

ELS PREMIS NOBEL

DE L'ANY 2008

SOBRE EL

PREMI NOBEL DE QUÍMICA

CONCEDIT A

OSAMU SHIMOMURA, MARTIN CHALFIE

I ROGER Y. TSIEN,

A CÀRREC D'ERNEST GIRALT,

DE L'INSTITUT DE RECERCA BIOMÈDICA

(PARC CIENTÍFIC DE BARCELONA)

I EL DEPARTAMENT DE QUÍMICA ORGÀNICA

DE LA UNIVERSITAT DE BARCELONA

RESUM

L'any 2008, els guanyadors del Premi Nobel de Química han estat el japonès Osamu Shimomura i els nord-americans Martin Chalfie i Roger Y. Tsien. En paraules de la Reial Acadèmia Sueca de Ciències, han guanyat el premi «pel descobriment i el desenvolupament de la proteïna fluorescent verda». La proteïna fluorescent verda (o GFP, de l'anglès *green fluorescent protein*) és un compost de massa molecular de 26.900 que es caracteritza per la seva intensa fluorescència verda (509 nm) quan s'il·lumina amb llum violada (395 nm). Va ser aïllada inicialment de la medusa *Aequorea victoria*, un hidroïdeu de les aigües costaneres del Pacífic nord-americà, i posteriorment s'ha trobat en altres organismes. La proteïna fluorescent verda està formada per 238 aminoàcids. La seva fluorescència és deguda a la formació de diversos enllaços covalents entre les cadenes laterals d'alguns d'aquests aminoàcids. És un fenomen espontani, que es coneix amb molta precisió, raó per la qual ha estat possible preparar variants de la GFP que emeten fluorescència en una àmplia gamma cromàtica. La importància de la proteïna fluorescent verda rau en la seva utilitat com a marcador en l'estudi de processos biològics. La seva baixa, o nul·la, fototoxicitat ha permès utilitzar-la en una àmplia varietat d'experiments tant *in vitro* com *in vivo*.

PARAULES CLAU: proteïna fluorescent verda, GFP, *green fluorescent protein*, fluorescència, *Aequorea victoria*, biomarcador, fototoxicitat.

ABSTRACT

In 2008 the winners of the Nobel Prize in Chemistry have been the Japanese Osamu Shimomura and American Martin Chalfie and Roger Y. Tsien. In the words of the Royal Swedish Academy of Sciences they have won the prize “for the discovery and development of the green fluorescent protein, GFP.” The green fluorescent protein is a compound of molecular weight 26,900 which is characterized by its intense green fluorescence (509 nm) when exposed to violet light (395 nm). It was initially isolated from the jellyfish *Aequorea victoria*, an hydrozoon from coastal waters of the American Pacific, and later found in other organisms. The green fluorescent protein consists of 238 amino acids. Its fluorescence is due to the formation of multiple covalent bonds between side chains of some of these amino acids. This is a spontaneous phenomenon, which now is known with great precision. For this reason, it has been possible to prepare variants of GFP that emit fluorescence in a wide color range. The importance of the green fluorescent protein lies in its usefulness as a marker in the study of biological processes. Its low, or no, phototoxicity has allowed to use it in a wide variety of experiments both in vitro and in vivo.

KEYWORDS: GFP, green fluorescent protein, fluorescence, *Aequorea victoria*, biomarker, phototoxicity.

La proteïna fluorescent verda presenta una estructura química que pot representar-se de manera senzilla (figura 1). Està formada per 238 aminoàcids, units segons una seqüència determinada. Els aminoàcids, que es representen amb una lletra de l'alfabet, són el que podríem dir-ne els maons que serveixen per construir l'estructura, no només d'aquesta proteïna, sinó de qualsevol de les que existeixen a la nostra biosfera. El

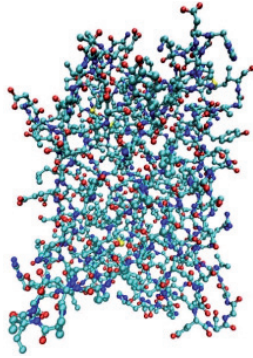


FIGURA 1. Estructura química de la proteïna fluorescent verda, segons un model de representació atòmica. El codi de colors és el següent: carboni (blau cel), nitrogen (blau marí), oxigen (vermell) i sofre (groc). Els àtoms d'hidrogen no hi estan representats.

34

nombre de proteïnes conegudes és enorme, però ridículament petit si el comparem amb el nombre de proteïnes que podem inventar i crear. El camp de la química de proteïnes és un camp magnífic per als que hi treballem, perquè les funcions de les proteïnes que s'han desenvolupat al llarg de l'evolució, i la perfecció de moltes d'aquestes funcions, ens permeten ser optimistes per a la possibilitat futura de poder crear proteïnes molt interessants en el laboratori.

En la figura 2 hi ha representats, amb un codi d'una lletra, cadascun dels 238 aminoàcids de la proteïna fluorescent verda. Normalment es fa servir un codi una mica més extens, que consisteix en un conjunt de tres lletres per a cada aminoàcid.

MSKGEELFTC	VVPVLVELDC	GEGEDATYC	DVNGQKFSVS	KLTLNFICTT	GKLPVPWPTL
VTTFSYGVQC	FSRYPDHMKQ	HDFFKSAMPE	GVVQERTIFY	KDDCNKTRA	EVKFEGDTLV
NRIELKGIDF	KEDCNLGHK	MEYNYNSHNV	YIMGDKPKNG	IKVNFKIRIN	IKDCSVQLAD
HYQQNTPIGD	GPVLLPDNHV	LSTQSALSKD	PNEKRDHMIL	LEFVTAARIT	HGMDELYK

FIGURA 2. Seqüència d'aminoàcids (1-238).

Un *aminoàcid* és una substància que té un grup àcid i un grup amina. Una *proteïna* és una cadena formada per la unió repetida d'uns aminoàcids amb uns altres, amb un enllaç anomenat *enllaç peptídic*, que uneix el grup àcid d'un aminoàcid amb el grup amina de l'aminoàcid veí. En el codi genètic hi ha vint aminoàcids codificats. En la figura 3, hi ha recollits els vint aminoàcids, amb el seu nom, el codi de lletres i l'estructura completa. Tots tenen el grup amina (és el grup NH_2) i un grup àcid (és el que apareix com a COOH). La diferència entre uns aminoàcids i els altres està en les cadenes laterals.

Amb una mirada ràpida, diríem que són molt semblants. Però, evolutivament, la vida al planeta Terra està basada en les seves diferències. És a dir, entre les cadenes laterals, que són l'element diferenciador, trobem grups àcids diferents: àcid aspàrtic, àcid glutàmic...; i també hi trobem grups bàsics, com ara la lisina. Els químics orgànics trigaren

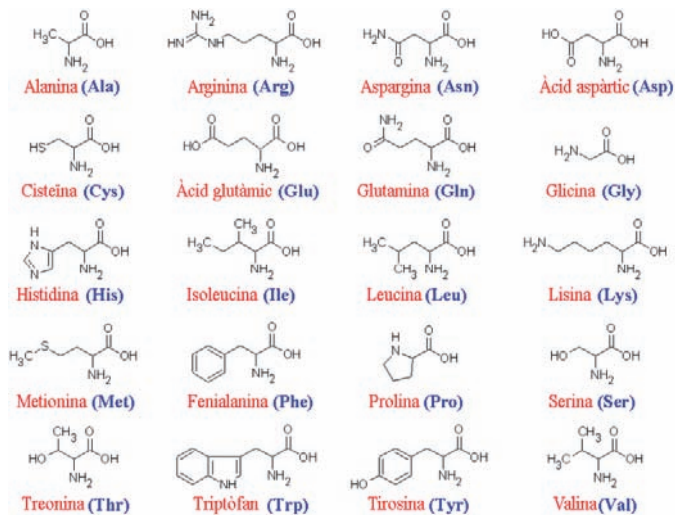


FIGURA 3. Aminoàcids essencials.

bastant de temps, fins al segle XX, a descobrir la base orgànica més forta de les conegudes, que és la guanidina. Però, evolutivament, el planeta ja feia moltíssims anys que l'havia obtingut. L'arginina inclou aquest grup a la seva cadena lateral: el guanidini. En els aminoàcids es reuneixen altres propietats de la química orgànica —podríem dir que són un resum de qual-sevol llibre de química orgànica. Són uns nucleòfils magnífics, presenten tautomeries i tenen molts tipus d'estructures possibles: aminoàcids aromàtics, alifàtics, lineals, ramificats... És a dir, es podria afirmar que aquests vint aminoàcids són el resum dels resums, i déu n'hi do el que es fa a la biosfera amb aquests vint aminoàcids.

Avui, però, no estem limitats a aquests vint aminoàcids, i en laboratoris com el nostre, en el qual sintetitzem proteïnes, disposem de centenars d'aminoàcids no naturals que ens permeten obtenir substàncies amb propietats interessants. Per exemple, que siguin molt resistents a la degradació per proteases i que puguin tenir un ús en biomedicina superior a les substàncies naturals.

La seqüència de la proteïna fluorescent (p. 30) mostra el que n'anomenem la *constitució*, és a dir, tots els àtoms que formen la proteïna i com estan connectats els uns amb els altres. Però per conèixer-ne bé l'estructura, d'aquesta i de qual-sevol proteïna, el que cal saber és l'arquitectura tridimensional de tots aquests àtoms. I això és el que es mostra en les figures següents, en què la mateixa proteïna ha estat visualitzada amb l'ordinador de diferents maneres.

La figura 4 permet veure la superfície de la proteïna. És una representació que s'anomena *de Connolly*, molt fàcil de generar computacionalment acostant-hi una esfera. En aquest cas, es fa servir una esfera d'1,4 Å de diàmetre, que és més o menys el d'una molècula d'aigua, i aleshores es mesura la superfície d'interacció d'aquesta esfera amb la proteïna. Però si el que volem és mirar endins i no quedar-nos a



FIGURA 4. Representació estructural de Connolly.

la superfície, si el que volem veure és on se situen tots els àtoms individuals, el que utilitzem és la representació que es mostra en la figura 1 (observeu el codi de colors). Els àtoms d'hidrogen no hi estan representats perquè aquesta imatge prové de dades experimentals de difracció de raigs X, de cristal·lografia de raigs X, i amb aquesta tècnica els àtoms d'hidrogen són molt poc densos i electrònicament no es poden apreciar.

Si intentem unir els àtoms d'un aminoàcid amb els de l'aminoàcid veí, comencem a intuir el que podríem anomenar la *topologia general de la proteïna* (figura 5).

Aquestes cintes, molt habituals en les representacions de química de proteïnes, s'anomenen *cadena beta* i, en la figura, es troben juntament amb uns cilindres que representen un altre tipus d'estructura, anomenada *hèlix alfa*, on la cadena peptídica està formant una estructura helicoidal. Què fa, però, que la proteïna fluorescent verda sigui especial? Doncs justament això, que és fluorescent.

Per esbrinar l'origen de la fluorescència d'aquesta proteïna, cal abans revisar l'origen del fenomen de la fluorescència en general. Tots sabem que la llum blanca, la llum d'ori-

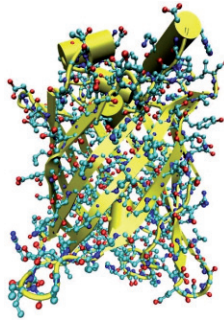


FIGURA 5. Representació de la proteïna verda fluorescent, sobreposada a les unions peptídiques de l'estructura atòmica.

gen solar, està formada per una sèrie de components de diferents longituds d'ona. El concepte *color* es considera avui un concepte biològic. El color no és més que la sensació; l'efecte biològic en el cervell té la incidència de la radiació electromagnètica de qualsevol d'aquestes freqüències sobre els fotoreceptors de les cèl·lules de la retina. Si nosaltres tenim una substància orgànica, aquesta substància pot tenir un gran nombre d'electrons de tipus π , i passa a ser una molècula que anomenem *molt conjugada*. Aquests tipus de substàncies, quan absorbeixen llum ultraviolada o visible, passen d'un estat electrònic fonamental a un estat excitat. Si realment la conjugació que experimenta la molècula és molt extensa, en absorbir llum visible pot passar d'un primer estat singlet a un segon estat d'excitació.

Una vegada la molècula es troba en estat d'excitació, pot passar a un altre estat excitat d'energia més baixa, i es produeix el fenomen de la *fluorescència*. Posteriorment, per tornar a l'estat fonamental, té lloc l'emissió d'un quàntum. El quàntum que s'ha emès té menys energia que el quàntum que s'ha absorbit. Una energia menor significa una freqüència de radiació més petita i, per tant, una longitud d'ona més gran.

Això és el que passa en el cas de la proteïna que ens ocupa. L'absorció inicial té lloc a 395 nm; això vol dir en la zona del blau a l'espectre, gairebé del violat; i l'emissió, en canvi, per a nosaltres ja és visible i de color verd (509 nm), en la zona límit entre el verd i el blau.

A la natura, la fluorescència no és l'únic fenomen d'emissió de llum. Hi ha sistemes que passen d'aquest estat excitat singlet a un estat excitat triplet, i, aleshores, s'emeten també llum i es parla de *fosforescència*. La diferència entre aquests dos fenòmens és de tipus cinètic. La fluorescència és un fenomen molt ràpid, mentre que la fosforescència és molt lenta, perquè aquesta transició normalment és lenta. Tots dos fenòmens s'anomenen *luminescència*. L'un és ràpid i l'altre és lent. Encara hi ha un altre fenomen pel qual les molècules poden emetre llum, i és el que s'anomena *quimioluminescència*, que és el que fan servir, per exemple, les cuques de llum. Aquí té lloc una reacció química i, com a conseqüència, apareix una substància en un estat excitat i s'emeten un quàntum de llum a la zona del visible.

I on és el *fluoròfor*? És a dir, on és la part de la proteïna verda fluorescent que emet aquesta llum? Quan es va descobrir la proteïna, tothom pensava que aquest fluoròfor seria una substància aliena a l'estructura de la proteïna. El terme tècnic que fem servir per anomenar-lo és *grup prostètic*. Per exemple, l'hemoglobina és roja no perquè els aminoàcids li donin un color roig, sinó perquè té una altra part, el grup hemo, que és un colorant que fa que sigui de color roig. Avui sabem que són els aminoàcids mateixos els que es combinen entre si per donar aquesta propietat de fluorescència. El fluoròfor es troba dins mateix de l'estructura proteica. El fluoròfor de la proteïna fluorescent verda es coneix sovint amb les sigles GFP (*green fluorescent protein*). Aquest fluoròfor de la GFP té una part que ve de la tirosina (que és l'aminoàcid que ocupa la posició 66 de la cadena), una altra part que ve de la

glicina (en la posició 67) i una altra part que ve de la treonina (en la posició 65).

En la figura 6 es pot observar la molècula tal com es biosintetitza. En aquesta etapa, la molècula encara no és fluorescent. Consisteix en una cadena formada per enllaços peptídics amb una cadena lateral amb els grups hidroxil, metil, alcohol, fenol, tirosina; també hi ha la glicina que només té un protó... La primera transformació que ocorre en la molècula precursora és una *ciclació*, és a dir, la formació d'un anell entre l'esquelet i una cadena lateral. Posteriorment, té lloc una segona ciclació, que dóna com a resultat un anell de cinc baules que, combinat amb l'aparició d'un doble enllaç i junt amb els dobles enllaços de la tirosina i d'altres grups, és el responsable de la fluorescència. Aquest, doncs, és el grup fluoròfor característic de la proteïna.

La història de l'estudi de la proteïna fluorescent verda comença amb el seu aïllament a la dècada dels seixanta pel científic japonès Osamu Shimomura. A banda d'ell, s'ha parlat de sis grups de recerca, que no són els únics, però sí els que em semblen més rellevants:

- 1) Aïllament de la GFP de la medusa *Aequorea victoria*. Osamu Shimomura, 1960-1970.
- 2) Aïllament i seqüenciació del cDNA de la GFP. Douglas Prasher, 1992.
- 3) Expressió de la GFP en l'*E. coli* i en el *C. elegans*. Martin Chalfie, 1994.

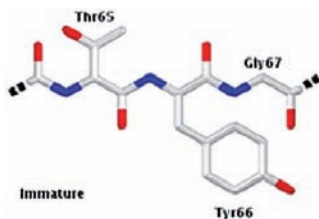


FIGURA 6. Estructura química del precursor del grup fluoròfor.

4) Estructura cristal·logràfica (raigs X) d'un mutant de GFP. James Remington, 1996.

5) Estructura cristal·logràfica (raigs X) de GFP salvatge. George Phillips, 1996.

6) Comprensió del mecanisme, GFP modificades. Roger Y. Tsien, 2002.

La Reial Acadèmia Sueca de Ciències va triar Osamu Shimomura, Martin Chalfie i Roger Y. Tsien per atorgar-los el Premi Nobel de Química de l'any 2008. Osamu Shimomura va aïllar la proteïna a partir d'una medusa (figura 7). En aquesta fotografia, el color blau de la medusa no correspon a la fluorescència de la GFP. La medusa és blava perquè és quimioluminescent; seria com una cuca de llum de color blau. Però, sobreposada, s'hi aprecia una fluorescència. Són els piquets de color verd que estan en una mena de corona, que són en realitat petits orgànuls, semblants a unes bossetes, que contenen la GFP, la proteïna verda fluorescent. Shimomura, que estava interessat en l'estudi de la quimioluminescència, quan va veure aquest color verd, va pensar que estava davant d'un nou tipus de quimioluminescència i es va proposar

41

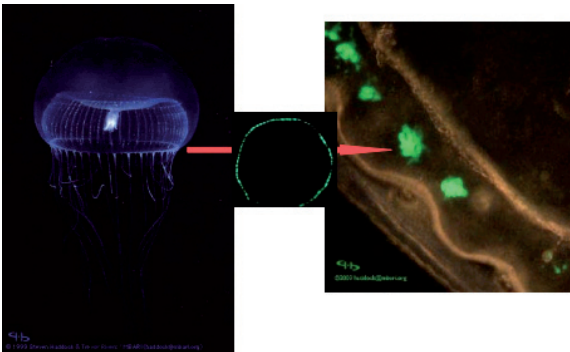


FIGURA 7. Imatge d'una medusa *Aequorea victoria* i detall de la fluorescència verda.

aïllar-ne la substància responsable. El problema que va trobar-se va ser que la quantitat present en cada medusa era petítíssima i va haver de fer servir desenes de milers de meduses per tal de poder aïllar una petita quantitat de proteïna.

Una vegada que va estudiar el sistema proteic, va entendre tot el procés. En la medusa hi ha una altra proteïna, que és l'aequorina (que també la va descobrir ell), que sí que és quimioluminescent, és a dir, transforma l'energia química en llum blava. Quan l'aequorina emet la llum blava, la GFP l'absorbeix i, al seu torn, emet fluorescència de color verd. O sigui, que tot el procés d'excitació i emissió té lloc dins del mateix animal.

L'any 1992, la biologia molecular ja ha fet aparició, l'enginyeria genètica és clarament una realitat, i aleshores Douglas Prasher es planteja clonar aquesta proteïna per poder seqüenciar-ne el DNA. Fins aquell moment, la seqüència d'aminoàcids (p. 34) encara no era coneguda. Se sabia que la substància era una proteïna, però no se'n coneixia res de l'estructura. Per determinar l'estructura d'una proteïna, hi ha dos mètodes, tots dos molt eficaços. El primer consisteix a agafar la proteïna i, per anàlisi química, esbrinar quins són els aminoàcids que la componen i en quin ordre hi són. L'altre mètode consisteix a fer el mateix però treballant amb l'àcid nucleic que codifica aquesta proteïna. Això és el que va fer Douglas Prasher. Va aïllar el DNA de la GFP i el va seqüenciar, però, malauradament, l'estudi el feia amb l'ajut d'un projecte finançat pels Instituts Nacionals de la Salut (NIH) i, al cap de dos anys, no l'hi van renovar i va haver d'aturar la seva recerca.

Dissortadament per a ell, quan ja tenia el DNA, el següent pas —que era el fantàstic— consistia a incorporar-ne una còpia, el cDNA, en un organisme, per exemple, un bacteri com l'*Escherichia coli*, i expressar la proteïna (sintetitzar-la en un organisme) no en la medusa, sinó en l'*E. coli*. I això és el que va fer Martin Chalfie, que va aprofitar-se de la generositat de Dou-

glas Prasher perquè li va demanar el seu clon, el cDNA, i Douglas Prasher l'hi va cedir. Així doncs, Martin Chalfie va ser el primer que va expressar la proteïna segons el que s'anomena *expressió heteròloga*, és a dir, la proteïna s'obté en un organisme viu diferent del de procedència. Ho va fer en dos organismes: en el bacteri *E. coli* i, més tard, en el cuc *Caenorhabditis elegans*.

Si el que tenim habitualment és un gen que codifica una proteïna determinada i el modifiquem genèticament, la modificació resultant tindrà lloc en un bacteri, un cuc, una mosca... Una modificació genètica consisteix en què, a més a més del gen que codifica la nostra proteïna d'interès, hi afegim una part que codifica una altra proteïna, per exemple, la GFP. Quan tota aquesta construcció s'expressa en l'ésser viu elegit, el que tindrem, en lloc de la nostra proteïna determinada, serà una proteïna híbrida: una fusió de la nostra amb la GFP. La importància d'aquest tipus d'exercici és la capacitat de fluorescència de la proteïna GFP.

Martin Chalfie va publicar els seus dos primers treballs en la revista *Science* el 1994 (www.conncoll.edu/ccacad/zimmer/GFP-ww/prasher.html). En la figura 8 es mostra una

43

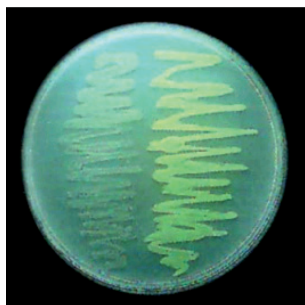


FIGURA 8. A la dreta, bacteris que contenen el plasmidi d'expressió de la GFP; a l'esquerra, una traça de bacteris *Escherichia coli*, el control negatiu. Les cèl·lules es van irradiar amb una font de radiació ultraviolada d'alta freqüència. (Font: Chalfie *et al.* (1994).)

de les figures del seu article, amb una placa de Petri amb medi de cultiu i el creixement de microorganismes. A l'esquerra s'observa una traça de bacteris *E. coli* que no han estat transformats amb la proteïna fluorescent; seria un control negatiu. I a la dreta hi ha una traça de bacteris fluorescents, ja que han estat modificats genèticament perquè expressin la proteïna GFP.

Passem a l'any 1996, quan ja es coneixia la seqüència de la proteïna verda fluorescent, però encara no se n'havien establert les estructures tridimensionals. James Remington, primer, i George Phillips, immediatament després, van ser els primers a descobrir i descriure l'estructura cristal·logràfica, és a dir, l'estructura tridimensional. El primer d'ells va utilitzar un mutant i el segon ja va fer servir la *GFP wild type*, la GFP salvatge. El descobriment d'aquesta estructura cristal·logràfica va permetre veure per primera vegada que el fluoròfor estava format per les cadenes de aminoàcids mateixes, la qual cosa constituïa una situació realment molt nova. Però la comprensió del mecanisme de formació d'aquest fluoròfor va ser resultat dels estudis de Roger Y. Tsien (2002).

Roger Y. Tsien és un científic que segueix en actiu a la Universitat de Califòrnia a San Diego. En la figura 9 es pre-



FIGURA 9. Mostres de proteïnes fluorescents produïdes al seu laboratori. (Font: www.tsienlab.ucsd.edu.)

senta una imatge amb mostres de diferents proteïnes produïdes al seu laboratori, entre aquestes la GFP, i d'altres que són resultat de modificacions. Totes aquestes proteïnes són fluorescents, però cadascuna produeix fluorescència a una longitud d'ona diferent i, per tant, té un color diferent. Aquestes eines, en mans dels biòlegs cel·lulars, són petites meravelles. En la part inferior de la figura es mostra un *divertimento* del seu laboratori. Torna a ser una placa. Cadascuna d'aquestes lletres és una colònia bacteriana que expressa una proteïna diferent i, per tant, mostra un color diferent. En el seu web, www.tsienlab.ucsd.edu, hi estan exposades les proteïnes que es fan servir per a diferents propòsits i diverses figures provinents d'articles publicats pel grup.

APLICACIONS DE LA GFP

A continuació descriurem un parell d'exemples d'aplicació de la GFP en diferents laboratoris. El primer exemple utilitza l'expressió d'una proteïna híbrida, en la qual una part és la proteïna d'interès i l'altra part és simplement la GFP. L'animal elegit per expressar la proteïna híbrida és el cuc *C. elegans* (figura 10). Del resultat de l'experiment es dedueix que on s'observen zones fluorescents s'expressa la β -integrina. Per què? Perquè la β -integrina és la proteïna el gen de la qual s'ha fusionat amb el gen de la GFP. Per això, allà on es localitza la GFP, sabem que també s'hi localitza la β -integrina. En aquest exemple visualitzem gairebé proteïnes individuals a nivell cel·lular.

Un altre exemple, en un nivell d'anàlisi superior al molecular, tracta del cervell d'una mosca, la *Drosophila melanogaster*. Tal com es veu en la figura 11, el color verd correspon a uns òrgans del cervell, anomenats *bossos bolet*, que són justament els que s'han etiquetat amb la GFP, i així s'aconsegueix localitzar-los.

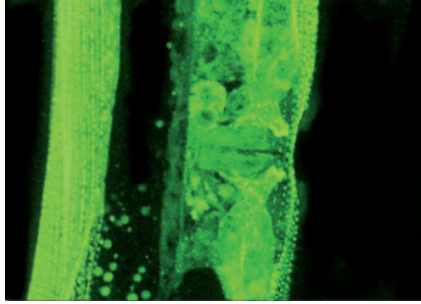


FIGURA 10. Imatge de microscòpia confocal que mostra la localització de la proteïna fluorescent en un fragment de teixit del nematode *Caenorhabditis elegans*.

En el camp de l'experimentació animal, dels models animals per a la recerca biomèdica, la GFP està tenint moltíssim impacte (www.conncoll.edu/ccacad/zimmer/GFP-wu/glowinggenes.html). Per exemple, el grup de Nicolaus Plesnilla, a la Universitat de Munic, disposa d'un model de barrera hematoencefàlica fantàstic. La barrera hematoencefàlica és el mecanisme que protegeix el cervell de la invasió de substàncies i organismes aliens. La base física d'aquesta barrera està constituïda per les cèl·lules endotelials dels vasos sanguinis, que mostren unes unions molt fortes, que fan que, a diferència del que passa a la resta de l'organisme, l'accés al cervell sigui molt difícil. Hi ha un experiment que va fer Paul Ehrlich fa més de cent anys i que ho exemplifica perfectament —en un estil una mica cru—: consisteix a injectar tinta xinesa a la cua d'un ratolí i obrir-lo en canal; tot el ratolí esdevé negre, excepte el cervell, que resta blanc.

Malgrat la dificultat, la barrera hematoencefàlica pot superar-se. En el nostre laboratori, una de les línies de recerca és dissenyar pèptids i petites proteïnes que siguin prou intel·ligents per travessar la barrera hematoencefàlica. L'objectiu és fer-les servir de llançadores de fàrmacs que normalment

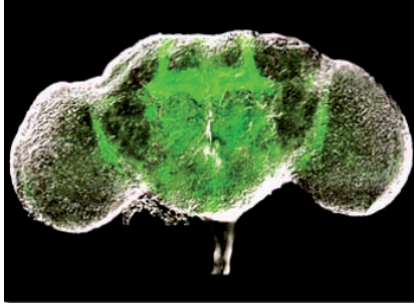


FIGURA 11. Imatge tridimensional de microscòpia confocal d'una secció de cervell d'un individu adult de *Drosophila melanogaster*. Els *bossos bolet* queden marcats per la GFP. (Font: www.olympusconfocal.com/gallery/index.html.)

no arribarien al cervell, perquè d'aquesta manera (de la mà de la llançadora) aconseguixin entrar-hi.

En el grup de Plesnila es treballa amb ratolins transgènics. En aquests, la barrera hematoencefàlica, és a dir, les cèl·lules endotelials dels capil·lars del cervell (i només aquestes), està etiquetada amb la GFP, i, per tant, la barrera hematoencefàlica esdevé fluorescent i es veu molt bé al microscopi, de color verd. Si la substància d'interès l'etiquetem amb un altre fluoròfor d'un color diferent, podrem saber immediatament si està davant de la barrera i no la travessa, si ja l'ha travessat o si, com, malauradament, passa de vegades, entra a la barrera però es queda allà i no la supera. Molts experiments de biologia, des de la biologia cel·lular fins a la biologia animal, es beneficien de la GFP.

47

Relacions personals amb la GFP

Quina és la meva relació amb la GFP? Pràcticament nul·la. En tot cas, molt indirecta. Us ho explicaré. A final dels anys vuitanta i començament dels noranta, la principal àrea d'in-

vestigació del nostre grup de recerca era intentar disposar de mètodes per sintetitzar pèptids, que són proteïnes petites. Els pèptids estan formats tan sols per quatre, cinc, deu o vint aminoàcids. Va ser Robert Bruce Merrifield, ja fa bastants anys, qui va aconseguir un mètode de síntesi dels pèptids en fase sòlida, i per aquest treball li van concedir el Premi Nobel de Química el 1984. Aquest mètode consisteix a agafar boletes de plàstic, de menys d'un mil·límetre de diàmetre, i ancorar-hi els aminoàcids. Un cop són a l'interior de les boletes, els anem unint químicament, l'un darrere l'altre, i quan ja tenim feta la cadena peptídica, és a dir, tota la petita proteïna, trenquem l'enllaç que separa el pèptid de la boleta de plàstic, purifiquem el pèptid obtingut i el caracteritzem.

L'any 1990, aquest mètode estava perfectament establert, però quan es va voler aplicar a l'obtenció de proteïnes, de seixanta o setanta aminoàcids, no va funcionar. El problema rau en el fet que les reaccions d'introducció d'un aminoàcid rere l'altre no donen rendiments del 100 %. Com que el producte no es pot purificar fins al final, quan alliberem el pèptid obtingut i el separem, apareixen moltes impureses, que són molt semblants al pèptid però que tenen imperfeccions. Pot ser que manqui un aminoàcid, o que hi hagi cadenes amb repeticions, etc., i això fa que resultin impossibles de separar. En resum, si el pèptid que volem sintetitzar és petit, el mètode funciona; però si és una proteïna, no.

En el nostre grup es va desenvolupar el mètode de síntesi convergent en fase sòlida. Per què convergent? Perquè el que es feia era sintetitzar fragments de la proteïna; purificar-los; deixar-los tots protegits, és a dir, preparats químicament, i tot seguit unir-los fins a formar la proteïna gran. Tots els passos tenien lloc en la fase sòlida. S'utilitzava una primera resina que servia per preparar els pèptids. Un cop sintetitzats, s'alliberaven de la resina, es purificaven i després s'utilitzava una segona resina per fer la proteïna.

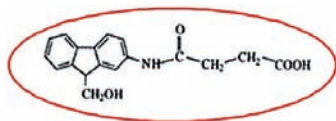


FIGURA 12. Estructura química que mostra la unió d'un aminoàcid a la resina.

Una de les resines que el nostre grup va descobrir per fer aquests pèptids protegits estava basada en una unió entre els aminoàcids i la resina (figura 12) que permetia d'obtenir els pèptids protegits d'una manera molt suau. Quina no seria la nostra alegria quan, quatre anys després d'haver descrit aquest mètode, ens vam trobar en la literatura que el grup japonès de Shumpei Sakakibara havia fet servir la nostra resina, amb una petita modificació de la nostra metodologia, per a la síntesi de la GFP.

La GFP és una proteïna molt gran. El màxim de grandària de proteïna que nosaltres havíem abordat era la síntesi d'una toxina d'escorpi de seixanta-quatre aminoàcids, i va requerir molts anys de feina de molta gent. En el grup de Sakakibara, que he visitat, van participar més de vint científics en aquest treball i la recerca fins a aconseguir sintetitzar la GFP va prendre prop de sis anys. En aquell moment, va representar el «rècord» de la síntesi de proteïnes i, fins i tot avui, continua essent la proteïna més gran que s'ha sintetitzat químicament amb un grau de puresa elevat. En la figura 13, hi ha la seqüència d'aminoàcids de la GFP que es mostrava a l'inici d'aquest text (part A) i tots els pèptids que es van sintetitzar en el grup de Sakakibara (part B).

Tots aquests pèptids es van obtenir utilitzant la resina que nosaltres havíem descrit, només amb una diferència. La unió dels pèptids, en lloc de fer-la novament en la fase sòlida, la feren en solució, és a dir, en la fase líquida. Aquest treball de Sakakibara, a part del seu mèrit intrínsec per al nivell de la química sintètica, va tenir un impacte important en el camp

(A) Primary structure of GFP



FIGURA 13. Seqüència d'aminoàcids de la GFP (A) i tots els pèptids sintetitzats pel grup de Sakakibara (B). (Font: Sakakibara *et al.* (1998).)

50

de l'estudi de la GFP. Va ser la primera demostració, i de fet continua essent l'única, de què per formar el fluoròfor no es necessita cap mena de reacció enzimàtica, ni cap mena de reactiu especial. És a dir, evolutivament, la proteïna ha aconseguit una seqüència d'aminoàcids en la qual, espontàniament, apareix el fluoròfor. La síntesi química mateixa ho demostra.

Quan finalitza la síntesi, la proteïna no és fluorescent, però deixant-la en un tampó de pH i de força iònica adequats, amb el temps, la fluorescència comença a aparèixer, i al final la proteïna esdevé fluorescent de manera estable (en poques hores). L'espectre d'absorció i l'espectre d'emissió de la proteïna sintètica són idèntics als de la proteïna natural. Això ens permet afirmar que no cal que en *E. coli*, ni en el *C. elegans* existeixi cap mena de sistema enzimàtic que acabi de transformar la seqüència de la GFP en la GFP fluorescent, sinó que la seqüència peptídica mateixa ja conté tota la informació perquè espontàniament es formi el fluoròfor.

Anys més tard, hem col·laborat amb el laboratori de Sakakibara. En el nostre grup es treballava amb l'inhibidor

natural de la catepsina L humana (P41icf). Es tracta d'una proteïna que no aconseguíem sintetitzar per mètodes tradicionals, ni podíem expressar-la en l'*E. coli* per enginyeria genètica. Cristina Chiva, estudiant de doctorat, va fer una estada d'uns quants mesos al Japó i, en col·laboració amb el grup de Sakakibara, va sintetitzar la proteïna (figura 14) i va determinar-ne l'estructura bidimensional fent servir, en aquest cas, no la difracció de raigs X, sinó la ressonància magnètica nuclear (RMN).

La bellesa de la GFP

Per finalitzar, comentaré un darrer exemple d'aplicació de la GFP, que s'ha triat no tant pel seu interès científic, sinó pel seu interès estètic. Es tracta de l'article «Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system», de Jean Livet *et al.*, publicat a *Nature* a final de 2007. Aquest treball tracta, com el seu nom indica, de la uti-

51

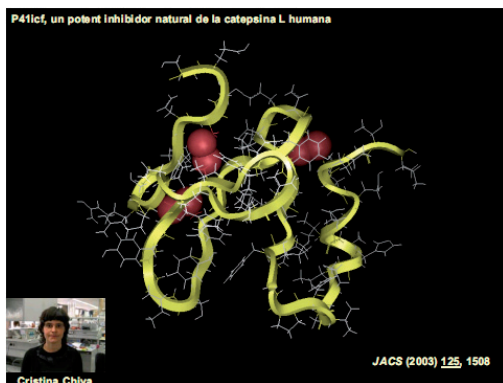


FIGURA 14. Estructura de l'inhibidor de la catepsina L humana (P41icf) i l'estudiant Cristina Chiva, responsable de la seva síntesi.

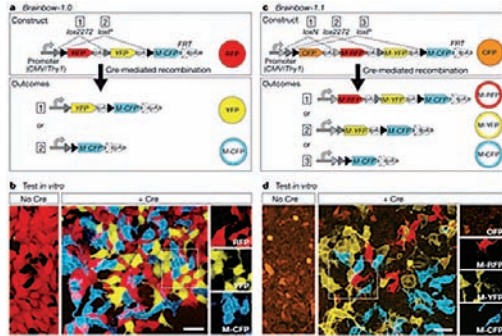


FIGURA 15. Una expressió combinatòria de proteïnes fluorescents de diferents colors. (Font: Livet *et al.* (2007).)

52

lització de tècniques de genètica en el camp del sistema nerviós central per fer el que els autors anomenen «una expressió combinatòria de proteïnes fluorescents de diferents colors». No descriuré en detall l'article; no podria perquè en absolut no sóc un expert en genètica. Em limitaré a mostrar com, amb tècniques de genètica combinatòria, expressen una proteïna, derivada de la GFP, de color vermell; o una altra de color groc, o una altra de color blau, però de manera que una mateixa neurona només expressa moltes còpies d'una de les proteïnes (figura 15). Així s'aconsegueix que cadascuna de les cèl·lules tingui un color diferent: hi ha cèl·lules vermelles, grogues i blaves. I totes les cèl·lules, en aquest cas concret, són idèntiques, és a dir, que aquesta tècnica permet fer un traçat del circuit neuronal amb molta facilitat.

Acabarem amb una altra visió del mateix sistema. En la figura 15, els colors eren de parxís (colors purs), però per obtenir colors pastel (si voleu passar de Van Gogh a Degas), cal fer una química combinatòria més subtil, en la qual el que s'expressi siguin barreges en diferents proporcions de les tres proteïnes. Així s'obtenen les precioses imatges de la figura 16.

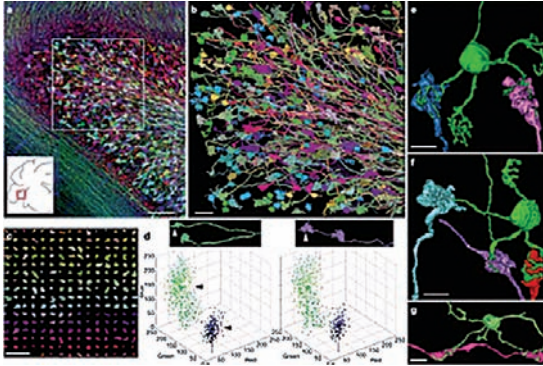


FIGURA 16. Mapatge de neurones individuals cerebrals mitjançant l'ús de diferents proteïnes fluorescentes. (Font: Livet *et al.* (2007).)

Bellesa i ciència van de braçet. També això ens ha aportat aquesta petita excursió pel món de la GFP i dels guardonats amb el Nobel de Química de l'any 2008: Osamu Shimomura, Martin Chalfie i Roger Y. Tsien.

BIBLIOGRAFIA

- CHALFIE, M.; TU, Y.; EUSKIRCHEN, G.; WARD, W. W.; PRASHER, D. C. (1994). «Green fluorescent protein as a marker for gene expression». *Science*, vol. 263, núm. 5148, p. 802-805.
- LIVET, J.; WEISSMAN, T. A.; KANG, H.; DRAFT, R. W.; LU, J.; BENNIS, R. A.; SANES, J. R.; LIGHTMAN, J. W. (2007). «Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system». *Nature*, vol. 450, p. 56-62.
- LLOYD-WILLIAMS, P.; ALBERICIO, F.; GIRALT, E. (1993). «Convergent solid-phase peptide synthesis». *Tetrahedron*, vol. 49, núm. 48, p. 11065-11133.

- ORMÖ, M.; CUBITT, A. B.; KALLIO, K.; GROSS, L. A.; TSIEN, R. Y.; REMINGTON, S. J. (1996). «Crystal structure of the *Aequoria victoria* green fluorescent protein». *Science*, vol. 273, núm. 5280, p. 1392-1395.
- PRASHER, D. C.; ECKENRODE, V. K.; WARD, W. W.; PRENDERGAST, F. G.; CORMIER, M. J. (1992). «Primary structure of the *Aequorea victoria* green-fluorescent protein». *Gene*, vol. 111, núm. 2, p. 229-233.
- RABANAL, F.; GIRALT, E.; ALBERICIO, F. (1995). «Synthesis and applications of a new base-labile fluorene derived linker for solid-phase peptide synthesis». *Tetrahedron*, vol. 51, núm. 5, p. 1449-1458.
- SAKAKIBARA, S.; NISHIUCHI, Y.; INUI, T.; NISHIO, H.; BÓDI, J.; KIMURA, T.; TSUJI, F. I. (1998). «Chemical synthesis of the precursor molecule of the *Aequorea* green fluorescent protein, subsequent folding, and development of fluorescence». *PNAS*, vol. 95, núm. 23, p. 13549-13554.
- YANG, F.; MOSS, L. G.; PHILLIPS, G. N. (1996). «The molecular structure of green fluorescent protein». *Nature Biotechnology*, vol. 14, p. 1246-1251.
- ZHANG, J.; CAMPBELL, R. E.; TING, A. Y.; TSIEN, R. Y. (2002). «Creating new fluorescent probes for cell biology». *Nature Reviews Molecular Cell Biology*, vol. 3, p. 906-918.

WEBS

- <<http://www.conncoll.edu/ccacad/zimmer/GFP-ww/prasher.html>>
- <<http://www.tsienlab.ucsd.edu>>
- <<http://www.conncoll.edu/ccacad/zimmer/GFP-ww/glowing-genes.html>>
- <<http://www.olympusconfocal.com/gallery/index.html>>