

# Point fort : au cœur de la jungle des protéines

Autor(en): **Meili, Erika / Roth, Patrick**

Objektyp: **Article**

Zeitschrift: **Horizons : le magazine suisse de la recherche scientifique**

Band (Jahr): - **(2006)**

Heft 70

PDF erstellt am: **15.05.2024**

Persistenter Link: <https://doi.org/10.5169/seals-551933>

## **Nutzungsbedingungen**

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

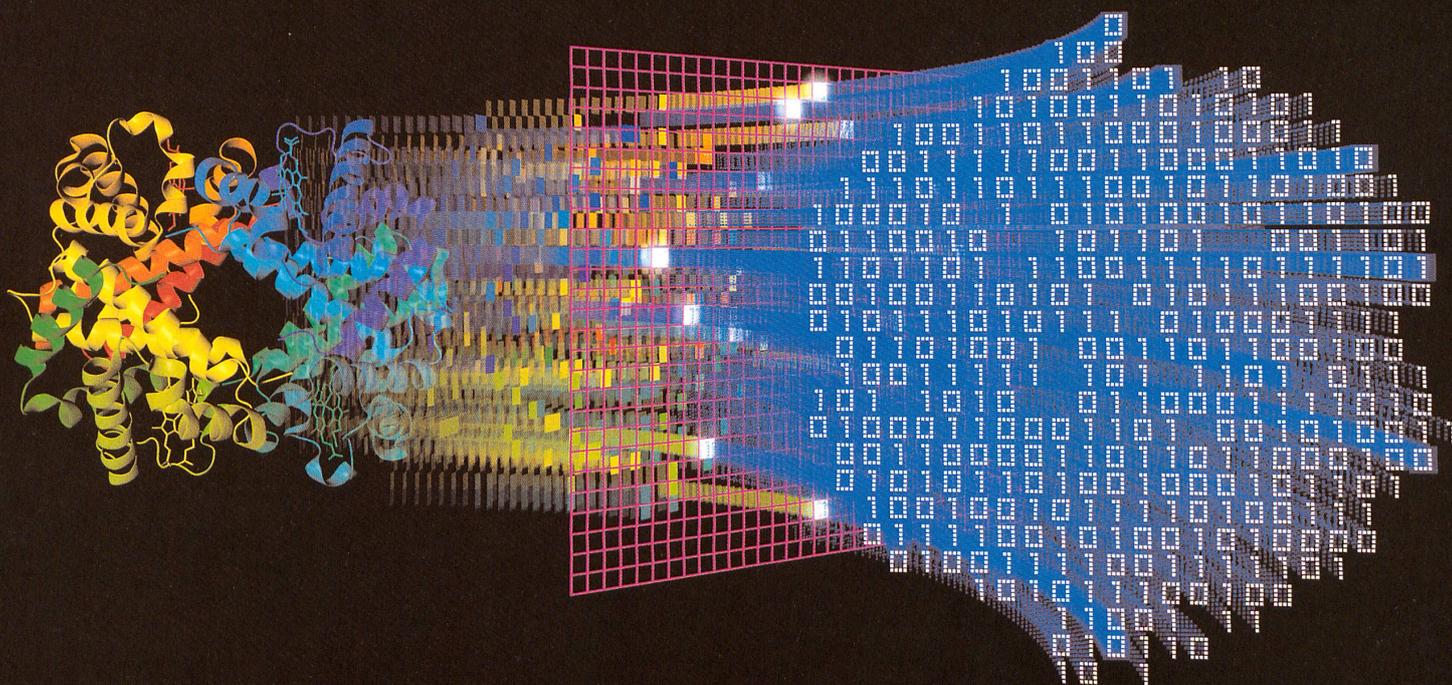
## **Haftungsausschluss**

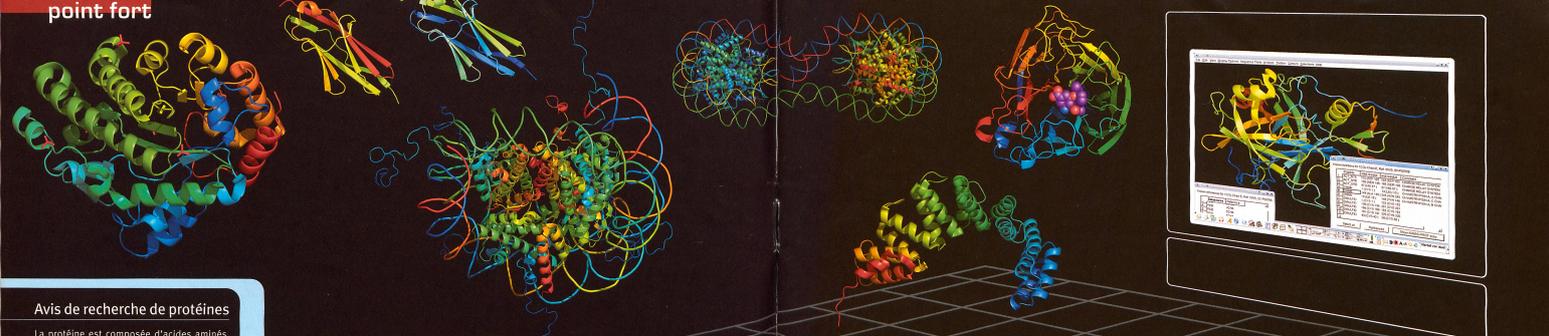
Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.



## Au cœur de la jungle des protéines

Au cours de son développement, le corps humain fabrique au moins 30000 protéines différentes. Leurs fonctions sont toutefois encore largement méconnues. Des banques de données sur les protéines fournissent une aide en permettant des comparaisons entre des protéines connues et des protéines qui viennent d'être découvertes. La banque de données la plus renommée dans ce domaine s'appelle Swiss-Prot et a été fondée, il y a vingt ans, par le biochimiste genevois Amos Bairoch. Photo: Severin Nowacki/Illustration: Mathias Bader





Avis de recherche de protéines

La protéine est composée d'acides aminés, petites molécules organiques contenant de l'azote. L'être humain a besoin de vingt acides aminés pour fabriquer ses protéines naturelles. Huit ne peuvent pas être produits dans son organisme et sont fournis par l'alimentation. Au cours de son développement, le corps humain forme au moins 30000 protéines différentes qui sont les éléments essentiels des cellules et déterminent

presque tous les processus vitaux. Il y a par exemple la protéine structurale qui confère solidité et élasticité aux cellules. Le collagène du tissu conjonctif et la kératine des cheveux en font partie. Plusieurs protéines influencent le mouvement des muscles. Certaines transportent des substances comme l'hémoglobine dans le sang qui alimente le corps en oxygène. D'autres encore transmettent des informations en tant qu'hormone d'un organe à l'autre ou servent d'anticorps contre les infections. Mais la majorité des protéines sont des enzymes, ce qui signifie qu'elles permettent ou accélèrent les réactions biochimiques; elles éliminent des substances nutritives, fabriquent de l'énergie, forment de nouveaux éléments de protéines ou encore doublent l'ADN lors de la division cellulaire.

La longueur des chaînes va de deux à mille acides aminés et les chaînes ayant une longueur de deux à cent acides aminés sont appelées peptides. On ne parle de protéine que lorsque la chaîne comprend plus de cent acides aminés. La plus grande protéine humaine est la titine, composée de 26920 acides aminés. Elle agit sur la stabilité et l'élasticité des muscles.

# Protéines à la mission encore inconnue

Les protéines nécessaires à la vie, mais beaucoup d'entre elles sont encore à peine connues. Il n'est pas seulement intéressant d'identifier la succession des éléments qui les composent mais aussi et surtout leur structure tridimensionnelle.

Par Erka Meili

Le fromage, les haricots, le poisson et la viande contiennent tous des protéines et ces aliments sont sains car ils livrent au corps humain les acides aminés lui permettant de fabriquer ses propres protéines.

Le corps a besoin des protéines, sortes de machines moléculaires qui peuvent servir de matériaux de construction, de transporteurs de signaux ou d'enzymes pilotant les réactions biochimiques. L'absence d'une protéine ou son mauvais fonctionnement peut perturber le développement ou engendrer une maladie comme Alzheimer ou le cancer. Pour comprendre un dysfonctionnement, il est nécessaire en premier lieu de connaître les fonctions des protéines. «En se basant sur un gène, on peut prédire les très nombreuses protéines d'un être vivant», explique Markus Grütter, professeur de biochimie à l'Université de Zurich. Mais on ne saura toutefois pas

encore ce qu'elles font. La structure tridimensionnelle d'une protéine, et en particulier sa structure fine, détermine sa fonction. «Si l'on veut savoir pourquoi une réaction chimique a lieu, on doit connaître la structure atomique de l'endroit où la protéine produit la réaction», souligne Markus Grütter. Il dirige le Pôle de recherche national «Biologie structurale» qui existe depuis cinq ans et auquel participent douze groupes de recherche de Zurich et Bâle. Leur but est de découvrir la structure et la fonction des protéines importantes.

Dans la pratique, la connaissance de la structure fine sert à créer, grâce à l'ordinateur, de nouvelles combinaisons (ce qu'on appelle le «drug design») qui se fixent à l'endroit de réaction de la protéine et ainsi la bloquent. Les scientifiques «dessinent» ainsi des substances en ciblant les principes actifs et en évitant en règle générale les effets

secondaires. «Aujourd'hui, l'industrie pharmaceutique fait toujours appel à l'analyse structurale des protéines, lorsqu'elle développe un nouveau médicament», relève le biochimiste qui a mis sur pied un département de recherche chez Ciba Geigy et Novartis avant de partir en 1997 à l'Université de Zurich. On trouve déjà quelques médicaments créés au moyen du «drug design». L'un des premiers était une petite molécule ralentissant la propagation du virus du sida: il bloque l'enzyme brisant les combinaisons de la protéine (protéase) de l'agent pathogène. De tels inhibiteurs de protéase entrent dans la composition des thérapies mixtes utilisées avec succès contre le sida. Les médicaments Relenza et Tamiflu sont d'autres exemples: ils inhibent une enzyme du virus de la grippe et évitent ainsi sa propagation dans le corps.

Le Glivec, un médicament contre le cancer, cible une protéine spécifique qui inhibe l'enzyme responsable d'une division cellulaire non contrôlée. La détermination de la structure des protéines exige une technologie pointue comme la cristallographie par rayons X, la spectroscopie de résonance magnétique nucléaire (RMN) ou encore le rayon-

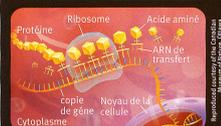
nement synchrotron. Un chercheur qui trouve une nouvelle protéine va d'abord s'enquérir de savoir si cette protéine a déjà été décrite pour d'autres êtres vivants pour ainsi prédire sa fonction.

Les chercheurs trouvent de telles similitudes dans les banques de données de séquences de protéines qui contiennent aussi des informations sur l'organisme, la structure et la fonction des protéines déjà étudiées. «La meilleure banque de données du monde est Swiss-Prot», estime Markus Grütter. Elle est très fiable car toutes les informations publiées dans des revues scientifiques sont vérifiées avant leur enregistrement. «Swiss-Prot propose des fonctions additionnelles très utiles pour les spécialistes en biologie structurale. Le programme «Swiss Modeller» cherche par exemple des séquences semblables et compare leurs structures. «Ce sont certes des modèles assez approximatifs, mais ils constituent de précieuses hypothèses de travail.»

Les appareils RMN, à rayonnement synchrotron ou à rayons X continueront toutefois à être utilisés. Certains pensent qu'on pourra bientôt déterminer la structure d'une protéine par ordinateur, note-t-il. Ce qui est peut-être possible pour des chaînes de protéines très simples ne comprenant qu'entre 100 ou 200 acides aminés. Mais les protéines fonctionnent souvent en combinaison et forment des ensembles complexes que

l'on ne peut pas simuler si facilement par ordinateur. Les chercheurs en biologie structurale ont encore du pain sur la planche.

De l'ADN à la protéine



Contenus dans le noyau de la cellule, les gènes abritent les plans de construction des protéines.

Pour fabriquer une protéine déterminée, il faut réaliser une copie du gène correspondant. Cette copie est extraite du noyau et transportée dans le cytoplasme.

La séquence d'ADN détermine l'enchaînement des brins élémentaires (acides aminés) de la protéine; un acide aminé déterminé correspond à trois éléments d'ADN. Les acides aminés appropriés sont transmis par des auxiliaires spéciaux, les ARN de transfert, et sont reliés ensemble par des protéines spécifiques, les ribosomes. La chaîne d'acides aminés est souvent coupée, pliée et parfois assemblée en un complexe protéique plus grand afin de pouvoir exercer sa fonction. Ce processus a aussi lieu grâce à des protéines très spécifiques.

# Encyclopédiste des protéines

Amos Bairoch est depuis plus de vingt ans pionnier dans la recherche sur les molécules biologiques à l'aide de la technologie de l'information. Sa carrière est le reflet du développement de la bioinformatique.

Par Patrick Roth, Photos Severin Nowacki

Adolescent, il était déjà attiré par la recherche spatiale. Et il était fasciné par l'idée d'étudier des formes de vie extraterrestres. En quoi ces dernières pourraient-elles se différencier de la biologie terrestre? C'est pour répondre à cette éventuelle interrogation qu'Amos Bairoch a choisi d'étudier la biochimie. Parallèlement, il s'est intéressé de près à l'informatique. « Mon père avait déjà une passion pour la technique, se souvient-il. Cet héritage a éveillé très tôt chez moi une fascination pour les calculatrices et les ordinateurs. »

## Peu reconnu au début

Alors qu'il était encore étudiant en biochimie à l'Université de Genève, Amos Bairoch a programmé le premier micro-ordinateur Apple II de la ville afin d'analyser des expériences biochimiques. « Le

concept de bioinformatique n'existait pas à l'époque et les étudiants qui jonglaient avec l'informatique étaient plutôt considérés comme des chercheurs ratés », relève le professeur de biologie structurale et de bioinformatique à l'Université de Genève. Agé aujourd'hui de 49 ans,

il a expérimenté toutes les phases de l'application de la technologie de l'information au domaine de la biologie moléculaire, depuis les débuts en utilisant de gros ordinateurs en time-sharing en passant par les premiers programmes avec une souris jusqu'à la diffusion globale des séquences de protéines les plus récentes sur le web. Il a aussi directement contribué au développement de nouveaux programmes d'analyse. Il est par ailleurs l'un des premiers chercheurs du monde à avoir mis en évidence la possibilité de décoder le

codé presque sans fin du vivant à l'aide de puces en silicium et de logiciels.

Au milieu des années 1980, Amos Bairoch a programmé pour son travail de doctorat trois applications totalement nouvelles qui ont beaucoup marqué le développement de la bioinformatique: le logiciel PC/Gene, un ensemble de programmes d'analyse de séquences protéiques, PROSITE, la bibliothèque des modèles de protéines, ainsi que Swiss-Prot, la banque de données de séquences protéiques qui fête cette année son vingtième anniversaire. Entre-temps, un ordinateur avec une connexion Internet est devenu pratiquement son seul instrument de travail, avec une pile impressionnante d'articles scientifiques sur les biomolécules récemment décodées. « Heureusement que, depuis le début, l'ensemble des connaissances sur les protéines n'a jamais augmenté plus vite que la capacité des disques durs d'un PC, précise le professeur avec un sourire amusé. Cet avantage me permet de continuer d'entretenir Swiss-Prot sur mon ordinateur portable. »

## Des souris et des hommes

A la question de savoir s'il se sent plutôt informaticien ou biologiste, Amos Bairoch répond sans hésiter: « Je suis biolo-

giste! » L'objectif de Swiss-Prot est de faciliter l'analyse et la comparaison des protéines pour les chercheurs. Depuis sa création, il y a vingt ans, la banque de données s'est muée en une encyclopédie électronique qui décrit pratiquement toutes les protéines actuellement connues dans le monde. La majorité des entrées et de loin, concernent les protéines humaines et celles de la souris. « Swiss-Prot nous livre des listes d'éléments qui rendent une collue vivante. Nous ne comprenons cependant pas encore comment le tout fonctionne, admet le scientifique. Nous n'en sommes toujours qu'au début. »

Spécialiste en histoire économique, le père du chercheur a travaillé dans diverses universités, contraignant sa famille à un certain nomadisme. D'origine française, Amos Bairoch a suivi sa scolarité en Belgique, en France et au Canada avant de s'installer en Suisse. La fidélité à Genève dont il a fait preuve tout au long de sa carrière est d'autant plus surprenante. « Mon travail pour Swiss-Prot et PC/Gene m'a lié à l'Université de Genève », confirme-t-il. Un engagement auprès d'une autre institution académique aurait supposé la fin du projet. Mais pour compenser sa sédentarité professionnelle, il entreprend durant ses loisirs de longues promenades dans la nature avec sa femme Martine qui enseigne le yoga. Ce travailleur invétéré se ressource en famille. « Durant les week-ends et les vacances, je suis capable de m'extraire totalement du monde de la recherche, avoue-t-il. Et j'oublie tout, tout de suite! » Chez lui, on ne parle pas de bioinformatique et aucun de ses trois enfants - Alice

(20 ans), Jonas (17) et Colin (13) - n'envisage de suivre ses traces.

## De l'idée à la concrétisation

« Créer une banque de données est un projet qui n'a pas de fin », souligne le professeur. Afin de pouvoir assurer la durabilité de Swiss-Prot, soutenu auparavant par le Fonds national, le projet a été institutionnalisé en 1998 dans le cadre de la création de l'Institut suisse de bioinformatique (ISB). Sous sa direction et aussi l'entre-temps grâce au soutien international, le groupe Swiss-Prot a pour objectif de rassembler à l'ISB toutes les connaissances sur les protéines, de les améliorer puis de les diffuser aux chercheurs du monde entier. Ce besoin d'échange et de partage est le fil rouge du travail d'Amos Bairoch. Un besoin qu'illustre aussi la boîte ouverte d'appétissants chocolats qui trône sur la table autour de laquelle se déroule l'entretien.

Amos Bairoch a contribué de façon significative à l'interprétation des plans de construction codés de la biologie terrestre. Sa contribution au développement de la bioinformatique a été honorée en 2004 par l'attribution du prix Latsis européen. Mais son ancienne passion est toujours présente. Avec des yeux pétillants, il commente les résultats des missions spatiales actuelles. Qu'il s'agisse des nappes phréatiques dissimulées sous le désert de Mars, de l'océan salé sous l'imposante calotte de glace d'Europe, la lune de Jupiter, ou encore des volcans de glace crachant du méthane de Titan, une lune énorme et exotique dans le lointain système de Saturne, il étudierait très volontiers ces prometteuses oasis de vie découvertes dans le système solaire. Une manière de rajouter éventuellement un chapitre sur les découvertes exobiologiques à son encyclopédie des protéines.

# Du mini-ordinateur à la banque de données géante

La banque de données internationale sur les protéines Swiss-Prot doit son excellente renommée à son annotation soignée des séquences de protéines et à ses logiciels performants. Par Patrick Roth, Photo Severin Nowacki

La base de la bioinformatique est très simple: les plans de construction du vivant reposant sur un code, qui peut être analysé et décodé de manière très efficace par les ordinateurs. Fort de cette découverte, le biochimiste Amos Bairoch de l'Université de Genève a saisi, en 1981 déjà, les séquences de protéines connues à l'époque afin de pouvoir les comparer et les analyser à l'aide de logiciels. Alors que sa banque de données comportait déjà 3900 entrées de séquences en 1986, le chercheur a décidé de la rendre accessible au public: Swiss-Prot était née. Le projet bénéficia du soutien du Fonds national suisse depuis 1993.

## Importante évaluation

En deux décennies, Swiss-Prot est devenue la plus grande source de connaissances sur les protéines et depuis sa fusion avec d'autres banques de données, elle est devenue membre du consortium UniProt. Elle est toujours domiciliée à l'Institut suisse de bioinformatique de l'Université de Genève, même si le recueil des informations s'est globalisé: il y a des équipes en Grande-Bretagne et aux États-Unis. Plus de 230 000 séquences ont déjà été enregistrées dans la banque de données. Elles décrivent la succession de plus de 84 millions d'acides aminés. Une nouvelle version de Swiss-Prot est diffusée toutes les deux semaines et chaque mise à jour comprend environ 1000 nouvelles protéines. L'évaluation des données est toutefois encore plus importante que leur

enregistrement. Ce commentaire (appelé aussi annotation) est fait manuellement, en s'appuyant sur les descriptions de protéines parues dans les articles scientifiques ainsi que sur les résultats de l'application de logiciels d'analyse des séquences de protéines. Ces données sont régulièrement mises à jour. Les chercheurs des différentes institutions ainsi que de l'industrie trouvent donc non seulement des séquences et des référen-

d'une protéine apparaissant dans d'autres banques de données sont ainsi rassemblées. Toutes les informations disponibles se trouvent sur des serveurs auxquels les personnes intéressées peuvent accéder librement par Internet. Afin de faciliter la navigation dans cet océan d'informations, un programme de recherche a été développé afin de permettre de parcourir la banque de données indexée. D'autres instruments sont conçus pour faciliter le



ces de publication dans Swiss-Prot, mais également des informations concernant la fonction de chaque protéine, des éléments fonctionnels de la protéine, des centres actifs, la forme tridimensionnelle entre autres. Certaines familles de protéines sont attribuées à des experts qui participent à l'actualisation des informations concernant «leur» famille de protéines.

Un autre objectif de Swiss-Prot est de limiter la redondance des données globales. Les différentes descriptions

travail de détective des biochimistes: la fonction d'une nouvelle protéine peut souvent être établie grâce la ressemblance de sa séquence avec celles d'autres groupes connus de protéines. De telles comparaisons ne sont toutefois possibles qu'au moyen de programmes informatiques très performants. Les protéines humaines les plus longues peuvent en effet comprendre des séquences allant jusqu'à 30 000 acides aminés. ■

[www.expasy.org](http://www.expasy.org)