

Instrumentation et biologie moléculaire à l'EMBL : quelques contributions récentes

Le développement intégré, au sein d'un même ensemble, de l'instrumentation et des méthodes d'analyse en biologie est l'un des objectifs initiaux du Laboratoire européen de biologie moléculaire (EMBL). Après quelques difficultés de démarrage, liées à la difficulté de trouver un langage commun entre biologistes, physiciens et spécialistes de la robotisation, cet objectif a été en partie atteint, avec la création de deux unités spécialisées, l'une dans le développement de l'instrumentation biochimique, l'autre dans celui de l'instrumentation physique. D'importantes réalisations sont ainsi à mettre à l'actif de l'EMBL dans le domaine des détecteurs de photons et d'électrons, de la robotisation des systèmes de synthèse de peptides et du traitement informatique du signal dans différentes méthodes d'imagerie.

Christian Boulin

ADRESSE

C. Boulin : *senior scientist, chef de groupe*.
Laboratoire européen de biologie moléculaire, Meyerhofstrasse, 1, Postfach 10.2209, 6900 Heidelberg 1, Allemagne.

Lors de la création du Laboratoire européen de biologie moléculaire (EMBL), une attention toute particulière avait été portée à l'« intégration », au sein de ce futur centre européen, d'une activité de développement d'instrumentation. Cela était sans nul doute une conséquence des contributions récentes de quelques physiciens et chimistes prestigieux à des succès en biologie moléculaire [1]. Associer sous un même toit, biologistes, physiciens et spécialistes en développement d'instrumentation était une idée louable. Dans la réalité, la mise en œuvre de ce concept a été

délicate pour de nombreuses raisons. Pour n'en citer que deux, il faut dire que trouver des scientifiques pouvant assurer le rôle d'interface n'est pas chose facile et qu'il existe généralement quelques différences au niveau des échelles de temps avec lesquelles le biologiste peut et veut voir progresser son travail de recherche et les temps de développement de nouvelles techniques de mesure ou de construction de prototypes d'instruments permettant de valider les idées de départ. Néanmoins, après une quinzaine d'années de fonctionnement, parfois difficile, on peut considérer qu'une grande partie du chemin a

RÉFÉRENCES

1. Thuillier P. Comment est née la biologie moléculaire. In : Blangy D, Brachet J, Bussard A, *et al.*, eds. *La Recherche en biologie moléculaire*. Paris : Éditions du Scuil, 1975 : 13-36.
2. Svendsen K, Koch MHJ, Boulin C, Gabriel A. A synchrotron radiation X-ray diffraction study of the crystallinity of tendon fibres. *Int J Biol Macromol* 1984 ; 6 : 298-302.
3. Bordas J, Koch MHJ, Clout PN, Dorrington E, Boulin C, Gabriel A. A synchrotron radiation camera and data acquisition system for time resolved scattering studies. *J Phys E Sci Instrum* 1980 ; 13 : 938-44.
4. Boulin C. Calibration Channel Unit (CCU) a 16 channels data acquisition module with autonomous memory. *Nucl Instr Meth* 1982 ; 201 : 251-9.
5. Boulin C, Dainton D, Dorrington E, *et al.* Systems for time-resolved X-ray measurements using one-dimensional and two-dimensional detectors : requirements and practical experience. *Nucl Instr Meth* 1982 ; 201 : 209-20.
6. Bartunik HD, Boulin C, Schwab H. Real-time reduction of area detector data by hardware (DACOM) for static and time-resolved crystallography, Colloque C5. *J Phys* 1986 ; 47 (suppl 8) : 157-66.
7. Leyendecker P, Vonthron N, Boulin C. Real-time digital display with adjustable persistence time for 2-D detectors, Colloque C5. *J Phys* 1986 ; 47 (suppl 8) : 167-73.
8. Boulin C, Kempf R, Koch MHJ, McLaughlin SM. Data appraisal, evaluation and display for synchrotron radiation experiments : hardware and software. *Nucl Instr Meth Phys Res* 1986 ; A 249 : 399-407.
9. Boulin C, Kempf R, Gabriel A, Koch MHJ. Data acquisition systems for linear and area detectors using delay line readout. *Nucl Instr Meth Phys Res* 1988 ; A 269 : 312-20.
10. Boulin C, Buldt G, Dauvergne F, *et al.* Anomalous scattering in membrane studies. In : Sweet RM, Woodhead AD, eds. *Synchrotron Radiation in Structural Biology*. New York : Plenum Publishing Corp, 1989 : 83-92.
11. Barbosa AF, Boulin C, Gabriel A, Golding F, Leyendecker P, Riekel C. A VME based data acquisition system for a 2D-gasfilled detector. *Nucl Instr Meth Phys Res* 1991 ; A 302 : 489-92.
12. Dayhoff MO, Barker WC, Hunt LT. Establishing homologies in protein sequences. *Meth Enzymol* 1983 ; 91 : 524-45.
13. Argos P. A sensitive procedure to compare amino acid sequences. *J Mol Biol* 1987 ; 193 : 385-96.
14. Kahn P, Cameron G. EMBL data library. *Meth Enzymol* 1990 ; 183 : 23-31.

été faite tant par les biologistes que par les instrumentalistes.

A l'heure actuelle, les activités d'instrumentation à l'EMBL sont organisées en deux unités : le programme d'instrumentation biochimique et celui d'instrumentation physique. Le programme d'instrumentation biochimique fournit des services aux scientifiques du laboratoire et travaille également à l'établissement de nouvelles méthodologies pour résoudre les problèmes posés par la biologie moléculaire moderne. Dans le passé, les progrès en biologie moléculaire ont souvent été freinés par le temps nécessaire pour déterminer et analyser les séquences des sous-unités constituant l'ADN ou les protéines. Aujourd'hui, de nombreuses machines permettent de déchiffrer ces informations de manière automatique et rapide. Un séquenceur d'ADN utilisant des marqueurs fluorescents a été développé par un groupe de ce programme. Cet instrument représente l'un des robots les plus performants pour la recherche de la séquence des gènes. Il permet la lecture des nucléotides à raison d'une trentaine par minute. Des systèmes encore plus rapides sont actuellement en cours de développement dans ce groupe. Les protéines et acides nucléiques de synthèse sont d'autres outils d'une utilité fondamentale en biologie moléculaire. Deux autres groupes du programme d'instrumentation biochimique travaillent donc sur le développement et l'amélioration de méthodes de synthèse rapides et performantes. D'autres innovations et activités de service concernent les méthodes de tri cellulaire et l'introduction de molécules dans des cellules soit par micro-injection soit par électroporation. Le programme d'instrumentation physique est également un programme prestataire de services pour l'ensemble des chercheurs de l'EMBL et mène concurremment, d'une part, le développement de nouveaux instruments et techniques, et, d'autre part, leurs applications à l'étude des structures biologiques. A présent, ce programme compte trois groupes travaillant dans les directions suivantes : (a) la microscopie électronique ; (b) la microscopie optique et (c) la micro-informatique et la détection. D'ici à la fin de l'année, un quatrième

groupe rejoindra le programme et consacrera ses activités aux applications de la microscopie à effet tunnel et à force atomique à l'étude d'échantillons biologiques dans le cadre d'un effort commun avec le programme de biologie cellulaire. Le groupe de microscopie électronique a pour objectifs principaux de perfectionner des microscopes existants de manière à améliorer leurs performances tant au niveau de la résolution spatiale que des limites de sensibilité et, simultanément, d'offrir de nouvelles méthodes d'analyse. Le groupe de microscopie optique a conçu et construit plusieurs microscopes confocaux à laser particulièrement adaptés aux problèmes posés par la biologie cellulaire. Ces microscopes permettent la reconstruction et la visualisation tridimensionnelle des échantillons grâce à un balayage dans la direction Z. La poursuite du développement de ces instruments se fait dans le but de les rendre plus rapides pour pouvoir suivre, par exemple, les mouvements d'organelles marquées à la fluorescéine à l'intérieur des cellules. Les activités du troisième groupe sont orientées vers cinq thématiques principales : la détection des rayons X, le support matériel et logiciel pour les systèmes d'acquisition et de commande en temps réel, la détection des électrons appliquée à la microscopie électronique, le traitement d'images et la mise en œuvre d'architectures avancées de calcul pour des problèmes d'exploitation de banques de données de protéines.

Dans ce qui suit, nous allons présenter de manière plus détaillée quelques développements entrepris au sein du groupe de micro-informatique et détection, et essayer de montrer leur utilité pour la recherche en biologie moléculaire. Nous commencerons par donner un aperçu de nos contributions à la radiation synchrotron dans le cadre de l'étude des cinétiques de complexes macromoléculaires. Nous aborderons ensuite un sujet important pour le futur de l'exploitation des banques de données : la construction d'un système multiprocesseur très puissant ayant une architecture adaptée aux algorithmes utilisés. La réalisation d'un détecteur et de son système d'acquisition rapide pour des applications en microscopie électro-

que sera décrite avant de passer à un autre sujet d'actualité concernant la robotique au laboratoire, avec une application à la synthèse automatisée de peptides à grande échelle. Pour terminer ce tour d'horizon des activités de notre groupe, nous présenterons deux exemples de traitement d'images concernant l'analyse de particules nanométriques en microscopie électronique et le recalage de coupes sériées en microscopie optique pour la localisation de domaines d'expression génétique.

Détection de rayons X (DESY et futur ESRF)

L'EMBL a une antenne à Hambourg (auprès de DESY : un accélérateur de particules) permettant l'utilisation de la radiation synchrotron pour des activités telles que la cristallographie des protéines et l'étude des systèmes non cristallins : émulsions, solutions, gels, membranes ou fibres. Les expériences de diffraction sur ces systèmes mènent à la compréhension des relations entre les modifications de structure et les fonctions des systèmes biologiques [2]. Ces études sont généralement réalisées avec des résolutions temporelles se situant entre quelques dizaines de micro-secondes et quelques centaines de secondes. A ces fins, nous avons investi beaucoup d'efforts dans la conception et la réalisation de systèmes d'acquisition puissants et intelligents capables de satisfaire les demandes des nombreux scientifiques européens venant effectuer leurs expériences sur nos installations [3-10].

Une part importante de l'instrumentation développée (systèmes de mesure de temps, de visualisation, de séquençage des expériences, logiciels...) a été dupliquée et installée dans les laboratoires nationaux possédant des sources de radiation synchrotron (France, Allemagne, Royaume-Uni). Ces transferts de technologies ou de techniques s'inscrivent parfaitement dans le cadre des objectifs initiaux de l'EMBL quant à la dissémination des travaux. De plus, le laboratoire a conclu récemment un accord de licence avec une société française (BioLogic, Grenoble, France) pour la production et la commercialisation des détecteurs et de

l'électronique associée. Nos travaux dans ce domaine se poursuivent afin de fournir des systèmes encore plus précis et rapides en vue d'améliorer les possibilités expérimentales actuelles et de pouvoir bénéficier pleinement des sources de radiation synchrotron futures telles que celle actuellement en construction à Grenoble (*European Synchrotron Radiation Facility*) [11].

Architecture des systèmes multiprocesseurs (banques de données)

La connaissance des séquences, des structures et des fonctions des protéines est un outil puissant utilisé très largement en biologie moléculaire. Alors que la détermination des séquences de protéines devient une chose relativement aisée grâce à de nouvelles techniques d'analyse, la détermination des structures tridimensionnelles, à l'échelle atomique, reste difficile dans la mesure où elle requiert la cristallisation de la protéine avant de pouvoir mettre en œuvre les techniques de diffraction à l'aide de rayons X.

Partant de l'idée que des morceaux relativement « similaires » dans deux séquences de protéines doivent avoir des fonctions identiques, de nombreux chercheurs ont développé des algorithmes de recherche de similarités, d'homologies et d'alignement de séquences [12]. A l'EMBL, une méthode extrêmement performante a été mise au point [13]. Malheureusement, les moyens de calcul conventionnels (mini-ordinateurs, stations de travail), s'ils permettent aisément la comparaison de deux séquences, impliquent des temps de calcul prohibitifs si l'on essaie de comparer une séquence à l'ensemble des séquences répertoriées dans une banque de données telle SWISSPROT [14]. Une première tentative, réussie, de ramener les temps de calcul à une dimension raisonnable a été faite au laboratoire en utilisant un réseau de 20 « transputers » T800 organisés en anneau, interfacé sur un micro-ordinateur Macintosh IICx [15]. Chaque processeur traite, en parallèle, une partie des informations et, grâce à un système performant de communication, peut envoyer les

résultats intermédiaires de proche en proche. Le gain en temps apporté par un tel système, comparé à un ordinateur Vax 8650 sur lequel l'algorithme avait été développé, est de l'ordre de 20.

A la suite de cet essai, nous avons étudié une autre solution, mettant en jeu moins de processeurs, mais utilisant une architecture différente, tenant compte des particularités de l'algorithme en ce qui concerne les possibilités de traitement en parallèle et en *pipe-line*. Comme élément de base du système multiprocesseur, nous avons utilisé un circuit spécialisé pour le traitement numérique en point flottant (Motorola DSP 96002). Ce circuit possède une puissance de calcul environ 20 fois supérieure à celle du transputeur et possède également une architecture à deux ports d'entrées/sorties, ce qui est d'une absolue nécessité pour éviter les problèmes de congestion au niveau des transferts importants de données entre les étages successifs d'une structure en *pipe-line*.

A l'heure actuelle, nous avons pratiquement terminé la conception de la carte de base et nous avons codé et simulé une part importante de l'algorithme. Notre objectif présent est de pouvoir fournir un ensemble intégré, matériel et logiciel, se connectant sur une station de travail classique ou un ordinateur personnel, et de mettre ainsi à la disposition d'un grand nombre d'utilisateurs les techniques performantes de comparaison de séquences.

Détection des électrons (microscopie électronique)

Il est établi que la microscopie électronique est un outil puissant pour la biologie moléculaire et tout particulièrement pour la biologie cellulaire. Une grande partie du travail est encore réalisée avec ce que l'on peut qualifier de microscopie conventionnelle. Des microscopes plus élaborés, tel le microscope à balayage en transmission (STEM, *scanning transmission electron microscope*), ont été conçus et réalisés dans les deux dernières décennies [16]. La formation de l'image dans un STEM est faite point par point lors du balayage de

l'échantillon par le faisceau d'électrons. L'image est généralement, obtenue à l'aide de deux détecteurs de fond noir et clair (assurant une sélection angulaire des électrons transmis) consistant en un assemblage de scintillateurs et de photomultiplicateurs convertissant les électrons de 100 KeV en lumière puis en un courant mesurable. L'avantage réel du STEM par rapport aux microscopes conventionnels réside essentiellement dans les principes de base de sa conception et dans les possibilités d'analyse qu'il offre (pertes d'énergie des électrons, spectroscopie de rayons X...). Un tel instrument est donc capable de produire simultanément plusieurs images d'un même échantillon, un avantage qui n'a que très peu souvent été exploité, principalement à cause des difficultés technologiques de réalisation de systèmes de détection appropriés [17]. Ce nouveau type de microscope, s'il a réussi à s'implanter de façon indiscutable dans des domaines comme la science des matériaux, n'a certes pas encore réussi à prouver son potentiel pour les applications en biologie. Parmi les différentes raisons à cet état de fait, il faut évidemment mentionner, entre autres, la dégradation de l'échantillon biologique par l'interaction avec les électrons, un phénomène classique en microscopie électronique. Plusieurs approches ont été envisagées pour contourner ces difficultés.

A l'EMBL, on a suivi une voie consistant à mettre en œuvre les possibilités offertes par les techniques cryogéniques, tant au niveau de la préparation des échantillons qu'au niveau du microscope lui-même. Cela a conduit à la construction d'un Cryo-STEM permettant le travail à très basse température afin de préserver au mieux la structure des échantillons [18].

Une autre voie pour améliorer les performances d'un tel microscope est de réduire les doses d'électrons nécessaires pour former une image et de disposer d'un détecteur de géométrie adaptée aux expériences menées. Dans ce but, nous avons mis au point un système de détection multi-éléments en silicium permettant de collecter la majorité des électrons transmis, comprenant un système de traitement digital capable d'engendrer

simultanément, en temps réel, plusieurs images et offrant des possibilités quasi illimitées au niveau de la configuration de détection et, par conséquent, au niveau des modes d'imagerie.

Un nouveau détecteur

Ce nouveau type de détecteur silicium a été développé initialement pour la mesure de particules chargées en physique des hautes énergies [19]. Cette technologie permet, par des procédés lithographiques similaires à ceux utilisés pour la confection de « puces » électroniques, d'obtenir des structures telles que des pistes, des rectangles ou même des formes plus complexes. Par l'utilisation de techniques adéquates de connexion, les îlots ainsi formés constituent des détecteurs individuels qui, reliés à des amplificateurs suivis de discriminateurs, permettent d'associer une impulsion électrique mesurable à chaque électron pénétrant le détecteur. Chaque électron atteignant le détecteur produit en fait $2,3 \times 10^4$ paires électron-trous dans le silicium, ce qui représente une quantité de charge extrêmement faible. La technologie des préamplificateurs hybrides développée au Centre européen pour la recherche nucléaire (CERN) permet pourtant d'extraire un signal rapide identifiant l'interaction d'un électron avec le détecteur [20].

L'utilisation combinée de ces techniques a donc permis la réalisation d'un détecteur de géométrie parfaitement adaptée aux problèmes biologiques, de faire des mesures directes du nombre d'électrons sans passer par des méthodes de conversion intermédiaire (on a donc un meilleur rapport signal/bruit) et par conséquent de travailler avec des doses d'électrons bien plus faibles qu'avec des détecteurs conventionnels.

Le détecteur que nous avons conçu est un ensemble de 30 anneaux concentriques divisés en quatre quadrants [21]. Il permet de collecter les électrons avec une efficacité voisine de 100 %. La *figure 1* montre le détecteur monté dans la partie mécanique réalisant l'adaptation sur le microscope. Les performances d'un tel multidétecteur sont impressionnantes quand on sait que chaque élément de détection est capable de mesurer

jusqu'à 6 millions d'électrons par seconde, ce qui représente un taux de comptage global avoisinant le milliard d'électrons par seconde pour l'ensemble du détecteur. La question qui se pose immédiatement est donc comment traiter tant d'informations sachant que le temps par point-image est de l'ordre de $4 \mu\text{s}$ et que l'on souhaite acquérir des images dont la taille maximale est de 4096×4096 points. En fait, ce qui nous intéresse est de créer des zones d'intégration radiales et annulaires, puis d'effectuer des combinaisons linéaires de ces informations, qui traduiront des caractéristiques physico-chimiques des structures observées [22].

Le prétraitement des données

Afin de limiter le volume des données à transférer entre l'électronique de comptage des électrons et l'ordinateur d'acquisition nous avons décidé de construire une unité de prétraitement des données brutes. En sortie, cette unité fournira jusqu'à un maximum de 8 points-images ce qui représente une réduction du flot de données d'un facteur 8 ou mieux dans le cas où on se limite à enregistrer moins d'images simultanément. La solution à un tel problème peut se trouver dans la mise en œuvre d'un ensemble de calcul puissant, comprenant plusieurs processeurs et fondé sur des concepts de traitement parallèle et de *pipe-line*, deux techniques généralement utilisées dans les super-ordinateurs [23]. La structure parallèle est relativement intuitive dans la mesure où le détecteur lui-même est segmenté en quadrants. La structure en *pipe-line* permet d'obtenir artificiellement un temps de traitement compatible avec le temps de point-image pour chaque ensemble de données. Les résultats des calculs, pour chaque point, sont en fait produits avec un temps de retard équivalent au nombre d'étages du *pipe-line*.

Pour construire un tel système, nous avons développé une carte processeur de base capable de lire une nouvelle donnée, de l'additionner à une valeur intermédiaire et d'envoyer en même temps un résultat partiel d'intégration pour le point-image précédent vers le prochain étage du *pipe-line*, à raison de 10 millions d'opérations par

seconde. Le dernier étage du système est également celui qui transfère les données ainsi réduites, c'est-à-dire les points-images, vers le système général d'acquisition et de commande du microscope. Ce système modulaire, réalisé dans un standard électronique et informatique, le VME, est connecté à une station de travail permettant l'archivage des images et l'intégration du microscope dans le réseau informatique local du laboratoire.

Le logiciel

L'autre aspect important pour la réussite d'un tel projet est le logiciel requis pour faire fonctionner l'ensemble du matériel. Le logiciel peut se décomposer en deux parties : l'interface avec l'utilisateur du microscope pour la définition des conditions d'exploitation et le traducteur permettant d'engendrer les programmes individuels destinés à chacun des pro-

cesseurs du système. L'interface avec l'utilisateur a été développée sur une station de travail de SUN Microsystems sous un environnement graphique (X11 et XView). Elle permet à l'utilisateur de définir interactivement et graphiquement, à l'aide de la souris, le choix des détecteurs et les différents modes de production d'images. Toutes les autres parties du logiciel telles que la traduction en langage machine des équations décrivant le mode d'imagerie, le téléchargement des programmes vers les processeurs, l'initialisation des tâches d'acquisition, sont invisibles à l'utilisateur. Nous avons estimé que cela faisait partie de la convivialité requise pour un tel système : l'avantage étant que l'utilisateur peut se concentrer sur son travail de microscopie plutôt que d'être obligé d'intervenir au niveau de l'ordinateur.

Afin de permettre à l'expérimentateur de juger rapidement de la qualité de

ses images, le système d'acquisition comprend également un module de calcul vectoriel rapide (CSP Inc., USA) fournissant une puissance de pointe de 80 millions d'instructions/s et 66 millions d'opérations/s sur des nombres réels. Ce processeur est principalement utilisé pour le calcul des transformées de Fourier rapides des images. Pour les traitements d'images plus ciblés ou plus complexes, nous avons installé le programme Khoros [24], qui offre l'avantage d'être un environnement de développement puissant faisant appel à des techniques de programmation visuelles.

Robotique (synthèse automatisée de peptides)

De nombreuses techniques utilisées quotidiennement en biologie moléculaire requièrent des manipulations et des procédures de caractère répétitif.

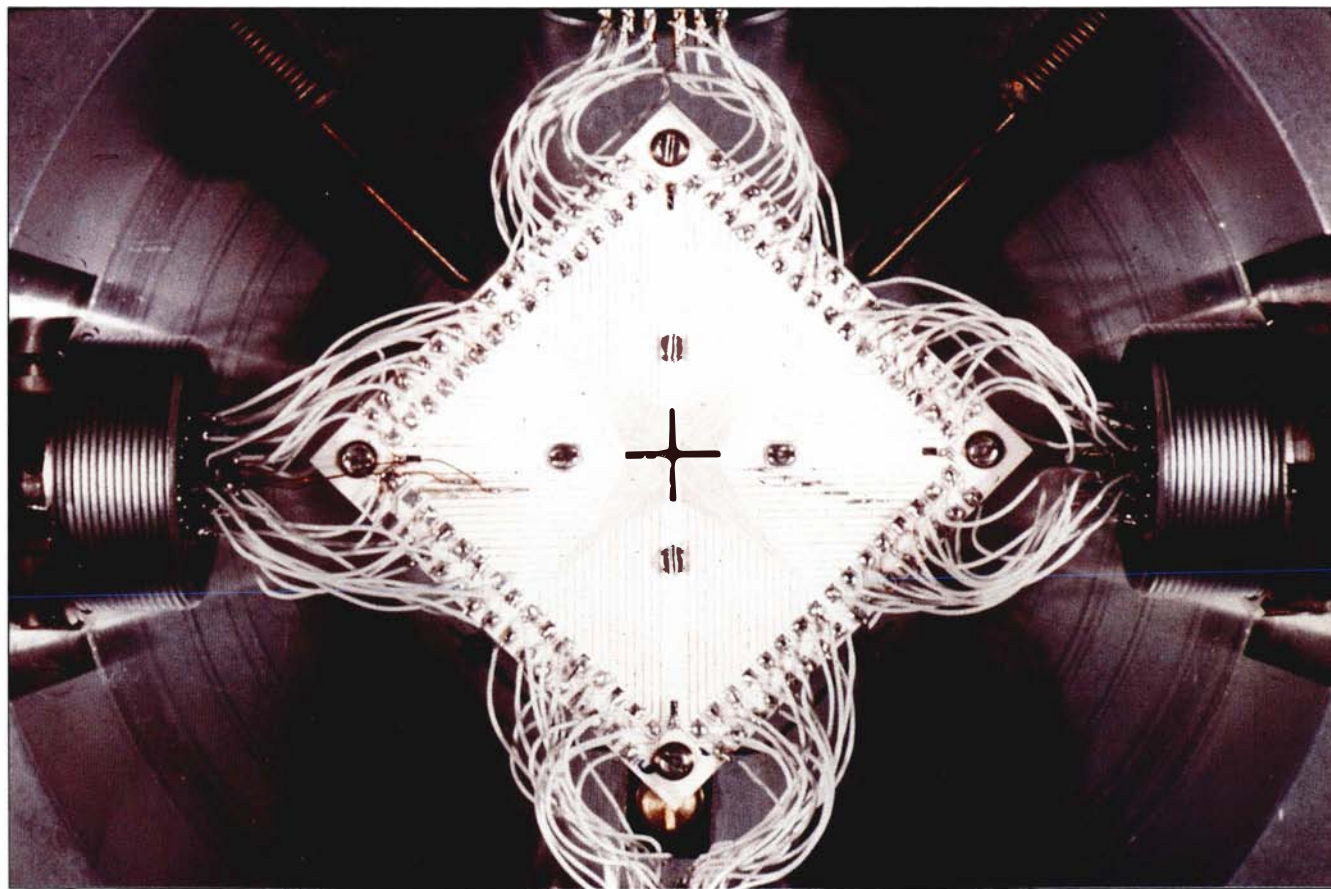


Figure 1. Photographie du détecteur silicium multi-éléments monté dans la partie mécanique d'adaptation sur le Cryo-STEM. Les électrons arrivent sur l'autre face.

Des protocoles tels que l'extraction et la purification de l'ADN, les réactions de séquençage, la production de bibliothèques de clones nécessitent des transferts répétitifs de microvolumes de réactifs et, éventuellement, des modifications cycliques de la température des échantillons. Ces techniques, demandant généralement beaucoup de temps et d'attention de la part du personnel chargé de leur exécution, sont des candidates de choix pour une automatisation à l'aide de robots. Nous avons donc entrepris un travail de développement de logiciel [25] avec comme objectif de fournir des outils suffisamment généraux pour être en mesure de couvrir un grand nombre de problèmes. La

synthèse de peptides est un exemple de mise en œuvre des outils logiciels développés.

Le peptide : un outil puissant pour le biologiste

Il y a encore quelques années, les peptides n'étaient connus des biologistes que sous la forme d'hormones naturelles ou de produits de dégradation enzymatique de protéines. La synthèse chimique, extrêmement lourde et coûteuse, n'était possible que dans les mains de chimistes experts. L'explosion des techniques associées à la synthèse peptidique au cours des dernières années a fait qu'aujourd'hui elle est devenue praticable également dans des laboratoi-

res n'ayant pas une spécialisation particulière dans cette technique.

Le domaine principal d'utilisation des peptides synthétiques est l'immunologie où l'on étudie la spécificité d'anticorps. Pour l'analyse d'épitopes, on « découpe sur le papier » une protéine antigénique en courts fragments se recouvrant et on synthétise, par la suite, l'ensemble de ces peptides. Ils sont alors analysés individuellement (à l'aide de tests ELISA) quant à leurs réactions avec des anti-sérums. Les peptides peuvent alors être utilisés soit pour des diagnostics à la place des agents pathogènes, soit pour donner des informations fondamentales sur la protéine dans le cadre de développements de vaccins synthétiques.

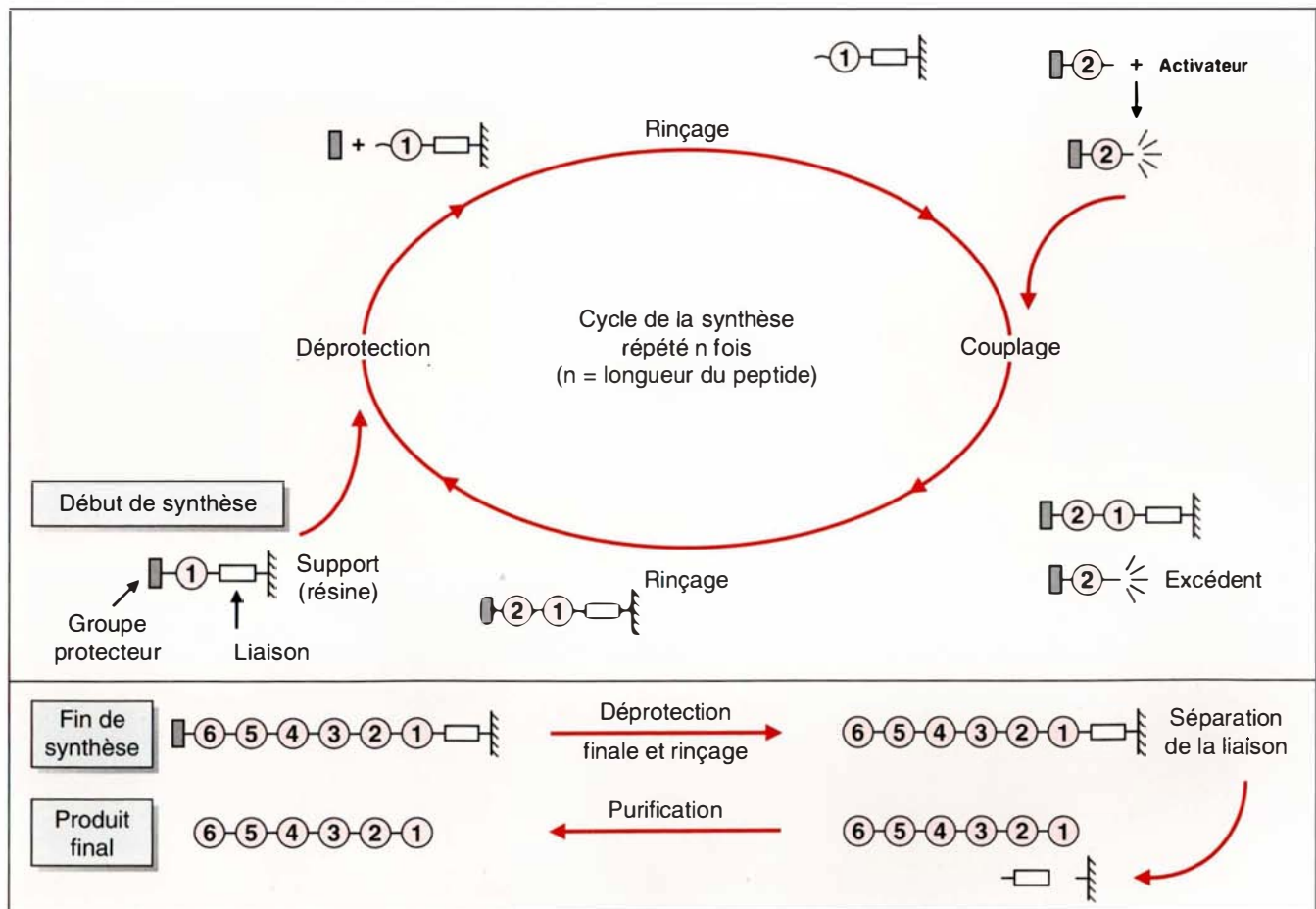


Figure 2. **Les différentes étapes de la synthèse peptidique.** La partie haute représente les cycles de construction des peptides sur un support solide. Le bas de la figure montre les différentes étapes nécessaires pour arriver au produit final.

La synthèse par la méthode de Merrifield

La méthode de synthèse la plus importante est encore la méthode en phase solide développée en 1963 par Merrifield. L'idée décisive était de faire croître la chaîne peptidique sur un support chimiquement stable pendant la synthèse. De ce fait, il est possible d'éliminer par simple lavage tous les produits chimiques requis pour les différentes réactions au cours des étapes successives. Les chaînes peptidiques croissantes sont attachées au support solide de façon covalente et n'en seront séparées qu'en fin de synthèse. La construction des chaînes peptidiques se réalise donc suivant un schéma simple et cyclique consistant à réaliser l'activation et le couplage d'un nouvel acide aminé, suivis de la déprotection de l'acide aminé nouvellement couplé comme on peut le voir sur la *figure 2*.

L'automatisation de la méthode

La synthèse peptidique est donc *a priori* une technique pouvant faire l'objet d'une automatisation poussée. La possibilité d'automatisation offre alors un autre avantage immédiat : celui de la synthèse en parallèle d'un grand nombre de peptides. Cela permet de répondre aux besoins croissants en peptides synthétiques. Les avantages de l'automatisation des procédures par rapport à une exécution manuelle sont indiscutables quand il s'agit de synthétiser simultanément quelques centaines de peptides : ils se traduisent par un gain au niveau des ressources humaines du laboratoire dans la mesure où le technicien peut s'occuper de d'autres travaux plus complexes pendant le déroulement de la synthèse (pouvant nécessiter de quelques heures à deux jours), d'une plus grande précision au niveau de la gestion des temps de couplage ou de réaction (ce qui, de toute évidence, se traduit par des produits de meilleure qualité) et finalement par un risque d'erreur de manipulation ramené à zéro.

Les contraintes principales d'un système automatisé de synthèse multiple concernent les points suivants : (a) une définition simple des séquences ; (b) le calcul précis des quantités de produits pour la synthèse ; (c) une documentation très complète ;

(d) une méthode rapide et efficace de couplage ; (e) des temps de cycle de synthèse courts ; (f) une méthode simple pour extraire les peptides synthétisés.

Le développement d'un automate permettant la synthèse simultanée de 48 peptides (de longueur inférieure ou égale à 30 acides aminés) et des logiciels associés a donc été entrepris en collaboration avec un autre groupe du Programme d'instrumentation biochimique, qui possédait déjà une expérience importante au niveau de la chimie impliquée [26].

La synthèse peptidique étant avant tout un problème de micromanipulation de produits chimiques sous forme liquide, l'instrument a été conçu et réalisé en utilisant un robot de pipetage du commerce (Gilson Medical Electronics S. A. Villiers-le-Bel, France). Cet équipement de base consiste en un robot cartésien, c'est-à-dire assurant le déplacement d'une aiguille en X, Y, Z, et d'un module de commande d'une seringue permettant l'aspiration ou la distribution de liquides. Une vue schématique du système est donnée par la *figure 3*.

L'élément central du système est le réacteur, conçu au laboratoire et construit par notre atelier de mécanique de précision, qui permet d'effectuer la synthèse en flux continu dans des petites colonnes de réaction (ce qui représente un atout important pour la qualité finale des produits synthétisés). Le support pour la synthèse se trouve dans les colonnes de réaction qui sont munies d'un filtre et connectées au bloc de microvannes. Ce bloc, qui sert de support pour les 48 petits réacteurs, permet également le contrôle de l'aspiration des produits réactifs et des produits de rinçage grâce à une connexion vers un système de vide. Les acides aminés, l'activateur ainsi que les autres réactifs se trouvent, en solution, dans des réservoirs individuels regroupés dans un conteneur. Au cours de la synthèse, les acides aminés et les réactifs sont prélevés dans les réservoirs respectifs, séparés entre eux par de petits segments d'air, et ensuite pipetés dans la colonne de réaction correspondante. L'activation des dérivés se fait donc au moment où les différents produits sont délivrés

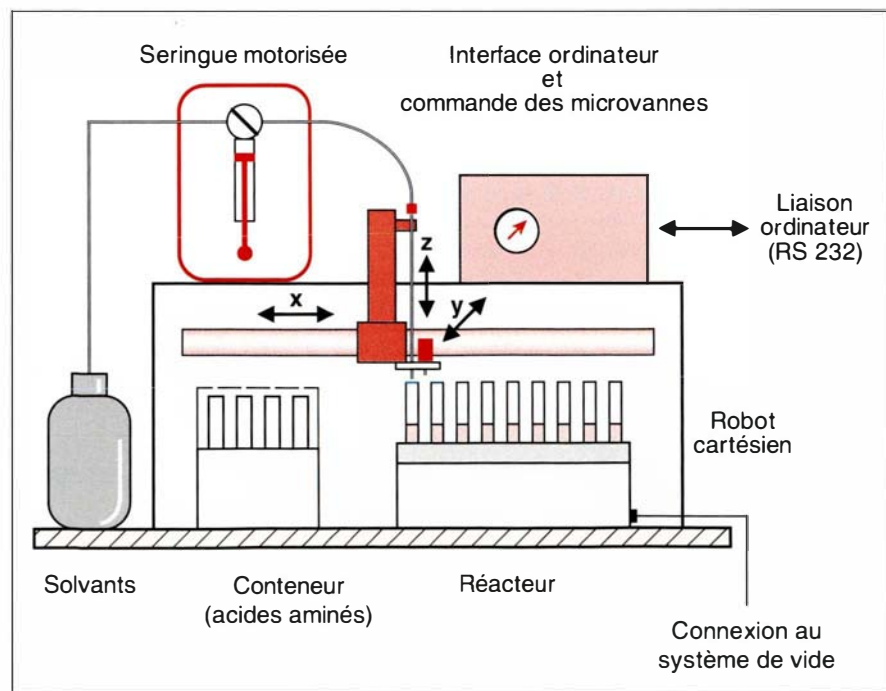


Figure 3. Représentation simplifiée des différents composants de l'automate de synthèse réalisé.

RÉFÉRENCES

15. Vogt G, Argos P. Searching for distantly related protein sequences in large databases by parallel processing on a transputer machine. *Comput Applic Biosci* 1992 ; 8 : 49-55.
16. Crewe AV, Langmore JP, Isaacson MS. Resolution and contrast in the scanning transmission electron microscope. In : Siegel B, Beaman DR, eds. *Physical Aspects of Electron Microscopy and Microbeam Analysis*. New York : Wiley, 1975 : 47-62.
17. Rose H. Nonstandard imaging methods in electron microscopy. *Ultramicroscopy* 1977 ; 2 : 251-67.
18. Jones AV, Homo JC, Unitt BM, Ebsster N. The CryoSTEM : a STEM with superconducting objective lens. *J Microsc Spectrosc Electron* 1985 ; 10 : 361-70.
19. Kemmer J. Improvement of detector fabrication by the planar process. *Nucl Instr Meth Phys Res* 1984 ; A 226 : 89-93.
20. Jarron P, Goyot M. A Fast current sensitive preamplifier (MSD 2) for the silicon microstrip detector. *Nucl Instr Meth Phys Res* 1988 ; A 269 : 156-62.
21. Boulin C, Epstein A. Parallel and pipelined front-end for multi-element silicon detectors in scanning electron microscopy. *IEEE Trans Nucl Sci* 1992 ; 39 : 226-31.
22. Dekker NH, de Lang H. Differential phase contrast in a STEM. *Optik* 1974 ; 41 : 4 : 452-546.
23. Epstein A. Multiprocessor system with parallel and pipelined architecture for a STEM quadrant detector. *Rev Sci Instrum* 1988 ; 59 : 2438-44.
24. Rasure J, Williams C, Argiro D, Sauer T. A visual language and software development environment for image processing. *J Imag Syst Techn* 1990 ; 2 : 183-99.
25. Frank R, Bossherhoff A, Boulin C, Epstein A, Gausepohl, Ashman K. Automation of DNA sequencing reactions and related techniques : a workstation for micromanipulation of liquids. *Biotechnology* 1988 ; 6 : 1211-3.
26. Gausepohl H, Kraft M, Frank R. *In situ* activation of Fmoc amino acids by BOP in solid phase peptide synthesis. In : Jung G, Bayer E, eds. *Peptides 1988*. Berlin : De Gruyter, 1989 : 241-3.
27. Olivo JC, Kahn E, Halpern S, Fragu P. Image registration and distortion correction in ion microscopy. *J Microscopy* 1991 ; 164 : 263-72.
28. Olivo JC. Nanometer-size metal particle analysis in microscopical biological images. SPIE/IS and T. *Image Process Algorithms Techn III* 1992 ; 1657 : 558-62.
29. Price M, Lazzaro D, Pohl T, et al. Regional expression of the homeobox gene NKX-2.2. in the developing mammalian forebrain. *Neuron* 1992 ; 8 : 241-55.

dans le réacteur. Après le temps de couplage (modifié en fonction du pas de synthèse par un coefficient multiplicateur défini à l'avance par l'utilisateur), on ouvre toutes les microvannes et on aspire les réactifs. Afin d'optimiser les temps de cycle par pas de synthèse, les opérations de rinçage se font par groupes de six colonnes de réaction grâce à un dispositif à microseringues multiples que nous avons rajouté sur le bras du robot. La distribution des produits de rinçage est faite par une suppression d'azote au niveau des conteneurs et par des microvannes contrôlant les différentes arrivées. Le contrôle des volumes pour ces produits est peu critique, et, de ce fait, il est réalisé par le contrôle des temps d'ouverture des microvannes.

Le développement du logiciel

Deux types de logiciel sont nécessaires pour de telles applications : (a) une interface utilisateur permettant de définir ce qu'on veut réaliser et (b) un ensemble de modules de gestion des tâches de commande du robot en fonction des conditions de la synthèse. Nous avons choisi un micro-ordinateur du type Macintosh (*Apple Computer, Inc.*) principalement

pour l'environnement graphique et la convivialité.

• L'interface utilisateur

Une attention particulière doit être portée à la manière dont l'opérateur va définir les peptides, les séquences et la manière de découper celles-ci. Pour cela, nous avons développé un module logiciel « éditeur d'acides aminés » permettant l'entrée et l'édition de peptides et de séquences à partir d'un clavier graphique, l'importation de fichiers (par exemple, des séquences contenues dans une banque de données) et la sauvegarde sur disque. La figure 4 montre un exemple de l'interface graphique utilisée.

Afin de préparer la synthèse dans des conditions optimales et de conserver l'ensemble des conditions et paramètres, nous avons également écrit un module permettant d'obtenir un « rapport » complet décrivant les conditions de la synthèse (les échelles, les temps...) sélectionnées et donnant à l'utilisateur, pour chacun des produits chimiques impliqués, les poids et volumes nécessaires pour la préparation des solutions. Ce module fournit également à l'utilisateur la liste de tous les peptides qui vont être synthétisés avec des indications telles que les

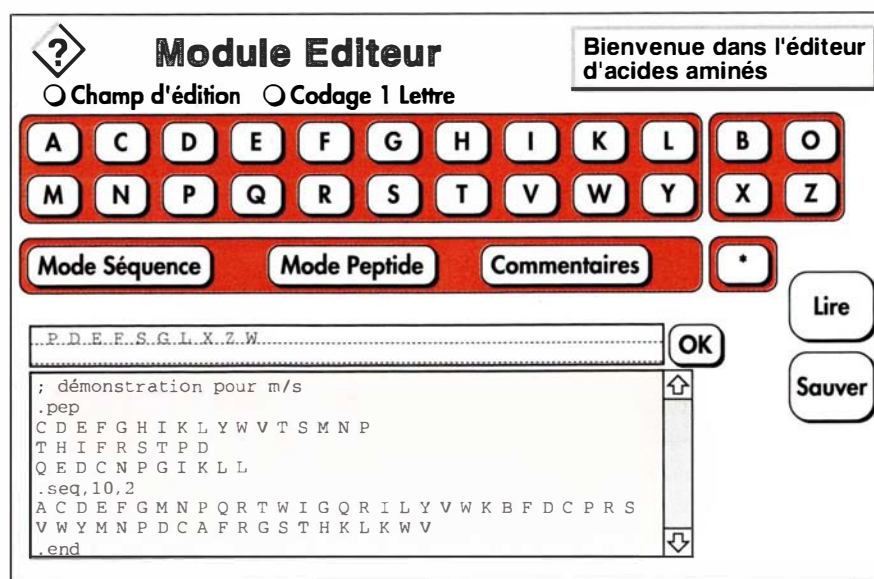


Figure 4. Copie d'écran du module éditeur de peptides et de séquences inclus dans le logiciel de synthèse peptidique.

masses moléculaires des produits et les coordonnées dans le réacteur. Ces aspects de gestion se sont avérés très importants et nous ont également conduits à écrire un logiciel pour la création automatique d'un journal de bord lors du déroulement de la synthèse.

La synthèse peptidique étant un domaine en pleine évolution, nous avons organisé notre logiciel de manière modulaire, ce qui donne à l'utilisateur une très grande souplesse tant dans le choix des procédures que dans la définition des paramètres.

• *La gestion des tâches du robot*

Le lien entre le robot et l'ordinateur de commande est assuré par une interface série standard (RS 232). L'exécution des opérations de base du robot est assurée par l'envoi de courts messages composés de caractères et de coordonnées absolues ou relatives. L'écriture de gros programmes avec de telles directives n'est pas très pratique. De même, la gestion des coordonnées réelles n'est pas très élégante, surtout dans un contexte de développement où l'on souhaite une souplesse maximale. Afin de contourner ces obstacles, nous avons défini un ensemble d'instructions correspondant aux actions requises pour le déroulement des processus. Un fichier de description de la géométrie des différents conteneurs et du réacteur permet de travailler en coordonnées symboliques ; les calculs des coordonnées réelles sont effectués au moment de l'exécution du programme. Durant le déroulement de la synthèse, l'utilisateur peut suivre à l'écran l'état d'avancement (type de séquence en cours, affichage du peptide traité...).

Cette machine est à présent disponible sur le marché suite à un accord entre un partenaire industriel (Abimed, Langenfeld, Allemagne) et le laboratoire : une autre manière de promouvoir les développements technologiques réalisés au LEBM.

Traitement d'images (microscopie)

Depuis quelques années, l'imagerie microscopique a connu une sorte de nouvelle naissance en biologie grâce à l'apport des nouvelles techniques de traitement numérique des images.

m/s n° 2 vol. 9, février 93

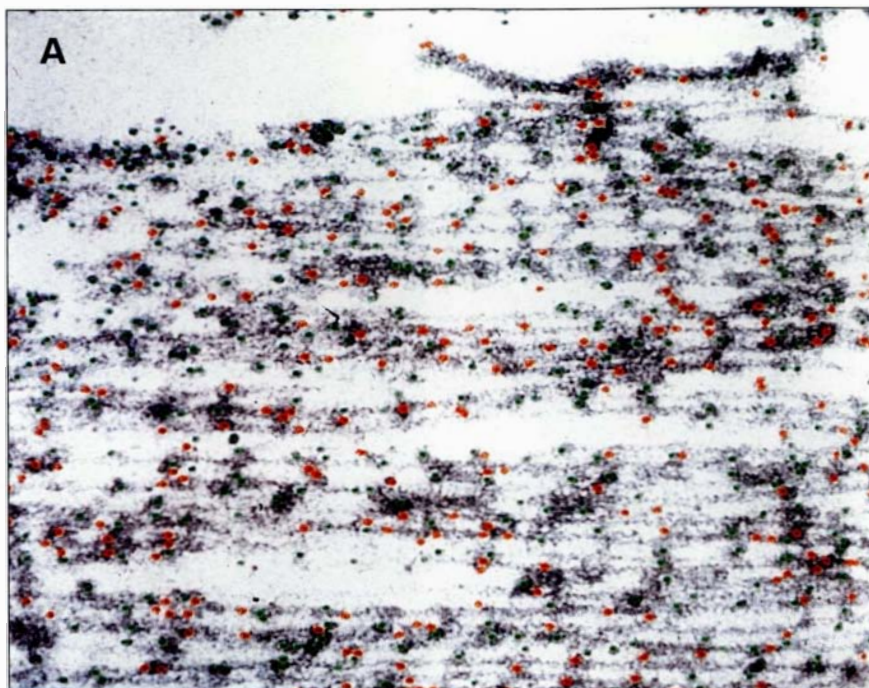


Figure 5. (A) *Image originale d'immuno-microscopie électronique d'un muscle d'insecte marqué avec un anticorps contre la troponine-H attaché à l'or colloïdal de 3 nm et un anticorps contre la troponine-T attaché à l'or colloïdal de 5 nm.* (B) *Identification et classification des deux marquages avec les anticorps. L'image résultante montre la détection et la classification des particules métalliques suivant leur taille (particules de 3 nm encerclées en rouge, particules de 5 nm encerclées en vert).*

Certes, l'analyse visuelle de photographies, d'écrans fluorescents, de cartographies à relevé manuel, a fourni des moyens de recherche précieux aux biologistes. Mais le traitement numérique offre un certain nombre d'avantages qui en font un outil privilégié : si, à la précision, la fiabilité et la rapidité, nous ajoutons la possibilité d'automatisation, il est facile de comprendre l'essor de cette discipline. Pour être complet, il convient de signaler que le traitement numérique ne se limite pas uniquement à l'exploitation des images mais qu'il est à la base même des toutes nouvelles méthodes d'imagerie comme la microscopie confocale, la microscopie à effet tunnel ou à force atomique qui révolutionnent actuellement la biologie.

Si la preuve de l'utilité du traitement numérique en biologie n'est certes plus à faire, il est néanmoins essentiel de fournir aux biologistes des outils parfaitement adaptés à leurs besoins et non pas d'essayer de faire coller à ces besoins des méthodes développées dans le cadre d'autres thématiques. C'est ainsi qu'un projet de traitement d'image destiné à offrir aux biologistes du logiciel adapté et convivial a été lancé dans notre groupe. Le développement de ce logiciel répond à la nécessité de fournir aux biologistes du LEBM des programmes ou des méthodes originales adaptés aux problèmes spécifiques soulevés par leurs recherches. Plusieurs modules logiciels ont été développés pour la restauration et l'amélioration des images ainsi que pour leur segmentation et leur interprétation. Dans la suite, nous présentons deux méthodes qui ont été utilisées en collaboration avec des groupes du LEBM : la première concerne le comptage de particules de taille nanométrique dans des images de microscopie électronique (en collaboration avec K. Leonard, du programme de biologie structurale) et la seconde, le recalage de coupes sériées (en collaboration avec M. Price du programme d'expression génétique).

Analyse de particules nanométriques

En immunomicroscopie électronique le marquage à l'aide de particules métalliques nanométriques est une

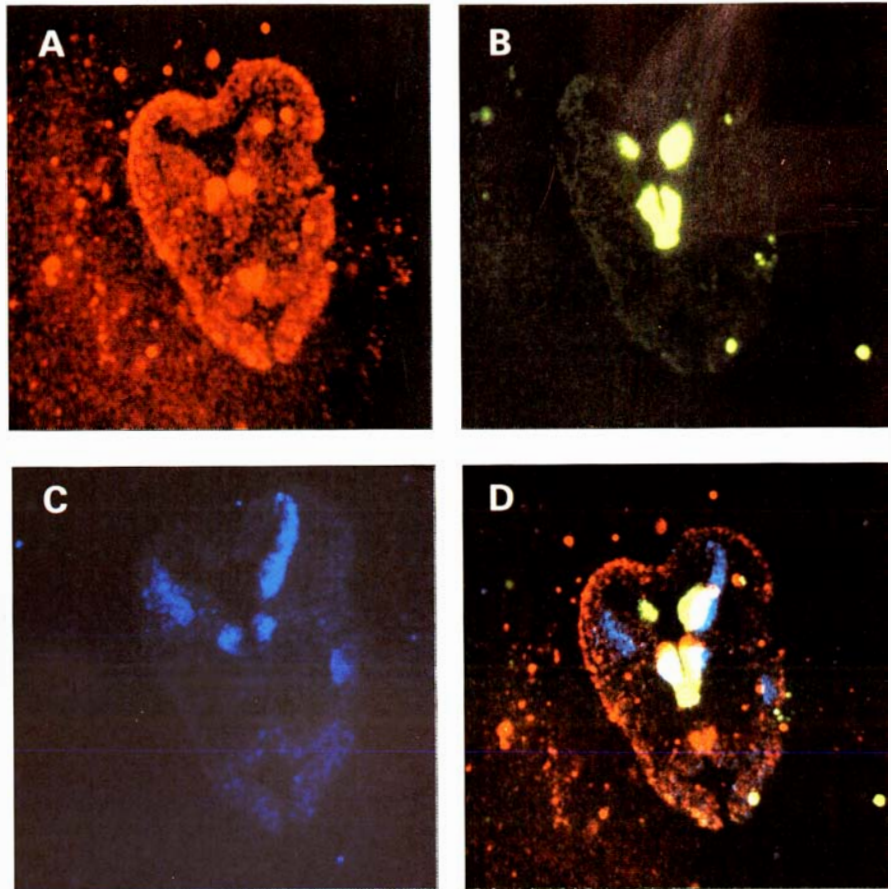


Figure 6. (A, B, C) Coupes sériées d'un cerveau de souris marquées pour identifier respectivement les domaines d'expression des gènes *Nkx-2.2*, *TTF-1* et *Dlx*. (D) Reconstruction à partir des trois coupes permettant de situer les positions relatives et les recouvrements des domaines d'expression. L'image finale, où les trois coupes sont affectées respectivement à une couleur suivant le schéma (rouge, *Nkx-2.2*), (vert, *TTF-1*) et (bleu, *Dlx*), permet de bien apprécier les positions relatives et les recouvrements de ces domaines d'expression.

technique bien établie, qui permet d'identifier la distribution d'antigènes dans des systèmes cellulaires. Pour marquer des antigènes de surfaces cellulaires, les cellules ou le tissu sont incubés, après d'autres étapes préliminaires, avec des complexes or colloïdal/immunoglobuline qui se fixent spécifiquement à des récepteurs de surface. Les particules métalliques sont alors visualisées en microscopie électronique à transmission, qui permet de localiser des particules de dimension nanométrique (supérieure à 2 nm). En utilisant des particules d'or de taille différente attachées à des marqueurs différents, il est possible d'obtenir, dans un même cliché,

des marquages multiples. Pour analyser ces images de façon quantitative, il est alors nécessaire de détecter et de caractériser automatiquement les particules. Dans la méthode que nous avons développée [27], cela est fait en deux étapes : un pré-traitement, destiné à rehausser l'information, suivi d'une extraction, qui réalise l'analyse. Le rehaussement consiste en une série de filtrages qui ont pour résultat de réduire le bruit et d'uniformiser les inhomogénéités. Le schéma d'extraction est lui-même divisé en deux étapes : localisation de tous les candidats possibles à être une particule, puis tri des seules véritables particules, étape qui elle-même contient une procédure

pour distinguer deux populations de particules de taille différente. Les résultats de la méthode sont illustrés dans la *figure 5a et 5b p. 187*. L'image originale est une image d'immunomicroscopie électronique de muscle d'insecte marqué avec un anticorps contre la troponine-H attaché à l'or colloïdal de 3 nm et un anticorps contre la troponine-T attaché à l'or colloïdal de 5 nm. L'image résultante montre la détection et la classification des particules métalliques suivant leur taille (particules de 3 nm encerclées en rouge, particules de 5 nm encerclées en vert). On pourra apprécier la qualité de la méthode qui n'a pas trouvé de particules artefactuelles dans la zone très inhomogène située dans la partie droite de l'image et qui a bien discriminé les deux populations d'or.

Recalage de coupes sériées

En biologie, l'autoradiographie est une technique de choix pour obtenir des informations détaillées sur la localisation de diverses substances marquées radioactivement. Combinée à l'autoradiographie, l'hybridation *in situ* permet de marquer et d'analyser les domaines d'expression spécifiques d'un gène dans, par exemple, des coupes de tissus ou d'organes. La visualisation en champ noir des grains d'argent produits par la sonde radioactive hybridée permet en effet de mettre en évidence les ARNm spécifiques d'un gène. Bien que très puissant, ce moyen d'analyse ne permet cependant l'étude de gènes que de façon individuelle. Nous avons donc développé une technique de superposition et de recalage destinée à la visualisation simultanée de coupes sériées et permettant l'étude des localisations respectives de plusieurs domaines d'expression.

La procédure consiste en une étape d'approximation interactive en temps réel suivie par une étape différée de raffinement. Pendant la première phase, on réalise la superposition en pseudo-couleurs d'une image préalablement acquise et servant de référence, et de l'image en direct provenant de la caméra montée sur le microscope. En déplaçant la lame observée, on peut, avec un peu de maîtrise, réaliser ainsi le recalage global entre les deux champs d'observa-

tion. Si les distorsions entre les deux lames sont importantes, on peut ensuite procéder en temps différé à une étape de recalage fin entre les images. Sur la base de points de référence choisis manuellement, une transformation affine suivie d'une transformation élastique sont successivement appliquées aux images de façon à mettre en correspondance des propriétés locales et des structures communes aux lames observées [28]. Cette procédure en deux temps permet de tenir compte et de corriger les deux types de déformation qui peuvent apparaître entre deux images, à savoir des différences globales dues à un défaut de positionnement des lames et/ou des distorsions locales causées par le processus de préparation des coupes sériées. Les résultats de la méthode sont illustrés dans la *figure 6* où est montrée la superposition d'images de trois coupes sériées provenant d'un cerveau embryonnaire de souris. Les trois lames ont été préparées et marquées pour révéler respectivement les domaines d'expression des gènes Nkx-2.2, TTF-1 et Dlx [29]. Lors de la superposition numérique, chacune des lames a été recalée sur la précédente. L'image finale, où les trois coupes sont affectées respectivement à une couleur suivant le schéma (rouge, Nkx-2.2), (vert, TTF-1) et (bleu, Dlx), permet de bien apprécier les positions relatives et les recouvrements de ces domaines d'expression.

Conclusion

Nous avons donc présenté, de manière simplifiée, quelques activités de notre groupe. Nous espérons avoir convaincu le lecteur de l'intérêt d'une activité d'instrumentation étroitement associée à un environnement de biologie moléculaire. L'atout majeur d'une structure comme celle du LEBM est très certainement, pour les instrumentalistes, de bénéficier d'échanges quasi-journaliers avec les biologistes et d'arriver ainsi à mieux saisir leurs besoins et leurs souhaits. Cela permet un choix plus aisé et certainement plus ciblé des objectifs de développement. Les transferts de technologie réalisés, vers des laboratoires nationaux ou vers l'industrie, sont très certainement une preuve de la justesse des choix effectués ■

Summary

Instrumentation and molecular biology at the EMBL : some recent contributions

When the European Molecular Biology Laboratory was founded, one of its original goals was the integration of biological, instrumental and methodological development. Although this integration has not always been easy during the last fifteen years, the Institute now runs two main branches of activities oriented towards biochemical and physical aspects. The activities of the microcomputing and detection group are focussed on the following areas : X-ray detection and data acquisition systems for fast time-resolved measurements, hardware and software support for real-time applications, electron detection applied to electron microscopy, image processing and the implementation of digital signal processing for computation, intensive problems such as protein sequence comparison with large databases. Some aspects of our work have been presented. These include : (a) the development of a novel silicon multi-element detector and its parallel and pipelined front-end ; (b) an application to robotize multiple peptide synthesis and finally some recent contributions to image processing.

TIRÉS A PART

C. Boulín.