

BIOLOGIA ESTRUCTURAL

MIQUEL COLL

*Institut de Recerca Biomèdica i Institut de Biologia Molecular de Barcelona, CSIC,
Parc Científic de Barcelona.*

Adreça per a la correspondència: Miquel Coll. Institut de Recerca Biomèdica i Institut de Biologia Molecular de Barcelona, CSIC, Parc Científic de Barcelona. Josep Samitier, 1-5. 08028 Barcelona. Adreça electrònica: *mcoll@ibmb.csic.es*.

RESUM

La biologia estructural ha arribat a la seva maduresa. Avenços tecnològics sense precedents han ajudat, en els últims anys, a assolir resultats espectaculars, tant pel que fa al nombre d'estructures de macromolècules resoltes com a la complexitat d'algunes d'aquestes. L'aplicació biotecnològica està ben establerta, per exemple, en el procés de desenvolupament de fàrmacs. Tot i així, la biologia estructural i, en particular, la cristallografia de macromolècules, es troba en un moment de reflexió. Podem seguir endavant com fins ara i esperar nous descobriments importants per a la biologia o cal buscar nous camins?

Paraules clau: biologia estructural, complex de proteïnes, genòmica estructural, difracció de raigs X, cristallografia.

STRUCTURAL BIOLOGY

SUMMARY

Structural biology has reached maturity. During recent years unprecedented technological advances have contributed to spectacular results both in the number of macromolecular structures solved and in the complexity of some of these. The biotechnological application of this branch of biology is well established, for example, in drug development. However, structural biology, in particular macromolecular crystallography, is now being reconsidered. Can we expect further discoveries with great impact in biology or should we look for new avenues?

Key words: structural biology, protein complex, structural genomics, X-ray diffraction, crystallography.

Francis Crick deia que si vols conèixer la funció t'has de mirar l'estructura (Crick, 1988). I és veritat: el fet de conèixer l'estructura molecular dels àcids nucleics i les proteïnes ha obert moltes portes per entendre com funcionen aquestes macromolècules, que es troben en la base de la vida. Però com s'ha de mirar l'estructura? Les tres tècniques més importants per determinar estructures moleculars en 3D són la difracció de raigs X (RX), la ressonància magnètica nuclear (RMN) i la microscòpia electrònica (ME). Cadascuna té els seus avantatges i inconvenients i cadascuna pot donar una visió diferent i complementària (vegeu la figura 1).

LA MICROSCÒPIA ELECTRÒNICA

Comencem per la darrera. La microscòpia electrònica (ME) i, concretament, la reconstrucció tridimensional de partícules a partir d'imatges de les mateixes amb orientacions diverses, proporciona un model fiable en 3D. Això es pot fer de dues maneres: mitjançant tinció negativa o criomicroscòpia. En la primera de les tècniques la mostra es tracta amb una sal d'un metall pesant que dispersarà molt més els electrons que no pas els àtoms de la proteïna, i proporcionarà així un contrast fort entre aquest i la mostra. La imatge que s'obté serà, per tant, el «negatiu» de l'espècimen en qüestió i tindrà un cert grau de distorsió. La resolució, atès el granulat mateix de la sal pesada, no passarà dels 15 o 20 Å. La reconstrucció final serà, doncs, una visió de la forma global,

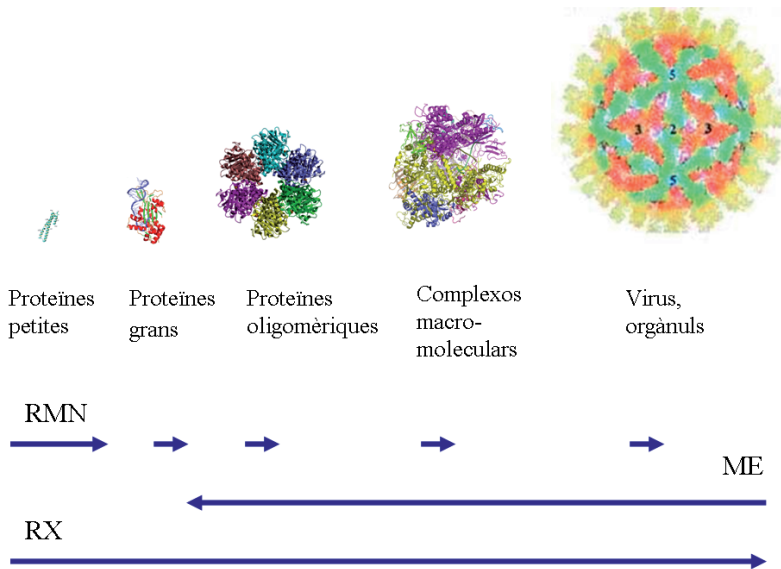


FIGURA 1. La microscòpia electrònica, la ressonància magnètica nuclear i la cristal·lografia de raigs X permeten analitzar l'estructura en 3D de les macromolècules. Les fletxes indiquen quina tècnica és adient segons la mida de la molècula o partícula. Les fletxes petites de la RMN indiquen que aquest mètode es pot emprar per determinar fragments o dominis de les proteïnes o acoblaments grans. D'esquerra a dreta: cofactor A de plegament de la tubulina (Guasch *et al.*, 2002), complex relaxasa/DNA (Guasch *et al.*, 2003), proteïna acobladora de la conjugació TrwB (Gomis-Rüth *et al.*, 2001), RNA-polimerasa II (Cramer *et al.*, 2001), virus de la llengua blava (Grimes *et al.*, 1998).

no gaire definida, de la partícula (vegeu la figura 2). Aquesta tècnica clàssica és relativament senzilla, ja que costa poc preparar les mostres, i encara es fa servir molt.

L'altra tècnica, la criomicroscòpia, és més moderna; proporciona imatges sense distorsió i una resolució major (vegeu la figura 13). La congelació de la mostra a la temperatura del nitrogen líquid o encara més baixa permet augmentar considerablement la radiació que pot suportar l'espècimen sense que es destrueixi i, per tant, es poden fer servir microscopis electrònics més potents, com els de canó d'emissió de camp (FEG), de 200 o 300 kV. La mostra, però, no es pot congelar de manera ordinària, ja que els cristalls de gel que es formarien amb l'aigua de la solució en què es troba la farien malbé. Aquesta aigua s'ha de vitrificar, congelant-la a gran velocitat. Les imatges amb mostres tractades d'aquesta manera tenen poc contrast, però estan més protegides i proporcionen reconstruccions fidedignes, sense distorsions. La mitjana de moltes imatges per compensar el poc contrast de la criomicroscòpia i assolir la reconstrucció 3D de les partícules es fa amb ordinadors potents. Es necessiten moltes projeccions diferents de la mateixa partícula, bé aprofitant que les partícules es troben diposi-

tades a la reixeta portamostres en diverses orientacions o bé inclinant la reixeta mateixa en diversos angles. No totes les mostres es deixen tractar bé criogènicament, i això és un factor limitant. També la mida de la partícula ho és, ja que el poc contrast fa encara més difícil o impossible treballar amb partícules petites. Però si tenen la mida apropiada i es comporten adequadament en vitrificar-les, els resultats són espectaculars: ens podem apropar als 4 o 5 Å de resolució, una resolució que permet començar a localitzar elements d'estructura secundària, com ara les hèlixs alfa (Henderson, 2004). Tot i així, aquests són casos favorables i encara estem lluny dels detalls que, de vegades, ens interessin: els àtoms. És clar, però, que la criomicroscòpia electrònica és una tècnica molt potent que ha revolucionat el camp i que, per exemple, complementada amb la difracció de raigs X permet obtenir estructures quasiatòmiques, com veurem més endavant.

Un altre desenvolupament prometedor en la microscòpia electrònica és la tomografia. També en aquest cas es fa una reconstrucció 3D amb ordinadors (un tomograma), però a partir d'imatges d'una sola partícula, en diversos angles. El problema és que la resolució és proporcional al rang

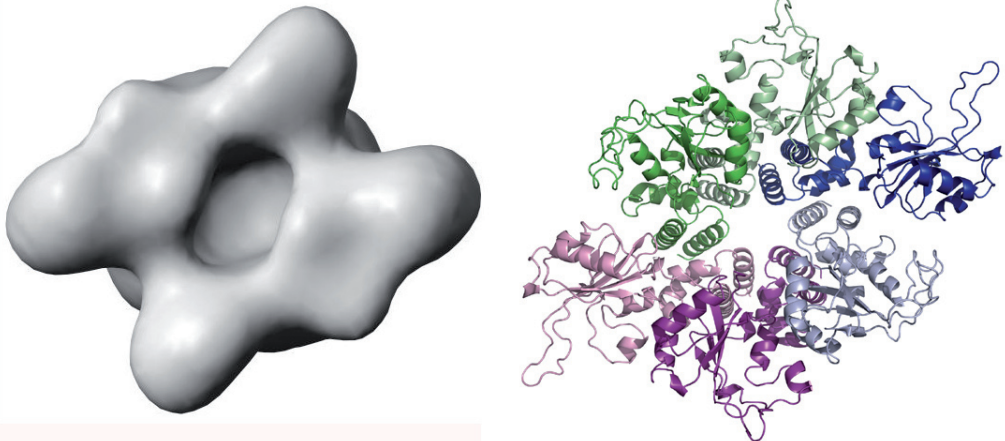


FIGURA 2. Proteïna iniciadora de la replicació RepB (Boer *et al.*, 2007). Esquerra: Reconstrucció en 3D a partir d'imatges de ME de tinció negativa. Dreta: Estructura de RX.

angular de rotació cobert, és a dir, que per tenir resolucions acceptables s'han de prendre imatges en molts angles. I, com més imatges es prenguin, més patirà la mostra. L'automatització recent dels microscopis ha permès fer aquest procés ràpidament, minimitzant-ne el dany electrònic. Grans orgànuls i cèl·lules senceres que no poden ser estudiats amb les tècniques microscòpiques abans esmentades poden ser, sorprenentment, visualitzats en 3D amb la tomografia electrònica (Nickell *et al.*, 2006).

LA RESSONÀNCIA MAGNÈTICA NUCLEAR

La ressonància magnètica nuclear (RMN), a diferència de la microscòpia electrònica, permet determinar la posició de gairebé cadascun dels àtoms que formen part d'una proteïna o àcid nucleic (Ferentz i Wagner, 2000). En aquest sentit els resultats són semblants a la cristal·lografia, però, a diferència d'aquesta, no necessita que la macromolècula formi part d'un cristall. L'experiment es fa en solució aquosa i, per tant, ens estalviem un coll d'ampolla important com és la cristallització, sempre difícil i imprevisible. Un altre avantatge de la RMN és que podem estudiar amb més facilitat les

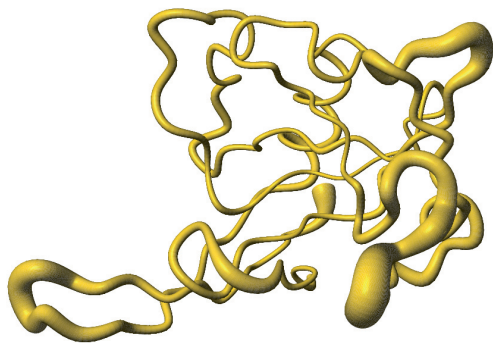


FIGURA 3. Estructura de l' α -sarcosina determinada per NMR (Perez-Canadillas *et al.*, 2002). La major o menor flexibilitat, segons la zona, està representada pel gruix variable de la cadena polipeptídica.

proprietats dinàmiques de les macromolècules (vegeu la figura 3). A més, amb aquesta tècnica es pot obtenir informació estructural de macromolècules que no tenen una única estructura sinó diverses en equilibri. Una aplicació clara d'aquesta capacitat és l'estudi del plegament de les proteïnes en què poden ser caracteritzats estats de plegament parcial. La limitació més gran de la RMN —i que no ho és pas en ME ni en RX— és la mida de la proteïna, que sol estar al voltant dels 20 o 30 kDa, encara que hi ha algun estudi que ha arribat fins als 80 kDa.

Per adquirir un espectre de RMN s'ha de posar la mostra dins d'un camp magnètic proporcionat per una bobina superconduïdora refredada amb heli líquid i s'ha d'irradiar amb polsos de radiofreqüència. La utilització de camps magnètics elevats representa un avenç molt important per a l'estudi de biomolècules, ja que dona lloc a un augment de la sensibilitat i de la dispersió i a una disminució de l'amplada dels senyals que permet abordar proteïnes més grans. En la RMN la magnitud del camp magnètic s'expressa com la freqüència de ressonància del ^1H i els aparells comercials arriben fins a 900 MHz, tot i que la instal·lació dels primers aparells d'1 GHz és imminent. A Catalunya hi ha un instrument de 800 MHz al Parc Científic de Barcelona.

A banda dels estudis dinàmics, una altra aplicació molt avantatjosa de la RMN és la caracterització de les interaccions de les proteïnes amb diferents lligands, ja siguin altres macromolècules o molècules petites com ara inhibidors o fàrmacs. Segons la tècnica que es faci servir (p. ex., STD), es pot caracteritzar aquesta interacció sense haver de determinar l'estructura completa de la proteïna. Això pot ser molt ràpid i, per tant, factible per fer cribratges de petites molècules respecte a la seva capacitat d'unir-se a una proteïna determinada, que pot ser una diana farmacològica. L'aplicació d'aquesta tècnica a la indústria farmacèutica és,

doncs, evident. La RMN es pot fer servir també per caracteritzar les interaccions febles entre proteïnes, útils, per exemple, per entendre fenòmens de regulació.

Malgrat que la RMN no requereix la formació de cristalls, la determinació de les estructures 3D de proteïnes no sol ser més ràpida que mitjançant RX, ja que, a més del marcatge de la proteïna amb ^{15}N i ^{13}C , es necessiten assignar els senyals de ressonància als nuclis atòmics individuals, en definitiva, a cada un dels aminoàcids de la proteïna. Aquest procés és certament laboriós, malgrat els recents desenvolupaments de programari d'assignació automàtica.

LA DIFRACCIÓ DE RAIGS X

Si accedim al Protein Data Bank (PDB), la base de dades en què es troben totes les estructures 3D de les macromolècules, veurem que hi ha arxivades, actualment, 48.181 estructures atòmiques. D'aquestes, el 85 %, han estat determinades per RX. Aquesta és, doncs, la tècnica d'anàlisi estructural que

ha proporcionat un major nombre d'estructures i segurament, si les coses no canvien molt, ho continuarà sent en el futur. Aquesta és, a més, la tècnica que proporciona major resolució, que en els casos més favorables s'apropa a l'àngstrom, la unitat de distància d'enllaç entre dos àtoms. El detall i, per tant, la informació que pot proporcionar una estructura 3D de RX, no admet comparacions (vegeu la figura 4). Tampoc no hi ha limitació en la mida de la proteïna, complex o partícula a estudiar... sempre que cristal·litzi.

El *boom* de la cristal·lografia de proteïnes, en el qual s'ha passat de les quatre-centes estructures de RX que hi havia dipositades al PDB a finals dels noranta a les quaranta mil que hi ha ara, s'ha degut a diversos factors (Dauter, 2006). En primer lloc, a la possibilitat d'obtenir proteïnes amb tècniques de DNA recombinant, sense haver de dependre de mètodes d'extracció a partir de teixits i fent accessibles, per clonatge i expressió en bacteris o en altres sistemes, quantitats suficients de proteïna per emprendre estudis estructurals. I això es pot fer amb tot tipus de proteïnes, per minoritàries que siguin a la cèl·lula.

La cristallització mateixa, tot s'ha de dir, no ha avançat gaire conceptualment (encara que els automatismes hagin facilitat les coses, com veurem més endavant). La coneguda tècnica de difusió de vapor, en la qual es concentra, a poc a poc, una solució de proteïna fins a fer-la precipitar ordenadament —el que és un cristall— segueix sent el sistema emprat majoritàriament. La cristallització, curiosament, no depèn de la mida de la macromolècula: hi ha enormes bacteriòfags que s'han cristallitzat (Abrescia *et al.*, 2004) i petites proteïnes d'un centenar d'aminoàcids amb les quals resulta impossible.

Es tracta d'un procés difícil de controlar, en el qual pot arribar a ser determinant la presència d'un sol aminoàcid a la superfície

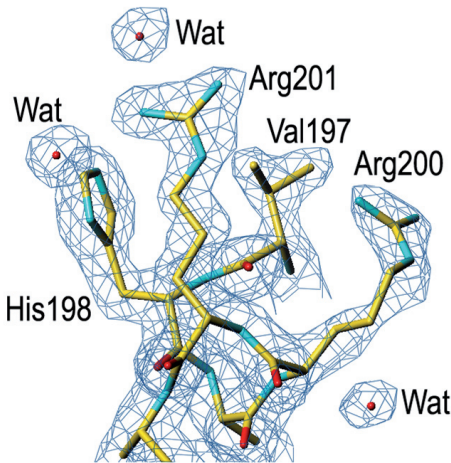


FIGURA 4. Densitat electrònica corresponent a una part de l'estructura de RX de l'activador transcripcional PhoB (Blanco *et al.*, 2002). L'alta resolució permet localitzar tots els àtoms de la proteïna i, fins i tot, la primera capa d'hidratació.

cie de la proteïna que faciliti o dificulti la interacció amb les molècules veïnes que formaran el cristall. Per descomptat, influeixen també la temperatura, el pH, les sals, el tampó, els additius, els detergents, el precipitant, el volum de les gotes de cristallització, la proporció entre proteïna i precipitant... Massa variables. I quan passa això no hi ha més remei que fer cribratges, en aquest cas de centenars o milers de condicions de cristallització. Si això tampoc no funciona cal provar proteïnes una mica diferents: homòlegs d'altres espècies o construccions una mica més llargues, o una mica més curtes, o amb alguns aminoàcids de la superfície mutats o modificats.

Un cop obtinguts els cristalls, que no passen d'uns centenars de micròmetres, s'han

d'irradiar amb un feix de raigs X. El cristall, més concretament els electrons dels àtoms continguts en aquest, dispersaran els raigs X incidents i, a causa del seu ordre en una xarxa, donaran lloc al fenomen de la difracció. Doncs bé, la construcció de sincrotrons dedicats, de tercera generació, on es generen raigs X d'altíssima intensitat, ha estat un dels factors determinants per a l'augment exponencial d'estructures resoltes. Però no ha estat només la brillantor del feix, que en un sincrotró és d'ordres de magnitud superior al d'un generador de raigs X convencional, el que ha estat determinant per poder resoldre milers d'estructures en pocs anys: també ho ha estat el fet que la radiació de sincrotró és sintonitzable. És a dir, podem variar l'energia o longitud d'ona dels raigs



FIGURA 5. Vista de les obres del sincrotró Alba l'octubre de 2007. Es poden veure les parets protectores de formigó entre les quals passarà l'anell, amb les sortides tangencials, una per a cada línia de llum.

X. Aquesta característica permet aprofitar l'anomenada *difracció anòmala* que tenen certs àtoms pesants i, així, amb un sol derivat pesant es pot resoldre l'estructura tridimensional. Aquest derivat es pot preparar directament quan expressem la proteïna si fem servir un medi en què l'aminoàcid metionina ha estat substituït per selenometionina. Aquesta s'incorpora a les proteïnes en lloc de l'aminoàcid natural i el seleni, un àtom que dispersarà els fotons anòmalament, serà útil per resoldre l'estructura.

Un altre factor clau per a l'acceleració de tot el procés de determinació d'estructures ha estat el desenvolupament de detectors bidimensionals —CCD als sincrotrons—, que permeten la recollida de dades de difracció, de milers de reflexions, en uns pocs minuts.

A Cerdanyola del Vallès s'està construint un sincrotró, anomenat ALBA, que inclou una línia de llum per a cristallografia de macromolècules i altres —microscòpia de raigs X, difracció de baix angle— també per a aplicacions biològiques (vegeu la figura 5). Amb un pressupost de 200 milions d'euros, serà la instal·lació científica més gran de la península i un instrument extraordinari per a la biologia estructural al nostre país. És clar, que si se sap aprofitar, també serà

una eina molt valuosa per a la indústria biotecnològica i farmacèutica.

Finalment, com en altres camps de la ciència, el desenvolupament d'ordinadors molt ràpids i de sistemes gràfics molt potents ha facilitat les coses als cristal·lografs, com així ha estat també per als biòlegs estructurals que fan servir ME i RMN.

LA GENÒMICA ESTRUCTURAL I LES TÈCNiques D'ALT RENDIMENT

Els avenços en la seqüenciació del genoma humà i la seqüenciació completa del genoma de diversos organismes, units a les tècniques de DNA recombinant, van obrir les portes a la determinació massiva d'estructures 3D de proteïnes (Grabowski *et al.*, 2007). Si s'han pogut seqüenciar tots els gens d'un organisme, per què no s'han de poder resoldre les estructures de totes les proteïnes que codifiquen aquests gens? A finals dels noranta es van començar a organitzar iniciatives i consorcis, inicialment als EUA i al Japó, per implementar tecnologies automatitzades que fessin possible aquest objectiu ambiciós. La tasca ha estat més difícil del que pensaven alguns i, certament, no s'ha pogut resoldre cap «genoma estructural» complet. La figura 6 mostra que de les més de cent mil dianes (proteïnes o dominis) dipositades públicament pels diferents consorcis de genòmica estructural arreu del món, menys d'un 10 % s'han aconseguit cristallitzar. Aquest percentatge anirà augmentant, ja que les dianes es dipositen des del bon començament dels projectes, però ara ja és clar que, ni de bon tros, ens apropem mai al 100 %. És cert però, que moltes proteïnes han estat resoltes per part d'aquests consorcis, concretament 5.608, però ja s'indica també que per a 27.393 proteïnes s'ha aturat la feina. Sobretot, el que s'ha resolt és el que algú ha anomenat, amb prou encert, *els fruits de les branques més bai-*

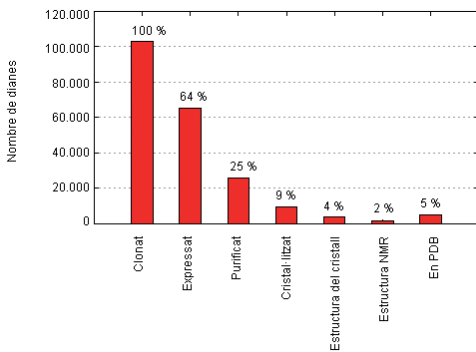


FIGURA 6. Resultats dels projectes de genòmica estructural segons el registre del Protein Data Bank l'octubre de 2007 (<http://www.rcsb.org/pdb>).

xes, és a dir, les proteïnes menys difícils, per exemple aquelles expressables fàcilment en forma soluble.

Un dels grans èxits de la genòmica estructural ha estat, però, el desenvolupament de tecnologies automatitzades o d'alt rendiment que s'han anat incorporat a la vida normal dels laboratoris estructuralistes encara que no sigui per a la determinació massiva d'estructures, sinó més aviat per a projectes molt focalitzats.

El primer punt d'automatització ha estat en el clonatge i expressió. Passar del tub Eppendorf i de les pipetes d'un sol canal a la placa de 96 i les pipetes multicanals o robots de dispensació de líquids requereix un esforç de reconversió organitzativa i també de diners. Però pot pagar la pena. De vegades, per resoldre amb èxit una sola proteïna. Hem dit que un coll d'ampolla del treball estructural és la cristallització. Un altre, abans que aquest, és l'expressió de la proteïna soluble. Moltes proteïnes s'expressen molt malament o s'agregen formant cossos d'inclusió d'on difícilment es poden recuperar. Els mètodes automàtics i paral·lelitzats permeten provar l'expressió de moltes construccions diferents de la mateixa proteïna, per dominis, truncant progressivament els extrems, etc. Aquesta metodologia portada

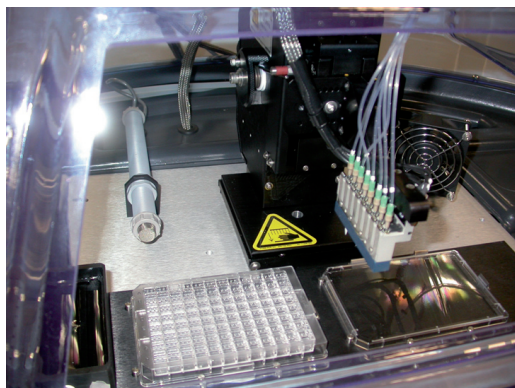


FIGURA 7. Placa de cristallització de noranta-sis pous en un robot dispensador de nanogotes.

a l'extrem és el que es coneix com a *rendiment ultraalt* (Hart i Tarendeau, 2006), en el qual es fan cribratges d'expressió de biblioteques de milers de construccions del gen en qüestió, truncats d'aminoàcid en aminoàcid pels dos extrems.

Un altre punt clau d'automatització ha estat la cristallització. Aquí també s'ha passat de les plaques de cristallització de vint-i-quatre pous, fetes a mà, a les de noranta-sis, fetes amb robots capaços de dispensar acuradament volums d'un centenar de nanolitres (vegeu la figura 7). Les gotes de cristallització s'han fet, doncs, més petites —p. ex., de 200 nl—, de manera que ara és possible fer cribratges amb un ordre de magnitud menys de mostra proteica. Un aspecte negatiu d'aquesta automatització és que els cristalls són també més petits i molts cops s'han d'escalar amb mètodes més manuals. La comprovació de les plaques també és un problema i fa necessari un altre robot per a l'emmagatzematge i la visualització de les gotes de cristallització.

També a la recollida de dades ha arribat l'automatització, amb robots per al muntatge dels cristalls al feix de les línies de llum de macromolècules, als sincrotrons. També, cal esmentar el programari que s'ha anat desenvolupant recentment, més amigable i automàtic en alguns aspectes, com és el traçat inicial de la cadena polipeptídica.

AIXÒ JA NO ÉS CIÈNCIA!

Amb tota aquesta automatització, consorcis milionaris de genòmica estructural i determinació massiva d'estructures, un cristallògraf, que actualment és a la Universitat d'Uppsala, va exclamar: «Si es pot fer de manera industrial, això ja no és ciència!» Semblaria, doncs, que des d'un punt de vista metodològic el problema ja està resolt i que no caldria més que posar força bruta per acabar resolent totes les estructures. Si

tot són només millores tecnològiques i està perfectament establerta la metodologia per resoldre estructures, on es troba el repte científic?

Bé, vist els resultats, fins ara més aviat minsos, dels projectes de genòmica estructural, és clar que moltes proteïnes que, aparentment, per la mida, haurien de ser fàcils de resoldre amb tota la tecnologia disponible, no acaben de poder-se expressar, purificar o cristallitzar. Només per això ja és evident que la tasca que queda per fer és ingent. Tot i així, es podria argumentar que és un problema purament de temps i que s'acabarà superant, només afinant la tecnologia d'alt rendiment. Sigui això cert o no, el fet és que hi ha àrees de la biologia estructural en les quals la genòmica estructural gairebé (encara) no ha entrat; àrees que es pot dir que estem només començant a explorar. A un parell d'aquestes ens referim a continuació.

EL MALSON DE LES PROTEÏNES DE MEMBRANA

Entre les proteïnes que certament no són fruits de les branques baixes de l'arbre es troben les proteïnes de membrana. Aquestes proteïnes són de gran interès, ja que fan moltes funcions essencials a la cèl·lula: controlen el tràfic d'ions i altres substàncies a través de la membrana, generen l'energia que necessita la cèl·lula en forma d'ATP, actuen com a antenes captadores de l'energia lumínica en la fotosíntesi, són receptors hormonals, constitueixen sensors de tot tipus que permeten a la cèl·lula relacionar-se amb el seu entorn, són part crucial de vies de transducció de senyals, etc. Precisament amb motiu d'algunes d'aquestes funcions hi ha moltes proteïnes de membrana que són dianes farmacològiques rellevants. Conèixer la seva estructura, a més d'ajudar a entendre com funcionen,

ajudaria molt al desenvolupament de fàrmacs nous.

Si tornem a mirar el PDB i les 48.181 proteïnes dipositades, resulta que només 280 són transmembranoses (133 úniques o no redundants). És evident que els progressos en aquest camp estan lluny del *boom* de les proteïnes solubles. Per què són tan difícils d'estudiar? El problema és que les proteïnes de membrana tenen dues cares i no se sap per on mirar-les; tenen parts hidrofòbiques i parts hidrofíliques. Per això són difícils de mantenir en solució: en dissolvents orgànics es desnaturalitzen, en aigua precipiten. Requereixen, doncs, un medi en el qual concorrin les dues característiques: polaritat i apolaritat. Això s'aconsegueix amb detergents que s'han d'afegir, en petites quantitats, a les solucions tamponadores, tant en la purificació com en la cristallització. Els avenços tecnològics, incloent-hi la robotització i la miniaturització dels cribratges que hem mencionat abans, també han estat de gran ajut per a les proteïnes de membrana, però els bons resultats s'obtenen aquí molt més a poc a poc. També, una tècnica alternativa a l'ús de detergents, com és la de cristallitzar en una fase lipídica cúbica, tractant d'imitar la bicapa cel·lular, ha servit en alguns casos, però certament no ha revolucionat el camp. Aquí queda molt per fer.

NO ESTAN SOLES

Moltes, si no la majoria, de les proteïnes, fan la seva funció interaccionant amb altres proteïnes o altres components cel·lulars, com ara els àcids nucleics. D'aquí ve l'enorme interès a conèixer com són aquests acoblaments o complexos macromoleculars. Els complexos poden anar des d'un parell de proteïnes a dotzenes d'aquestes, barrejades o no amb àcids nucleics. El ribosoma, una partícula o orgànul ben caracteritzat, està format, en els bacteris, per cinquanta-cinc

proteïnes i tres molècules de RNA. Malgrat la seva complexitat, la seva estructura 3D ha estat resolta, tant per ME com per RX (vegeu la figura 8), i ha donat lloc a una allau

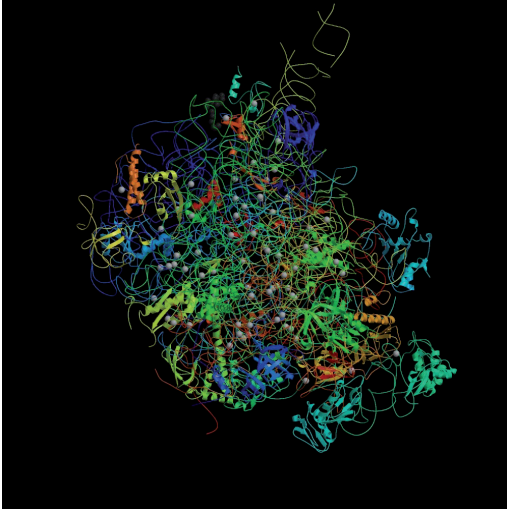


FIGURA 8. Estructura 3D de la subunitat gran del ribosoma, determinada per RX (Ban *et al.*, 2000).

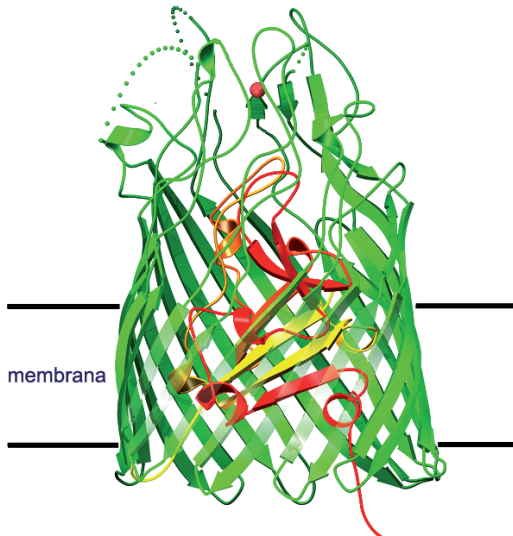


FIGURA 9. L'estructura del transportador transmembranós FepA, determinada per RX; mostra un plegament de tipus barril- β , amb les cadenes β inclinades travessant amunt i avall la membrana cel·lular (Buchanan *et al.*, 1999).

d'informació estructural i funcional sense precedents (Ban *et al.*, 2000; Korostelev *et al.*, 2006; Wimberly *et al.*, 2000). Semblaria, doncs, que tot complex fóra abordable.

Com en el cas de les proteïnes de membrana, ja la preparació d'aquests complexos és un problema. En principi es podrien preparar les parts per separat, expressant els components en algun dels sistemes coneguts, i després ajuntar-los. Això pot arribar a ser impossible, ja que aquests components tot sols ben sovint no són solubles perquè presenten grans zones d'interacció proteïna-proteïna més aviat hidrofòbiques, o simplement perquè no es pleguen apropiadament sense els seus *partners*. Hi ha estratègies per superar això. Una és la de produir subcomplexos, és a dir, complexos formats només per algunes parts o dominis que es puguin expressar solubles i ajuntar-los apropiadament. Algunes d'aquestes estratègies poden ser molt sofisticades (vegeu la figura 10) i requereixen un coneixement exhaustiu de la diana. El treball es torna de nou artesà, cas per cas, i fuig dels mètodes automàtics o d'alt rendiment. Fins i tot si aquestes parts es poden arribar a expressar,

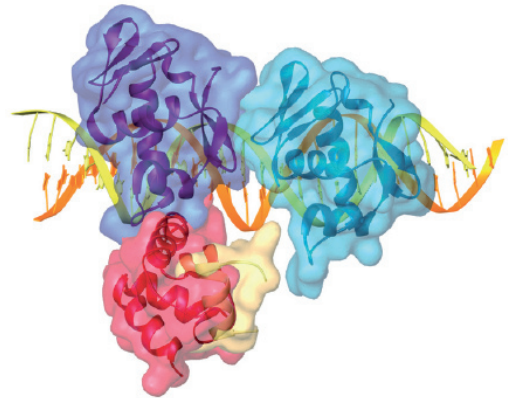


FIGURA 10. Estructura del subcomplex d'activació transcripcional PhoB/ σ_4 - β -flap/DNA determinat per RX. σ_4 i β -flap són parts de les subunitats σ^{70} i β , respectivament, de la RNA-polimerasa, i es van expressar com una quimera (en una mateixa cadena polipeptídica) per poder obtenir-les solubles i cristallitzar el complex (Blanco *et al.*, 2007).

reconstruir el complex i arribar a l'estequirmetria correcta pot ser molt difícil.

Una altra alternativa per preparar complexos és expressar els diversos components a l'hora. Això es pot fer clonant-los cadascun en vectors diferents —p. ex., infectant cèl·lules d'insecte amb diversos baculovirus—, o tots junts en un mateix vector policistrònic.

Si els complexos o acoblaments són grans i prou estables es poden purificar directament de la cèl·lula. Òbviament, això requereix cultius molt grans i protocols de purificació prou sofisticats i específics per a aquell complex concret per tal que se'n mantingui la integritat. Un cas espectacular, premiat amb el Nobel de Química de 2006 (Coll, 2006), ha estat l'estructura de la RNA-polimerasa II de llevat, resolta per RX a alta resolució (vegeu la figura 11; Cramer *et al.*, 2001). La purificació d'aquest enzim amb dotze subunitats requereix cultius de llevat de centenars de litres i l'ús de grans fermentadors.

Per a complexos làbils, recentment s'ha desenvolupat un nou mètode per preparar-

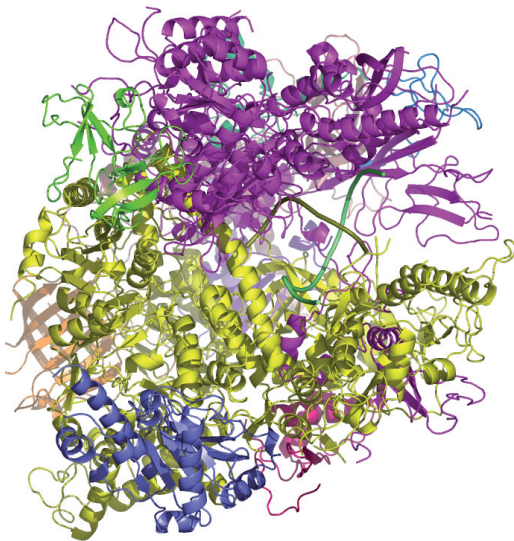


FIGURA 11. Estructura de la RNA-polimerasa II de llevat, resolta per RX (Cramer *et al.*, 2001).

los, conegut com a *purificació d'afinitat en tàndem* (TAP). En aquest mètode es posen dues etiquetes d'afinitat en tàndem en un dels components del complex i es fa passar primer per una primera columna d'afinitat i, després de tallar la primera etiqueta, per una segona columna (vegeu la figura 11). Les condicions d'elució, tall, etc., són suaus, per no desfer el complex.

Hi ha complexos o orgànuls molt coneguts, com ara el ribosoma, el proteosoma, el centrosoma, etc., però la cèl·lula n'està farcida de molts altres. Un estudi fet per investigadors de l'EMBL i la Biotech Cellsome en va trobar 491 fent servir la tècnica TAP i espectrometria de masses (Gavin *et al.*, 2006). Caracteritzar estructuralment, ni que fos una part petita, d'aquest «complexosoma», serà un esforç científic considerable que necessitarà el desenvolupament de noves tècniques preparatives, que espero que veurem florir en els propers anys.

Si caracteritzar aquests complexos més o menys estables és i serà difícil, encara ho serà més caracteritzar estructuralment tot tipus de complexos transitoris que es donen en les vies o xarxes de senyalització o regulació. Aquestes interaccions múltiples i transitòries són, de fet, les que fan que la cèl·lula funcioni com a tal, com el sistema sofisticadament regulat que és. Abordar-les des d'un punt de vista estructural és un ca-

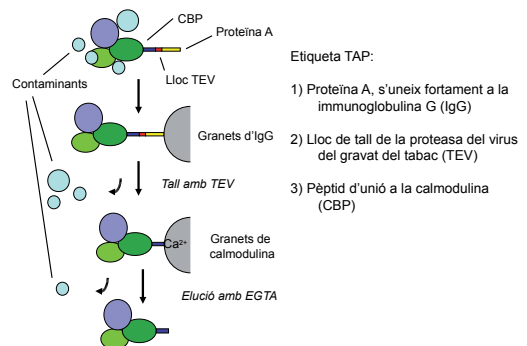


FIGURA 12. Esquema de la purificació d'afinitat en tàndem d'un complex de tres proteïnes.

mí que pot semblar adient per a visionaris sense els peus a terra però que ja s'està reclamant fa un temps (Harrison, 2004; Beltrao *et al.*, 2007; Abad-Zapatero, 2007). I ja té nom: biologia estructural de sistemes.

BARREIG DE TÈCNIQUES

Els grans complexos, siguin perdurables o transitoris, estables o làbils, requeriran per al seu estudi estructural l'ús de diverses tècniques alhora. Per exemple la barreja de ME i RX, que ja ha donat tan bons resultats en l'estudi dels virus, s'haurà de generalitzar a altres partícules i complexos. En aquesta barreja els components (més petits) d'una partícula són resolts per RX a molt alta resolució i després s'encaixen dins dels volums de ME de més baixa resolució però abastant la partícula sencera (vegeu la figura 13). Amb aquest encaix acabem tenint un bon model del tot amb les parts a alta resolució, un model quasiatòmic.

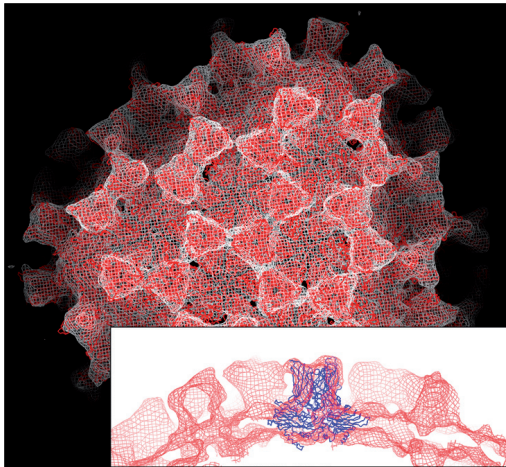


FIGURA 13. Estructura d'una càpsida del virus aviar IBDV, determinada per criomicroscòpia electrònica (Coulibaly *et al.*, 2006). En el requadre inferior es mostren tres proteïnes de la càpsida, determinades a alta resolució per RX, encaixades en una de les torretes del volum de ME, que resulta en una estructura quasiatòmica.

En els propers anys veurem com aquesta idea va uns passos més enllà, i s'intentaran encaixar també els complexos i partícules dins d'imatges de tota la cèl·lula, imatges globals que podran provenir de la tomografia, de la microscòpia confocal o d'altres tècniques. Ja es veu que la biologia estructural i la biologia cel·lular s'hauran d'aproximar (Harrison, 2004) encara que la distància tecnològica i conceptual entre aquestes disciplines sembla avui encara molt gran. Al final del camí hi ha la foto detallada, a resolució atòmica, de tota la cèl·lula.

Hem parlat de foto, però en realitat el que necessitem és una pel·lícula. Per entendre realment com funcionen els components de la cèl·lula, haurem d'esbrinar inexorablement el moviment de les proteïnes, dels complexos, de les màquines i motors moleculars i el desplaçament de tots dins de la cèl·lula. L'espai d'aquest article no ens permet ara de perdre'ns en els avenços de les diverses tècniques per analitzar aquesta dinàmica —per exemple, tècniques d'anàlisi de partícula única, o mètodes computacionals. Ara només em cal mencionar que aquestes metodologies s'hauran d'integrar, d'una manera o altra, en l'anàlisi estructural del futur.

Un futur que no serà pas avorrit per als biòlegs estructurals.

AGRAÏMENTS

Vull expressar el meu agraïment a Miquel Pons per haver revisat la part de RMN d'aquest article, a Oscar Llorca i Roeland Boer per la figura 2, a Marta Bruix per la figura 3, a Alex G. Blanco per les figures 4 i 9, a Johann Deisenhofer per la figura 8, a Joan Pons per la figura 12 i a Albert Canals per corregir el text.

BIBLIOGRAFIA

- ABAD-ZAPATERO, C. (2007). «Notes of a protein crystallographer: quo vadis structural biology». *Acta Crystallogr. D. Biol. Crystallogr.*, 63: 660-664.
- ABRESCIA, N. G.; COCKBURN, J. J.; GRIMES, J. M.; SUTTON, G. C.; DIPROSE, J. M.; BUTCHER, S. J.; FULLER, S. D.; SAN MARTIN, C.; BURNETT, R. M.; STUART, D. I.; BAMFORD, D. H.; BAMFORD, J. K. (2004). «Insights into assembly from structural analysis of bacteriophage PRD1». *Nature*, 432: 68-74.
- BAN, N.; NISSEN, P.; HANSEN, J.; MOORE, P. B.; STEITZ, T. A. (2000). «The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution». *Science*, 289: 905-920.
- BELTRAO, P.; KIEL, C.; SERRANO, L. (2007). «Structures in systems biology». *Curr. Opin. Struct. Biol.*, 17: 378-384.
- BLANCO, A. G.; SOLA, M.; BERNÚES, J.; COLL, M. (2007). «Structural basis for RNA polymerase recruitment and transcription activation by response regulator PhoB: the σ_4 / β -flap / PhoBE / *Pho Box* DNA Complex». [En prensa]
- BLANCO, A. G.; SOLA, M.; GOMIS-RÜTH, F. X.; COLL, M. (2002). «Tandem DNA recognition by PhoB, a two-component signal transduction transcriptional activator». *Structure*, 1: 1-20.
- BOER, R.; RUIZ-MASÓ, J. A.; GÓMEZ-BLANCO, A.; LOPEZ-BLANCO, J. R.; CHACÓN, P.; USÓN, I.; GOMIS-RÜTH, F. X.; ESPINOSA, M.; LLORCA, O.; SOLAR, G. DEL; COLL, M. (2007). «The structure of streptococcal plasmid replication initiation protein RepB is a hexameric ring with mobile DNA recognition and catalytic domains». [En prensa]
- BUCHANAN, S. K.; SMITH, B. S.; VENKATRAMANI, L.; XIA, D.; ESSE, L.; PALNITKAR, M.; CHAKRABORTY, R.; HELM, D. VAN DER; DEISENHOFER, J. (1999). «Crystal structure of the outer membrane active transporter FepA from *Escherichia coli*». *Nat. Struct. Biol.*, 6: 56-63.
- COLL, M. (2006). «El transcriptor de ADN». *El País* (11 d'octubre): 41.
- COULIBALY, F.; CHEVALIER, C.; GUTSCHE, I.; POUS, J.; NAVAZA, J.; BRESSANELLI, S.; DELMAS, B.; REY, F. A. (2006). «The birnavirus crystal structure reveals structural relationships among icosahedral viruses». *Cell*, 120: 761-772.
- CRACK, F. H. C. (1988). *What mad pursuit*. Londres: Weindelfeld and Nicolson.
- CRAMER, P.; BUSHNELL, D. A.; KORNBERG, R. D. (2001). «Structural basis of transcription: RNA polymerase II at 2.8 Å resolution». *Science*, 292: 863-1876.
- DAUTER, Z. (2006). «Current state and prospects of macromolecular crystallography». *Acta Crystallogr. D. Biol. Crystallogr.*, 62: 1-11.
- FERENTZ, A. N.; WAGNER, G. (2000). «NMR spectroscopy: a multifaceted approach to macromolecular structure». *Quarterly Reviews of Biophysics*, 33: 29-65.
- GAVIN, A. C.; ALOY, P.; GRANDI, P.; KRAUSE, R.; BOESCHE, M.; MARZIOCH, M.; RAU, C.; JENSEN, L. J.; BASTUCK, S.; DÜMPFELFELD, B.; EDELMANN, A.; HEURTIER, M. A.; HOFFMAN, V.; HOEFERT, C.; KLEIN, K.; HUDAK, M.; MICHON, A. M.; SCHEIDER, M.; SCHIRLE, M.; REMOR, M.; RUDI, T.; HOOPER, S.; BAUER, A.; BOUWMEESTER, T.; CASARI, G.; DREWES, G.; NEUBAUER, G.; RICK, J. M.; KUSTER, B.; BORK, P.; RUSSELL, R. B.; SUPERTI-FURGA, G. (2006). «Proteome survey reveals modularity of the yeast cell machinery». *Nature*, 440: 631-636.
- GOMIS-RÜTH, F. X.; MONCALIÁN, G.; PÉREZ-LUQUE, R.; GONZÁLEZ, A.; CABEZÓN, E.; CRUZ, F. DE LA; COLL, M. (2001). «The bacterial conjugation protein TrwB resembles ring helicases and F1-ATPase». *Nature*, 409: 637-641.
- GRABOWSKI, M. K.; JOACHIMIAK, A.; OTWINOWSKI, Z.; MINOR, W. (2007). «Structural genomics: keeping up with expanding knowledge of the protein universe». *Curr. Opin. Struct. Biol.*, 17: 347-353.
- GRIMES, J. M.; BURROUGHS, J. N.; GOUET, P.; DIPROSE, J. M.; MALBY, R.; ZIENTARA, S.; MERTENS, P. P.; STUART, D. I. (1998). «The atomic structure of the bluetongue virus core». *Nature*, 395: 470-478.
- GUASCH, A.; ALORIA, K.; PÉREZ, R.; CAMPO, R.; AVILA, J.; ZABALA, J. C.; COLL, M. (2002). «Three-dimensional structure of human tubulin chaperone cofactor A/p14». *J. Mol. Biol.*, 318: 1139-1149.
- GUASCH, A.; LUCAS, M.; MONCALIÁN, G.; CABEZAS, M.; PÉREZ-LUQUE, R.; GOMIS-RÜTH, F. X.; CRUZ, F. DE LA; COLL, M. (2003). «Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC». *Nat. Struct. Biol.*, 10: 1002-1010.
- HARRISON, S. (2004). «Whither structural biology». *Nat. Struct. & Mol. Biol.*, 11: 12-15.
- HART, D. J.; TARENDEAU, F. (2006). «Combinatorial library approaches for improving soluble protein expression in *Escherichia coli*». *Acta Crystallogr. D. Biol. Crystallogr.*, 62: 19-26.
- HENDERSON, R. (2004). «Realizing the potential of electron cryo-microscopy». *Q. Rev. Biophys.*, 37: 3-13.
- KOROSTELEV, A.; TRAKHANOV, S.; LAURBERG, M.; NOLLER, H. F. (2006). «Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements». *Cell*, 126: 1065-1077.
- NICKELL, S.; KOFLER, C.; LEIS, A. P.; BAUMEISTER, W. (2006). «A visual approach to proteomics». *Nat. Rev. Mol. Cell. Biol.*, 7: 225-230.
- PÉREZ-CANADILLAS, J. M.; GUENNEUGUES, M.; CAMPOS-OLIVAS, R.; SANTORO, J.; MARTÍNEZ DEL POZO,

- A.; GAVILANES, J. G.; RICO, M.; BRUIX, M. (2002). «Backbone dynamics of the cytotoxic ribonuclease alpha-sarcin by ^{15}N NMR relaxation methods». *J. Biomol. NMR*, 24: 301-316.
- WIMBERLY, B. T.; BRODERSEN, D. E.; CLEMONS, W. M. JR; MORGAN-WARREN, R. J.; CARTER, A. P.; VONRHEIN, C.; HARTSCH, T.; RAMAKRISHNAN, V. (2000). «Structure of the 30S ribosomal subunit». *Nature*, 407: 327-339.