dans la réplication, la recombinaison et la réparation de l'ADN	3

European Bioinformatics Institute Proteome Analysis Group at EBI	Cambridge, Royaume-Uni	Base de données de protéomes	1
Gerstein Lab (Yale)	New Haven, Connecticut	Mouvement macromoléculaire; Génomique comparative	3
Gygi Lab (Harvard)	Boston, Massachusetts	Technologie de spectrométrie de masse; Interactions protéine-protéine; Analyse de protéome	2
Harvard Medical School Institute of Proteomics	Boston, Massachusetts	Système automatisé d'analyse protéomique; Isolation de protéines;	1
Hogle Lab (Harvard)	Boston, Massachusetts	Cristallographie à rayons X; Étude des protéines reliées aux virus	1
Hunt Lab (U. of Virginia)	Charlottesville, Virginie	Spectrométrie de masse; Fonction de protéines	2
Institute for Systems Biology	Seattle, Washington	Interactions complexes entre génomique, ARN, protéomique, réseaux et voies biochimiques	3
Kuriyan Lab (Rockefeller U)	New York, New York	Biologie structurelle des signaux cellulaires de transduction de réparation de l'ADN	2
Levitt Lab (Stanford)	Stanford, Californie	Prédiction de structure de protéines, Dynamique moléculaire	4
Linnaeus Centre for Bioinformatics (Uppsala U)	Uppsala, Suède	Algorithmes et Base de données; Étude de génomes; Expression et fonction; Structure et design	2
Lipper Center for Computational Genetics (Harvard)	Boston, Massachusetts	Génomique structurelle 3D du cancer; Modifications, quantifications et localisation de protéines; Structure et interaction; Signaux et modélisation de système	3
Michael Barber Centre for Mass Spectrometry (UMIST)	Manchester, Royaume-Uni	Spectrométrie de masse	1
Moult Group (Center for Advanced Research in Biotechnology)	Rockville, Maryland	Repliement de protéines; Prédiction de structure et fonction de protéines; Bio-informatique pour génomique structurelle	3
Munich Information Center for Protein Sequences	Munich, Allemagne	Étude de génomes (levure, plante, humain); Classification de protéines; Base de données	2
National Center for Biotechnology Info (NCBI) Structure Group	Rockville, Maryland	Comparaison de structure entre protéines; Algorithmes pour repliement de protéines; Visualisation 3D de structure	3
NJCST Initiative in Structural Genomics and Bioinformatics (Rutgers)	Newark, New Jersey	Étude des fonctions de gènes par la structure 3D de protéines, Spectroscopie RMN; Cristallographie à rayon X	2
Pande Group, Stanford University (Folding@Home)	Stanford, Californie	Repliement de protéines; Dynamiques distribuées	3
Program in Proteomics and Bioinformatics (U of Toronto)	Toronto, Ontario	Programme universitaire	N/A
Protein and Peptide Group (EMBL)	Heidelberg, Allemagne	Séquençage de protéines; Spectrométrie de masse; Développement d'outils d'analyse	1

Protein Characterization and Proteomics Laboratory (University of Cincinnati)	Cincinnati, Ohio	Spectrométrie de masse	1
Protein Information Resource (Georgetown U)	Washington, DC	Base de données de protéines	1
Protein Research Foundation (Osaka, Japan)	Osaka, Japon	Base de données de protéines	1
Protein Sequence Analysis Group (U of Manchester)	Manchester, Royaume-Uni	Site non-disponible	N/A
Protein Structure Initiative	Programme gouvernemental américain	Structure de protéines, Regroupement de protéines en famille	2
Protein Structure Prediction Center (Lawrence Livermore National Lab)	Livermore, Californie	Prédiction de structure de protéines; Modélisation comparative	2
Protein Function Group (U of Liverpool)	Liverpole, Royaume-Uni	Rôle des protéines dans la communication chimique; Protéines musculaires	2
Proteomics and Bioinformatics Group (Samuel Lunenfeld Research Institute)	Toronto, Ontario	Base de données de réseaux d'interactions bio moléculaires (BIND); Repliement de protéines; Visualisation 3D; Protéines de signalisation intracellulaires	4
Sali Lab (Rockefeller)	New York, New York	Modélisation comparative de protéines; Structure 3D	2
Scheraga Lab (Cornell)	Ithaca, New York	Repliement de protéines; Réactivité; Interaction protéine et petites molécule	3
Swiss Institute of Bioinformatics (SIB)	Genève, Suisse	Ressource logiciel; Développement de bases de données intégrées	2
Tubingen Univeristy Proteome Lab	Tubingen, Allemagne	Gels 2D électrophorèse	1
UCLA-DOE Lab of Structural Biology (U California-Los Angeles)	Los Angeles, Californie	Structure de protéines; Signaux de transduction; Dynamique; Interactions protéine-protéine; Expression protéique	3
Wagner Lab (Harvard)	Boston, Massachusetts	Protéines d'activation de transcription et traduction; Protéines d'apoptose	2
Yonsei Proteome Research Center (Korea)	Séoul, Corée	Identification de protéines	1

Compagnies

Nom de la compagnie	Localisation	Principales activités	Cote
ACE BioSciences	Odensee, Danemark	Développement d'une base de données sur les réseaux d'interaction protéique de microorganismes pathogènes	2
Accelrys	Leeds, Royaume-Uni; Madison, Wisconsin	Logiciels de modélisation et de visualisation; Méthodes de prédiction de comportement moléculaire; Outils d'analyse et de data mining pour informations chimiques et biologiques	3
Activx Biosciences	La Jolla, Californie	Se concentre sur les activités protéiques dans la cellule (sondes chimiques, quantification de protéines dans conditions normales et pathologiques)	2
Amersham Pharmacia Biotech	Piscataway, New Jersey; Freiburg, Allemagne	Analyse de protéines par gels 2D d'électrophorèse; Identification de cibles thérapeutiques	1
Applied Biosystems	Compagnie multinationale Canada : Streetsville, Ontario	Développement de matériels de technologie de pointe (spectromètre, biochromatographe, logiciels)	1
Aurelium Biopharma	Laval, Québec	Site non-disponible	N/A
Astex Technology Ltd	Cambridge, Royaume-Uni	Structures protéiques; Cristallographie à rayon X	2
Beyond Genomics	Waltham, Massachusetts	Biologie des systèmes : intégration de données moléculaires sur gènes et protéines (utilisation de spectromètres de masse, d'appareils RMN)	3
Biacore	Uppsala, Suède	Développement de technologie d'analyse d'interaction biomoléculaire	2
BioBridge Computing	Lund, Suède	Logiciels d'analyse de données pour spectrométrie de masse	1
BioRad	Hercules, Californie	Logiciels bio-informatiques pour protéomique (gels 2D, spectromètre); Système d'imagerie	2
BioSignal Packard	Montréal, Québec	Génie génétique et cellulaire; Matériel biotechnologique pour étude de molécules	2
BioTools	Edmonton, Alberta	Logiciels pour data mining et visualisation de données biologiques; PepTool -> Analyse de séquences protéiques	2
Biovation	Aberdeen, Écosse	Immunologie; Ingénierie des protéines, particulièrement les anti-corps	2

Bruker Daltonics	Bremen, Allemagne; Billerica, Massachusetts	Technologie de spectrométrie de masse; Logiciels d'analyse	1
Caliper	Mountain View, Californie	Développement de puces LabChip® qui miniaturisent, intègrent et automatisent des processus de laboratoire	1
Caprion Pharmaceuticals	Montréal, Québec	Fonction de protéines dans la cellule; Analyse à haut-débit; Spectrométrie de masse; Bioinformatique	2
Celera Genomics	Rockville, Maryland	Logiciels bio-informatiques (visualiser, parcourir, analyser); Services en ligne de base de données; Analyse protéomique en complément de séquence de génomes	2
Cellzome	Heidelberg, Allemagne	Caractérisation de complexes protéiques cellulaires; Cartographie de voies biochimiques d'interactions de protéines; Spectrométrie de masse, Outils bio-informatiques	3
Ciphergen Biosystems	Fremont, Californie	Puces à protéines	1
Cognia	New York	Logiciels bio-informatiques : base de données pour expression génomique, régulation des gènes	2
Compugen	Jamesburg, New Jersey	Logiciel d'analyse d'images de gels 2D d'électrophorèse; Puces à ADN	1
CuraGen	New Haven, Connecticut	Logiciels bio-informatiques (séquençage de génomes, identification de gènes, protéomique, pharmacogénomique); Base de données en ligne	2
De Novo Pharmaceuticals	Cambridge, Royaume-Uni	Recherche virtuelle de nouveaux médicaments	2
Discovery Partners International	San Diego, Californie	Ensemble intégré de compagnies offrant produits et services visant à accélérer le processus de découverte de médicaments	2
DoubleTwist	Oakland, Californie	Logiciels d'analyse de séquences protéiques (familles de protéines, interactions)	2
Ecopia Biosciences	St-Laurent, Québec	Recherche en génomique (gènes responsables de molécules particulières); Logiciels bio-informatiques d'analyse	2
EraGen Biosciences	Madison, Wisconsin	Plate-forme bio-informatique pour l'analyse de séquence incorporant bases de données et outils d'analyse	1
Europroteome	Hennigsdorf, Allemagne	Analyse protéomique en oncologie (préparation d'échantillon, identification de marqueurs, gels 2D, spectrométrie de masse, cristallographie)	1

GeneData	Bâle, Suisse; Martinsried, Allemagne; Daly City, Californie	Développement de logiciels pour la gestion et l'analyse de données en génomique et protéomique	2
GeneFormatics	San Diego, Californie	Protéomique structurelle (détermination de fonctions et structure de protéines)	3
GeneBio	Genève, Suisse	Gestion des bases de données du SIB; Logiciel d'analyse de gels 2D	1
GeneProt	Evanston, Illinois; Genève, Suisse	Identification, découverte et caractérisation de protéines; Base de données sur les protéines;	3
Genomics Solutions	Ann Arbour, Michigan	Analyse de gels 2D; Intégration de données; Digestion de protéines; Imagerie	1
Genops Bioinformatics	Tujunga, Californie; R&D : Burnaby, BC	Logiciels bio-informatiques d'analyse : algorithmes, statistiques, base de données biologiques	2
Gentech	Antibes, France	Biotechnologie végétale; Interactions protéine-protéine; Logiciels bio-informatiques	2
Glaucus Proteomics	Bunnik, Pays-Bas	Développement d'un processus pour l'identification et la caractérisation de protéines de maladie spécifique	2
GPC Biotech	Martinsried, Allemagne; Waltham, Massachusetts	Cartographie d'interactions protéine-protéine; Purification de protéines et profilage d'expression; Identification et validation de cibles; Identification de récepteurs extracellulaires et de ligands	2
Hybrigenics	Paris, France	Cartographie et analyse d'interaction protéine-protéine; Domaines d'interaction; Outils bio-informatiques	3
IBM Life Sciences Business Unit	White Plains, New York	Développement d'infrastructures pour la recherche; Logiciels d'intégration de données; Gestion de connaissances; Calcul haute performance	1
Informax	Bethesda, Maryland	Logiciels d'analyse de séquence et de structure de protéines; Gestion de bases de données; Analyse de gels 2D	2
Inpharmatica	Londres, Royaume-Uni	Analyse d'information sur la séquence, la structure et la fonction de protéines; Outils chimio-informatiques d'analyse de molécules et de médicaments	3
Inphogene Biocom	Burnaby, Colombie-Britannique	Commercialisation de bases de données sur la génomique et pharmacogénomique de maladies cardio-pulmonaires et neurologiques ainsi que du cancer; Logiciels bio-informatiques d'analyse; Collaboration avec industries pharmaceutiques pour recherche en pharmacogénomique	2

Integrative Proteomics	Toronto, Ontario	Détermination de structures et de fonctions des protéines; Design de molécules thérapeutiques	4
Kendrick Labs	Madison, Wisconsin	Laboratoire offrant services de gels 2D d'électrophorèse	1
Kinetek Pharmaceuticals	Vancouver, Colombie-Britannique	Recherche de médicament agissant sur les voies de signalisation contrôlant des processus dans l'inflammation et le cancer; Biologie moléculaire; Analyses de criblage; Modélisation moléculaire	3
Kinexus Bioinformatics	Vancouver, Colombie-Britannique	Détermination de la structure, la composition et les fonctions des protéines kinases	2
Large Scale Biology	Vacaville, Californie	Base de données (Index des protéines humaines); Marqueurs protéiques; Biopuces à protéines; Développement pharmaceutique (toxicologie, protéines associées aux pathologies)	2
LumiCyte	Fremont, Californie	Marqueurs protéiques; Biopuces; Profils moléculaires de protéines	2
Lynx	Hayward, Californie	Analyse systématique de protéines; Identification et quantification; Spectrométrie de masse; Gels 2D d'électrophorèse	1
Matrix Science	Londres, Royaume-Uni	Spectrométrie de masse et logiciels	1
MDS Proteomics	Toronto, Canada	Identification, sélection et validation de cibles protéiques (récepteurs et voies cellulaires en oncologie); Annotation du protéome humain	2
MediChem	Woodrigde, Illinois	Expression protéique; Cristallisation de protéines; Détermination de structure;	2
Merck Frosst	Kirkland, Québec	Utilisation de gels 2D d'électrophorèse et spectrométrie de masse pour découvrir des marqueurs de maladies; Analyse de données par la bio-informatique	2
Myriad Genetics	Salt Lake City, Utah	Interactions protéine-protéine; Spectrométrie de masse;	1
Nonlinear Dynamics	New Castle, Royaume-Uni; Durham, Caroline du Nord	Logiciels d'analyse de gels 2D et de data mining	1
Oxford Glycosciences	Abingdon Oxon, Royaume-Uni	Identification et isolement de protéines; Corrélation protéine-pathologie;	1
Phylos	Lexington, Massachusetts	Développement d'une plate-forme technologique exploitant les biopuces à protéines	1

Protein Pathways	Los Angeles, Californie	Algorithmes de prédiction de fonctions de gènes et d'intégration de nouveaux gènes dans processus biochimiques connus; Interactions protéine-protéine	2
Proteologics	Rehovot, Israël	Site en construction	N/A
Proteom	Cambridge, Royaume-Uni	Étude des principes des interactions protéine-protéine; Design de protéines; Sélection *in silico* de cibles thérapeutiques	3
Proteome	Beverly, Massachusetts	Outils d'exploitation de connaissances génomiques et protéomiques (bases de données, revue de la littérature)	2
Proteom Factory	Berlin, Allemagne	Découverte de marqueurs de maladie; Criblage à haut-débit; Identification et validation de cibles thérapeutiques	1
Proteome Sciences	Surrey, Royaume-Uni	Découverte de marqueurs de maladie et de cibles thérapeutiques (cancer, désordre neurologique, maladies cardio-vasculaires, obésité, diabète)	1
Proteome Systems	North Ryde, Australie	Logiciels d'identification de protéines; Analyse d'images de gels; Analyse de spectres de masse	2
ProteoMetrics	New York, New York; Winnipeg, Manitoba	Logiciels de support à l'analyse à grande échelle de protéomes; Développement d'Intranet spécifique à la recherche en protéomique	2
Protiveris	Rockville, Maryland	Biopuces à protéines	1
Scanalytics	Fairfax, Virginia	Logiciels d'analyse d'image (imagerie scientifique, analyse de gels 2D, microscopie digitale)	1
SciMagix	Redwood Shores, Californie	Infrastructures et logiciels d'analyse reliés à l'imagerie scientifique	1
SignalGene	Montréal, Québec	Identification de gènes; Identification et validation de cibles thérapeutiques;	1
Signature Bioscience	Hayward, Californie	Catégorisation structurelle et fonctionnelle de cibles thérapeutiques protéiques	2
Structural Bioinformatics	San Diego, Californie	Prédiction de structure des protéines, Spectroscopie à RMN, Cristallographie à rayons X; Logiciels d'analyse	2
Structural GenomiX	San Diego, Californie	Détermination de structure 3D de protéines à l'aide de séquence génomique; Purification de protéines; Cristallisation; Diffraction à rayons X	2
SynX Pharma	Mississauga, Ontario	Identification et caractérisation de protéines	1
Syrrx	San Diego, Californie	Détermination de structure des protéines; Utilisation de la structure pour identification de médicaments	2
Thermo Finnigan	San Jose, Californie	Équipement de spectrométrie de masse et de chromatographie	1

Wita Proteomics	Teltow, Allemagne	Analyse de protéines par gels 2D	1
Xenon Genetics	Vancouver, Colombie-Britannique	Identification de gènes et de cibles thérapeutiques	1
Xerion Pharmaceuticals	Martinsried, Allemagne	Identification de protéines reliées à des maladies; Spectrométrie de masse	1
Zyomyx	Hayward, Californie	Biopuces à protéines	1

Conférences

Quarante-deux conférences reliées à la génomique, à la protéomique et à la bio-informatique ont été recensées dans cette section, couvrant la période de janvier 2000 à avril 2002. Sans représenter un éventail exhaustif des rassemblements scientifiques du globe, cette liste se veut toutefois un échantillon assez complet pour permettre une analyse de la situation actuelle et des enjeux de la recherche. On peut déduire des sujets abordés lors de ces conférences, l'état de la recherche et des thèmes qui sont explorés dans les laboratoires des universités et des industries.

Si on compile les lieux où se déroulent les conférences, on remarque une concentration plus importante des rencontres aux États-Unis. En fait, 26 des conférences énumérées y ont lieu, ce qui confirme la présence d'un pôle majeur en protéomique et en génomique en Amérique du Nord. Des rapports des gouvernements canadien et européens faisaient déjà état de cette situation[14] [15]. Le nombre de conférenciers participant aux colloques américains est aussi plus grand et il y a davantage d'Américains sur les comités scientifiques. Un positionnement rapide des industries américaines dans le domaine ainsi qu'un réseau académique s'étant intéressé relativement tôt à la génomique et à la protéomique leur ont donné des avantages indéniables.

En observant les principaux thèmes abordés, on constate que la principale activité pour le moment en génomique est l'annotation de génomes. C'est-à-dire que l'on cherche à identifier tous les gènes des génomes séquencés, à trouver leur fonction et les protéines qu'ils expriment. Du côté de la protéomique, la recherche se situe davantage du côté de la collecte de données : spectrographie de masse, cristallographie à rayons X et puces à protéines. On développe de nouvelles techniques afin d'étudier les protéines et on cherche à automatiser l'identification

[14] BIOTECanada, *PASSER DE L'INFORMATION À LA CONNAISSANCE : la capacité du Canada en matière de bio-informatique*, Industrie Canada, Ottawa, mars 2000

[15] Sérusclat F, *Génomique et informatique : l'impact sur les thérapies et sur l'industrie pharmaceutique*, Office parlementaire d'évaluation des choix scientifiques et technologiques, Paris, octobre 1999, 204 pages

de l'ensemble des peptides des protéomes. La section de la protéomique de la conférence Beyond the Genome 2001 qui a eu lieu en juin dernier a abordé les thèmes des interactions protéine-protéine, de l'identification et de la quantification des protéines, de l'assignation de fonctions et de l'annotation de protéines. Ces priorités représentent avec justesse l'état des connaissances. Par exemple, en génomique, les techniques de séquençage sont maintenant très rapides et on dispose de plusieurs génomes qui n'attendent que d'être analysés. Le goulot d'étranglement de la génomique se situe au niveau de la compréhension de ces séquences où il devient maintenant essentiel d'analyser les protéines produites par les gènes. À l'opposé, la protéomique n'a toujours pas atteint cette maturité. Les principaux problèmes auxquels s'attaquent les chercheurs se situent au niveau de la caractérisation des protéines. On cherche à connaître plus facilement la séquence des acides aminés, la structure de ces macromolécules et à déterminer leurs interactions.

Ce retard de la protéomique sur la génomique est tout à fait compréhensible. Les protéines sont beaucoup plus difficiles à analyser que les acides nucléiques et les technologies nécessaires pour en sous-tirer des informations sont plus complexes. Les organisateurs de la rencontre du HUPO (HUman Proteom Organization) ont exprimé cette réalité en intitulant leur congrès : « Genes were easy ». De plus, la protéomique est une science de l'ère post-génomique : c'est grâce aux questions restées sans réponse en génomique que la protéomique devient une science déterminante pour les progrès de la biologie.

La reconstitution de systèmes biologiques est aussi plus souvent effectuée en génomique. Plusieurs conférences ont traité de la régulation des gènes et du développement de pathologies génétiques alors qu'on se limite encore aux interactions protéine-protéine quant il s'agit de protéomique. Il y a bien eu une conférence sur les réseaux biochimiques (voir Computation of biochemical pathways and genetic networks), mais celle-ci a tout de même laissé une place importante à la génomique. Cette situation est peut-être due au manque de données complètes sur les interactions protéine-protéine ou à l'absence d'outils performants.

Toutefois, il ne fait aucun doute que la modélisation des voies biologiques deviendra un champ plus important dans les prochaines années. Depuis janvier 2001, on fait maintenant une distinction entre protéomique structurelle et fonctionnelle. La protéomique structurelle concerne l'étude des protéines individuelles (séquence et structure) alors que la protéomique fonctionnelle s'attarde aux fonctions, aux interactions et aux réseaux. L'intérêt de l'industrie pharmaceutique pour ces connaissances n'aura d'autre choix que de guider la recherche dans cette direction.

À propos des applications de la protéomique et de la génomique, la découverte de médicaments ou de cibles thérapeutiques constituent, dans les conférences listées, la presque totalité des objectifs visés pour le moment. Malgré le fait que les découvertes découlant de l'étude des protéomes auront aussi des impacts sur les sciences de l'environnement, sur l'agronomie et en biologie évolutive, la pharmaceutique demeure l'industrie qui y investit massivement. Les espoirs fondés sur la protéomique sont grands. Les protéines pourront servir de marqueurs pour le diagnostic de maladies et pour évaluer leur évolution (monitoring). On pourra aussi prédire la toxicité d'un médicament, ce qui raccourcira certaines étapes de tests lors du développement de médicaments.

Parmi les sujets récurrents, la détermination de la structure des protéines et de la prédiction de leur repliement sont toujours des problèmes qui préoccupent les scientifiques.

En ce qui concerne la bio-informatique spécifiquement utilisée à la protéomique, les bases de données sur les protéines constituent un important sujet de discussion. Plus exactement, on s'interroge sur la façon de gérer les données sur les structures des protéines et leurs interactions. Intégrant les connaissances en génomique, des efforts sont aussi faits pour tenter de cartographier les interactions et pour modéliser les processus biologiques ainsi que les comportements d'une cellule.

Bioinformatics 2002 — 4-7 avril 2002
Society for Bioinformatics in the Nordic Countries — Bergen, Norvège

Genome tri-conference — 23 février-1 mars 2002
Cambridge Healthtech Institute — Santa Clara, CA

Gene functional analysis — 28 février- 1 mars 2002

- SYSTEM BIOLOGY
- MECHANISM OF GENE EXPRESSION
- MONITORING PROTEIN INTERACTION MAPPING IN VIVO AND IN VITRO
- METABOLIC PROFILING
- COMPUTATIONAL CHALLENGES TO GENE EXPRESSION INTERPRETATION
- CHEMICAL GENOMICS

Protein structure and function — 10-14 février 2002
Lorne, Australie

- STRUCTURAL GENOMICS
- PROTEIN LOCALISATION
- PROTEOMICS : GLOBAL VERSUS TARGETED APPROACHES

Human Proteom Project — 9-11 janvier 2002
Cambridge Healthtech Institute — San Diego, CA

Pacific Symposium on Biocomputing — 2-8 janvier 2002
Kaua'i, HW

Protein informatics

12-13 novembre 2001
San Diego, CA

Cambridge Healthtech Institute

- BIOINFORMATICS TECHNOLOGY TO CROSS REFERENCE PROTEIN INFORMATICS WITH GENOMIC DATABASES
- SEQUENCE DATA OF PROTEIN FRAGMENTS BY MASS SPECTROMETRY AND IDENTIFICATION OF THESE FRAGMENTS USING MORE REMOTE RELATIONSHIPS
- CONSTRUCTION AND MANAGEMENT OF INTERNATIONAL PROTEIN STRUCTURAL DATABASES
- PROTEIN PROFILING AND CHARACTERIZATION DATA HANDLING
- DATA THAT ELUCIDATES THE RELATIONSHIP BETWEEN STRUCTURES AND FUNCTIONS OF BIOLOGICAL MACROMOLECULES BY X-RAY CRYSTALLOGRAPHY, LARGE SCALE MOLECULAR SIMULATION AND STRUCTURAL BIOINFORMATICS
- PROTEIN STRUCTURE DATA HANDLING AND STORAGE
- STRUCTURAL BIOINFORMATICS COVERING MOLECULAR MODELING AND DESIGN
- PROTEIN ARRAY AND CHIP DATA HANDLING
- DEVELOPMENT OF NEW ALGORITHMS AND SOFTWARE FOR LARGE SCALE SIMULATION CALCULATIONS BY PARALLEL COMPUTERS
- PROTEIN-PROTEIN INTERACTION DATA AND LIBRARIES
- PROTEIN STRUCTURE DATA DETERMINATION BY X-RAY CRYSTALLOGRAPHY AND DEVELOPMENT OF AUTOMATIC ANALYSIS SYSTEMS
- PROTEIN EXPRESSION DATABASES

Bioinformatics and Bioengineering

4-6 novembre 2001
Washington DC

IEEE

- BIO-MOLECULAR AND PHYLOGENETIC DATABASES, QUERY LANGUAGES, INTEROPERABILITY, BIO-ONTOLOGY AND DATA MINING
- IDENTIFICATION AND CLASSIFICATION OF GENES
- SEQUENCE SEARCH AND ALIGNMENT
- PROTEIN STRUCTURE PREDICTION AND MOLECULAR SIMULATION
- MOLECULAR EVOLUTION AND PHYLOGENY
- FUNCTIONAL GENOMICS
- DRUG DISCOVERY
- GENE EXPRESSION ANALYSIS

The role of protein structure prediction in the post-genomic era

25-27 octobre 2001
Cornell Theory Center — Ithaca, NY

Functional genomics

9-10 octobre 2001
Cambrigde Healthtech Institute — Cambridge, MA

- MECHANISM OF GENE EXPRESSION : ALTERNATIVE SPLICING
- GENOTYPE-PHENOTYPE CORRELATION : ADRESSING THE VALIDATION OF PHENOTYPE BOTTLENECK
- METABOLIC PROFILING AND SYSTEMS BIOLOGY
- FUNTIONAL PROTEOMICS
- TARGET DISCOVERY AND SCREENING

Artificial Intelligence and Heuristic Methods for Bioinformatics

1-11 octobre 2001
NATO Advanced Studies Institute — San Miniato, Italie

International Proteomics Conference "Proteomics at the edge"

30 septembre-4 octobre 2001
Pacific Rim Conference — Canberra, Australie

- PROTEOME / TRANSCRIPTOME
- PROTEOMIC TECHNOLOGY 1 – ANALYTICAL PROBLEMS AND SOLUTION
- PROTEOMIC TECHNOLOGY 2 – ALTERNATIVES TO 2-DE
- PROTEIN MODIFICATIONS
- PROTEIN INTERACTIONS
- INTERPRETING THE PROTEOME – DATA HANDLING AND BIOINFORMATICS
- PROTEOMIC APPLICATION 1 – BIOLOGY (INCLUDING PLANTS) AND HEALTH
- PROTEOMIC APPLICATIONS 2 – BIOTECHNOLOGY AND ECONOMIC WEALTH

"Next Generation" Technologies for High-Throughput Proteomics

24-25 septembre 2001
Global Business Research Ltd San Franscisco, CA

- FUNCTIONAL PROTEOMICS
- PROTEIN ARRAYS / BIOCHIPS
- PROTEIN-PROTEIN INTERACTIONS / MAPPING
- STRUCTURAL PROTEOMICS
- PROTEIN EXPRESSION
- PROTEIN BIOINFORMATICS
- MASS SPECTROMETRY: WHAT'S THE STATE-OF-THE-ART?

Bioinformatics : From inference to predictive models

19-24 août 2001
Gordon Research Conferences Tilton, NH

- SURVEYING MOLECULAR INFORMATION IN CELLS
- BIOPHYSICS OF MOLECULAR INTEREACTION
- NEW BIOLOGICAL INSIGHTS FROM PRIMARY MEASUREMENTS
- DESIGNING DATA RELATIONS FOR OPTIMAL UNDERSTANDING
- DEDUCTION OF BIOLOGICAL NETWORKS
- PREDICTIVE ALGORITHMS AND COMPUTATIONAL METHODS
- MODELING SINGLE CELLS
- SPATIAL AND WHOLE-GENOME MODELING
- CROSS-SCALE MODELING

7th International Congress on Amino Acids and Proteins

6-10 août 2001
Vienne, Autriche

15th Symposium of the Protein Society

28 juillet-1 août 2001
Philadelphie, PA

Meeting of the Federation for Biochemical and Molecular Biology

30 juin-5 juillet 2001
Lisbonne, Portugal

Computation of biochemical pathways and genetic networks

21-22 juin 2001
Bioinformatics and Computational Biochemistry Group — Heidelberg, Allemagne

- TOPOLOGY OF BIOCHEMICAL NETWORKS
- GENETIC NETWORKS
- DYNAMICS OF BIOCHEMICAL PROCESSES
- SIMULATION OF BIOCHEMICAL PATHWAYS

Beyond the genome 2001

17-22 juin 2001
Cambridge Healthtech Institute — San Francisco, CA

Bioinformatics and genome research — 17-19 juin 2001

- ALGORITHMS FOR MINING GENE EXPRESSION DATA
- GENE PREDICTION METHODS KEYNOTE ADDRESS
- PATTERN RECOGNITION IN GENE CHIP DATA
- USING BIOINFORMATICS TO FIND USEFUL TOOLS
- IDENTIFICATION OF REGULATORY PATHWAYS

In silico biology — 19-20 juin 2001

- NETWORKS
- CELLS AND SYSTEMS REPRESENTATION
- TARGET PRIOTIZATION AND DRUG DEVELOPMENT

Proteomics — 21-22 juin 2001

- PROTEIN-PROTEIN INTERACTIONS
- PROTEIN IDENTIFICATION AND QUANTIFICATION
- PROTEIN FUNCTION AND ANNOTATION
- APPLICATIONS

Modeling of protein interactions in genomes

16-19 juin 2001
Ilya Vakser et Sandor Vajda — Charleston, SC

- STRUCTURE PREDICTION AND BINDING SIMULATION
- ENERGETICS AND PROTEIN STRUCTURE-FUNCTION RELATIONSHIPS
- PROTEIN – SMALL MOLECULE INTERACTIONS
- INTERACTING AND NONINTERACTING PROTEINS

Proteomics – Delivering New Routes to Drug Discovery

14-17 mai 2001
IBC Conferences Philadelphia, PA

- FUNCTIONAL GENOMICS
- STRUCTURAL GENOMICS
- PROTEOMICS ANALYSIS TECHNOLOGIES
- STRATEGIC ALLIANCES
- PROTEOMICS APPLICATIONS
- BIOINFORMATICS FOR PROTEOMICS

Proteomics : Science, Technology and Applications

23-24 avril 2001
Barnett International Philadelphia, PA

- DRUG DISCOVERY AND TECHNOLOGY
- THERAPEUTICS
- PROTEOMIC SOLUTIONS

International Conference on Computational Molecular Biology

22-25 avril 2001
Montréal, QC

- GENOMICS,
- MOLECULAR SEQUENCE ANALYSIS,
- RECOGNITION OF GENES AND REGULATORY ELEMENTS,
- MOLECULAR EVOLUTION,
- PROTEIN STRUCTURE,
- STRUCTURAL GENOMICS,
- GENE EXPRESSION,
- GENE NETWORKS,
- DRUG DESIGN,
- COMBINATORIAL LIBRARIES,
- COMPUTATIONAL PROTEOMICS,
- STRUCTURAL AND FUNCTIONAL GENOMICS

Human Genome Meeting 2001 19-22 avril 2001
Human Genome Organization Édinbourg, Écosse

- GENOME COMPLEXITY
- GENE REGULATION
- FROM DNA TO PROTEIN
- FUNCTIONNAL GENOMICS
- MUTATION DETECTION
- GENE EXPRESSION AND REGULATION
- MODEL ORGANISMS
- HIGH-THROUGHPUT TECHNOLOGIES
- COMPLEX TRAITS
- GENETICS OF COMPLEX TRAITS
- MEDICAL GENOMICS
- MAPPING AND SEQUENCING

Protein expression 5-6 avril 2001
Cambridge Healthtech Institute McLean, VA

- EXPRESSION ANALYSIS AND SCIENCE
- HIGH-THROUGHPUT APPROCHES FOR PROTEIN EXPRESSION
- ADVANCES IN PROTEIN PRODUCTION
- IMPROVED PURIFICATION TECHNOLOGIES

Human Proteom Project : "Genes were easy" 2-4 avril 2001
Cambridge Healthtech Institute McLean, VA

- HISTORY
- FUNDING AND FOSTERING
- SCOPE AND SCALE
- FINANCIAL IMPLICATIONS
- CURRENT PROTEOMIC EFFORTS AND LESSONS
- PATENT AND INSURANCE ISSUES
- POTENTIAL SOLUTIONS

Bioinformatics 2001 29 mars-1 avril 2001
Society for Bioinformatics in the Nordic Countries Skövde, Suède

- MOLECULAR EVOLUTION AND GENOME COMPARISONS
- DATA MINING
- LIFE PROCESS MODELLING AND SIMULATION

- ALGORITHMS
- PROTEOMICS ANS EXPRESSION ANALYSIS

Protein Microarray Technology - From Proteomics Discovery to Diagnostics : Expectations and Limitation
21-23 mars 2001
IBC Conferences
San Diego CA

- DIAGNOSTICS : PROTEIN-PEPTIDE-ARRAYS
- PROTEOMICS : PROTEIN-PEPTIDE-ARRAYS

Proteomics : Revolutionnary technologies for drug target and biomarker identification
1-2 mars 2001
IBC Conferences
Bâle, Suisse

- MASS SPECTROMETRY
- AUTOMATION AND HIGH THROUGHPUT
- FUNCTIONNAL PROTEOMICS
- PROTEO-INFORMATICS
- PROTEIN ARRAYS

The New Biology : Technologies for resolving macromolecular communications
24-27 février 2001
Association of Biomolecular Ressource Facilities
San Diego, CA

Integrated bioinformatics : High-throughput interpretation of pathways and biology
24-26 janvier 2001
Cambridge Healthtech Institute
Zurich, Suisse

- COMPUTATIONAL ANALYSIS OF GENE EXPRESSION DATA
- GENE FUNCTION PREDICTION
- GENE PROFILING FOR TARGET IDENTIFICATION
- PROTEIN EXPRESSION
- COMPUTATIONAL GENOMICS

Pacific Symposium on Biocomputing

3-7 janvier 2001
Mauna Lani, HW

- HUMAN GENOME VARIATION: LINKING GENOTYPES TO CLINICAL PHENOTYPES
- NATURAL LANGUAGE PROCESSING FOR BIOLOGY: TERM EXTRACTION, INFORMATION RETRIEVAL, INFORMATION EXTRACTION, AND CORPUS ANNOTATION
- GENOME, PATHWAY AND INTERACTION BIOINFORMATICS
- PHYLOGENETICS IN THE POST-GENOMIC ERA
- HIGH PERFORMANCE COMPUTING FOR COMPUTATIONAL BIOLOGY
- DISORDER AND FLEXIBILITY IN PROTEIN STRUCTURE AND FUNCTION
- DNA STRUCTURE, PROTEIN-DNA INTERACTIONS, AND DNA-PROTEIN EXPRESSION
- STRUCTURES, PHYLOGENIES, AND GENOMES: THE INTEGRATED STUDY OF PROTEIN EVOLUTION
- BIOETHICS, FICTION SCIENCE, AND THE FUTURE OF MANKIND

Research informatics : Optimizing and integrating informatics for drug developments

29-30 novembre 2000
Cambridge Healthtech Institute
Philadelphie, PN

- GAINING KNOWLEDGE FROM GENOMICS SOURCES
- INTEGRATING DATA FOR TARGETS
- INTEGRATING CHEMISTRY AND BIOLOGY

Transcriptome 2000 : De la génomique fonctionnelle à la biologie des systèmes

6-9 novembre 2000
Institut Pasteur
Paris, France

- 25 ANS DE RECHERCHE AVEC LES ADNc
- CLONAGE ET SÉQUENÇAGE D'ADNc
- REGROUPEMENT D'ADNc ET ANNOTATION DU GÉNOME
- L'ANALYSE DU TRANSCRIPTOME
- TRANSCRIPTOMES, PROTÉOMES ET BIOLOGIE DES SYSTÈMES
- APPLICATIONS EN BIOLOGIE, BIOTECHNOLOGIE ET MÉDECINE
- PERSPECTIVES FUTURES
- QUESTIONS ÉTHIQUES LÉGALES ET ÉCONOMIQUES

Protein structure 26-27 octobre 2000
Cambridge Healthtech Institute McLean, VA

- PROTEIN STRUCTURE DETERMINATION
- COMPUTATIONAL PROTEIN STRUCTURE
- PROTEIN FUNCTION ANALYSIS
- TARGET SELECTION AND DRUG DESIGN

Beyond the genome 2000 19-23 juin 2000
Cambridge Healthtech Institute San Francisco, CA

Bioinformatics and genome research

- SOFTWARE ENGINEERING AND DATA MANAGEMENT
- INTERPRETATION OF EXPRESSED GENES
- STRUCTURAL GENOMIC

In silico biology

- DEVELOPING FUNCTIONAL UNDERSTANDING FROM GENETIC DATA: LINKING GENES TO PATHWAYS TO DISEASES
- ORGANS AND SYSTEMS
- CELLS
- TARGET

Proteomics

- PROTEIN-PROTEIN INTERACTION STUDIES
- PROTEIN IDENTIFICATION AND ANALYSIS
- DATABASES AND INFORMATICS
- PROTEOMIC APPLICATIONS

Genes, proteins and computers 26-28 avril 2000
Collaborative Computational Project Chester, United Kingdom

- FUNCTIONAL GENOMICS / PROTEOMICS
- STRUCTURAL GENOMICS
- SEQUENCE ANALYSIS / PHYLOGENY

International Conference on Computational Molecular Biology

8-11 avril 2000
Human Genome Center, Université de Tokyo — Tokyo, Japon

Protein discovery technologies

3-4 avril 2000
Cambridge Healthtech Institute — Boston, MA

- PHAGE DISPLAY IN ANTIBACTERIAL DRUG DISCOVERY
- AUTOIMMUNITY AND INFLAMMATION
- ONCOLOGY
- ANTIBODY AND PEPTIDE LIBRARIES

Protein expression, Protein production

8-10 mars 2000
Cambridge Healthtech Institute — Arlington, VA

- PROTEIN EXPRESSION
- FUSIONS AND PURIFICATION
- PROTEIN ANALYSIS

Integrated bioinformatics

19-21 janvier 2000
Cambridge Healthtech Institute — Zurich, Suisse

- INTEGRATION OF DATA
- GENE EXPRESSION ANALYSIS
- INFORMATICS OF MICROARRAYS
- USE OF INFORMATICS IN PROTEIN SEQUENCES, STRUCTURES AND FUNCTION ANALYSIS
- DATA VISUALIZATION

Pacific Symposium on Biocomputing

4-9 janvier 2000
Oahu, HW

- PROTEIN EVOLUTION AND STRUCTURAL GENOMICS
- PROTEIN STRUCTURE PREDICTION IN BIOLOGY AND MEDICINE
- MOLECULES TO MAPS: TOOLS FOR VISUALIZATION AND INTERACTION IN COMPUTATIONAL BIOLOGY
- MOLECULAR NETWORK MODELING AND DATA ANALYSIS
- DATA MINING AND KNOWLEDGE DISCOVERY IN MOLECULAR DATABASES
- IDENTIFICATION OF COORDINATED GENE EXPRESSION AND REGULATORY SEQUENCES
- NATURAL LANGUAGE PROCESSING IN BIOLOGY
- COMPUTER-AIDED COMBINATORIAL CHEMISTRY & CHEMINFORMATICS
- APPLICATION OF INFORMATION THEORY IN BIOLOGY
- HUMAN GENOME VARIATION: ANALYSIS, MANAGEMENT AND APPLICATION OF SNP DATA

Pacific Symposium on Biocomputing

4-9 janvier 1999
Mauni Lani, HW

- GENE EXPRESSION AND GENETIC NETWORKS
- DATA MINING AND KNOWLEDGE DISCOVERY IN DATABASES
- COMPUTER MODELING IN PHYSIOLOGY : FROM CELL TO TISSUE
- INFORMATION THEORIC APPROACHES TO BIOLOGY
- MOLECULES TO MAPS : TOOLS FOR VISUALIZATION AND INTERACTION
- COMPUTER-AIDED DRUG DESIGN
- PROTEIN STRUCTURE PREDICTION
- DISORDER IN PROTEIN STRUCTURE AND FUNCTION

Les protéines au Canada

L'intérêt de la recherche en protéomique et en génie protéique n'a pas échappé aux scientifiques canadiens. Plusieurs biochimistes, biologistes moléculaires et microbiologistes du Canada investiguent la complexité du monde des protéines. Les problèmes à résoudre sont très nombreux et une grande diversité de projets est possible. De plus, la recherche dans ce domaine est actuellement grandement encouragée car le potentiel des découvertes qui seront faites en étudiant les protéines attire le financement, tant public que privé.

Afin de donner un aperçu de la recherche qui se fait actuellement au Canada, les prochaines sections présentent différents projets ou individus qui s'intéressent d'abord aux protéines, mais aussi à des sujets connexes, la génomique et la bio-informatique.

La première section énumère les chaires de recherche canadiennes dont l'objet de recherche concerne les protéines. Ensuite, une liste des scientifiques canadiens ayant davantage d'implication en bio-informatique a été rassemblée. La troisième section est à propos d'autres chercheurs dont les travaux sont subventionnés par le CRSNG. Un réseau des centres d'excellence en génie protéique, PENCE, existe au Canada et ses thèmes de recherche sont présentés. Finalement, certains projets qui ont été choisis pour être financés par Génome Canada seront décrits.

Chaires de recherche

Le Programme des chaires de recherche du Canada a été créé afin de permettre aux universités canadiennes d'atteindre l'excellence dans le domaine de la recherche et de devenir des centres de recherche de classe mondiale dans l'économie du savoir à l'échelle internationale. Certaines chaires ont été attribuées à des scientifiques qui font de la recherche sur les protéines. Leurs travaux sont variés et ils s'attaquent à des problématiques différentes.

Diverses approches sont possibles pour faire l'étude des protéines. On peut étudier les protéines au point de vue moléculaire, comme on le fait à la chaire de Hue Sun Chan à l'Université de Toronto, qui tente de décrypter la complexité du repliement des chaînes de protéines. Aussi, certains mécanismes cellulaires peuvent être scrutés pour déterminer les protéines qui y sont actives, comme l'équipe David W. Andrews, de l'Université de McMaster, qui analyse la mort cellulaire ou apoptose.

Bien sûr, la détermination de la structure des protéines occupe une place de choix. Ce peut être fait à l'aide du rayonnement synchroton comme Louis Delbaere, de l'Université de la Saskatchewan, qui se sert de cette technologie pour identifier et caractériser les protéines responsables du métabolisme des sucres, ou à l'aide de la cristallisation à rayons X, utilisée par le groupe de Michael James, à l'Université d'Alberta, qui cherche de nouveaux médicaments basés sur la structure de certaines protéines et aussi par Zongchao Jia de l'Université Queen's qui étudie le fonctionnement de protéines uniques qui dictent aux cellules leur comportement. De même que Bryan Sykes, lui aussi de l'Université d'Alberta, se sert de la spectroscopie à résonance magnétique nucléaire pour comprendre la fixation de certaines molécules aux protéines. C'est cette dernière technologie que Lewis Key, de l'Université de Toronto, tente de perfectionner afin d'obtenir la structure très complexe de protéines particulières dont le portrait n'a pas été saisi jusqu'à présent.

David Thomas, de l'Université McGill, s'intéresse quant à lui à l'envoi de signaux biochimiques et à certaines protéines spécialisées, dites chaperons, qui surveillent le repliement d'autres protéines. Il y a aussi Michael Walsh, de l'Université de

Calgary qui analyse le tissu musculaire lisse et ses protéines régulatrices. Toujours à l'Université d'Alberta, Richard Rachubinski emploie une approche protéomique pour comprendre le fonctionnement des peroxysomes en dépistant et analysant rapidement un grand nombre de protéines présentes dans cet organite. À l'Université de Montréal, Michel Bouvier examine une famille spécifique de protéines, des récepteurs qui jouent un rôle de contrôle dans plusieurs processus biologiques.

Sans porter spécifiquement sur les protéines, approfondir la recherche sur la molécule d'ARN et sur son épissage alternatif, comme tentent de le faire Richard Collins, à l'Université de Toronto, et Benoît Chabot, à l'Université de Sherbrooke, pourra aider à comprendre le processus encore mystérieux de synthèse des protéines. Finalement, mentionnons les travaux de Stephen Michnick, de l'Université de Montréal, qui a développé une technique utilisant les interactions entre protéines afin de déterminer la fonction de certains gènes.

Tous les scientifiques mentionnés ci-haut sont titulaires d'une chaire de recherche du Canada. Les profils des titulaires, disponibles sur le site Internet suivant : http://www.chaires.gc.ca, ont été inclus dans le présent rapport dans les pages qui suivent.

David W. Andrews
Chercheur/consultant en biotechnologie
(905)525-9140
provost@mcmaster.ca

Objet de la recherche :	Rôle des protéines dans le contrôle de la mort cellulaire
Importance de la recherche :	Recherche sur le cancer : mise au point de nouveaux médicaments qui visent les membranes; thérapies de bio-ingénierie
Chaire :	McMaster University

Comprendre la mort cellulaire

Imaginez-vous un instant dans un monde où les scientifiques savent pourquoi les cellules saines meurent prématurément ou découvrent comment accélérer la mort des cellules cancéreuses. Que de progrès importants pourraient être réalisés dans le traitement du cancer. Ce n'est que le commencement. Une fois qu'ils auront percé le mystère des cellules à la base du cancer, les scientifiques pourront faire de même pour de nombreuses autres maladies.

La chaire de recherche à McMaster University fournira à David Andrews les ressources nécessaires pour poursuivre ses expériences portant sur les protéines qui régulent la mort cellulaire, aussi appelée apoptose. Elle lui permettra également d'examiner de plus près les modifications qui se produisent dans les membranes cellulaires, ainsi que les anomalies dans la façon dont ces membranes participent à l'échange de protéines à l'intérieur et à l'extérieur des cellules et le rôle qu'elles jouent dans un grand nombre de maladies. La compréhension de ces mécanismes pourrait fournir des renseignements essentiels à la lutte contre la fibrose kystique, la maladie d'Alzheimer et certains types d'encéphalite.

Au cours des 10 dernières années, plus de 30 000 articles ont été publiés sur la mort cellulaire et les moyens de la prévenir. Les chercheurs sont d'avis que l'étude de l'apoptose permettra de faire des progrès déterminants en matière de biologie du développement, d'immunologie et de résistance aux médicaments, en plus d'expliquer l'évolution des altérations causées par les carcinogènes. Les travaux de recherche de M. Andrews sur les liens entre les protéines et la mort cellulaire pourraient apporter d'importants éléments de réponses à la progression des tumeurs et à la résistance acquise aux médicaments.

Le chercheur s'intéresse tout particulièrement aux protéines qui sont ciblées après s'être intégrées dans les membranes. Il s'emploie à déterminer si elles atteignent leur destination prévue ou si elles se dégradent avant d'y arriver. Ses recherches antérieures sur les signaux entre les cellules ont jeté les bases de ses travaux actuels et de ses efforts pour comprendre comment la mort cellulaire est régulée.

Grâce à son travail de consultant aux États-Unis, au Canada et au Royaume-Uni, M. Andrews maintient des liens avec l'industrie biotechnologique, liens qui pourraient s'avérer précieux au fur et à mesure que ses recherches progresseront et que des applications pratiques verront le jour.

Hue Sun Chan
Professeur agrégé, Biochimie
(416) 978-2697
chan@arrhenius.med.utoronto.ca

Objet de la recherche :	Décryptage de la complexité du repliement des chaînes de protéines
Importance de la recherche :	Complément essentiel des projets portant sur le génome pour la mise au point de nouveaux médicaments et de nouvelles thérapies fondés sur ce nouveau savoir
Chaire :	University of Toronto

Décodage du « deuxième code génétique »

Si Hue Sun Chan était scout, il chercherait à obtenir l'insigne du plus haut niveau pour les nœuds. Un insigne indiquant une compréhension de nœuds qui non seulement sont d'une complexité incroyable — ils se nouent et se dénouent eux-mêmes — mais qui sont aussi microscopiques. Chan ne pense toutefois pas aux nœuds que l'on apprend à faire au cours des réunions de la troupe scoute. Il veut comprendre le phénomène que constituent les repliements et les boucles des nœuds essentiels de la vie : les protéines.

On compare souvent la solution du problème posé par le repliement des protéines au décodage du deuxième code génétique. La grande différence, c'est que le décodage du repliement des protéines est encore plus difficile. Le code génétique est constitué d'une séquence de quatre molécules de base disposées de façon normalisée — la fameuse « hélice double ». Les protéines, elles, peuvent comporter plus d'une vingtaine d'éléments constituants (les acides aminés) et présentent une variété apparemment sans fin de formes repliées.

Les chercheurs ont toujours essayé de décoder le repliement par rapport à la séquence des acides aminés d'une protéine. Chan est le chef de file d'une nouvelle méthode d'analyse qui permet d'examiner la structure globale de la protéine comme déterminant de sa forme repliée.

Dans le contexte de la recherche en cours, l'équipe du biochimiste poursuivra cette recherche de base au niveau le plus fondamental, soit en caractérisant la physique du repliement des protéines. La recherche produira un compte rendu quantitatif des forces énergétiques infinitésimales en jeu dans le repliement et le dépliement de ces molécules. On produira des modèles mathématiques du repliement des protéines et l'on testera ensuite ces modèles en fonction de données expérimentales. Chan prévoit documenter ainsi les états de transition du repliement et du dépliement, ainsi que la géométrie de la protéine repliée, sa topologie et sa thermodynamique.

Cette tâche herculéenne réalisée sur une échelle infinitésimale sera rendue possible grâce à l'appui de la chaire de recherche du Canada que reçoit Chan. Les résultats issus de ce travail de pionnier ont de nombreuses applications cliniques et pharmaceutiques au Canada et partout dans le monde. L'efficacité d'un grand nombre de médicaments thérapeutiques dépend de la compréhension exacte du repliement complexe d'une protéine.

Louis Delbaere
Professeur et chef de département, biochimie
(306) 966-4360
louis.delbaere@usask.ca
Accomplissements : Déterminer la structure moléculaire d'une enzyme clé menant à la production de glucose.

Objet de la recherche :	Utilisation d'un rayonnement synchrotron pour déterminer la structure et les fonctions des protéines.
Importance de la recherche :	La structure et les fonctions des protéines pourraient mener à des thérapies novatrices pour guérir certains troubles de santé sérieux tels que le diabète.
Chaire :	University of Saskatchewan

Faire la lumière sur les structures moléculaires

Les protéines constituant les voies des fonctions métaboliques clés, telle la réponse immunitaire, font de bons candidats pour de nouvelles catégories plus efficaces de pharmacothérapies. Cependant, avant que ces médicaments ne puissent être conçus, les chercheurs doivent en savoir plus sur le fonctionnement des protéines par l'étude approfondie de leur structure moléculaire.

À titre de détenteur de la chaire de recherche du Canada en biochimie structurelle, Louis Delbaere étudiera les caractéristiques de plusieurs protéines importantes pour le métabolisme. Le chercheur utilisera beaucoup le premier Centre canadien de rayonnement synchrotron, actuellement en construction à l'University of Saskatchewan. Les intenses sources de faisceaux générées par le synchrotron permettent de déterminer la structure de molécules biologiques complexes plus rapidement qu'avec n'importe quelle autre méthode.

Au cours des dernières années, M. Delbaere a fait partie des partisans les plus fervents du projet synchrotron. Il a aidé les organisateurs du projet à trouver les fonds nécessaires et il a travaillé avec ses collègues à mettre en place un espace de recherche intéressant à Saskatoon. Parmi les molécules qu'il analysera à l'aide du synchrotron se trouve une protéine associée au diabète non insulino-dépendant, l'un des troubles métaboliques communs les plus sérieux au monde.

Bien que les symptômes de ce déséquilibre du glucose dans le sang soient bien connus, toutes les structures des protéines responsables de cette condition n'ont pas encore été étudiées. M. Delbaere a déjà travaillé avec une enzyme bactérienne qui a démontré plusieurs caractéristiques intéressantes. En étudiant l'enzyme mammalienne grâce au synchrotron, le chercheur espère en apprendre plus sur la nature de ces protéines et démontrer le potentiel même de l'enzyme et s'en servir de modèle pour la création d'une thérapie éventuelle qui permettrait de traiter ce type de diabète.

Michael N.G. James
Professeur, Département de biochimie
(780) 492-4550
michael.james@ualberta.ca
Accomplissements : G. Malcolm Brown Award, Canadian Federation of Biological Sciences; détenteur d'un brevet.

Objet de la recherche :	Détermination et modélisation de la structure des protéines.
Importance de la recherche :	Nouveaux traitements médicamenteux pour l'hypertension, et potentiellement pour l'hépatite C; autres traitements médicamenteux.
Chaire :	University of Alberta

Médicaments sur mesure

Les protéines font l'objet de nouvelles recherches de pointe en médecine. De plus en plus, les scientifiques croient que ces macromolécules complexes recèlent la clé de nouveaux traitements médicamenteux et nous permettront de maîtriser la réaction naturelle de l'organisme à l'infection et à la maladie.

Pour percer les mystères des protéines, il faut pouvoir se représenter leur structure. C'est à cet égard que les recherches de Michael James s'avèrent cruciales. Ce professeur de l'University of Alberta est un spécialiste de la cristallographie des protéines par rayons X. Aidé des plus récents progrès technologiques dans ce domaine, il projette un faisceau étroit de rayons X sur des cristaux pour obtenir un motif de diffraction. L'analyse assistée par ordinateur de ce motif lui permet de reconstituer la forme des molécules. Les enzymes étant la cible de nombreux nouveaux médicaments en cours d'élaboration, l'aspect de ces molécules constitue un facteur clé dans le processus de mise au point.

M. James a déjà fait d'importantes découvertes dans le domaine de la biochimie. Son laboratoire a été le premier à produire une représentation graphique des protéinases des virus de l'hépatite A et de la poliomyélite. Ce genre d'analyse structurale est essentielle pour comprendre comment ces enzymes interagissent avec les molécules auxquelles elles s'attaquent, tout comme la compréhension de cette interaction est essentielle pour déterminer comment contrôler l'action des protéines.

L'octroi de cette chaire à Michael James lui permettra d'asseoir sa réputation comme l'un des plus éminents spécialistes de la biologie structurale en Amérique du Nord. Les travaux qu'il continuera de réaliser pour mieux faire comprendre les fonctions des protéines ouvriront la voie à de nouveaux traitements médicamenteux; de plus, cette chaire lui permettra de s'adjoindre des biologistes qui deviendront le pivot du programme canadien de recherche sur les fonctions des protéines.

Zongchao Jia
Professeur agrégé de biochimie
(613)533-6277
jia@post.queensu.ca

Objet de la recherche :	Les protéines d'importance biologique et leur rôle dans le développement du cancer et des troubles immunitaires
Importance de la recherche :	La découverte de la structure et des fonctions des protéines permettra de mieux comprendre le cancer, les maladies cardiovasculaires, le diabète et l'ostéoporose
Chaire :	Queen's University at Kingston

Les protéines en 3D

Des réactions chimiques se produisent profondément à l'intérieur de la structure cellulaire de nos organismes. Ces réactions contrôlent des événements biologiques et font partie du processus qui régule le métabolisme, l'expression génique et la division cellulaire, de même que la croissance, le développement, la locomotion, l'apprentissage et la mémoire.

Importants agents chimiques, les protéines phosphatase et kinase jouent un rôle capital dans ce processus. Les protéines agissent comme des interrupteurs dans le système de communication qui régit la façon dont les cellules acheminent l'information et incitent l'organisme à passer à l'action.

Le professeur Zongchao Jia étudie diverses protéines d'importance biologique. Ces dernières possèdent des propriétés uniques qui peuvent avoir des répercussions sur le développement du cancer. Grâce à ses recherches, le professeur tente de comprendre le fonctionnement des protéines. Si les chercheurs peuvent découvrir la façon dont les protéines dictent aux cellules leur comportement, ils pensent être en mesure de trouver un moyen d'intervenir dans la progression de nombreuses maladies ou même d'en empêcher l'apparition.

Selon le chercheur, le caractère exceptionnel de ces protéines pourrait permettre aux médecins de dépister et de diagnostiquer des affections physiologiques liées à ces protéines. Afin d'examiner de plus près ces protéines nouvellement découvertes, il les cristallisera et aura recours aux rayons X afin de créer des structures en trois dimensions de haute résolution. Cette « photographie atomique » détaillée des protéines facilitera la compréhension de leurs propriétés et apportera d'importants renseignements qui permettront de réaliser des progrès en biotechnologie.

Cette technique appelée la cristallographie prend de plus en plus d'importance et a déjà aidé le chercheur à étudier les protéines antigel. Ces protéines protègent de nombreux organismes en les aidant à survivre à des températures sous le point de congélation. La connaissance de leur mode de fonctionnement fournira des renseignements importants sur les conditions environnementales qui permettent à certaines espèces, comme la tordeuse des bourgeons de l'épinette, de survivre à l'hiver.

La chaire de recherche permettra à M. Jia de perfectionner l'utilisation de la cristallographie et de poursuivre ses importants travaux de recherche, confirmant ainsi sa réputation dans le domaine de la cristallographie des protéines.

Lewis Kay
Professeur, génétique médicale, microbiologie et biochimie
(416) 978-0741
kay@pound.med.utoronto.ca

Objet de la recherche :	Nouvelle technique d'identification des constituants biochimiques des protéines complexes
Importance de la recherche :	Comme ces protéines sont en enroulement complexe, il est plus difficile d'en établir la composition; si on apprend à mieux les connaître, on sera mieux renseigné sur la fonction génétique et ses applications possibles en santé et en médecine.
Chaire :	University of Toronto

Cartographier la carte — les protéines enroulées sont la clé de la fonction génétique

Comme cette carte routière tout usée que l'on jette dans la boîte à gants d'une voiture, les protéines liées à nos gènes n'ont rien d'un paquet bien ficelé. Déjà complexes, les constituants de ces molécules sont en enroulement subtil dans les trois dimensions. La disposition en est non seulement compliquée, mais déconcertante et proprement insaisissable pour un grand nombre d'instruments mis au point en vue de l'étude des structures moléculaires.

Pourtant, il importe que nous comprenions et appréciions la structure de ces protéines. Ce que nous en apprendrons nous fera mieux voir comment les systèmes vivants croissent et se développent, tout en permettant d'améliorer les techniques de protection de ces mêmes systèmes contre la maladie. Il pourrait en résulter d'importantes percées pour l'industrie pharmaceutique, ainsi que l'établissement des bases biotechnologiques de production de nouvelles matières techniques.

Toutefois, si nous nous mettons à explorer les fonctions de ces protéines, il nous faut améliorer les outils à notre disposition. C'est justement ce qu'a fait Lewis Kay. Il a affiné les techniques de résonance magnétique nucléaire (RMN), d'où la possibilité de donner à un système bien établi une meilleure capacité de cerner une configuration moléculaire par ses propriétés magnétiques.

Il n'y a que depuis dix ans qu'il nous est possible d'appliquer les techniques RMN à l'étude de molécules d'une taille extrême comme les protéines génétiques. Ces molécules émettent généralement des signaux en recouvrement, ce qui empêche tout à fait d'en préciser les formes. Grâce à ses travaux novateurs, M. Kay a pu trouver des façons d'« étiqueter » les parties moléculaires d'un intérêt particulier de manière à pouvoir indirectement en dégager la forme complète.

Brian D. Sykes
Professeur et chef, département de biochimie
(780) 492-3357
brian.sykes@ualberta.ca
Accomplissements : Boursier, Royal Society of London

Objet de la recherche :	Systèmes biologiques; fonction des protéines; utilisation de la spectroscopie à résonance magnétique nucléaire
Importance de la recherche :	Compréhension des protéines et comment les médicaments et d'autres substances se lient à elles. Création possible d'un nouveau traitement pour diverses maladies.
Chaire :	University of Alberta

Les processus protéiques

Percées médicales. La soif de découvertes est palpable chez les chercheurs, dans le secteur de la biotechnologie et dans le milieu médical. Sans oublier chez les patients.

De nombreux chercheurs croient maintenant que les protéines donnent le plus d'espoir d'effectuer des percées qui révolutionneront le traitement des maladies.

Depuis 25 ans, soit depuis son retour au Canada après avoir étudié à Harvard, Brian Sykes utilise la spectroscopie à résonance magnétique nucléaire pour étudier la structure et les fonctions des protéines. Ces appareils d'imagerie de haute technologie lui permettent d'examiner la relation entre les protéines et les gènes. L'un de ses projets consiste à comprendre le mécanisme de fixation du calcium qui régit la contraction des muscles. Ces travaux devraient aider les chercheurs à comprendre comment des sensibilisateurs du calcium interagissent avec les protéines cardiaques.

Où cela mènera-t-il? En découvrant comment les protéines fonctionnent, le professeur et d'autres chercheurs espèrent finir par mettre au point des médicaments et d'autres thérapies qui permettront d'inhiber ou de réguler l'activité des protéines qui déclenche des réactions chimiques dans le corps humain.

La présente chaire de recherche permettra au professeur Sykes de se concentrer sur la résolution de problèmes biologiques en s'appuyant sur ses réalisations antérieures. Le nouvel angle? Cette fois, il travaillera avec des protéines complexes et des systèmes plus importants que dans certaines de ses précédentes expériences.

Une grande partie des travaux du professeur de l'University of Alberta visera également à mettre au point de nouvelles techniques de résonance magnétique nucléaire adaptées aux besoins de cette recherche. Les progrès rapides des capacités informatiques et des méthodes de calcul, conjugués à la capacité de cette intelligence artificielle de traiter les grandes quantités de données fournies par la spectroscopie à résonance magnétique nucléaire, ont permis une explosion de l'information sur les structures des protéines. Il est essentiel que M. Sykes reste à l'avant-garde des nouvelles technologies qui sont à sa disposition pour faire avancer sa compréhension de ce domaine crucial.

David Thomas
Professeur et président, Département de biochimie
(514) 398-2973
dthomas@med.mcgill.ca
Accomplissements : Découverte de la machinerie moléculaire des cellules qui pourrait être responsable de diverses affections dégénératives.

Objet de la recherche :	Examen des voies de transmission des signaux et des mécanismes de repliement suivis par les protéines des cellules.
Importance de la recherche :	Les voies biochimiques suivies par les cellules pourraient devenir la cible de nouveaux traitements d'affections.
Chaire :	McGill University

Signaux et chaperons : apprendre le langage des protéines

Les cellules de tous les organismes communiquent régulièrement de l'information sous forme de signaux biochimiques qui, en général, ordonnent des réactions physiologiques internes aux cellules face à des changements avoisinants. De pareils signaux sont en outre associés avec l'apparition de plusieurs genres d'affections inflammatoires et dégénératives comme le cancer et les infections virales et bactériennes. Les chercheurs s'intéressent désormais à la nature et à la fonction de ces mécanismes de signalisation dont l'identification pourrait tracer de nouvelles voies pour traiter ces affections au niveau moléculaire.

À titre de titulaire de la Chaire de recherche du Canada, David Thomas analysera les activités de ce qu'il appelle les machines moléculaires, c'est-à-dire les composants cellulaires qui interagissent sur les plans physique et fonctionnel. Ses compétences seront mises à profit au Centre de génomique et de protéomique de Montréal, une installation de pointe de McGill University dont la mise sur pied devrait être achevée en 2002, de même que dans le cadre des activités de recherche intégrée qui seront effectuées au Complexe des sciences de la vie de McGill.

En outre, ses domaines d'intérêt comptent une catégorie de protéines spécialisées qu'on appelle chaperons qui surveillent le repliement effectué par les protéines pour prendre la forme qui leur permet précisément de remplir leurs fonctions physiologiques. Ces protéines jouent un rôle important dans le cadre de problèmes médicaux comme la maladie d'Alzheimer ou les troubles du système immunitaire, un rôle que pourrait cibler une pharmacothérapie novatrice.

Les travaux de M. Thomas et de ses collaborateurs reposent sur des recherches d'avant-garde portant sur un chaperon moléculaire appelé calnexine qui fait office de mécanisme de contrôle de la qualité ayant pour objet des protéines mutantes repérées dans les cellules. Lorsque ce mécanisme ne fonctionne pas correctement, le mauvais repliement des molécules peut entraîner des affections dégénératives.

Michael P. Walsh
Professeur, biochimie et biologie moléculaire
(403) 220-3021
Walsh@ucalgary.ca

Objet de la recherche :	Le tissu musculaire lisse et ses protéines régulatrices; définition de la fonction des protéines; schématisation de la base moléculaire de l'hypertension et d'autres maladies liées à la fonction cellulaire du tissu musculaire lisse
Importance de la recherche :	Comprendre les processus de signalisation intervenant dans des maladies comme le cancer et l'hypertension. Possibilités de percées pour de nouveaux types de médicaments.
Chaire :	The University of Calgary

Le tissu musculaire lisse

Les biologistes moléculaires croient de plus en plus qu'une meilleure connaissance du processus de communication cellulaire est la clé qui permettra de contrôler la progression de certaines maladies. C'est ce qui a amené Michael Walsh à concentrer ses travaux sur le tissu musculaire lisse. Les parois des vaisseaux sanguins sont formées en grande partie de tissu musculaire lisse dont les cellules régularisent la tension artérielle par leur capacité de contraction et de relâchement, contrôlant ainsi le diamètre des vaisseaux. Une face de la paroi d'un vaisseau sanguin est tapissée d'une couche unique de cellules, appelées cellules endothéliales, qui régularisent la fonction du tissu musculaire lisse. Toute anomalie dans le fonctionnement des cellules endothéliales et du tissu musculaire lisse peut favoriser l'apparition de diverses maladies, dont l'hypertension, le cancer, la coronaropathie, l'insuffisance rénale terminale et le diabète sucré.

À l'University of Calgary, le laboratoire de Michael Walsh se concentre sur les voies de signalisation qui régularisent la contraction des muscles lisses. S'ils trouvent les mécanismes défectueux de ce processus, les scientifiques pourraient tenter de contrôler la progression ou même de prévenir l'apparition de ces maladies. Cette recherche repose principalement sur l'utilisation de technologies associées à l'étude des protéines et sur la façon dont celles-ci interagissent avec les cellules et réagissent aux signaux perçus à leur surface. Comme il s'agit d'un sujet de recherche prédominant en biologie et en microbiologie, l'attribution de cette chaire de recherche aidera à établir Calgary comme centre d'excellence dans ce domaine. Les découvertes de Michael Walsh arrivent à point nommé, compte tenu des progrès réalisés simultanément en technologie et dans le projet du génome humain. La technologie des puces à ADN devrait à la longue permettre aux chercheurs de déterminer les anomalies ou les altérations dans le mode d'expression des gènes. Si l'on peut schématiser le lien entre ces gènes et certaines maladies, il sera possible de mettre au point de nouveaux médicaments et traitements qui stopperont la progression d'une maladie ou qui mettront hors circuit les « interrupteurs » qui l'activent avant qu'elle puisse porter atteinte à l'organisme qu'elle attaque.

Richard A. Rachubinski
Professeur et directeur, département de biologie cellulaire
(780) 492-9868
doug.owram@ualberta.ca

Objet de la recherche :	Protéomique : Dépistage rapide et analyse d'un grand nombre de protéines. L'étude des peroxysomes : structure à l'intérieur d'une cellule qui remplit des fonctions essentielles dans le métabolisme lipidique.
Importance de la recherche :	Certains troubles génétiques, y compris le syndrome de Zellweger et les troubles néonatals, seraient causés par un assemblage anormal des peroxysomes. Mise au point de solutions et de thérapies biomédicales
Chaire :	University of Alberta

Dépistage des protéines

La biologie moléculaire se concentre de plus en plus sur l'assemblage et la fonction des compartiments de membranes à l'intérieur des cellules appelés organelles. Les peroxysomes sont des organelles qui participent au métabolisme lipidique et qui sont essentielles au développement humain normal et à la physiologie.

Le laboratoire de Richard Rachubinski a joué un rôle déterminant dans la découverte révélant que l'assemblage des peroxysomes se produit en plusieurs étapes. En déterminant ces étapes et en établissant les fonctions des peroxysomes, ce professeur de l'University of Alberta a permis d'avoir une meilleure compréhension des anomalies moléculaires qui causent les troubles tels que le syndrome de Zelleweger et les dysfonctionnements néonatals.

L'objectif à long terme de cette recherche complexe est de comprendre les bases génétiques de ces maladies. Les réponses que fournira Rachubinski aideront les scientifiques à élaborer des thérapies géniques ou à mettre au point de nouveaux médicaments afin de contrer ces dysfonctionnements protéiques.

Comprendre les peroxysomes n'est que l'un des aspects de la recherche dite protéomique. La protéomique comporte aussi le dépistage et l'analyse rapide d'un grand nombre de protéines. Une fois que toutes les protéines auront été cataloguées, les biologistes moléculaires espèrent qu'ils auront fourni des éléments clés permettant de réaliser des percées médicales dans le traitement du cancer et des accidents cérébrovasculaires, de même que des maladies cardio-vasculaires, inflammatoires et infectieuses. La protéomique a aussi des applications commerciales directes pour la mise au point de médicaments destinés au bétail, ainsi que de pesticides, de fongicides et d'herbicides. De plus, en travaillant en collaboration avec l'industrie pharmaceutique, Rachubinski et son équipe composée de brillants étudiants diplômés mettent à profit leurs connaissances pour repérer des candidats à la vaccination et de nouvelles cibles pour les médicaments.

Les chercheurs spécialisés en biologie cellulaire ont longtemps débattu la fonction et le développement des peroxysomes – et leurs théories s'opposaient. Avec les preuves qu'il a apportées, Rachubinski a fourni l'évidence réconciliant des observations qui semblaient auparavant contradictoires. En lui attribuant cette chaire de recherche, nous voulons assurer la poursuite de son important travail sous les auspices de l'University of Alberta.

Michel Bouvier
Professeur titulaire, Département de biochimie
(514) 343-6372
bouvier@bcm.umontreal.ca
Accomplissements : Chaire Hans Seyle (1997) Scientist Scholarship (CMRC) 1994

Objet de la recherche :	Étude des RCPG, une famille de protéines qui jouent un rôle clé dans le contrôle de plusieurs processus biologiques
Importance de la recherche :	Améliorera l'efficacité de médicaments existants et permettra l'élaboration de nouveaux médicaments
Chaire :	Université de Montréal

Coup de pouce aux médicaments

La moitié des médicaments d'ordonnance vendus en Amérique du Nord ont pour cible une même famille de protéines, les récepteurs couplés aux protéines G (RCPG). Par leur localisation à la surface des cellules, ces récepteurs jouent un rôle clé dans le contrôle de processus biologiques aussi divers que la neurotransmission, le métabolisme, la croissance cellulaire, les réponses immunitaires et inflammatoires, l'olfaction et la vision. Le fonctionnement des RCPG influence une foule de maladies et si nous les connaissions mieux, il serait possible d'améliorer nombre de médicaments existants et d'en élaborer de nouveaux, plus efficaces.

Michel Bouvier, professeur à l'Université de Montréal, est un des grands spécialistes mondiaux des RCPG. Ses travaux ont contribué au développement de plusieurs des concepts qui forment aujourd'hui la base de notre connaissance de ces protéines. Il s'est notamment fait connaître en remettant en question des notions que la science tenait pour acquises. Il a ainsi ouvert de nouveaux champs de recherche. Ses nombreux contacts dans l'industrie biopharmaceutique montréalaise et mondiale faciliteront les transferts technologiques et l'élaboration rapide de nouveaux médicaments. La chaire de recherche que monsieur Bouvier dirigera étudiera notamment les mécanismes moléculaires qui régissent l'efficacité des médicaments agissant via les RCPG.

Le marché mondial des médicaments qui ciblent les RCPG avait une valeur de 100 milliards de dollars en 1998.

Richard Collins
Professeur, génie médical et microbiologie
(416) 978-3541
rick.collins@utoronto.ca

Objet de la recherche :	Compréhension du comportement de la protéine ARN
Importance de la recherche :	Cette protéine étroitement liée à l'ADN pourrait jouer un rôle comme agent chimique dans la conception d'un grand nombre de nouvelles applications médicales ou industrielles.
Chaire :	University of Toronto

Un retour à la base en biologie

Moins illustre que sa cousine chimique plus complexe, l'ADN, la protéine ARN est pourtant considérée comme une des composantes de base en biologique. En fait, les scientifiques pensent que l'apparition de l'ARN sur terre a frayé la voie à l'ADN appelé à devenir la pierre angulaire biochimique de la vie comme nous la connaissons.

Il faut ajouter que, si l'ADN est bien connu comme molécule porteuse du code génétique, l'ARN joue un rôle primordial dans l'usage que fait l'organisme de cette information. Voilà pourquoi les chercheurs se sont mis à s'intéresser davantage au comportement de l'ARN. Ils essayent de comprendre quelles parties précises de cette protéine complexe sont les sièges de réactions chimiques. Ils s'attachent en outre à ses spires subtiles, ce qui pourrait nous renseigner plus en détail sur la nature des fonctions que soutiennent ces molécules.

Richard Collins a découvert à l'ARN de nouvelles caractéristiques qui nous indiquent que cette protéine pourrait, par les réactions qu'elle entretient, se prêter à des applications médicales ou industrielles. Il reste qu'avant de pouvoir passer aux actes, il faudra répondre à une foule de questions au sujet du lien entre la fonction et la structure de l'ARN.

Comme titulaire de cette chaire canadienne de recherche, M. Collins sera en quête des réponses à donner à ces questions. S'appuyant sur une batterie de techniques d'investigation, ses collègues et lui veulent acquérir des renseignements bien plus fondamentaux sur la composition et la fonction de l'ARN. La polyvalence de la démarche fait partie intégrante de leur stratégie, où ils se guideront sur les résultats déjà obtenus pour orienter les futurs travaux. M. Collins décrit l'aboutissement éventuel de l'entreprise comme rien de moins qu'une véritable mine de détails susceptibles d'ériger l'ARN en un des agents chimiques les plus importants au XXIe siècle.

Stephen Michnick
Professeur agrégé, biochimie
(514) 343-5849
stephen.michnick@umontreal.ca
Accomplissements : Une nouvelle technique pour évaluer rapidement la fonction d'un gène dans une cellule.

Objet de la recherche :	Analyse de la fonction des gènes en étudiant l'interaction des protéines à l'intérieur des cellules.
Importance de la recherche :	En apportant rapidement des indices sur la fonction des gènes, cette stratégie de recherche devrait réduire une partie du travail volumineux qu'est la poursuite de l'analyse du génome humain.
Chaire :	Université de Montréal

Le génome humain réexaminé : le vrai travail commence!

Maintenant qu'ils ont décodé les quelque 3000 séquences constituant le génome humain, les chercheurs doivent déterminer la fonction de ces molécules et le lien complexe qui unit les produits de chaque gène, travail beaucoup plus exigeant. Mis en présence d'un déluge éventuel d'informations, Stephen Michnick a adopté une nouvelle approche expérimentale pour interpréter la fonction des gènes à partir de l'interaction des protéines et des autres molécules biologiques dans les cellules vivantes. Ses travaux apporteront rapidement des indices au sujet de l'évolution du génome, du rôle des gènes et peut-être même une idée de la nature des processus des maladies.

À titre de détenteur de la chaire de recherche du Canada en génomiques intégrées, M. Michnick utilisera cette technique pour déterminer les protéines participant à l'envoi de signaux biochimiques à l'intérieur des cellules et entre les cellules. Cette stratégie unique est applicable à n'importe quelle cellule humaine ou animale ce qui permet une compréhension initiale du fonctionnement de la cellule.

Le chercheur a été acclamé comme un scientifique éminent et a effectué des travaux hautement ciblés qui ont une grande influence sur la communauté scientifique. Son analyse d'un médicament immunosuppresseur a mené à une meilleure compréhension de la base structurelle de l'effet de ce médicament. De même, il a étudié le mécanisme d'activation d'un récepteur d'hormones particulier, une catégorie générale de récepteurs cellulaires que l'on retrouve dans tout le corps.

Benoit Chabot
Professeur
(819) 564-5295
b.chabot@courrier.usherb.ca
Accomplissements : L'un des grands spécialistes mondiaux de l'épissage alternatif de l'ARN; auteur de plusieurs découvertes qui ont changé l'orientation de la discipline ou jeté une nouvelle lumière sur celle-ci.

Objet de la recherche :	L'épissage alternatif de l'ARN et la biologie des télomères
Importance de la recherche :	Les résultats aideront à comprendre comment 40 000 gènes produisent des centaines de milliers de protéines différentes. De plus, une meilleure connaissance de la régulation de la longueur des télomères pourrait mener à de nouvelles stratégies de traitement du cancer.
Chaire :	Université de Sherbrooke

Synthèse de molécules servant de messagers dans le corps humain

M. Benoit Chabot est l'un des très rares spécialistes de l'ARN (acide ribonucléique) de réputation internationale travaillant au Canada. La molécule d'ARN aide l'ADN à créer les protéines. Dans ce processus, l'ARN transmet les « instructions » de l'ADN. C'est la synthèse de cet « ARN messager » ou ARNm qui intéresse particulièrement M. Chabot.

Celui-ci est un spécialiste mondialement reconnu dans le domaine de l'épissage alternatif de l'ARN, processus biologique crucial pour l'expression génétique dans la plupart des espèces animales, y compris l'homme. Étant donné que l'épissage alternatif de l'ARN permet de créer des types de molécules d'ARNm différents à partir d'un même gène, il engendre une diversité de fonctions et de structures protéiques essentielles aux organismes complexes. Comme plus de la moitié des gènes humains subissent l'épissage alternatif de leur pré-ARNm, ce mécanisme détermine l'identité de plus de 90 p. 100 de toutes les protéines humaines.

Les travaux de M. Chabot portent l'une des marques d'excellence les plus enviées, la capacité de voir en avant et de suivre de nouvelles voies lorsque des découvertes imprévues sont faites.

Par exemple, en 1998, il est passé de la recherche de l'étude de l'ARN à des études sur les mécanismes qui contrôlent les télomères (les extrémités des chromosomes). Ses recherches ont été longues, mais les résultats importants puisqu'elles ont permis de découvrir une nouvelle façon d'influencer la longueur des télomères. Comme ceux-ci contrôlent la durée de vie des cellules et que ce contrôle est modifié dans les cellules cancéreuses, la découverte de M. Chabot aura d'importantes conséquences médicales.

M. Chabot se concentre présentement sur une protéine appelée hnRNP A1, qui est abondante dans le noyau des cellules de mammifères en croissance active. En plus de participer au contrôle de l'épissage alternatif de l'ARN, cette protéine joue un rôle important dans la formation des télomères.

Un groupe d'investisseurs a déjà pris un intérêt dans une entreprise qui exploite les découvertes de M. Chabot sur les télomères. Cette entreprise prévoit transformer ces résultats en applications commerciales afin de traiter le cancer et le vieillissement.

Les chercheurs canadiens en bio-informatique

Le développement du trio bio-informatique, génomique et protéomique au courant de la dernière décennie a amené de plus en plus de chercheurs à s'intéresser à ces domaines. Les professeurs de science informatique ont été sollicités pour des problèmes particuliers de la bio-informatique : traitement et stockage de données, algorithmes, intelligence artificielle, logiciels d'analyse, de recherche et de visualisation de données. De leur côté, les biochimistes et les biologistes, à l'aide de nouveaux outils, sont intéressés à tirer des connaissances de toute l'information génomique maintenant à leur disposition.

À travers le pays, les scientifiques universitaires tentent de résoudre des problèmes ou de répondre à des questions jusque là restées sans réponse. Voici une liste de ces personnes qui sont impliquées dans des travaux de recherche impliquant la bio-informatique. Leur nom provient d'une liste disponible sur Internet à l'adresse suivante : http://www.bioinformatics.ca/people/canada.phtml. Le nom des chercheurs dont les travaux sont spécifiques aux protéines sont soulignés en caractères gras.

Ken Barker	Departement of Computer Science – University of Calgary	Données en bio-informatique; Intégration de bases de données; Technologies d'entreposage de données
Anne Bergeron	Lacim et Département d'informatique, UQÀM	Algorithmes en bio-informatique
Anthony Bonner	Computer Science, University of Toronto Molecular Design and Information Technology (MDIT) Centre	Bases de données; Intelligence artificielle; Cartographie et séquençage de génomes
Fiona Brinkman	Molecular Biology and Biochemistry – Simon Fraser University	Pathogénomique et évolution des pathogènes; Localisation de protéines bactériennes intracellulaires; Annotation de génome

Daniel Brown	Computer Science, University of Waterloo	Génomique comparative; Algorithmes d'analyse de séquences
Gertraud Burger	Biochimie – Université de Montréal, Organelle Genome Megasequencing Project	Séquençage de génome d'eucaryotes; Génomique comparative et évolutive; Base de données intégrées
Gregory Butler	Computer Science, Concordia University	Architecture logiciel, Orienté-objet, Systèmes "knowledge-based", Calcul scientifique
Carol Cass	Department of Oncology – University of Alberta	Biologie des membranes; Transport et métabolisme des nucléotides; Chimie du cancer
Hue Sun Chan	**Department of Biochemistry – University of Toronto**	**Énergétique du repliement des protéines**
Anne Condon	Computer Science – UBC	Design et analyse d'algorithmes; Calcul biomoléculaire; Prédiction de la structure secondaire des ARN
Jamie Cuticchia	Research Institute for the Sick Children	Bases de données génomiques; Intégration de bases de données
Doug Demetrick	Faculty of Medicine – University of Calgary	Gènes de régulation de cycles cellulaires dans les cancers du sein et du pancréas
Nadia El-Mabrouk	Département d'Informatique – Université de Montréal, Laboratoire de Biologie Informatique et Théorique (LBIT)	Réarrangement de gènes dans le génome; Prédiction de motifs biologiques structurés
Michael Ellison	Department of Biochemistry – University of Alberta	Système régulateur de la cellule : Réseau protéolytique dépendant à l'ubiquitine
Patricia Evans	Computer Science – University of New Brunswick	Algorithmes de comparaison des ARN
Michael Fellows	Computer Science – University of Victoria	Complexité paramétrées et design d'algorithmes bio-informatiques
Janice Glasgow	Computer Science – Queen's University	Analyse de scènes moléculaires; Analyse de Microarray

Mark Glover	**Department of Biochemistry – University of Alberta**	**Étude de protéines impliquées dans le contrôle d'expression génique et dans la réparation d'ARN**
Randy Gloebel	Department of Computing Science – University of Alberta	Représentation de savoir; Intelligence artificielle
Lev Goldfarb	Computer Science – University of New Brunswick	Représentation symbolique; Data mining; Bases de données biomoléculaires
Paul Gordon	Dalhousie University	Visualisation de données biologiques
Roy A Gravel	Faculty of Medecine – University of Calgary	Maladies métaboliques génétiques
Russ Greiner	Computing Science – University of Alberta Bioinformatics Research Group	Intelligence artificielle; Machine learning
Mike Hallett	Computer Science – McGill University, McGill Center for Bioinformatics	Algorithmes de génomique comparative; Évolution de génome; Développement de logiciels
Christopher Hogue	**Samuel Lunenfeld Research, MDS Proteomics**	**Biomolecular Interaction Network Database (BIND); Repliement de protéines; Outils de visualisation 3D;**
Thomas Hudson	Montreal Genome Center, McGill University	Génomique; Génétique de traits complexes; Puces à ADN
Michael James	**Department of Biochemistry – University of Calgary**	**Mécanisme hydrolytique de protéinases**
Frank R Jirik	Faculty of Medicine – University of Calgary	Répresseur de tumeurs
Igor Jurisica	Ontario Cancer Institue – University of Toronto	Support décisionnel; Cristallisation de protéines à haut-débit; Identification de protéines
Ben F Koop	Biology – University of Victoria	Séquençage à grande échelle; Analyse de séquences; Analyse phylogénétique

Anthony J Kusalik	Computer Science – University of Saskatchewan	Modélisation et prédiction de structure d'ARN
Clement Lam	Computer Science – Concordia	Recherche combinatoire à grande échelle; Algorithmes
Franz Lang	Biochimie – Université de Montréal, Organelle Genome Megasequencing Project	Génomique comparative et évolutionniste; Modélisation comparative d'ARN; Assemblage et annotation de génomes; Bases de données; Phylogénétiques moléculaires
Gilles Lapointe	Biochimie – Université Laval	Génétique moléculaire et bio-informatique; découverte de gènes (plantes, foresterie, humain)
Peter Lee	Montreal Genome Center, Université McGill	Analyse de données des microarray; Réseau d'expression génique; Classification fonctionnelle; Visualisation d'informations
Roger Lévesque	Biologie médicale –Université Laval	Génomique fonctionnelle; contrôle moléculaire d'agents pathogènes
Ming Li	Mathematics – University of Waterloo	Logiciels bio-informatiques
Liang Li	**Department of Chemistry – University of Alberta**	**Spectrométrie de masse analytique; Caractérisation de macromolécules, Analyse protéomique**
François Major	Département d'Informatique – Université de Montréal Laboratoire de Biologie Informatique et Théorique (LBIT)	Modélisation tri-dimensionnelle d'ARN
Vladimir Makarenkov	Département des sciences biologiques – Université de Montréal	Reconstruction d'arbres phylogénétiques; Génomique comparative; Algorithmes et analyse statistique; Développement logiciel
Jean Morissette	Centre de recherche du CHUL	Génétique; Localisation de gène; Expression de gène

Michael Murphy	**Microbiology and Immunology – UBC**	**Comparaison structurelle de familles de protéines; Identification de réseaux métaboliques**
Raymond Ng	Computer Science – UBC	Data mining; Analyse comparative; Calculs parallèles à haute performance
Francis Ouellette	Center for Molecular Medecine and Therapeutics – UBC	Bases de données; Identification de gènes, Biomolecular Interaction Network Database (BIND); Génomique comparative; Pathogénomique
Frédéric Pio	**Biochemistry – Simon Fraser University**	**Apoptose de cellules cancéreuses; Caractérisation de structure 3D de complexes d'interactions protéine-protéine et protéine-ADN**
Richard Pon	Faculty of Medecine – University of Calgary	Synthèse chimique de l'ADN, ARN et oligonucléotides
Gil Privé	Ontario Cancer Institute – University of Toronto	Biologie structurale; Protéines membranaires; Identification moléculaire
Peter Rehse	CHUL Research center – Université Laval	Biologie structurale d'enzyme de répression de stéroïde
Stephen Robbins	Faculty of Medecine – University of Calgary	Signaux extracellulaires transmis au noyau pour contrôles des processus biologiques tels la croissance et la différenciation des cellules eucaryotes
David Sankoff	Mathématique – Université de Montréal	Algorithmes d'analyse de génomes; Phylogénie
Christoph W Sensen	Faculty of Medecine – University of Alberta	Annotation automatique de génomes et outils bio-informatiques
Brian Sykes	**Department of Biochemistry – University of Alberta**	**Structure, dynamique et fonction de protéines**
Peter Tieleman	**Department of Biological Sciences – University of Calgary**	**Protéines membranaires; Modélisation; Simulation moléculaire**
Marcel Turcotte	Computer Science – University of Ottawa	Découverte automatisée de motifs structuraux

Projets de recherche subventionnés par le CRSNG

Le Conseil de recherches en sciences naturelles et en génie est un organisme canadien qui subventionne un grand nombre de projets scientifiques. La protéomique, la génomique et la bio-informatique est un domaine où la recherche est actuellement grandement encouragée. Le potentiel des retombées de la recherche effectuée dans ce secteur est désormais reconnu et le CRSNG a développé plusieurs programmes pour soutenir les équipes universitaires.

Par exemple, des subventions particulières sont destinées aux projets de recherche en génomique. Dans le cadre de ce programme, une aide financière peut être accordée à des projets dont les sujets d'études se rapportent à la bio-informatique, à la génomique fonctionnelle, à la cartographie et au séquençage des génomes, à la détermination du patrimoine génétique, à la protéomique et au développement de technologie nécessaire à la recherche en génomique. De même, le génie protéique a été identifié comme un des domaines-cibles pouvant être l'objet d'un projet stratégique financé par le CRSNG.

Parmi les projets retenus dans la dernière année par le CRSNG, certains concernaient les protéines. Ces projets proviennent de plusieurs catégories : génomique, génétique moléculaire et du développement, et biologie cellulaire. Cette section du rapport présente une description des travaux des chercheurs responsables de ces projets, tirée de leur site web. Professeurs de biologie, de biochimie, de médecine, de microbiologie, de biologie moléculaire, de chimie et même de physique, la description de leurs intérêts et de leurs objectifs de recherche nous aide encore une fois à saisir l'étendue des possibilités des projets sur les protéines. Avant de présenter chacun des projets, en voici un résumé.

Carole Beaulieu et François Shareck s'intéressent aux streptomycètes, des bactéries retrouvées dans le sol qui participent à la dégradation de substances complexes. Les deux chercheurs font une étude protéomique du micro-organisme : Mme Beaulieu travaille sur des spécimens virulents qui causent la gale commune de la pomme de terre; M Shareck fait une étude moléculaire de la structure d'enzymes particulières

synthétisés par les streptomycètes afin d'en tirer des applications biotechnologiques. Les recherches de Normand Brisson portent sur l'expression des gènes chez les végétaux lors de leurs réponses de défense. Il étudie les protéines végétales qui jouent un rôle dans les changements métaboliques qui activent certains gènes lors de réactions de protection.

Avec une approche protéomique et génomique, le professeur Majambu Mbikay tente d'élucider le rôle de l'insuffisance de certains types de protéines qui interviendraient dans l'obésité et l'infertilité. Le groupe de Michael Siu veut améliorer les techniques d'analyse protéomique et pour y arriver, il élabore de nouvelles méthodes automatisées pour le séquençage et l'identification de protéines.

D'autres chercheurs examinent les étapes relativement incomprises menant à la synthèse des protéines à partir de l'ADN. Devakanand Mangroo étudie les mécanismes de transport de l'ARN de transfert du noyau vers le cytoplasme. Paul Romaniuk et son équipe se concentrent sur les liaisons entre l'ARN ribosomal et les protéines ainsi que sur la régulation du processus de transcription par les protéines et des molécules appelées doigts de zinc chez les eucaryotes. Kevin Wilson de son côté s'intéresse à la régulation du processus de traduction chez les bactéries par les interactions entre les ribosomes et l'ARN messager.

Afin de mieux comprendre la nature des protéines, Frédéric Pio étudie leur structure pour reconnaître les spécificités des domaines qui régissent les interactions entre les protéines, principalement les interactions impliquées dans l'apoptose et le cancer. Toujours en lien avec le cancer, l'équipe de Richard Blouin caractérise une protéine de la famille des kinases qui joue un rôle dans la croissance cellulaire. Jim Lepock s'intéresse quant à lui à la structure et aux interactions moléculaires des protéines dans des conditions particulières. Plus spécifiquement, il étudie leur dénaturation lors de chocs thermiques, il cherche à détecter les transitions et à déterminer la dynamique des molécules sous l'effet de la chaleur.

Étudier les fonctions des protéines n'a pas toujours pour but de figurer le rôle de protéines nouvellement identifiées, un des espoirs du génie protéique est aussi de pouvoir assembler des molécules qui accompliront une tâche bien précise. Pour atteindre cet objectif, Thor Borgford cherche à comprendre les principes qui relient la structure à la fonction des protéines. Il manipule des gènes afin qu'ils génèrent des protéines inconnues qu'il étudie par la suite. Les travaux d'Arthur Szabo sont semblables.

Au laboratoire de Richard Béliveau, une des problématiques que l'on approfondit est le rôle de la P-glycoprotéine dans la barrière sang-cerveau qui contre l'action de la chimiothérapie lors de cancer du cerveau. Finalement, les travaux de Nils Petersen concernent le comportement des membranes cellulaires et les macromolécules pouvant interagir avec celles-ci. Ceci signifie d'étudier les protéines qui interviennent dans la communication entre cellules, la transduction de signal, l'adhésion et le mouvement de cellules ainsi que la respiration.

La diversité des projets qui ont été brièvement présentés ci-haut illustre plusieurs réalités de l'étude des protéines. Premièrement, on remarque la versatilité des fonctions des protéines. Impliquées dans presque tous les processus se déroulant dans les organismes vivants, les protéines assument différentes responsabilités qui sont réalisables par les innombrables possibilités de configuration de leur structure. On peut étudier une enzyme qui métabolise un glucide, un récepteur à la surface d'une membrane cellulaire, une hormone qui envoie des signaux ou même une protéine régulatrice qui contrôle l'expression génétique à l'intérieur du noyau cellulaire.

Il est aussi important de bien faire la distinction entre la protéomique et le génie protéique. Ce sont des approches distinctes qui correspondent à des problématiques différentes. La protéomique peut se définir comme l'analyse systématique de l'expression protéique d'une cellule ou d'un tissu. C'est ce que plusieurs compagnies font présentement en essayant d'identifier et de caractériser le plus grand nombre possible de protéines à l'aide de différentes technologies, telles la spectrométrie de masse. L'emphase est mise sur le protéome, c'est-à-dire sur la

reconnaissance des protéines qui le composent et les interactions entre ses macromolécules. De l'autre côté, le génie protéique est davantage à la recherche de relations entre la structure et les fonctions des protéines, impliquant la biochimie, la génétique moléculaire et la biophysique. Il s'agit d'utiliser les connaissances de la structure et des fonctions des protéines pour faire le design de nouvelles thérapies, des vaccins, des outils de diagnostic, des enzymes, des produits industriels et agricoles. On essaiera de comprendre pourquoi deux protéines interagissent ensemble, quelles particularités de leur structure leur permettent de se lier. Le problème du repliement des protéines est aussi un des intérêts du génie protéique.

Les termes protéomique et génie protéique sont relativement récents, la limite qui les sépare n'est pas toujours aussi bien tranchée et les scientifiques ont chacun leur propre définition. De plus, des avancements dans l'un des domaines a immédiatement des répercussions sur l'autre. Par exemple, la protéomique peut isoler une nouvelle protéine dont la structure énigmatique intéressera le génie protéique. À l'opposé, l'information acquise par la recherche en génie protéique sur les fonctions d'une famille de protéines facilite la catégorisation des scientifiques oeuvrant en protéomique.

Le centre d'intérêt de la recherche peut être aussi différent. Certains s'intéresseront au protéome entier d'une bactérie alors que d'autres étudieront quelques protéines spécifiques intervenant dans des contextes particuliers. Le but peut être de comprendre un comportement d'un certain type de cellule comme l'apoptose ou comme la prolifération cellulaire ou bien se restreindre au rôle d'une protéine active dans un de ces comportements.

Les champs de recherche sont nombreux. Voici les fiches des chercheurs qui ont été brièvement présentés et dont les travaux se distinguent entre eux.

Projets en génomique 1999-2000

Carole Beaulieu
Professeure agrégée en biologie
(819)821-8000 poste 2997
c.beauli@courrier.usherb.ca

Objet de la recherche :	Protéomes de Streptomyces scabies et S coelicolor
Chaire :	Université de Sherbrooke

Thème de recherche

- Microbiologie-biotechnologie : phytopathologie

Objectifs de ses recherches

Ses travaux visent à comprendre les mécanismes de virulence des actinomycètes causant la gale commune de la pomme de terre et à identifier les facteurs influençant les différentes étapes de la pathogénèse. Une meilleure connaissance de la biologie des agents pathogènes permettra d'entreprendre des programmes adaptés de lutte contre la maladie.

D'autres actinomycètes agissent comme agents de lutte biologique et protègent les cultures végétales contre différents agents pathogènes. Ses travaux visent à élucider les mécanismes associés à la protection des plantes.

Projets de recherche en cours

- Gale commune
- Streptomycètes
- Pomme de terre
- Interactions plante-micro-organisme
- Lutte biologique

PHYTOPATHOLOGIE ET ÉCOLOGIE MICROBIENNE

Les actinomycètes présentent de nombreux intérêts au niveau industriel, médical et environnemental. Ces bactéries sont abondantes dans le sol et ce sont d'ailleurs elles qui donnent à la terre son odeur caractéristique. Grâce à leur capacité à dégrader des substances complexes, les actinomycètes jouent un rôle essentiel dans l'écologie des sols. De plus, la production d'antibiotiques et d'enzymes lytiques par les actinomycètes influence les interactions entre les micro-organismes du sol. Certaines espèces d'actinomycètes sont, par contre, nuisibles en agriculture, car elles sont responsables de la gale commune de la pomme de terre. Dans le laboratoire de la professeure Beaulieu, on s'intéresse au rôle des actinomycètes lors du compostage de la matière organique, à l'utilisation des actinomycètes comme outil de lutte biologique et on étudie différents aspects des mécanismes de virulence des actinomycètes phytopathogènes.

source : http://www.usherb.ca/SCES/BIO/beacar.html

Normand Brisson
Professeur titulaire en biochimie
(514)343-6984
brison@bcm.umontreal.ca

Objet de la recherche :	Clonal selection and mapping of the functionnal relationships among plant genes with a protein fragment complementation strategy, en collaboration avec Stephen Michnick
Chaire :	Université de Montréal

Principaux intérêts de recherche

ACTIVATION DE L'EXPRESSION DES GÈNES PENDANT LA RÉPONSE DE DÉFENSE CHEZ LES PLANTES

Les plantes ont élaboré au cours de leur évolution des mécanismes très efficaces de résistance aux maladies. Ainsi, lors d'une infection, il se produit des changements métaboliques importants chez la plante qui ont pour effet d'arrêter la progression du pathogène. Son groupe s'intéresse aux phénomènes reliés à l'induction de ces changements métaboliques. En particulier, il cherche à comprendre les mécanismes qui mènent à l'induction de la transcription de gènes spécifiques à la réponse de défense. Il a démontré que la transcription du gène PR-10a, un gène dont l'expression est induite chez plusieurs plantes lors de la réponse de défense, est liée à la fixation du facteur nucléaire PBF-1 à une petite région du promoteur de PR-10a. De plus, la fixation de PBF-1 à l'ADN est elle-même contrôlée par un mécanisme de phosphorylation/déphosphorylation, démontrant l'action de protéines kinases et de protéines phosphatases dans l'activation de PR-10a. Ainsi, il a récemment démontré l'implication d'un homologue de la protéine kinase C dans l'activation de la réponse de défense. Il poursuit présentement les travaux sur la caractérisation des voies de transduction du signal et des mécanismes d'activation de la transcription pendant la réponse de défense chez les végétaux.

Article sur Normand Brisson à l'adresse suivante :
http://www.forum.umontreal.ca/numeros/2000_2001/forum_01_05_07/article10.html

source : http://www.bcm.umontreal.ca/FMPro?-db=bottin.fp3&-lay=Detail&-format=prof.cdm&ID=B000018&-find

Majambu Mbikay
Professeur associé en médecine
mmbikay@ohri.ca

Objet de la recherche :	Elucidating the roles of proprotein convertases in obesity : a genomics and proteomics combined approach
Chaire :	Université d'Ottawa

Intérêts de recherche

- Obésité
- Diabète
- Reproduction
- Infertilité
- Immunocontraception

Principales activités de recherche

This laboratory studies the physiological, biochemical and cellular consequences of genetic deficiencies of proteolytic maturation of precursor proteins due to inactivation of proprotein convertase (PC) genes in cultured cells and in transgenic animals.

Project 1. Infertility due to lack of the testicular PC4.

Findings:

- PC4-deficient male mice lack the active form of testicular prohormones and sperm cell-surface proteins needed for sperm capacitation and egg-binding.
- In the female, PC4 deficiency also causes a macrophage dysfunction that adversely affects fertility.

Current research objectives :

- to identify other spermatozoon PC4 substrates and examine their function in fertilization;
- to analyse the effects of the macrophage dysfunction on intraovarian cytokine production, folliculogenesis, ovulation as well as embryo implantation.

Project 2. Obesity and hypogonadism due to lack of the neuroendocrine convertase PC1.

Findings :

- PC1-nullizygosity is embryonic-lethal in mouse whereas hemizygosity confers a survival advantage during embryogenesis.

Current research objectives :

- to rescue of the lethal phenotype by genetic complementation;
- to analyse the processing of precursors to appetite and satiety neuropeptides as well as body mass regulation and fertility in the surviving mice.

Project 3. Adipocyte, cholesterol and fatty acid metabolic dysfunctions induced by convertase inhibitors.

Findings:

- 3T3-L1 mouse preadipocytes expressing a general proprotein convertase inhibitor are resistant to adipogenic signals in culture.

Current research objectives:

- to elucidate the cellular pathways that are impaired due to lack of convertase activities
- to validate the role of proprotein convertases imfor adipogenesis by targeted expression of the inhibitor in the adipocytes of transgenic mice;
- to develop a genetic model for the study of the biological functions of the newly discovered convertase SKI-1. This enzyme activates transcriptional factors that regulate expression of genes involved in the cellular uptake and biosynthesis of cholesterol and fatty acids.

source : http://www.lri.ca/profiles/mbikay.htm

François Shareck
Professeur en microbiologie
(450)686-5501
Francois.Shareck@inrs-iaf.uquebec.ca

Objet de la recherche :	Étude protéomique de streptomyces
Chaire :	Institut Armand-Frappier INRS

Les streptomycètes sont des bactéries saprophytes du sol qui sécrètent toutes les enzymes nécessaires à la dégradation de la matière lignocellulosique telles les xylanases, les cellulases et les mannanases. Les gènes codant pour la synthèse de ces protéines extra-cellulaires ont été clonés et leur séquence nucléotidique a été établie. Ceci permet d'étudier au niveau moléculaire la structure de ces enzymes ainsi que la régulation de leur synthèse afin de développer différentes applications biotechnologiques.

Régulation du complexe xylanolytique de *Streptomyces lividans*

Ce microorganisme sécrète trois xylanases : une acétyl-xylane estérase et deux arabinofuranosidases, afin de dégrader efficacement le xylane en oligosaccharides assimilables. Les gènes respectifs forment un régulon discontinu et chacun est soumis à une régulation spécifique. Au cours de la croissance, l'induction des différents gènes est séquentielle et en fonction de l'apparition des produits d'hydrolyse du xylane issus de l'activité enzymatique. Cette régulation particulière est suivie par l'analyse des ARN messagers spécifiques ainsi que par la présence des enzymes sécrétées dans le milieu. Ce domaine de recherche est présentement subventionné par le CRSNG (dépenses courantes).

Étude de la répression catabolique chez *Streptomyces lividans*

La production de xylanases par *S. lividans* est soumise à une forte répression catabolique exercée par le glucose. Afin d'élucider ce phénomène, d'une part le Pr Shareck recherche des protéines qui se lient à l'ADN des régions promotrices de ces gènes, dans le but de cloner le gène correspondant par des expériences de génétique inverse. D'autre part, il effectue des travaux de mutagénèse dirigée au niveau des promoteurs, afin d'obtenir des mutants non-réprimés.

Structure moléculaire et fonction de la xylanase A de *Streptomyces lividans*

Un vaste projet de mutagénèse dirigée a permis de produire plus de 25 xylanases A mutantes. Chacune possède des propriétés biochimiques modifiées telles que le mécanisme d'hydrolyse, les constantes d'affinité pour le substrat, de vélocité, etc... La cristallisation et la détermination de la structure tridimensionnelle de l'enzyme a été établie en collaboration avec l'Université d'Alberta. Ceci a permis d'approfondir les connaissances quant à la fonction des différents acides aminés essentiels composant le site catalytique.

Analyse de la structure moléculaire de l'opéron *bxl* de *Streptomyces lividans*

Le séquençage de l'opéron *bxl* nous a permis d'identifier toutes les protéines nécessaires à l'utilisation des disaccharides, produits de l'hydrolyse du xylane par les xylanases, soient la ß-xylosidase ainsi que le système de trois perméases. De plus, deux gènes de régulation codant la biosynthèse de protéines se liant à l'ADN ont été découverts dans le voisinage adjacent. Ces protéines ont été purifiées et sont présentement à l'étude à l'aide de la technique d'altération de la mobilité électrophorétique ainsi que celle de protection de l'ADN à la DnaseI. Ces travaux établiront les bases moléculaires de la régulation du complexe xylanolytique tant au niveau de la répression catabolique exercée par le glucose qu'au cours de l'induction par le xylane. Enfin, les gènes de régulation sont étudiés spécifiquement par des expériences de dislocation génétique afin d'identifier leur rôle respectif.

source : http://www.inrs-iaf-microbiotech.uquebec.ca/Shareck_F_Recherche.html

Michael Siu
Professeur en chimie
(416)650-8021
kwmsiu@yorku.ca

Objet de la recherche :	Novel and automated methodologies for protein sequencing and identification ; their application to proteome analysis
Chaire :	York University

The Siu group uses mass spectrometry to perform research in the bioanalytical and biophysical area. Instrumentation includes prototypes of the ABI SCIEX QSTAR and API 3000, API 365, API III, TAGA 6000, Voyager DE STR, and separation apparatus such as HPLC, capillary electrophoresis instrument and ion mobility spectrometers. High-performance multi-processor computers are available for calculating energetics and structures of the ions under study. The theoretical areas of research are carried out in collaboration with Professor A.C. Hopkinson. Biological projects are performed in collaboration with Professors R.E. Pearlman and J.C. McDermott of the Department of Biology.

Intérêts de recherche

- Collision cross sections of peptide ions: their measurement using ion mobility and comparison with calculated values from candidate structures
- Proton affinities and metal-ion affinities of peptides using the kinetic method and threshold collision-induced dissociation measurements and calculations using density functional theory
- Mechanisms of protonated and metallated peptide ion fragmentation
- Molecular radical cation of peptides.
- Novel mass spectrometric hardware.
- Protein identification, sequencing and quantification

source : http://www.chem.yorku.ca/profs/siu/home.html

Projets en génétique moléculaire et développementale 2000-2001

Devakanand Mangroo
Professeur en chimie et biochimie
(519)824-4120 poste 3432
mangroo@chembio.uoguelph.ca

Objet de la recherche : Mechanism of protein initiation without formylation of the innitiator methionyl-tRNA

Chaire : Guelph University

Intérêts de recherche

Ins and Outs of Nucleocytoplasmic tRNA Export

In eukaryotes, the DNA replication and transcription machineries are located in the nucleus whereas the protein synthesis apparatus is confined to the cytoplasm. This compartmentalization necessitates the bi-directional transport of a variety of macromolecules between the nucleus and cytoplasm. Regulation of the nucleocytoplasmic transport process is an important strategy used by eukaryotic cells to control many vital biological processes such as gene expression, and cell division and differentiation. Thus, components of the nuclear import/export processes are considered as potential targets for the development of therapeutics against cancers.

A major goal of our research is to elucidate the mechanism by which tRNA is transported from the nucleus to the cytoplasm. Genetic and biochemical studies established that nucleocytoplasmic export of tRNA is protein-mediated and involves components which are distinct from those required for export of other classes of RNAs. However, the details are poorly understood and only a small number of these proteins have been identified. We use the yeast *Saccharomyces cerevisiae* as a model eukaryote to identify and characterize proteins involved in the nuclear tRNA export process. The goal is to utilize this information together with genomic analyses and our yeast system to obtain the identities of the mammalian proteins. To identify and characterize the *S. cerevisiae* proteins, we use a multidisciplinary approach, combining genomic analyses with genetic, biochemical and cell biological techniques.

Recently, we used a yeast three-hybrid selection system to isolate the genes of proteins that interact with tRNA. This screen resulted in the identification of two novel proteins designated Tex1p and Tex2p. Our characterizations indicate that both proteins are involved in the nuclear tRNA export process, and that the function of Tex1p is essential whereas that of Tex2p is not. Presently, we are investigating the function of Tex1p and Tex2p in nuclear tRNA export.

Initiation of Protein Synthesis without Formylation of the Initiator Methionyl-tRNA

Formylation of the initiator methionyl-tRNA, catalysed by methionyl-tRNA formyltransferase, has long been regarded as essential for initiation of protein synthesis in all eubacteria. Recently, we show that this process is, in fact, dispensable in *Pseudomonas aeruginosa*. Disruption of the chromosomal methionyl-tRNA formyltransferase gene in *P. aeruginosa* resulted only in a moderate decrease in the rate of cell growth whereas in *Escherichia coli* cell growth was severely impaired. The ability of the *P. aeruginosa* mutant strain to grow was not due to an additional copy of the methionyl-tRNA formyltransferase gene or to N-acylation of the methionyl moiety by a group other than formyl. These results indicate that *P. aeruginosa* can carry out formylation-independent initiation of protein synthesis, using the non-formylated methionyl-tRNA. Therefore, the dogma that eubacteria require formylation of the initiator methionyl-tRNA for initiation of protein synthesis may have been an invalid generalization of results obtained with *E. coli*.

The identity of the components that facilitate the use of non-formylated methionyl-tRNA in initiation in *P. aeruginosa* is not known. To identify the components, we are using expression cloning to identify their genes. We are also investigating whether other eubacteria have the capacity to initiate protein synthesis with the formyl-methionyl-tRNA and/or methionyl-tRNA. These and other studies should facilitate the understanding of the mechanism responsible for initiation without formylation.

source : http://www.chembio.uoguelph.ca/biochemgroup/mangroo/

Frédéric Pio
Assistant Professeur en biologie moléculaire et biochimie
(604)291-5660
fpio@sfu.ca

Subvention d'équipement	Protein purification system for crystallization studies of protein complexes
Chaire :	Simon Fraser University

Intérêts de recherche :

Our laboratory is focused towards the three dimensional structural characterization of new protein-protein and protein-DNA interaction complexes involved in cancer and apoptosis. Our goal is to identify the critical interactions that modulate affinity and specificity between the different domains. To perform such studies we are using multidisciplinary approaches. They include:

- **New protein complex identification using genomic databases, bioinformatics, gene "hunting", computer modeling, and protein-protein interaction assays.**
- **Minimal interaction domain definition necessary for complex formation using Molecular Biology, Protein Chemistry and Binding Studies.**
- **Purification of the minimal domain and Reconstitution in vitro of the complex using Protein expression and Purification. Structural characterization of such complexes using biophysics and macromolecular x-ray crystallography.**

source : http://www.sfu.ca/mbb/mbb/faculty/pio/pio.html

Paul Romaniuk
Professeur en biochimie
(250)721-7088
pjr@uvic.ca

Objet de la recherche : Studies of Xenopus 5S rRNA binding proteins

Chaire : Victoria University

Molecular basis of nucleic acid-protein interactions involved in the regulation of gene expression; structure-function relationships in oncogenes

Intérêts de recherche :

The research in this laboratory is organized around three main themes involving nucleic acid structure and function. One major focus is on the structure of RNA, and the structure and function of specific RNA-protein complexes. This group has concentrated their efforts on the 5S ribosomal RNA from Xenopus laevis, and the specific complexes this RNA forms with transcription factor TFIIIA, with storage protein p43, and with ribosomal protein L5. By making targeted mutations in the 5S RNA, and studying the effects that those mutations have on the RNA structure by solution probing methods, the group has devised a graphic model for the three-dimensional structure of this RNA. They have used this model as the starting point to study the specific nature of the three 5S RNA-protein complexes in Xenopus. Our study of the 5S RNA-TFIIIA complex is complete and the effects that over 60 specific mutations in the 5S RNA have on TFIIIA binding indicate that TFIIIA uses an array of weak "non-specific" interactions defined in the context of the three-dimensional shape of the RNA to achieve a high degree of specificity. Comparative studies of the 5S RNA-p43 and 5S RNA-L5 complexes are currently in progress.

The second major focus in the lab concerns the interaction of zinc finger proteins with nucleic acids, and the role of these proteins in transcriptional regulation in eukaryotes. Xenopus TFIIIA not only forms a specific complex with 5S RNA, it also binds specifically to the internal promoter of the eukaryotic 5S RNA gene and activates the expression of this gene. The nucleic acid binding domain of TFIIIA consists of 9 "zinc finger" motifs. Romaniuk's group is presently doing mutagenesis of both the internal promoter of the 5S RNA gene, and the zinc fingers of TFIIIA to characterize at a detailed molecular level, the nature of this interaction. The p43 storage protein of Xenopus also has 9 zinc fingers in its RNA binding domain, but unlike TFIIIA, this protein binds only to 5S RNA, and not the 5S RNA gene. The group is carrying out "swapping" experiments to identify how differences in the amino acid sequences of these two proteins can restrict interaction only to RNA.

source : http://web.uvic.ca/biochem/bioc/Faculty/romaniuk/romaniuk.html

Kevin Wilson
Professeur en biochimie
(780)492-9241
k.s.wilson@ualberta.ca

Objet de la recherche :	Ribosomal protein-mRNA interactions regulating bacterial translation
Chaire :	Victoria University

Intérêts de recherche :

Our research is directed at understanding fundamental mechanisms of translation conserved in all organisms. The central enzyme of translation is the ribosome, a large and complex assembly of RNA and protein components. The structure of the ribosome has recently been determined at atomic resolution by X-ray crystallography, and provides a detailed framework for understanding the biochemical mechanisms of protein synthesis. Our research is currently focused on understanding two aspects of translational initiation important for regulating gene expression:

1. Specific, conserved ribosomal proteins autoregulate their own translation by binding to specific folded structures in the mRNAs that encode the ribosomal protein components of the ribosome. These regulatory ribosomal proteins also bind to specific structures in the rRNA components of the ribosome. We are investigating the common structural features in the mRNA and rRNA that are recognized by these proteins and responsible for translational autoregulation.
2. We are also studying the mechanism of translation initiation itself, taking advantage of the recently determined ribosome structural information. In particular, we are investigating the role of protein initiation and elongation factors in the complex and varied mechanism of translational initiation leading to a productive elongation complex.

source : http://www.biochem.ualberta.ca/Biochem/Faculty/Wilson/Wilson.html

Projets en biologie cellulaire 2000-2001

Richard Béliveau
Professeur en biochimie, Directeur du laboratoire d'oncologie moléculaire de l'Hôpital Ste-Justine
(514)987-3000 poste 8551
richard.beliveau@uqam.ca

Objet de la recherche :	Études de la barrière sang-cerveau : rôle de la P-glycoprotéine
Chaire :	UQÀM

Principaux projets en cours :

- Oncogénèse et méthylation des protéines ras: rôle dans le rein
- Étude de la perméabilité de la barrière sang-cerveau pour divers médicaments: rôle de la P-glycoprotéine dans la chimiothérapie des tumeurs cérébrales
- Systèmes de co-transport de la membrane luminale du rein
- Étude moléculaire de l'adaptation rénale à l'hypophosphatémie
- Cytosquelette, cancer et protéines Rho: implication dans l'angiogénèse

Orientation des recherches :

- Interaction médicaments-membranes: rôle de la P-glycoprotéine des capillaires cérébraux dans la résistance à la chimiothérapie.
- Études structure-fonction des protéines
- Protéines G et mécanismes moléculaires de transduction dans l'oncogénèse
- Études moléculaires de la réabsorption rénale
- Pharmacologie moléculaire antiangiogénique

Pour davantage d'informations, consulter le site :
http://www.er.uqam.ca/nobel/dep_chim/prof/beliveau.htm

source : http://www.grtm.umontreal.ca/beliveau.html

Richard Blouin
Professeur en biologie
(819)821-8000 poste 2062
rblouin@courrier.usherb.ca

Objet de la recherche :	Molecular bases of cell growth and differenciation
Chaire :	Université de Sherbrooke

Objectifs de ces recherches

Le professeur Blouin et son équipe s'intéressent aux mécanismes moléculaires régissant le contrôle de la prolifération et de la différenciation cellulaire. Ceci est particulièrement important puisque la perte de la régulation de ces deux fonctions entraîne la transformation cellulaire et l'apparition de certaines formes de cancer. Chez les mammifères, l'orchestration de la croissance et de la différenciation cellulaire s'appuie en partie sur un réseau complexe de protéines possédant une activité kinase spécifique de la tyrosine et/ou de la sérine/thréonine. Ces protéines jouent un rôle crucial dans la réception de différents messagers chimiques au niveau de la surface cellulaire et la transmission de signaux à l'intérieur de la cellule.

L'équipe du professeur Blouin a récemment amorcé la caractérisation d'un nouveau membre de la famille des protéines kinases connu sous le nom de ZPK/DLK/MUK. Grâce à des études de biologie cellulaire et moléculaire, elle a pu mettre en évidence que cette protéine kinase pourrait potentiellement exercer une fonction importante dans le développement, la fonction et le maintien de nombreux organes et tissus. D'autres études de même nature ont aussi permis de démontrer le rôle de ZPK/DLK/MUK dans la régulation de la croissance cellulaire. Les projets de recherche actuellement en cours dans le laboratoire ont pour objectifs principaux d'apporter des éclaircissements sur la fonction de cette nouvelle protéine. Pour ce faire, des méthodes d'analyses diversifiées (microscopie, immunologie, biochimie, biologie cellulaire et moléculaire) sont utilisées pour: 1) élucider les mécanismes d'activation de ZPK/DLK/MUK; 2) identifier les protéines avec lesquelles ZPK/DLK/MUK interagit dans la cellule; et 3) mesurer les effets de l'expression de ZPK/DLK/MUK sur la croissance et la différenciation cellulaire. L'importance de ces études est énorme si on considère la possibilité que le dérèglement de l'activité de cette protéine est susceptible d'empêcher le déroulement normal du développement et le maintien de l'homéostasie cellulaire.

Projet de recherche en cours

- Analyse des mécanismes d'activation de la protéine kinase ZPK/DLK/MUK
- Détermination de l'identité des partenaires et des substrats de la protéine kinase ZPK/DLK/MUK
- Effets de l'expression de la protéine kinase ZPK/DLK/MUK sur la croissance et la différenciation cellulaire

source : http://www.usherb.ca/SCES/BIO/bloric.html

Thor Borgford
Assistant Professeur en biologie moléculaire et biochimie
(604)291-3571
borgford@sfu.ca

Objet de la recherche :	Protein engineering a toxin molecule
Chaire :	Simon Fraser University

Intérêts de recherche :

The research program of this laboratory is concerned with relationships between the structure and the function of proteins (enzymes). A multidisciplinary approach (commonly referred to as "protein engineering") is taken to unravelling these relationships which involves molecular genetic, biophysical and biochemical techniques. As an example, we use techniques in recombinant DNA and site-directed mutagenesis to manipulate genes and the sequences of genes which encode proteins of interest. Subsequently, changes in the function or activity of the protein can be interpreted in terms of the original changes made to sequence and structure.

Presently, there are three major projects underway in the laboratory, focusing on three different protein/enzyme systems. They are: 1) a study of ion binding in the skeletal, calcium-binding, protein troponin-C, 2) substrate specificity and editing "fidelity" in the isoleucyl-tRNA synthetase, 3) substrate specificity in members of a family of serine proteases derived from *Streptomyces griseus*. The unifying objective of these apparently disparate studies is to derive a better understanding of the details of molecular recognition and enzyme catalysis.

A representative example of recent studies in this laboratory involves troponin-C. We have shown that the ion affinity of a set of binding sites can be tailored, or fine-tuned, by single amino acid substitutions at the amino terminal end of helix-G. The analysis of variant proteins shows, indirectly, that α-helices in this protein are in an equilibrium between folded and unfolded states and the position of the folding equilibrium is shifted by metal ions.

By understanding the principles which relate protein structure and function we hope eventually to have the ability to design proteins with novel functions and activities.

source : http://www.sfu.ca/mbb/mbb/faculty/borgford/borgford.html

Jim Lepock
Professeur en physique, Directeur du laboratoire de biophysique des protéines
(519)888-4567 poste 2214
lepock@sciborg.uwaterloo.ca

Objet de la recherche : Irreversible denaturation of proteins during heat shock

Chaire : University of Waterloo

Principales activités de recherche :

The effects of elevated temperature (hyperthermia or heat shock) on cells, cellular organelles (e.g. membranes and nuclei), and proteins are being investigated to determine the direct effects of heat on biological material and to use thermal analysis as a means of studying the structure and interactions of biological molecules. Temperatures in the range used for cancer therapy are being employed. The techniques of differential scanning calorimetry (DSC), phase-modulation fluorescence spectroscopy, and electron spin resonance spectroscopy are being employed to detect transitions and to determine the dynamics of biological molecules during heating. Specific projects include: 1) the use of DSC to study protein denaturation and other transitions in living mammalian cells, 2) the study of the mechanism of protection and repair of thermolabile proteins by a class of proteins synthesized in response to heat and other stresses (heat shock proteins), 3) the effect of specific amino acid substitutions on the stability and thermodynamics of unfolding of proteins, 4) the influence of the mutations responsible for amyotrophic lateral sclerosis on the stability and folding of human Cu^{2+}, Zn^{2+} - superoxide dismutase, and 5) the thermal denaturation of the Ca^{2+} ATPase (a calcium transport protein).

source : http://www.science.uwaterloo.ca/physics/people/faculty/lepock/research_activities.html

Nils Petersen
Professeur en chimie et en biochimie
(519)661-2111 poste 86309
petersen@julian.uwo.ca

Objet de la recherche : Protein clustering and intermolecular interactions in biological membranes

Chaire : University of Western Ontario

Principales activités de recherche :

Research in the our group is focused on understanding intermolecular interactions in biological membranes. We study the dynamics and distribution of molecules within the membrane as a means of understanding their function in events such as cell-cell communication, signal transduction, adhesion and locomotion of cells and breating. There are four major themes:

1) Studies of the effect of shape and size of molecules on the diffusion and clustering of molecules in model membranes. We combine a synthetic program with biophysics, since we make molecules of specific shapes and sizes and study their dynamic behaviour as a function of temperature and lipid composition. So far we have shown that the diffusion is controlled by the size and the shape of the molecule right at the membrane surface. The size or shape in the aqueous or in the bilayer interior is of no consequence.

2) Studies of the distribution and extent of aggregation and interaction of membrane receptors in living cells. This work is based on analysis of laser scanning confocal microscopy images to estimate the density of protein clusters and their sizes. We have developed specific tools - Image Correlation Spectroscopy and Image Cross-correlation Spectroscopy for this purpose.

There are three sub-themes:

a) the function of growth factor receptors and their interaction with cellular substructures
b) the effect of integrin receptor function on assembly of coated pits
c) the interaction of GPI anchored proteins with lipid rafts

3) Studies of the mechanical properties of adhering cells. This work utilizes combinations of modern imaging tools such as confocal microscopy, atomic force microscopy, near-field scanning optical microscopy and interfacial force microscopy to study the mechanical properties of cell structures of known biochemistry. The purpose is to understand the forces that give cells shape and allow them to move.

4) Studies of lung surfactants from healthy and dysfunctional animals. The purpose is to understand how the particular lipids and proteins present in lung surfactants affect the function and malfunction of the surfactant films in lungs. The objective is to establish the phase behaviour of the systems, to identify the composition and physical characteristics of different phases and to establish the mechanism whereby these components create a low surface tension environment dynamically.

source : http://www.uwo.ca/chem/Tpetersonn.htm

Arthur Szabo
Professeur en chimie
(519)884-0710 poste 2129

Objet de la recherche :	Protein function studies with targeted unnatural amino acids
Chaire :	Wilfrid Laurier University

Intérêts de recherche :

The field of research of my laboratory can best be described as Biophysical Chemistry. The properties of peptides and proteins and other biomolecules including membranes and nucleic acids are investigated in order to elucidate the molecular basis of their activity. The primary objective of the research in proteins is to understand the interrelationship between protein structure, dynamics and function. In order to achieve this goal a variety of optical spectroscopic methods are used. The main expertise in the laboratory is centered on the use of fluorescence spectroscopy, especially picosecond laser time-resolved fluorescence techniques. This spectroscopic expertise is combined with biochemical and molecular biological procedures to provide a strong synergism in the research program.

The research program has three interwoven themes. One is to build a better understanding of the structural features such as local interactions, environment, and dynamics which influence the variety of fluorescence parameters arising from aromatic amino acids. This information content has been significantly increased measuring the fluorescence decay from single crystals of proteins.

The second aspect is the investigation of different proteins and the elucidation of features which describe their specific functionality. An important aspect of this work is to prepare mutants having a single tryptophan located in different segments of the protein by site directed mutagenesis. Hence changes occurring in the different segments associated with the protein function can be mapped out. One protein of interest is the enzyme tryptophanyl t-RNA synthetase, a key enzyme involved in the specific recognition of tryptophan and subsequent acylation of t-RNAtrp. Using analogs such as 7azatryptophan (7AW) and 5-hydroxytryptophan (5HW) and 4-fluorotryptophan (4FW) as substrates structural changes occurring upon activity can be determined.

The third particularly challenging theme is to determine the local molecular details of interacting segments of proteins. Examples include hormonal peptide-receptor complexes; protein inhibitors of proteolytic enzymes; and multi protein complexes. We have demonstrated the feasibility of biosynthetic incorporation of tryptophan analogs into proteins, and have been leaders in showing the enhanced information that such proteins provide.

source : http://www.cs.uwindsor.ca/units/chem/gradbroc/Arthur.Szabo.html

Réseau canadien des centres d'excellence en génie protéique

Créé en 1990, le réseau des centres d'excellence en génie protéique, dont l'acronyme est PENCE inc en anglais, est une compagnie de recherche et de transfert technologique spécialisée dans les technologies des protéines. Ce réseau basé à Edmonton relie les scientifiques canadiens aux différents partenaires industriels et gouvernementaux dans le développement de technologies qui ont de multiples applications. Dix-huit universités, hôpitaux et instituts de recherches sont impliqués, ce qui signifie près de 200 chercheurs et employés regroupés en 60 équipes. Au Québec, l'Université de Montréal et l'Université McGill ainsi que l'Institut de Recherches Cliniques de Montréal et le Centre de Recherche en Calcul Appliqué (CERCA) sont des membres de PENCE.

Les expertises sont variées, mais les collaborations de font à l'intérieur de 7 thèmes de recherche.

Facteurs de croissance, Récepteurs et Transduction de signal

Le but principal des recherches de ce thème est de développer des composés capables de moduler des récepteurs spécifiques ou d'influencer la transduction de certains signaux ayant un impact sur des thérapies pour le cancer, les allergies et l'inflammation. Voici les chercheurs travaillant sur ce thème :

Directeur : John Schrader (UBC)
Tony Pawson (Mt Sinai)
Gary Brayer (UBC)
Robert Hogues (U of A)
Emil Pay (U of T)
Alice Vrielink (McGill)
Bryan Sykes (U of A)
Ian Clark-Lewis (UBC)
Lewis Kay (U of T)
Mike Tyers (Mt Sinai)
Hermann Ziltener (UBC)

Protéinases de maladie

Ce deuxième thème regroupe trois projets portant sur des familles différentes d'enzymes : les convertases, les protéases et les Calpain (*calcium-activated neutral protease*). La recherche sur ces enzymes peut mener au développement d'outils

thérapeutiques pour le cancer, l'artériosclérose, la maladie d'Alzheimer et certaines infections virales. Les chercheurs participant à ces projets sont :

Directeur : Nabil Seidah (Institut de recherches cliniques de Montréal)
Michel Chrétien (Ottawa)
Peter Ottensmayer (U of T)
Zangchou Jia (Queen's)
Richard Ménard (Montréal)
Feng Ni (McGill)
Peter Davies (Queen's)
Mirek Cygler (McGill)
Claude Lazure (IRCM)
John Mort (McGill)

Technologie de dimérisation

Les projets de ce thème sont basés sur l'étude d'un motif de protéine constitué de l'enroulement de deux hélices alpha (structure secondaire particulière d'un assemblage d'acides aminés). Simple et régulier, ce motif constitue un excellent modèle pour l'étude des interactions protéine-protéine et du repliement des protéines. Ces études permettront le design, la synthèse et la production d'inhibiteurs aux propriétés spécifiques pour plusieurs applications médicales. L'équipe de recherche de ce thème est composée des scientifiques suivants :

Directeur : Randall Irvin (U of Alberta)
Robert Hogues (U of A)
Mark Glover (U of A)
Jean Gariépy (U of T)
Maureen O'Connor-McCourt (McGill)
David Bundle (U of A)
Brian Sykes (U of A)
Bruce Lennox (McGill)

Design d'ingénierie de peptides et protéines

L'objectif des recherches de ce thème est l'étude de la structure et fonctions de protéines particulières ayant des applications en biotechnologies. Actuellement, on étudie des protéines ayant des propriétés antigel et d'autres ayant une action antibactérienne. Voici les chercheurs travaillant sur ce thème :

Directeur : Peter Davies (Queens)
Cyril Kay (U fo A)
Brian Sykes (U of A)
David Wishart (U of A)
Ronald Kluger (U of T)
Zangchou Jia (Queen's)
Maureen O'Connor-McCourt (McGill)
Robert Hogues (U of A)
Ronald McElhaney (U of A)
Emil Pai (UBC)
R Hill (U of T)
Uri Saragovi (McGill)

Technologies à base de carbohydrate

Ce thème est à propos de protéines qui se lient aux carbohydrates. Elles sont impliquées dans les contacts cellule à cellule ainsi que dans l'invasion d'hôtes par des pathogènes. Les chercheurs contribuant à ce projet sont :

Directeur : Anthony Warren (UBC)
Ole Hindsgaul (U of A)
Jim Rini (U of T)
Dave Rose (McGill)
Charles Haynes (UBC)
Lawrence McIntosh (UBC)
Robert Hodges (U of A)
Bruce Lennox (U of A)
Harry Schachter (Hospital for Sick Children)
David Bundle (U of A)
Lewis Kay (U of T)
Gary Brayer (UBC)
Douglas Kilburn (UBC)
Stephen Withers (UBC)
Randall Irvin (U of A)
M O'Connor-McCourt (McGill)
Peter Ottensmeyer (U of T)

Enzymes microbiennes pour industries

Deux types de protéines, les xynalases/cellulases et les laccases, sont l'objet des recherche de ce thème. Ces enzymes sont utilisés dans la modification de fibres dans les industries du textile et du papier. Les chercheurs participant à ces projets sont :

Directeur : Grant Mauk (UBC)
Gary Bauer (UBC)
Lawrence McIntosh (UBC)
Anthony Warren (UBC)
Peter Ottensmayer (U of T)
Michael Paice (PAPRICAN)
Douglas Kilburn (UBC)
John Saddler (UBC)
Stephen Withers (UBC)
David Rose (McGill)

Bio-informatique et reconnaissance bio-moléculaire

La recherche en bio-informatique à PENCE a deux aspects : la résonance magnétique nucléaire et des méthodes informatiques, deux outils qui sont devenues essentiels au génie protéique. La technologie RNM a connu des développements récents qui permettent d'étudier de très grosses protéines et des complexes protéiques avec une meilleure précision. Ainsi, on peut comprendre comment les protéines reconnaissent et lient leurs ligands apparentés. Les méthodes informatiques englobent la manipulation et l'analyse de données sur les séquences protéiques. On développe des techniques informatiques de design de médicaments à

partir de structures tridimensionnelles de ligands et de récepteurs. Les chercheurs visent aussi à développer des logiciels améliorant la productivité de la recherche sur les protéines. L'équipe de recherche de ce thème est composée des scientifiques suivants :

Directeur : Brian Sykes (U of A)
Janice Glasgow (Queen's)
John Gunn (Montréal)
Benoît Roux (Montréal)
E Purisma (BRI)
Lawrence McIntosh (UBC)
Suzanne Fortier (Queen's)
François Major (Montréal)
Lewis Kay (U of T)
David Wishart (U of A)
Stephen Withers (UBC)

Le CERCA participe aux recherches de ce thème du réseau de génie protéique. Travaillant à la section chimie pharmaceutique du CERCA, le professeur de chimie de l'Université de Montréal John Gunn fait de la modélisation de structure de protéines sur ordinateur. Il développe aussi des modèles et des algorithmes visant à prédire la structure tridimensionnelle à partir de la séquence d'acides aminés. Des projets en bio-informatique sont aussi en cours au CERCA. On développe des programmes pour sonder les banques de données dans le but de collecter des informations sur les propriétés de protéines et pour incorporer des données expérimentales aux simulations. Les autres partenaires industrielles de PENCE pour ce thème sont les compagnies bio-informatiques BioTools et Molecular Mining ainsi que la compagnie pharmaceutique MerckFrost.

Projets de Génome Canada

L'intérêt du génome et des connaissances qui en découlent n'est plus à prouver, du moins pas aux différents gouvernements. Plusieurs programmes ont déjà été exposés dans ce document qui démontrent la volonté du gouvernement fédéral de voir le Canada jouer un rôle important en génomique et en protéomique. Toutefois, il n'a pas seulement agi par ses organismes déjà établis comme le CRSNG. Il a créé Génome Canada, une société sans but lucratif qui pour mission de coordonner la recherche en génomique au Canada afin qu'il en devienne un des leaders mondiaux.

Le gouvernement provincial du Québec a lui aussi fondé un organisme de recherche et de développement en génomique et en protéomique, Génome Québec, dont la mission est de développer ce secteur, de jouer un rôle structurant et mobilisateur et d'assurer les débouchés commerciaux de la recherche.

Avec le lancement de ces organismes, d'importants budgets leur ont été alloués. Des concours ont eu lieu et des projets ont été choisis. Les projets se rapportant aux protéines sont présentés ci-après. Les descriptions sont tirées du site de Génome Canada.

Victor Ling, Professeur de biochimie et de biologie moléculaire, UBC
Titre du projet : Génomique du cancer

Victor Ling et ses collègues développent des moyens pour détecter la façon dont les cellules se transforment en tumeurs malignes aux toutes premières étapes du cancer en identifiant les structures altérées des gènes et des protéines. Ces structures altérées résultent en mutations génétiques qui sont caractéristiques de ces premiers stages de la transformation. Le but de l'étude est de faire une analyse génétique sur de petits nombres de cellules et de déterminer les mutations et les gènes altérés qui distinguent les tissus normaux des premiers stages de cancers. La recherche porte sur les tumeurs des poumons, des seins, de la prostate, du système gastro-intestinal, buccales, lymphoïdes et myéloïdes. L'équipe entend décrire les moyens d'expression des gènes avant et après que les cellules commencent à croître et à se diviser.

Janet Rossant, Professeure de génétique moléculaire et médicale, U of T
Titre du projet : Génomique fonctionnelle et protéomique des organismes modèles

La protéomique – qui est l'étude des protéines et de leur fonction – est la prochaine vague critique de recherche en génomique. Le projet du professeur Rossant consistera à déterminer et à caractériser les complexes protéiniques et les interactions entre protéines, et à étudier la fonction des protéines dans divers organismes, allant des bactéries aux cellules de levure, des nématodes (vers), et des souris génétiquement modifiées.

Jack Greenblatt, Professeur de biologie moléculaire, U of T
Titre du projet : Installations de base de technologie protéomique

Ce projet permettra de créer des installations de base de ressources protéomiques pour identifier les protéines et permettre aux chercheurs de développer de nouvelles technologies d'analyse des fonctions protéiniques. L'identification de base des protéines sera située dans trois installations de l'Avenue Université, à l'Université de Toronto, à l'Hôpital pour enfants et au Samuel Lunenfeld Research Institute. L'installation d'analyse des fonctions protéiniques sera construite dans le cadre d'un projet coopératif de l'University Health Network et de l'Université de Toronto.

John J.M. Bergeron, Directeur du département de biologie cellulaire, McGill
Titre du projet : Réseau de Montréal de pharmaco-protéomique et de génomique structurelle

Le but de ce projet est de construire une installation pour permettre aux chercheurs d'étudier les fonctions et la structure des gènes et des protéines. L'installation serait équipée principalement pour cartographier, identifier et caractériser des protéines. Une première composante du projet, la cartographie cellulaire, portera sur la cartographie et la caractérisation des protéines et sur la façon dont elles s'associent avec d'autres parties de la cellule. La deuxième composante consistera à utiliser des dosages biologiques pour cartographier les interactions entre protéines. La troisième permettra de cartographier et d'analyser les structures de protéines clés. Le projet regroupera des chercheurs qui utilisent des technologies de biochimie, de biologie cellulaire, de génomique, d'ingénierie, de biopuce, d'analyse des séquences de protéines et de cristallographie par rayons-X, parmi d'autres disciplines et technologies innovantes.

Howard Bussey, Professeur de biologie, McGill
Titre du projet : Projets de génomique fonctionnelle utilisant des organismes modèles

Le professeur Bussey étudiera des gènes et des protéines sur des organismes modèles, à grande échelle. L'équipe de l'Université McGill caractérisera 5 000 mutations de *S. cerevisiae* et croisera des lignées de mutants pour déterminer la létalité potentielle. Les chercheurs utiliseront des gènes de levure, de mouche des fruits et de vers pour étudier les interactions entre les protéines.

Thomas J. Hudson, Professeur de génétique humaine, McGill
Titre du projet : Génétique régulatrice

Le professeur Hudson et son équipe, de l'Université McGill, examineront les polymorphismes régulateurs du génome humain – gènes qui contrôlent la susceptibilité aux maladies communes. Ils utiliseront trois approches différentes pour trier 1 000 gènes possibles. Lorsqu'ils auront trouvé les gènes requis, ils les valideront et étudieront leurs fonctions en les incorporant à des souris transgéniques. L'équipe construira aussi une base de données spécialisées pour enregistrer leurs données et les mettre à la disposition du public sur Internet.

Fernand Labrie, Professeur d'endocrinologie, Laval
Titre du projet : Atlas des profils de génomique de l'action de stéroïdes

Le professeur Labrie et son équipe étudieront les modifications de l'expression des gènes qui se produisent lors de traitement par les stéroïdes. Ils détermineront les mécanismes qui se déclenchent lorsque des hormones stéroïdes sont présentes et absentes et lorsqu'elles interviennent dans l'action d'autres hormones. Ils identifieront les protéines contrôlées par des hormones et cartographieront les types de cellule responsables des changements de transcription – le processus selon lequel l'ARN est synthétisé à partir de l'ADN. En utilisant des techniques bio-informatiques nouvelles, les chercheurs relieront entre elles les ensembles de données et l'information sur la fonction des protéines, de manière à pouvoir relier ces protéines à l'emplacement des chromosomes. L'équipe mettra au point un ensemble de logiciels de base de données, d'analyse et de visualisation pour intégrer les données de diverses sources. Elle développera aussi des modèles et des outils pour tester les modèles d'évaluation des gènes candidats qui jouent un rôle dans les stéroïdes régulateurs.

Thomas J. Hudson, Professeur de génétique humaine, McGill
Titre du projet : Centre d'excellence en génomique de Montréal

Le Centre d'excellence en génomique de Montréal est un organisme formé de groupes de recherche scientifique de l'Université McGill, de l'Université de Montréal, du Centre hospitalier de l'Université de Montréal et de l'Institut de recherche en biotechnologie du CNRC. Le gros de l'équipement et de l'administration de cette installation se trouvera sur le campus de l'Université McGill. Les autres établissements se chargeront d'autres activités clés. Parmi eux figurent un consortium technologique, un réseau de bio-informatique et une composante qui offrira des services de génotype, de séquençage, de production de biopuces, d'analyse de spectrométrie de masse et des services d'informatique destinés à des groupes extérieurs aux projets à grande échelle.

Les cinq derniers projets sont québécois et ils sont aussi subventionnés par Génome Québec.

La bio-informatique dans les universités canadiennes

Université de Montréal

Programme de baccalauréat en bio-informatique : Automne 2001

Programme d'études supérieures en bio-informatique : Automne 2002

Université McGill

Mineur en biologie moléculaire computationnelle : Automne 2001 pour étudiants en biologie et biochimie

Mineur en biochimie computationnelle : En planification pour étudiants en informatique

Programme de baccalauréat en bio-informatique : En planification

Option bio-informatique pour études graduées : En planification

University of Toronto

Programme d'études supérieures en bio-informatique et protéomique

University of Alberta

Aucun programme spécifique, mais plusieurs professeurs impliqués dans un groupe de recherche

University of Calgary

Programme d'études supérieures en biotechnologie médicale avec cours en bio-informatique et génomique

Queen's University

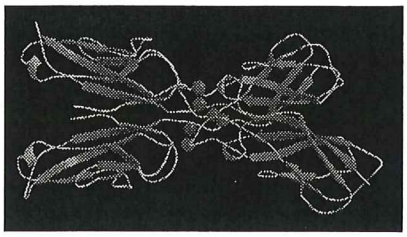
Programme d'études supérieures en informatique avec cours en bio-informatique (CISC-875)

University of Waterloo

Programme de baccalauréat en bio-informatique

University of British Colombia

Programme de baccalauréat en microbiologie avec cours en bio-informatique



Conclusion

Les découvertes qui sont maintenant à la portée des scientifiques grâce aux protéines sont multiples et suscitent beaucoup d'enthousiasme. Ce champ de recherche est vaste et il y a une grande quantité de projets possibles. On commence à disposer des outils nécessaires, mais un grand nombre d'entre eux restent à inventer. De plus, le travail qui reste à faire est colossal.

Les sciences de la protéomique et du génie protéique sont récentes et elles en sont à leurs premiers balbutiements. Toutefois, avec tous les moyens qui sont actuellement mis en oeuvre, certaines découvertes-clés devraient être faites au cours des prochaines années qui auront des impacts importants. Quelques chercheurs s'interrogent sur la pertinence d'un projet mondial étudiant le protéome semblable à celui sur le génome. Ils considèrent que les moyens actuels ne permettent pas la mise en place d'une telle initiative. Néanmoins, d'autres sont d'avis que lorsque le projet du génome humain a commencé en 1988, la recherche en génomique était aussi peu développée que l'est la protéomique aujourd'hui et que c'est la concertation mondiale qui a permis l'éclosion des techniques et les progrès phénoménaux.

Projet mondial ou pas, il est toutefois indéniable qu'il y a actuellement un momentum important dans la recherche sur les protéines. Dans plusieurs pays comme le Canada, ce secteur et celui de la bio-informatique sont considérés névralgiques pour le futur de l'industrie des biotechnologies. Le Québec particulièrement, qui compte 40% de entreprises canadiennes dans le secteur de la biotechnologie, doit assurer le développement de compétences dans ces nouveaux domaines pour que ses entreprises demeurent compétitives mondialement.

Est-il donc envisageable qu'un projet de recherche sur les protéines puisse prendre place à l'École Polytechnique ? Il est de mon avis que cette initiative serait possible. L'ingénieur, qu'il soit en chimie, en physique ou en informatique, peut

contribuer à plusieurs projets. La variété de la recherche actuelle suggère que plusieurs avenues sont encore inexplorées et la vision du génie pourrait offrir une nouvelle perspective inattendue. Il est à noter que le génie protéique sied davantage aux ingénieurs que la protéomique. Néanmoins, il demeure qu'un bon nombre d'applications bio-informatiques peuvent être développées pour les deux sciences des protéines. Les ingrédients à rassembler sont une bonne idée, des partenaires qu'ils soient de l'industrie, biologistes ou biochimistes universitaires, et la motivation de découvrir un monde surprenant par son organisation, sa dynamique et sa complexité.

Annexe 1

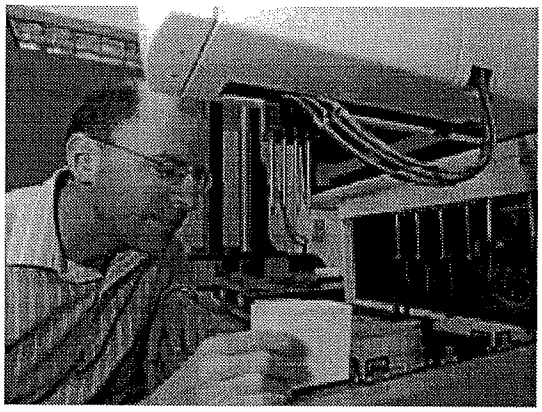
cyberpresse.ca

Dimanche 6 mai 2001

Le génome humain sur puce informatique

Luc Dupont
AGENCE SCIENCE-PRESSE

Quand le Projet international génome humain fut lancé en 1988, on savait que cette «botte de foin biologique» qu'est notre génome, contenait trois milliards de briques élémentaires (qu'on appelle les «paires de base»). On savait en outre que l'on trouverait, dissimulées en elle, 100000 petites «aiguilles» - aujourd'hui ramenées à 30000 -, nos gènes!



Photothèque La Presse

Le directeur du Centre de génomique de Montréal, le Dr thomas Hudson, aussi directeur adjoint du réputé Centre de génomique du Massachussetts Institute of Technology à Boston.

Ce qu'on ne savait pas, c'est comment on arriverait à démêler ces montagnes de chiffres et de données.

Étonnamment en effet, «on avait, à ce moment-là, bien peu idée du temps que ça prendrait et des stratégies techniques qu'il faudrait adopter pour mener à bien cette énorme entreprise», déclarait à *Science-Presse*, il y a deux ans(1), le directeur du Centre de génomique de Montréal, le Dr Thomas Hudson, aussi directeur adjoint du réputé Centre de génomique du Massachusetts Institute of Technology à Boston.

Mais alors... comment y est-on arrivé si vite? Grâce à deux innovations: la robotisation des grands laboratoires, qui entraîna une accélération foudroyante des procédés de décodage; et la croissance exponentielle de la puissance informatique, tout au cours des années 1990, qui a rendu possible le traitement de milliards et de milliards de données biologiques. La bio-informatisation était née.

Sans cette bio-informatique, point de génome. Du moins, pas en dix ans...

Imaginez: pour gagner son pari, le Projet génome humain (Human Genome Organisation ou HUGO) aura jonglé avec plus de 7 milliards de données. Celera Genomics, la société privée qui a aussi accouché d'une carte du génome, a enregistré 50 térabytes d'informations, l'équivalent de 80000 disques compacts! Et ce n'est que le début!

Le siècle de la bio-informatique

Le terme bio-informatique (ou biologie computationnelle) remonte au milieu des années 1980. Il désigne l'ensemble des applications de l'informatique aux sciences de la vie: le tri, l'analyse, la structuration, la transmission des données par Internet, etc.

Maintenant que la plupart de nos gènes ont été localisés (ce qui constituait la première phase du décodage du génome), il revient aux scientifiques d'identifier le rôle de chacun d'eux et l'expression de leurs diverses

protéines. Dans cette seconde phase, les analyses sont devenues infiniment plus fines et les calculs, infiniment plus complexes. C'est pourquoi, pour les généticiens, la nécessité de s'asseoir sur une assise bio-informatique est maintenant devenue absolue.

Or, le Conseil de la science et de la technologie du Québec, dans un avis rendu public en janvier, lance un sérieux avertissement: il y a actuellement pénurie de bio-informaticiens chez nous. Nos 30 à 50 spécialistes, selon une estimation de Jean Morissette, membre du comité scientifique, deviendront vite insuffisants.

Le problème n'est pas qu'au Québec: la soudaineté de l'émergence de la génomique a entraîné une pénurie mondiale de bio-informaticiens.

Nous devons, «en toute priorité, mettre en place, d'ici septembre 2001, un programme intensif de formation-perfectionnement en bio-informatique au Québec», indique le document *La bio-informatique au Québec: un levier essentiel du développement des bio-industries*. «Nous devons permettre, si possible dès l'hiver 2002, à une dizaine de boursiers d'acquérir une formation supérieure en bio-informatique à l'extérieur du Québec... Tout va se jouer dans les deux ou trois prochaines années.»

«Tout», c'est la compétitivité du Québec dans le domaine de la génomique. «Tout», c'est ce secteur extrêmement névralgique, la biotechnologie, puisque le Québec rassemble 42% des firmes canadiennes!

L'odyssée de Jean Morissette

Jean Morissette fut l'un des pionniers mondiaux de la bio-informatique. Entre 1991 et 1996, il a agi comme bio-informaticien principal de l'équipe française Généthon, dans le projet de cartographie génétique. Or, c'est Généthon qui a pondu la première carte génétique au monde, un ensemble de quelque 5000 marqueurs (ou repères de gènes) sur l'ensemble du génome. Et ce sont ces repères qui ont permis l'édification même de l'immense cathédrale que fut le Projet génome humain!

«La bio-informatique est effectivement apparue comme une nécessité dans les équipes engagées directement dans le Projet génome humain, parce que c'est là qu'on s'est mis à générer des données biologiques à haut débit», raconte ce mathématicien-informaticien-généticien, aujourd'hui au service du Centre de recherche en endocrinologie moléculaire et oncologie du CHUL (Centre hospitalier de l'Université Laval). Morissette avait créé alors de toutes pièces, pour Généthon, une «chaîne de montage informatique de production de données biologiques, c'est-à-dire un puissant réseau d'ordinateurs intégrant moteurs de calcul et bases de données, pour «l'usine à marqueurs» qu'allait vite devenir ce labo.

«Jusque-là, en 15 ans, tous les laboratoires du monde mis ensemble n'avaient produit que 2000 marqueurs, dit-il. Dans notre seul labo, on en produira 5000 en 5 ans!»

Jusque-là, on n'avait pas l'habitude de penser chiffres quand on pensait biologie: «Ce qui était nécessaire en mathématiques, en astrophysique et dans tous ces domaines qui génèrent des masses de données, l'est devenu aussi en biologie. La bio-informatique est née là.»

«Mais quand j'ai commencé en 1970, on parlait plutôt de bio-statistiques», poursuit Jean Morissette. Le scientifique travaille alors pour le Réseau de médecine génétique du Québec (RMGQ), où il structure les bases de données dans lesquelles sont stockés les résultats de tests génétiques subis par les quelque 100000 bébés qui naissent alors annuellement. «On travaillait avec des cartes et des rubans perforés que l'on faisait analyser par télétype sur des ordinateurs situés à l'extérieur!»

Entre 1972 et 1975, il fait à Londres des études en mathématiques et génétique. De retour au pays, jusqu'à sa participation à Généthon, il travaillera encore 15 ans au RMGQ.

Aujourd'hui, comme de nombreux généticiens et bio-informaticiens du monde entier, M. Morissette est engagé

dans la deuxième phase du décodage du génome humain, celle où on interroge le rôle de chaque gène et la fonction de leurs protéines (la protéomique). À ce stade, les techniques d'investigation sont devenues fascinantes: l'une des méthodes consiste à prélever des cellules de l'un des organes d'un corps (la rate, par exemple) afin de déterminer lesquels, parmi nos 30000 gènes, interviennent précisément dans le fonctionnement de cet organe.

«Pour ce faire, on pose les tissus sur des biopuces, explique M. Morissette. Les biopuces sont des outils à mi-chemin entre le vivant et l'électronique. Ces instruments peuvent rassembler, sur une toute petite surface, l'équivalent de 10000 gènes ou séquences codantes. Or, mis en leur présence, les gènes présents dans le tissu X (la rate, par exemple) vont s'exprimer (s'identifier) sur la biopuce. Ainsi, en faisant l'exercice sur tous les tissus de toutes les parties du corps, on arrivera un jour à réaliser un nouveau type de cartographie: l'atlas de l'expression des gènes. Cet atlas localisera, organe par organe, le lieu d'action de chaque gène.

Sauf que pour traiter toutes ces combinaisons, comparer ces millions d'exercices, il faut de l'informatique à plein régime.

Internet et compagnie

Pour Pierre Lepage, bio-informaticien de la nouvelle génération, qui travaille au tout nouveau Centre de génomique de Montréal (situé à l'Hôpital général), Internet est également un outil-clé au service de la bio-informatique.

«J'y passe la moitié de mon temps», dit-il. Et pour y faire quoi? «Pour parcourir les grandes bases de données, où les résultats du séquençage entier du génome humain sont accessibles publiquement et gratuitement! Prenez des sites comme celui du NCBI (National Center for Biotechnology Information): on y retrouve la localisation des quelque 30000 gènes humains. C'est là une base de travail essentielle pour nous qui sommes en possession de séquences géniques issues de personnes souffrant de diverses maladies. Dans ces bases, on peut aller chercher les régions des chromosomes qui nous intéressent, les télécharger pour fins de comparaison et d'analyse avec celles de nos malades, et identifier ainsi les possibles mutations à l'origine d'une maladie.»

Certains n'hésitent pas à affirmer que la bio-informatique sera au coeur de tous les grands développements de la biologie du XXIe siècle. Tellement qu'on emploie de plus en plus l'expression IN SILICO, par opposition à IN VIVO et IN VITRO.

«On a une grosse industrie en bio-pharmaceutique au Québec et, de son virage ou non en génomique, dépendront nos besoins - plus ou moins grands - en bio-informaticiens, conclut Jean Morissette. C'est pourquoi il faut se préparer à faire face à cette demande en mettant sur pied les programmes de formation universitaire requis.» L'Université McGill offre déjà une option en bio-informatique à l'intérieur de son doctorat en biologie. De son côté, l'Université Laval a mis sur pied un comité chargé de structurer, pour l'automne prochain, un début de programme.

Annexe 2

cyberpresse.ca

NOUVEAUTÉS TECHNOS Le vendredi 03 août 2001

Biologie et informatique: les nouveaux inséparables

André Mondoux
La Presse

La complétion du génome humain, le répertoire complet de tous les gènes dont nous sommes formés, a certes marqué une étape cruciale dans l'histoire des sciences de la vie. Cependant, il ne s'agit que d'un début. En effet, les chercheurs ont maintenant devant eux la lourde tâche d'explorer, notamment, comment gènes et protéines interagissent et quels rôles ils jouent dans la cause des maladies. Pour ce faire, des billions et des billions de possibilités devront être explorées, une à une, une tâche incroyablement longue et complexe.

Voilà pourquoi les ordinateurs seront des outils indispensables aux développements futurs de la génétique, qui en retour aura un impact majeur sur tous les secteurs des sciences de la vie, qu'il s'agisse de la médecine, de la pharmacologie la recherche fondamentale et les biotechnologies. Pour ce faire, une «nouvelle» science, issue de la convergence de l'informatique et de la biologie, a vu le jour depuis le début des années 90, la bio-informatique (*bioinformatics* en anglais).

Plusieurs définitions formelles de la bio-informatique ont été proposées. Cependant, l'usage courant du terme désigne la plupart du temps l'utilisation d'ordinateurs à des fins de recherche pour stocker et récupérer des quantités phénoménales d'information; sur ce plan, la bio-informatique se traduit souvent par l'utilisation de puissantes bases de données. Le terme sert également à décrire les efforts pour analyser et prédire la composition et la structure de molécules biologiques. En ce sens, les puissants outils informatiques de modélisation tridimensionnelle sont généralement considérés comme faisant également partie de la bio-informatique.

Plusieurs initiatives de bio-informatique ont été mises sur pied au cours des dernières années. Le 30 mai dernier, IBM a annoncé à Ottawa la création de blueprint WORLDWIDE, un organisme sans but lucratif qu'elle financera de concert avec MDS Proteomics, une firme spécialisée dans la recherche sur les protéines dans le but d'accélérer la productivité de l'industrie pharmaceutique. Plusieurs organismes canadiens de recherche, comme le Centre national de recherche, l'Institut de recherche sur le cancer et l'Institut de génétique gravitent autour de l'organisme.

Le but de blueprint WORLDWIDE est de faciliter le regroupement et le partage des informations publiques sur la recherche en biologie moléculaire. Au coeur de ces efforts se trouve la base de données BIND (*Biomolecular Interaction Network Database*), composées d'un réseau de serveurs IBM exécutant des applications UNIX et hébergeant des bases de données DB2 d'IBM.

Bien que l'initiative blueprint WORLDWIDE soit sans but lucratif, l'expérience acquise par les participants en matière d'application de technologies informatiques en biologie sera bien pratique pour oeuvrer dans un futur marché qui représentera, selon certains observateurs, plus de 40 milliards US d'ici les trois prochaines années...

L'Amérique étant ce qu'elle est, on ne sera également pas surpris de constater que ces efforts sans but lucratif ont une contrepartie bien mercantile. Plusieurs firmes privées ont été mises sur pied pour tirer profit des services nécessaires à la bio-informatique. Par exemple, DoubleTwist (nommée en l'honneur de la célèbre double hélice de la séquence d'ADN) est une firme qui a pour mandat d'offrir à la communauté scientifique un

accès à des sources d'informations privées. La firme utilise des agents de recherche pour automatiquement colliger des informations et interpréter les données selon des paramètres déterminés par le chercheur. Elle offre également plusieurs outils de visualisation et de modélisation spécifiquement conçus pour faciliter les recherches effectuées sur l'ADN.

De puissants outils

De toute évidence, la bio-informatique exige une formidable puissance de traitement. Voilà pourquoi, dès décembre 1999, IBM a annoncé qu'elle injectait 100 millions de dollars US dans un projet visant à créer le superordinateur le plus rapide du monde. Du nom de Blue Gene (!), le superordinateur sera conçu pour résoudre les grands problèmes fondamentaux posés par ce que les chercheurs nomment le «repliement des protéines».

Une protéine est constituée de chaînes d'acides aminés reliés entre eux. La protéine se replie sur elle-même en une forme tridimensionnelle extrêmement complexe qui détermine sa fonction. Percer les secrets de ce repliement ouvrira la voie à des découvertes spectaculaires.

Sur papier, Blue Gene est certes impressionnant. Le superordinateur utilisera la méthode du traitement informatique dite «massivement parallèle», une approche qui consiste à harnacher simultanément la puissance de traitement d'un grand nombre de processeurs. Blue Gene pousse cette conception de l'informatique à son extrême: plus de un million de processeurs seront mis à contribution, chacun ayant sa propre capacité de mémoire. Dans son ensemble, Blue Gene aura une capacité de traitement d'un petaflop par seconde, soit 1 000 000 000 000 000 (un trillion) d'opérations par seconde. Blue Gene sera ainsi 1000 fois plus puissant que la machine Deep Blue, qui a battu le champion mondial d'échecs Garry Kasparov en 1997, et environ deux millions de fois plus puissant que les ordinateurs de bureau que nous utilisons couramment.

Et tout ça, ne l'oublions pas, ne représente que les débuts de la bio-informatique...

Références

Monographies

Académie des sciences, *Développement et applications de la génomique*, Éditions Technique et Documentation, Paris, 1999, 231 pages

BIOTECanada, *PASSER DE L'INFORMATION À LA CONNAISSANCE : la capacité du Canada en matière de bio-informatique*, Industrie Canada, Ottawa, mars 2000

Conseil de la science et de la technologie, *La bio-informatique au Québec : un levier essentiel du développement des bio-industries*, Gouvernement du Québec, Québec, janvier 2001, 48 pages

Conseil de la science et de la technologie, *Actes du colloque, Sciences et technologies : des visées d'avenir*, Gouvernement du Québec, Québec, janvier 2001, 142 pages

Düx P et Moille F, *Bioinformatics : an inventory and analysis of recent developments in bioinformatics and related areas of research and development*, Institute Prospective Technological Studies, Séville, septembre 1999, 43 pages

Rashidi H et Buehler L, *Bioinformatics basics*, CRC Press, New York, 2000, 185 pages

Sérusclat F, *Génomique et informatique : l'impact sur les thérapies et sur l'industrie pharmaceutique*, Office parlementaire d'évaluation des choix scientifiques et technologiques, Paris, octobre 1999, 204 pages

Périodiques

Chevasssus N, *La course aux protéines*, Science et Vie, janvier 2001, p 74

Karow J, *Reading the book of life*, Scientific American Explore!, 12 février 2001

Brown K, *The Human genome business today*, Scientific American, juillet 2000, p 50

Howard K, *The bioinformatics gold rush*, Scientific American, juillet 2000, p 58

Ezzell C, *Beyond the human* genome, Scientific American, juillet 2000, p 64

Leutwiler K, *Interview with Stuart Kauffman*, Scientific American, 5 juin 2000

Spengler S, *Bioinformatics in the information age*, Science, vol 287, 18 février 2000, p 1221

Attwood T, *The Babel of bioinformatics*, Science, vol 290, 20 octobre 2000, p 471

Roos D, *Bioinformatics – Trying to swim in a sea of data*, Science, vol 291, 16 février 2001, p 1260

Gershon D, *Bioinformatics in a post-genomics age*, Nature, vol 389, 25 septembre 1997, p 417

Abott A, *A post-genomic challenge: learning to read patterns of protein synthesis*, Nature, vol 402, 16 décembre 1999, p 715

Vukmirovic O, Tilghman S, *Exploring genome space*, Nature, vol 405, 15 juin 2000, p 820

Eisenberg D, Marcotte E, Xenarios I, Yeates T, *Protein function in the post-genomic era*, Nature, vol 405, 15 juin 2000, p 823

Lockhart D, Winzeler E, *Genomics, gene expression and DNA arrays*, Nature, vol 405, 15 juin 2000, p 827

Pandey A, Mann M, *Proteomics to study genes and genomes*, Nature, vol 405, 15 juin 2000, p 837

Butler D, *Are you ready for the revolution?*, Nature, vol 409, 15 février 2001, p 758

Bottomley S, *Rapid searching of sequences databases*, Drug discovery today, vol 4, octobre 1999, p 482

Butte A J, *Challenges in bioinformatics : infrastructure, models and analytics*, TRENDS in biotechnology, vol 19, mai 2001, p159-160

O'Donovan C, Apweiler R et Bairoch A, *The human Proteomics initiative (HPI)*, TRENDS in biotechnology, mai 2001, p 178-181

Tsoka S, Ouzounis C A, *Recent developments and future directions in computational genomics*, FEBS Letters, vol 480, septembre 2000, p 42-48

Samson C E, Smith C A, *Computer applications in biomolecular applications : bioinformatics and genome projects*, Biochemical education, vol 28, mai 2000, p 127-131

Ono T, Hishigaki H, Tanigami A, Takagi T, *Automated extraction of information on protein-protein interactions from the biological literature*, Bioinformatics, vol 17, février 2001, p 155-161

Bader G D et Hogue C, *BIND – A data specification for storing and describing biomolecular interactions, molecular complexes and pathways*, Bioinformatics, vol 16, mai 2000, p 465-477

Paton N W et al., *Conceptual modeling of genomic information*, Bioinformatics, vol 16, mai 2000, p 548-557

Rechenmann F, *From data o knowledge*, Bioinformatics, vol 16, mai 2000, p 411

Marcotte E M, Xenarios I, Eisenberg D, *Mining literature for protein-protein interactions*, Bioinformatics, vol 17, avril 2001, p 359-363

Achard F, Vaysseix G et Barillot E, *XML, bioinformatics and data integration*, Bioinformatics, vol 79, février 2001, p 115-125

Journaux

Arcand D, *Quelques 270 millions alloués à la recherche génomique*, La Presse, 4 avril 2001, p A12

Cournoyer L, *Bio-informaticien, une profession méconnue*, La Presse, 10 février 2001, p I7

Dupont L, *Le génome humain sur puce informatique*, La Presse, 6 mai 2001, p C5

Mondoux A, *Biologie et informatique : les nouveaux incontournables*, La Presse, 3 août 2001

Corinne M, *L'informatique face à la logique du vivant*, Le Monde, 25 avril 2001, p 01

Ducruet C, *La bio-informatique, outil indispensable*, Les Echos, 7 février 2001, p 50

Barcelo Y, *L'avenir de la génomique est dans la protéomique*, Les Affaires, 10 mars 2001, p 33

Barcelo Y, *Une carte qui reste vraiment à exploiter*, Les Affaires, 10 mars 2001, p 34

Barcelo Y, *Le milieu de la recherche optimiste comme jamais*, Les Affaires, 10 mars 2001, p 36

Barcelo Y, *Coup d'œil sur quelques grands projets de recherche*, Les Affaires, 10 mars 2001, p 37

Barcelo Y, *Rien ne serait possible sans la bio-informatique*, Les Affaires, 10 mars 2001, p 38

ÉCOLE POLYTECHNIQUE DE MONTRÉAL
3 9334 00299146 9

École Polytechnique de Montréal
C.P. 6079, Succ. Centre-ville
Montréal (Québec)
H3C 3A7