

Ideas para un proyecto de cristalografía: identificando proteínas propicias para cristalizar

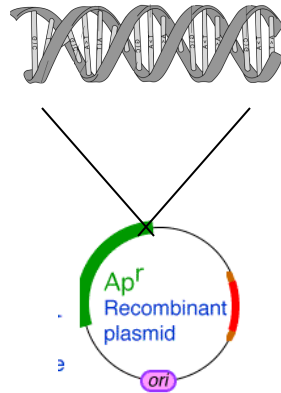
Nuria Sánchez Puig
Instituto de Química
Departamento de Química de Biomacromoléculas



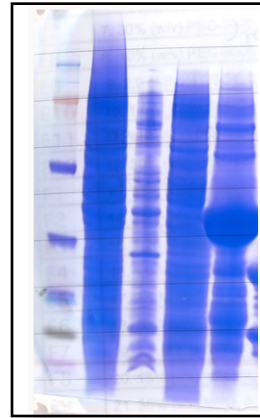
Fuente natural



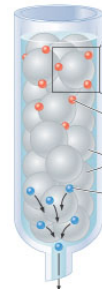
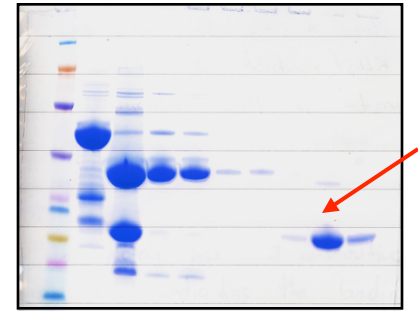
Clonación
del gen

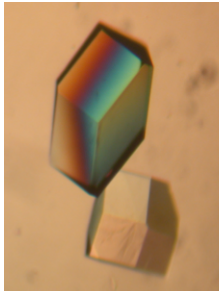


Expresión de la
proteína en un
sistema heterólogo

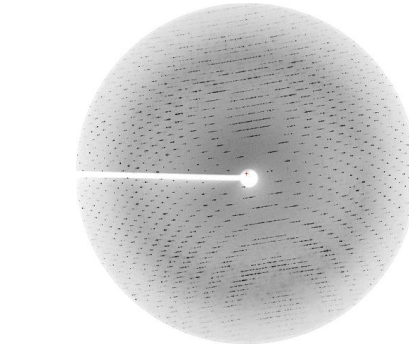


Obtención de
proteína pura

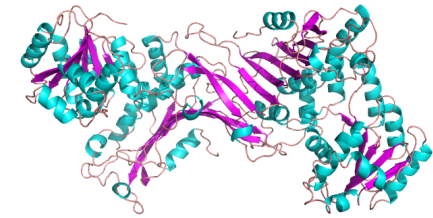




Obtención de
cristales de
proteína



Colecta de datos de
difracción de rayos X



Análisis y
resolución
estructural

Factores químicos

- Agente precipitante
- pH, pI, metales, detergentes...

Factores físicos

- Temperatura
- Metodología ...

Factores Bioquímicos

Pureza de la muestra
Concentración
Modificaciones
postraduccionales ...

≥ 1 año

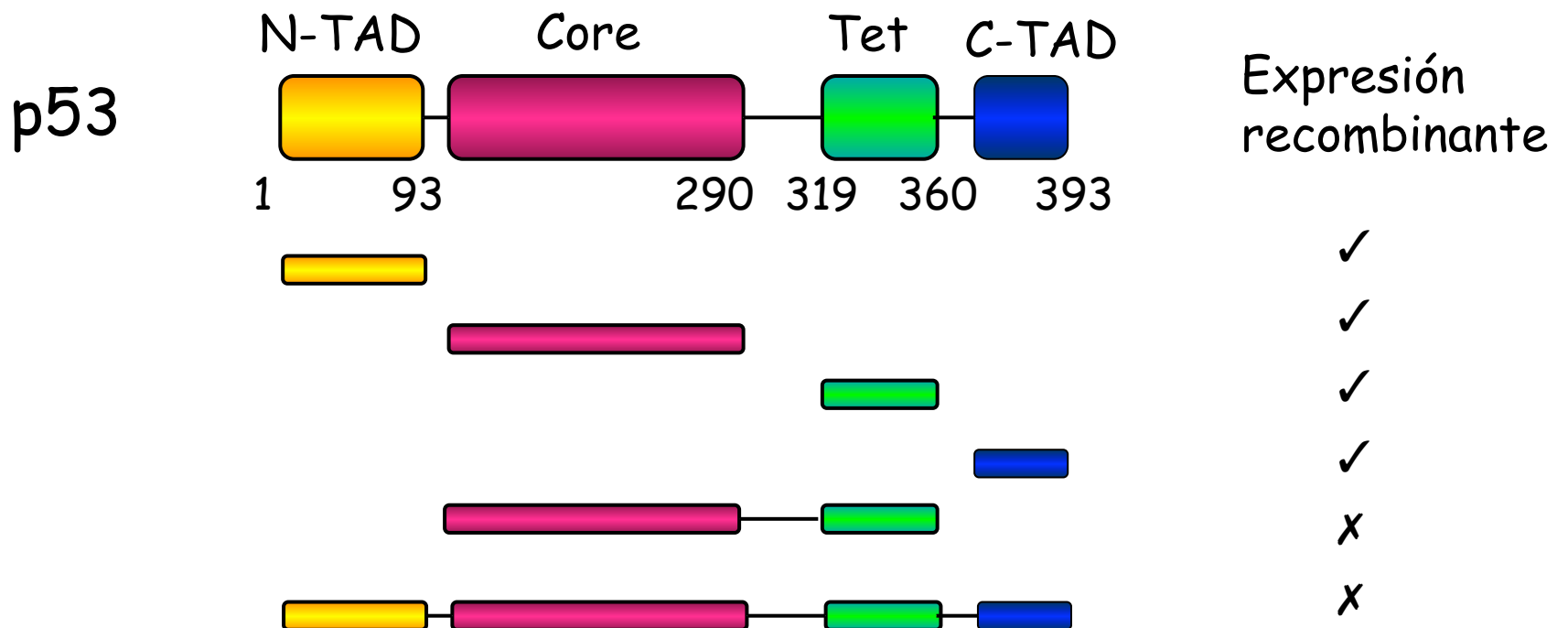
... es un proceso largo, costoso y laborioso

Como elegir inteligentemente un sistema adecuado de sobre-expresión

- ¿Cuál es el origen de la proteína quiero expresar?
- ¿En qué compartimento celular se encuentra naturalmente? Citosol, ER, peroxisomas
- Necesita de modificaciones post-traduccionales ó puentes de disulfuro para ser funcional ó para plegarse
- Tiene una arquitectura organizada en dominios
- Tiene regiones grandes desestructuradas o es una proteína intrínsecamente desestructurada

Tiene mi proteína dominios estructurales que se pueden expresar separadamente

CATH (Protein Structure Classification) www.cathdb.info
Pfam (Protein families database) <http://pfam.sanger.ac.uk/>



Tiene mi proteína regiones desordenadas o de baja complejidad - eliminalas!!

Segmentos transmembranales

TMHMM: <http://www.cbs.dtu.dk/services/TMHMM/>

Hebras entrelazadas

COILS: http://www.ch.embnet.org/software/COILS_form.html

PairCoils:

<http://groups.csail.mit.edu/cb/paircoil/cgi-bin/paircoil.cgi>

Peptidos señal

SIG-Pred:

http://bmbpcu36.leeds.ac.uk/prot_analysis/Signal.html

PrediSi: <http://www.predisi.de>

¿Por qué es importante saber si tu proteína es un proteína intrínsecamente desestructurada antes de cualquier experimento?

**Cristalizar proteínas intrínsecamente
desestructuradas
ó**

**Proteínas con fragmentos largos desordenados
son muy difíciles de cristalizar!!**

¿Cómo predecir si mi proteína se trata de una proteína intrínsecamente desordenada?

1. Predecir su contenido de desorden - Programas en la red.
2. Predecir su contenido de estructura secundaria.
3. Analizar su composición de amino ácidos - detectar un alto contenido de residuos promotores de desorden.
4. Evaluar su Carga neta media $\langle R \rangle$ y su Hidrofobicidad media $\langle H \rangle$.

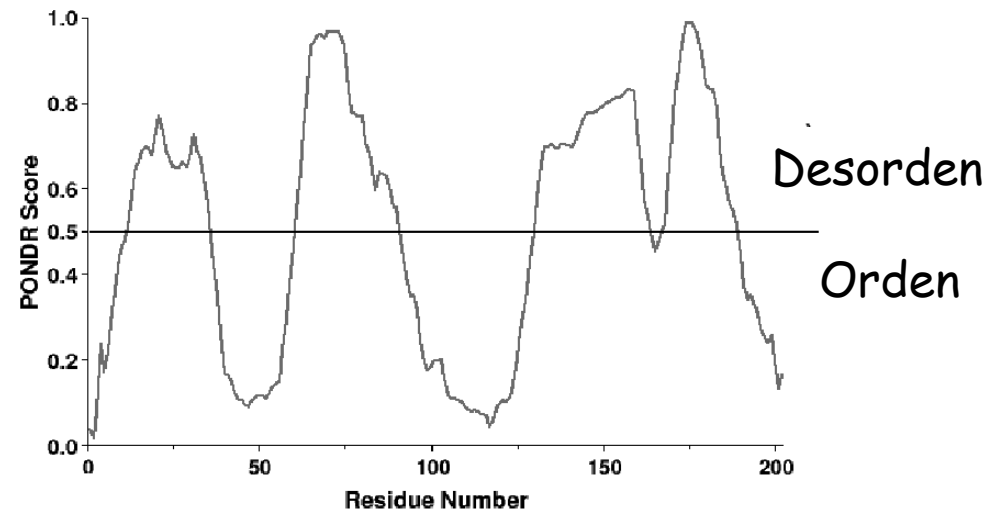
Programas para predecir el contenido de desorden de una proteína

Programa	Dirección URL
PONDR-FIT™	http://www.disprot.org/pondr-fit.php
DisEMBL™	http://www.disprot.org/
DISOPRED2	http://bioinf4.cs.ucl.ac.uk:3000/disopred/
DRIP-PRED	http://www.sbc.su.se/~maccallr/disorder/
DISpro	http://scratch.proteomics.ics.uci.edu/
FoldIndex©	http://bip.weizmann.ac.il/fldbin/findex
GlobPlot2	http://globplot.embl.de/
IUPred	http://iupred.enzim.hu/index.html
PONDR®	http://www.pondr.com/
RONN	http://www.strubi.ox.ac.uk/RONN
SPRITZ	http://protein.cribi.unipd.it/spritz/
FoldUnfold	http://skuld.protres.ru/~mlobanov/ogu/ogu.cgi
DisProt	http://www.ist.temple.edu/disprot/Predictors.html

1. Predicción de desorden usando PONDR®VLXT

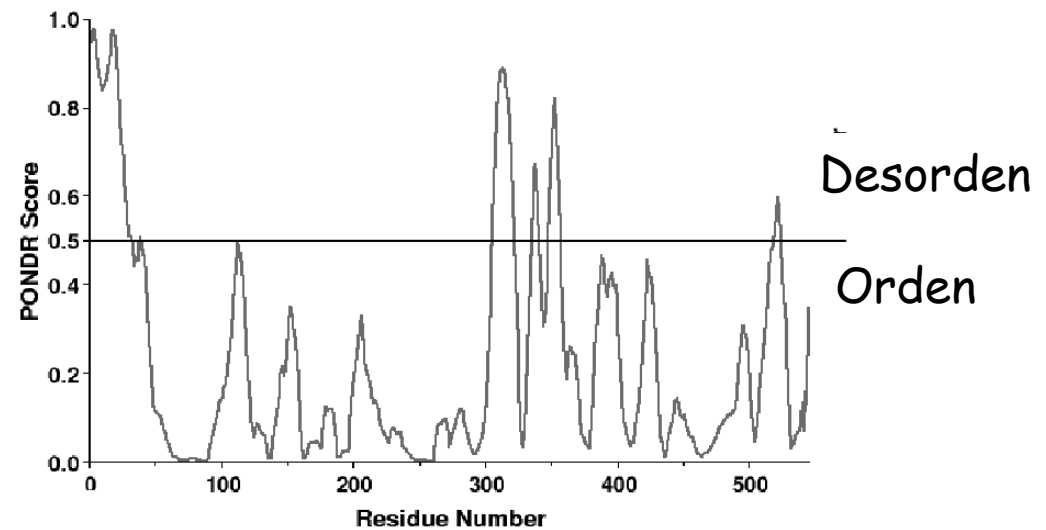
Proteína A

% Desorden = 53.96



Proteína B

% Desorden = 12.66



2. Predicción del contenido de estructura secundaria usando el servidor JPred3

```
1                                     40
MATLIYVDKENGEPEGTRVVAKDGLKLGSGPSIKALDGRSQ
-EEEEEEE-----EE-----EEE

41                                     80
VSTPRFGKTFDAPPALPKATRKALGTVNRATEKSVKTKGP
E-----HHHHHH-----

81                                     120
LKQKQPSFSAKMTEKTVKAKSSVPASDDAYPEIEKFFPF
-----EEE-----HHH----

121                                    160
NPLDFESFDLP EEHQIAHLPLSGVPLMILDEERELEKLFQ
-----EEE-----EEEE-HHHHHHHHHH

161                                    200
LGPPSPVKMPSPPWESNLLQSPSSILSTLDVELPPVCCDI
-----HHHHHHHH-----EEEE-

201
DI
--
```

Proteína A se predice un 28%
de estructura al azar

Proteína B se predice un
52% de estructura
secundaria

Si la estructura de una proteína está codificada en su secuencia de amino ácidos, entonces igualmente debe ocurrir para la falta de estructura.

Residuos promotores de **desorden**

Arg, Lys, Glu, Pro y Ser

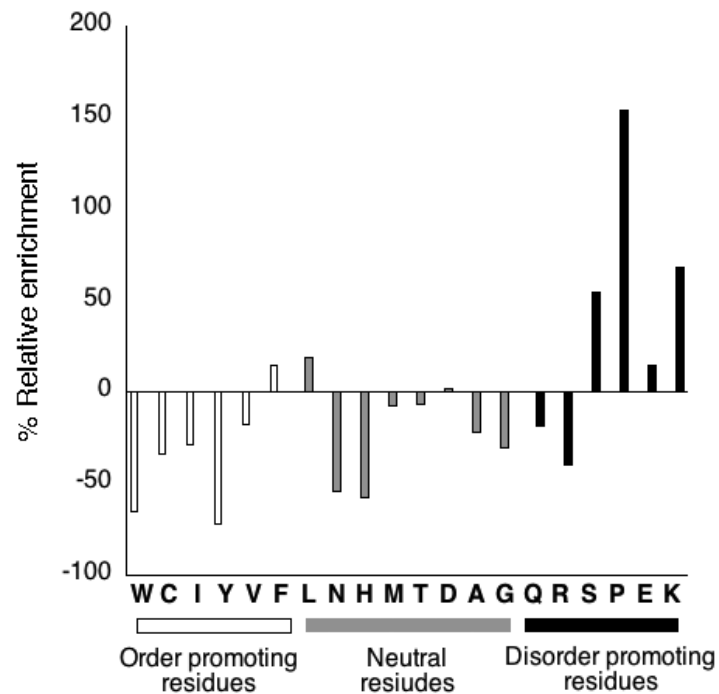
Residuos promotores de **orden**

Cys, Trp, Tyr, Ile y Val

Dunker, A.K., et. l. *J Mol Graph Model* (2001) 19:26-59
Williams, R.M., et. Al. *Pac Symp Biocomput* (2001) 89-100

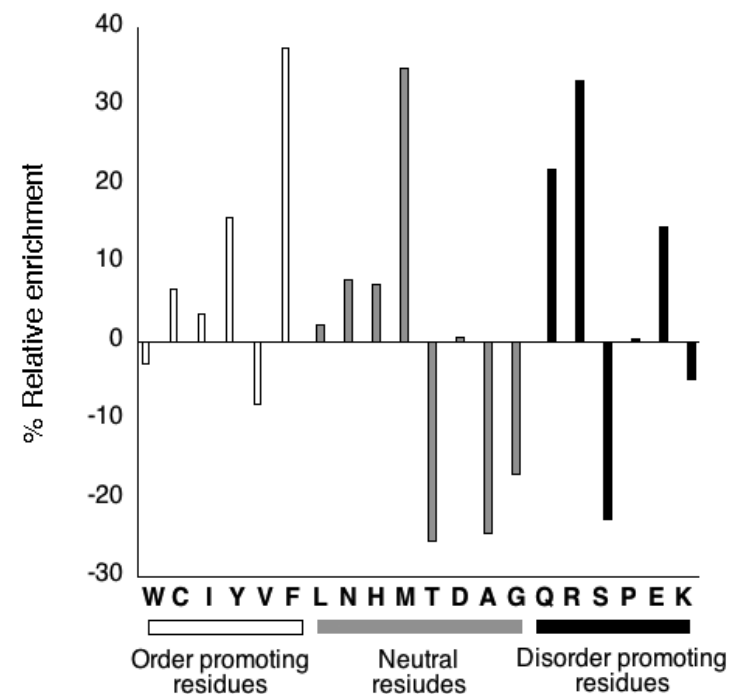
3. Análisis de la desviación en la composición de aminoácidos respecto a aquella de proteínas en la base de datos SCOP

A



Proteína A contiene 2 veces más residuos promotores de desorden Pro, Lys and Ser

B



Proteín B rica en residuos promotores de orden.

4. Valores de hidrofobicidad media y carga neta media

Proteína A

Hidrofobicidad media

$$\langle H \rangle = 0.42$$

Carga neta media

$$\langle R \rangle = 0.091$$

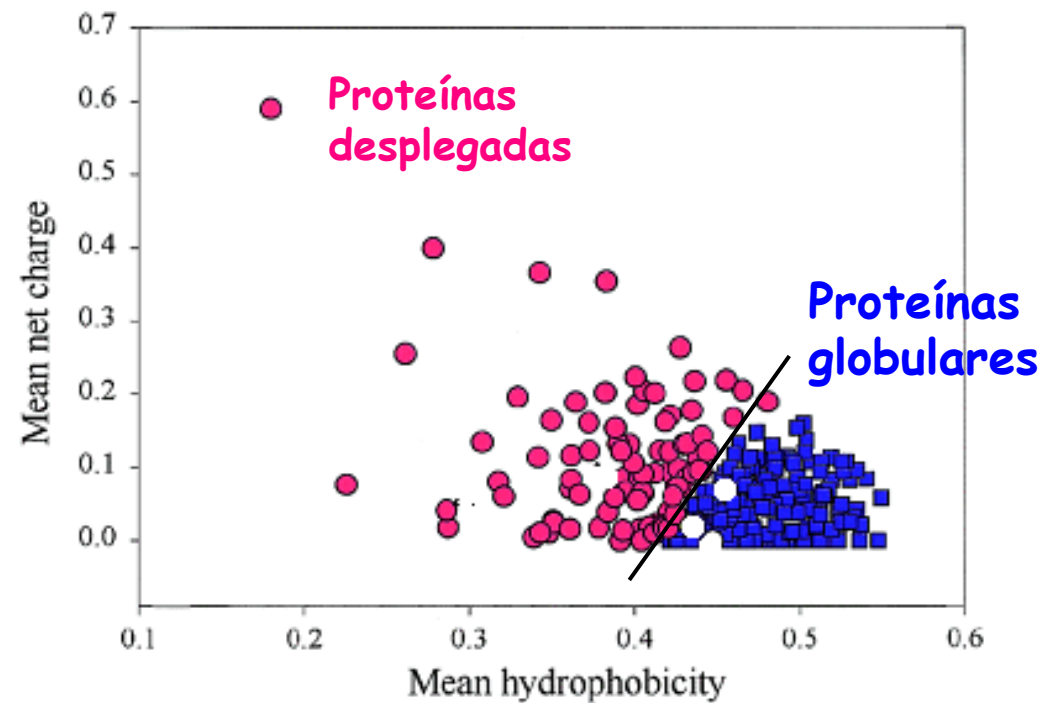
Proteína B

Hidrofobicidad media

$$\langle H \rangle = 0.45$$

Carga neta media

$$\langle R \rangle = 0.0091$$



$\langle R \rangle \uparrow$ mayor repulsión entre aminoácidos cargados

diferencia absoluta entre el número de residuos cargados positiva y negativamente dividido entre el total de residuos.

$\langle H \rangle \downarrow$ no favorece la formación de un núcleo hidrofóbico que pliega a la proteína

es la suma de las hidropatías normalizados de cada residuo dividido por el número total de residuos en la proteína.

Proteína A

Carece de
estructura
secundaria y
terciaria

Apunta a ser una
proteína
intrínsecamente
desestructurada

Proteína B

Tiene estructura
secundaria y terciaria

Esta plegada

Tiene una conformación
globular

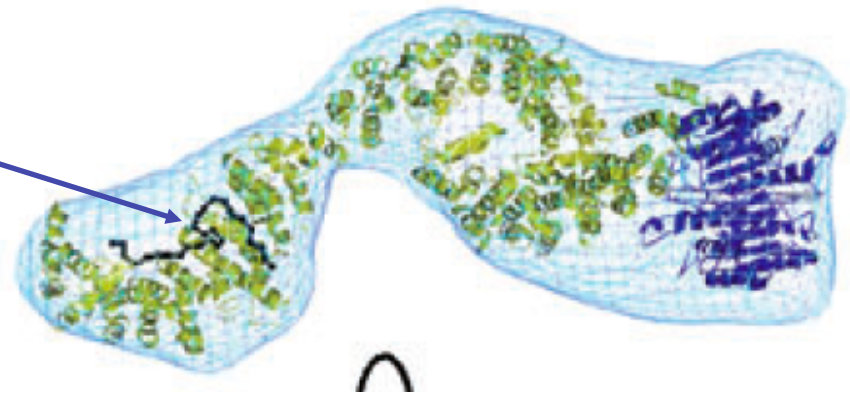
Proteína A - Securina

Es una proteína humana intrínsecamente desordenada que participa en la regulación del ciclo celular en la transición de metafase-anafase. Su sobreexpresión promueve la formación de tumores.

Sánchez-Puig, N. et. al., Prot Sci, 2005. Csizmok, V. et. al., JACS, 2008

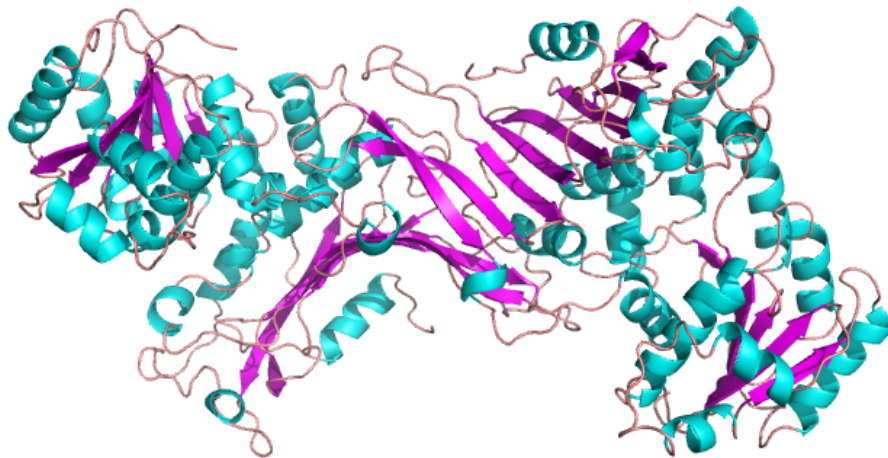
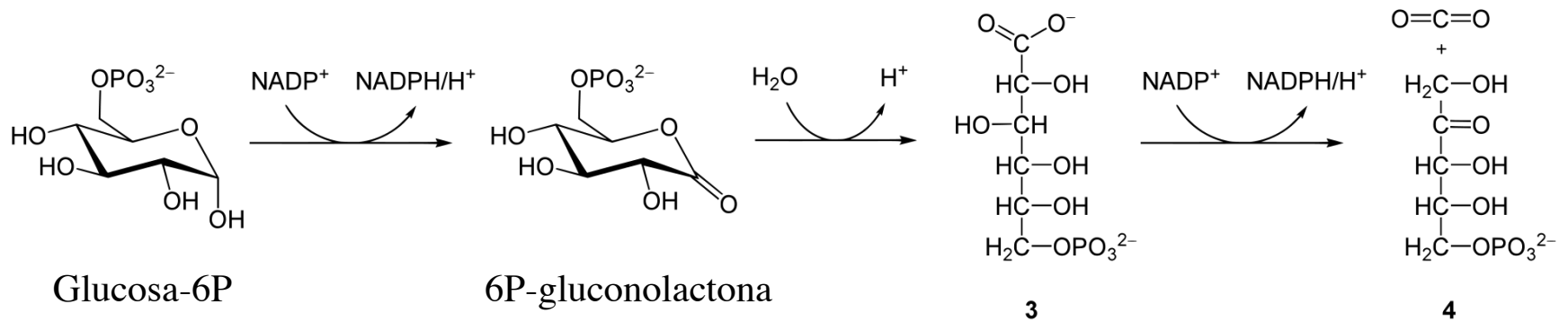
Modelo de unión de **securina**
a **separase**

Viadiu, H. et. al., Nat Struc Mol
Biol, 2005



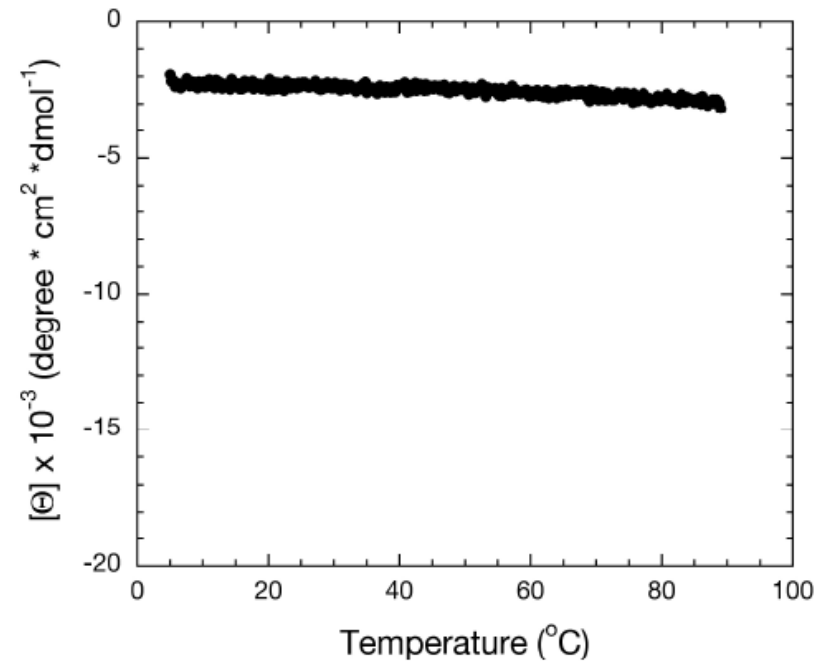
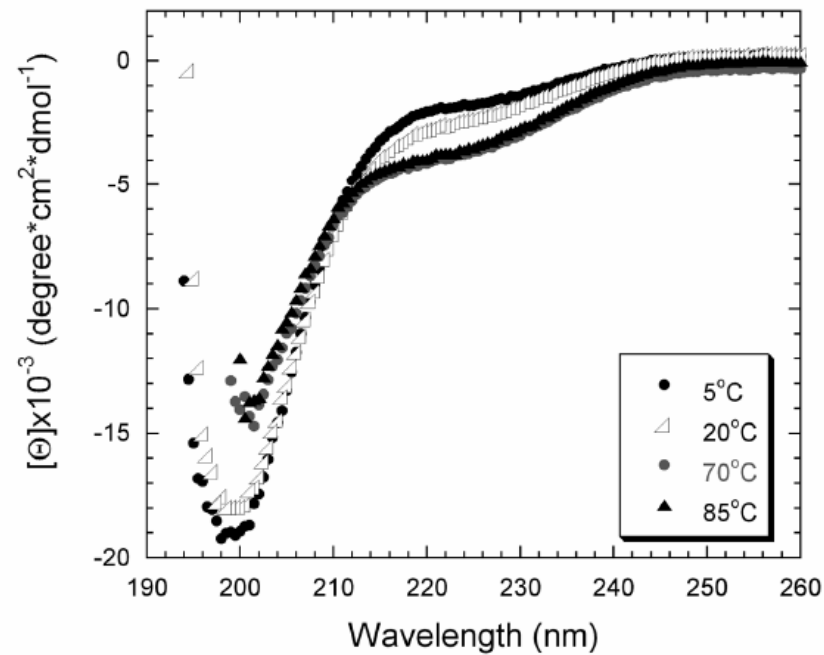
Proteína B - Glucosa 6-fosfato deshidrogenasa

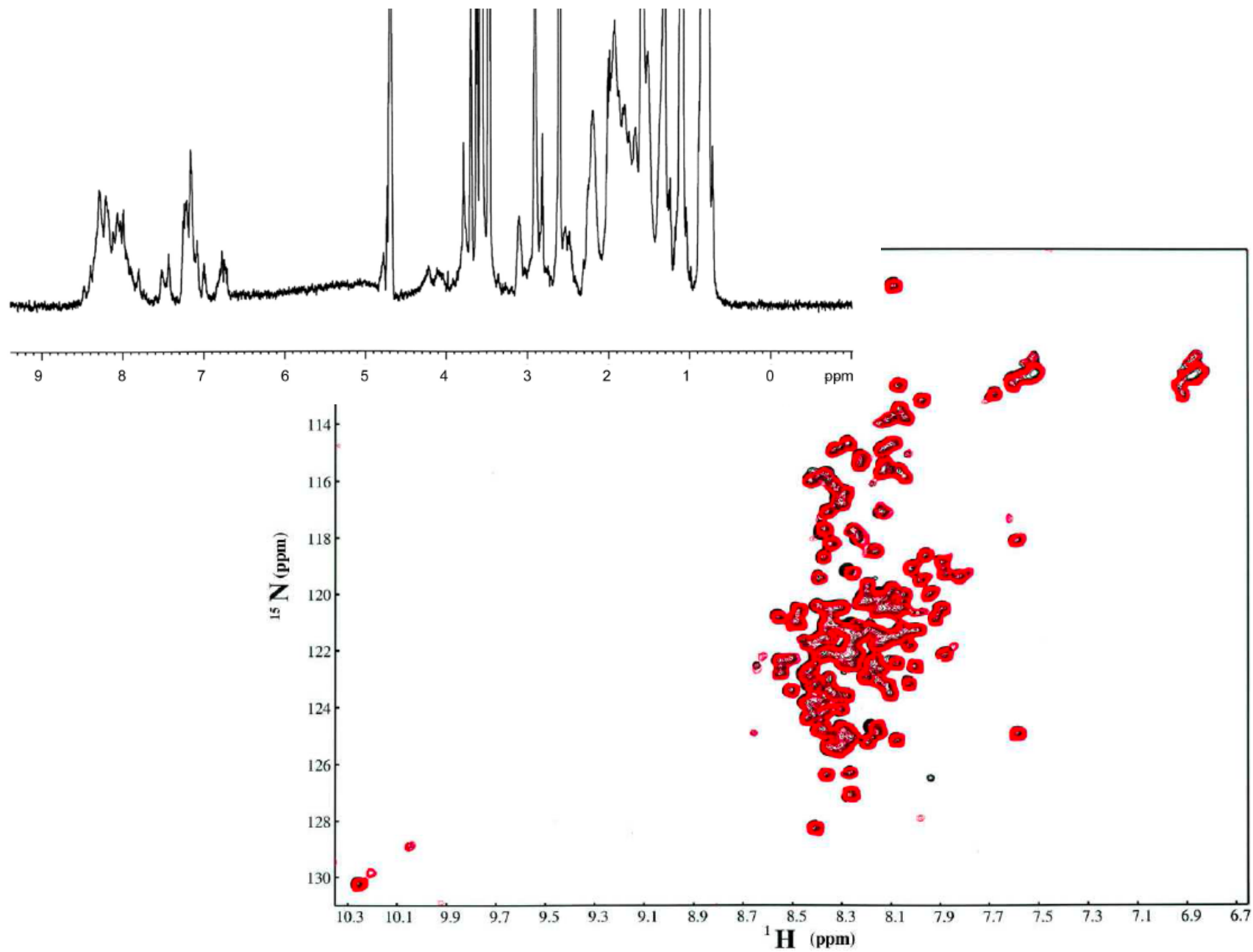
Una enzima que cataliza el primer paso en el ciclo de las pentosas y de quien ya existe su estructura cristalográfica.



PDB: 2BHL

Cómo se ve una proteína intrínsecamente desestructurada





¿Por qué es importante saber si tu proteína es un proteína intrínsecamente desestructurada antes de cualquier experimento?

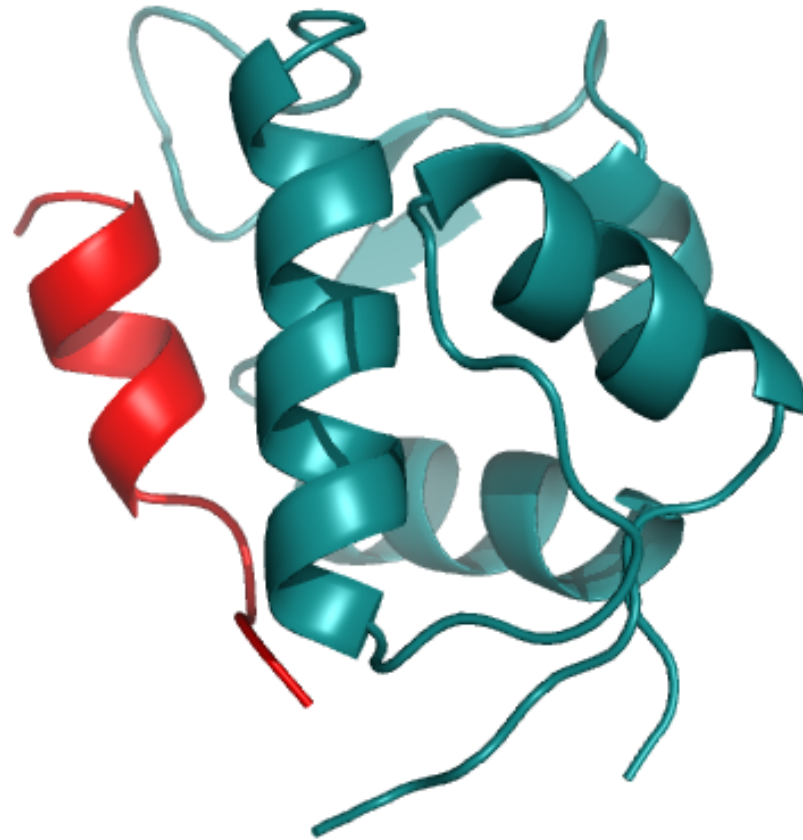
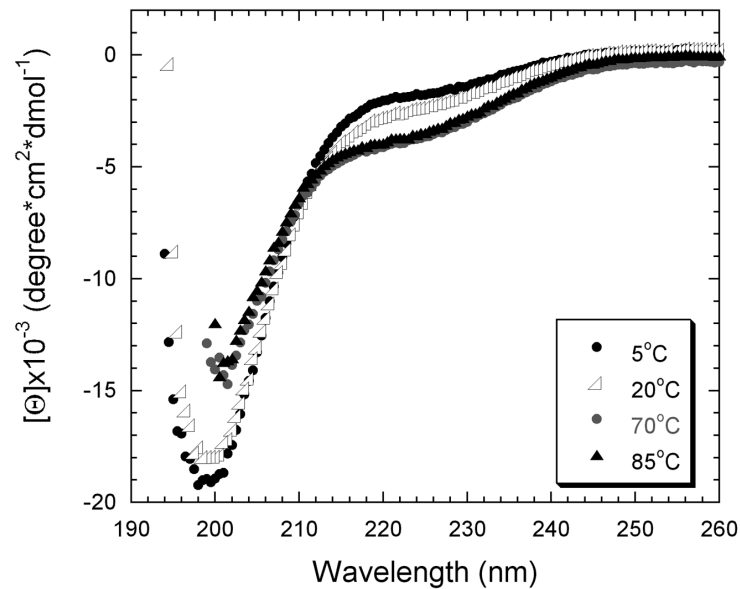
**Cristalizar proteínas intrínsecamente
desestructuradas
ó**

**Proteínas con fragmentos largos desordenados
son muy difíciles de cristalizar pero no imposibles!!**

¿Qué se puede hacer para sobrepasar este problema?
**Cristalizar dicha proteína en presencia de una
biomolécula e.g. Proteína con quien interaccione.**

p53 N-TAD

Dominio intrínsecamente
desestructurado



PDB: 1YCR

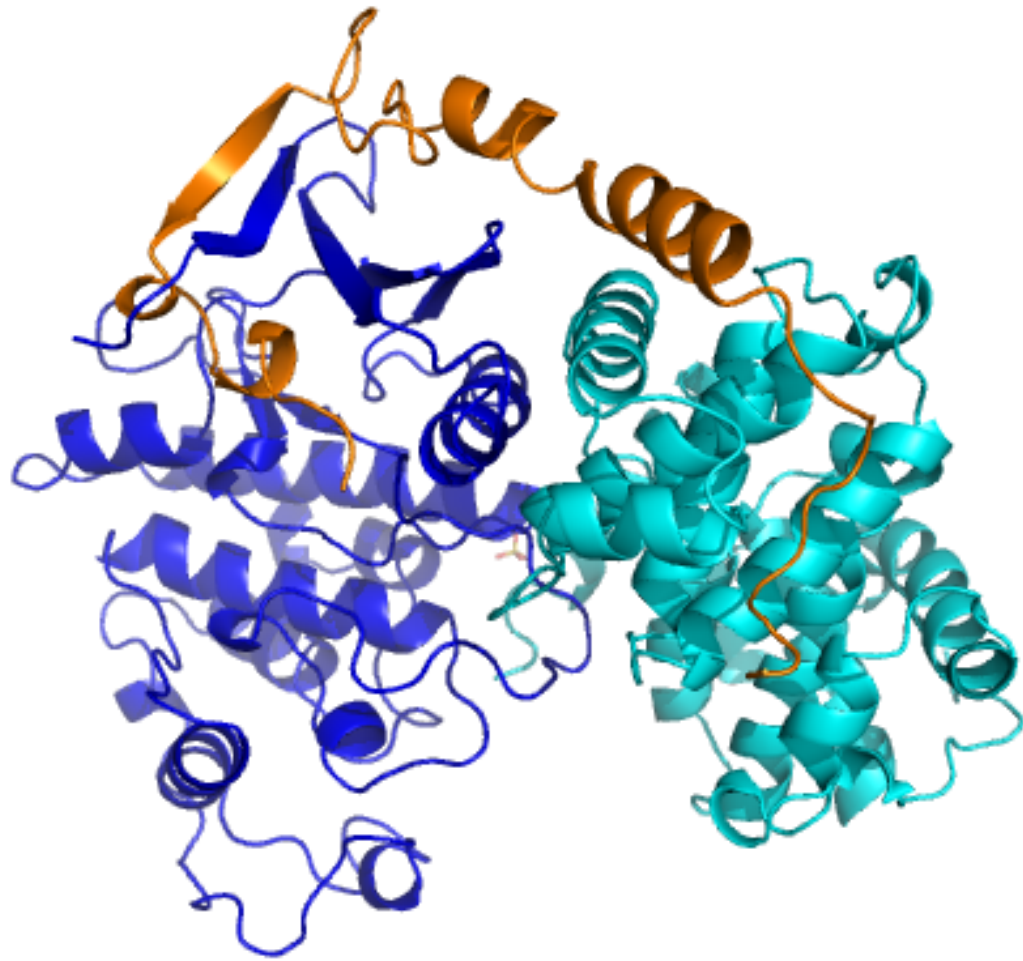
Mdm2 dominio
N-terminal

Kussie, P.H., et. al., (1996) Science **274**: 948-953

P27 (23-105aa)

Cdk2

Cyclin A



Russo, A.A., et al., *Nature* (1996) 382: 325-331
Pdb: 1JSU

nuclein

Calsequestrin-1

p65

DisProt: Base de datos de proteínas
desordenadas

Número de proteínas desordenadas - 645

Número de regiones desordenadas - 1388

like

Caldesmon

factor

subunit 1

se

El número de desorden en las proteínas es aumenta con la complejidad de los organismos.

Genoma	% Desorden
Eucariontes	25-40%
Archaea	21%
Bacteria	1-8%

La idea de que “La estructura de una proteína es fundamental para su función” solo es la mitad de la historia.

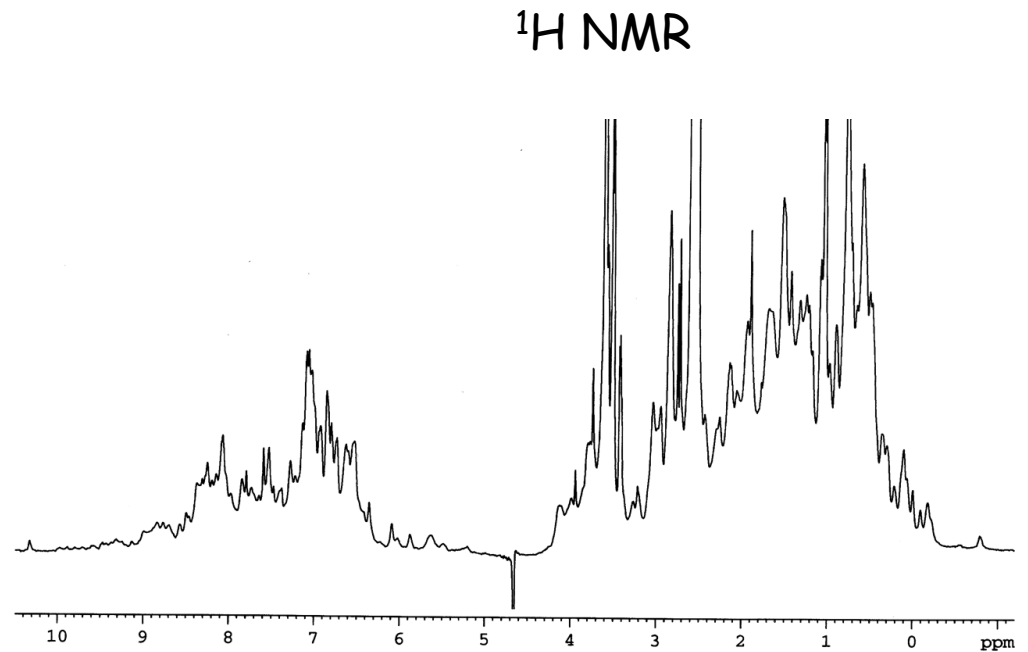
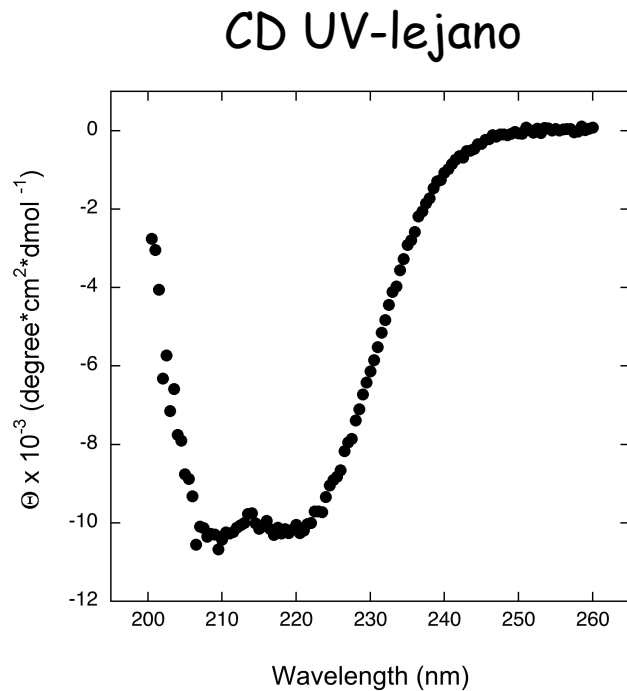
Existen muchas proteínas carentes de una estructura tridimensional definida y son funcionales.

Sin embargo....

Tener una proteína que este plegada y tenga estructura globular tampoco es garantía que cristalice

... que cristalice fácilmente

Proteína Bcp1



La proteína tiene estructura secundaria (hélices α) y estructura terciaria

Cristalización de Bcp1 en batch > 1500 condiciones
usando concentraciones 5 mg/ml y 10 mg/ml



Precipitados amorfos y algunos cristales de sal

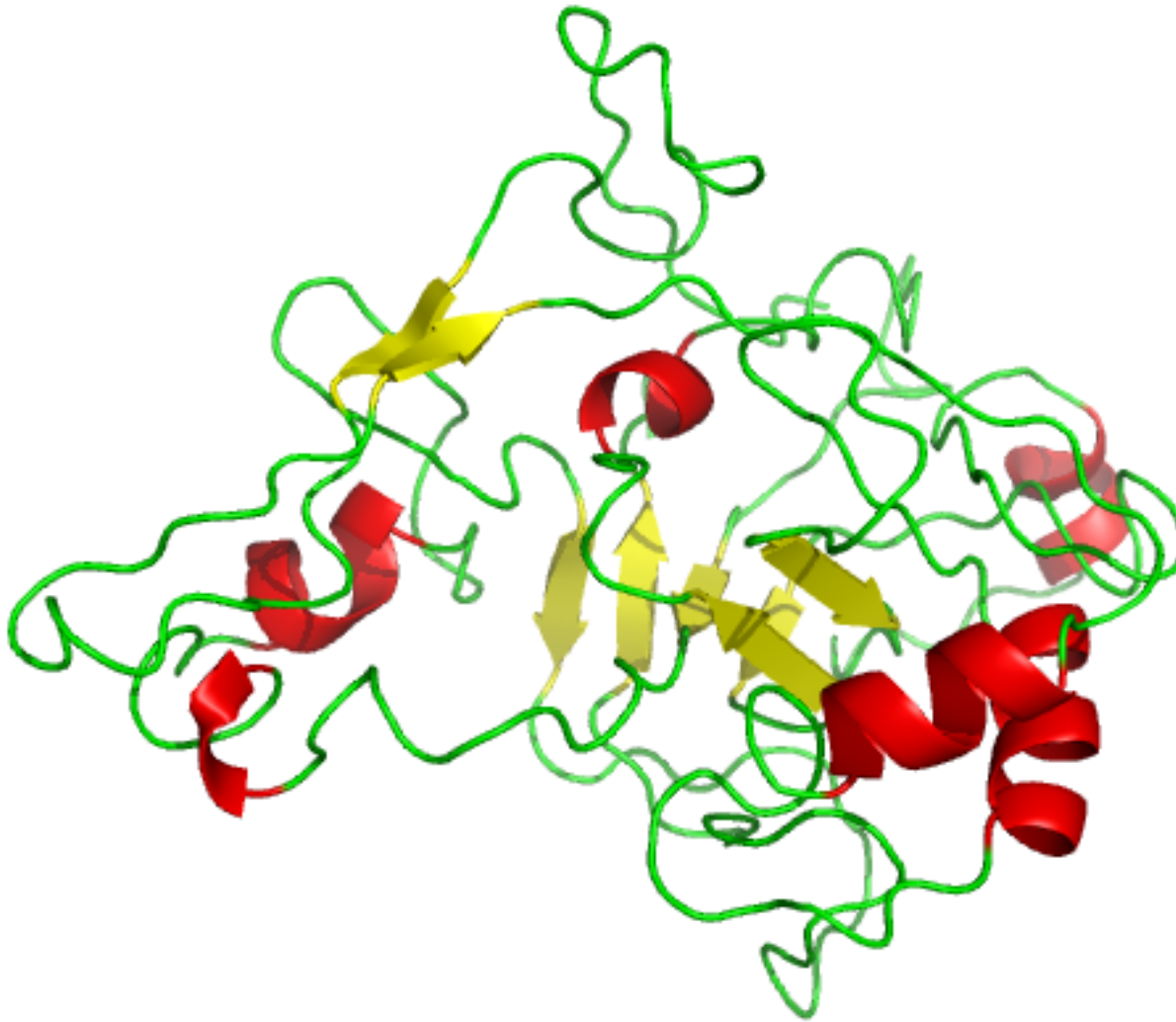
Programas para predecir la factibilidad de que una proteína cristalice

Programa	Dirección URL
SVMCRYST	http://www3.ntu.edu.sg/home/EPNSugan/index_files/svmcrys.htm
MetaPPCP	No disponible
CRYSTALP2	http://biomine-ws.ece.ualberta.ca/CRYSTALP2.html
CRYSTALP	No disponible
SECRET	http://mips.helmholtz-muenchen.de/secret/secret.seam
XtalPred	http://ffas.burnham.org/XtalPred-cgi/xtal.pl
ParCrys	http://www.compbio.dundee.ac.uk/xtal/cgi-bin/input.pl
OB-score	http://www.compbio.dundee.ac.uk/obscore/
XANNpred	http://www.compbio.dundee.ac.uk/xtal/cgi-bin/xannpred_in.pl

Predicción de cristalizabilidad de Bcp1

Ortólogos de Bcp1	XtalPred Score
<i>Homo sapiens</i>	Very difficult
<i>Mus musculus</i>	Difficult
<i>Gallus gallus</i>	Difficult
<i>Drosophila melanogaster</i>	Difficult
<i>Arabidopsis thaliana</i>	Difficult
<i>Anopheles gambiae</i>	Difficult
<i>Danio rerio</i>	Difficult
<i>Candida glabrata</i>	Very difficult
<i>Schizosaccharomyces pombe</i>	Difficult
<i>Saccharomyces cerevisiae</i>	Very difficult

Modelo estructural de Bcp1 realizado en Rosetta



Dr. Rogelio Rodríguez Sotres

A nivel de la proteína ¿Qué se puede hacer para mejorar la posibilidad de que cristalice?

Ingeniería de proteínas

- Remover regiones desestructuradas
- Disminuir la heterogeneidad
eliminar modificaciones post-traduccionales
oligomeros
- Identificar dominios estructurales
- Utilizar proteínas acarreadoras
proteínas de fusión
- Mejorar el empaquetamiento cristalino
introducir mutaciones para disminuir la entropía
de la superficie

Conclusiones

Análisis de la secuencia primaria de una proteína provee una gran cantidad de información útil acerca de la proteína en cuestión y ayuda a la toma de decisiones en un proyecto de investigación.

¡¡solo toma un par de horas hacerlo y no cuesta!!

Proteínas intrínsecamente desestructuradas

Proteínas que se caracterizan por la total ausencia de estructura rígida y compacta

Otros nombres:

Nativamente desnaturalizadas,
Nativamente desordenadas,
intrínsecamente desordenadas.

Función de las proteínas intrínsecamente desordenadas

- Transducción de señales - frecuentes en sitios con modificaciones postraduccionales.
- Regulación del ciclo celular
- Replicación
- Expresión genética - transcripción y traducción, interacción con DNA y RNA
- Ensamblaje de estructuras celulares

- Implicadas en desarrollo de enfermedades neurodegenerativas como la enfermedad de Alzheimer, el Síndrome de Down, la Distrofía miotónica y la enfermedad de Parkinson.

Características de las proteínas intrínsecamente desordenadas

- Tamaño variable 50 -1,000 residuos.
- Baja complejidad en la secuencia primaria.
- Tienen varias proteínas blanco con las que interaccionan.
- Algunas presentan transiciones del estado desplegado al plegado.

Intrínseco - que le es propio

Nativo - natural ó innato

Desnaturalizada - desordenada

Conjunto de conformaciones no rígidas de proteínas entre las que se incluyen conformaciones parcialmente plegadas e.g globulo fundido.

Desplegada - desestructurada

Proteínas desordenadas que carecen de estructura rígida