

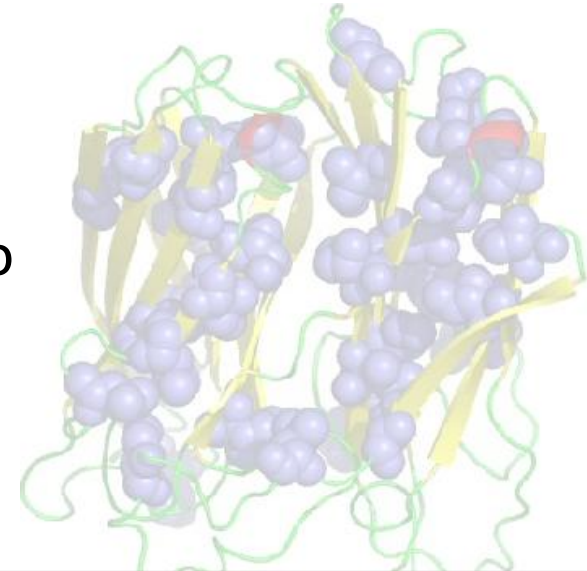
# Problema 9: Diseño de Proteínas

Tomado de: *The Ten Most Wanted Solutions In Protein Bioinformatics* por Anna Tramontano

## Seminario de Bioinformática

Jorge Hernán Victoria Moreno

Octubre 23 de 2009

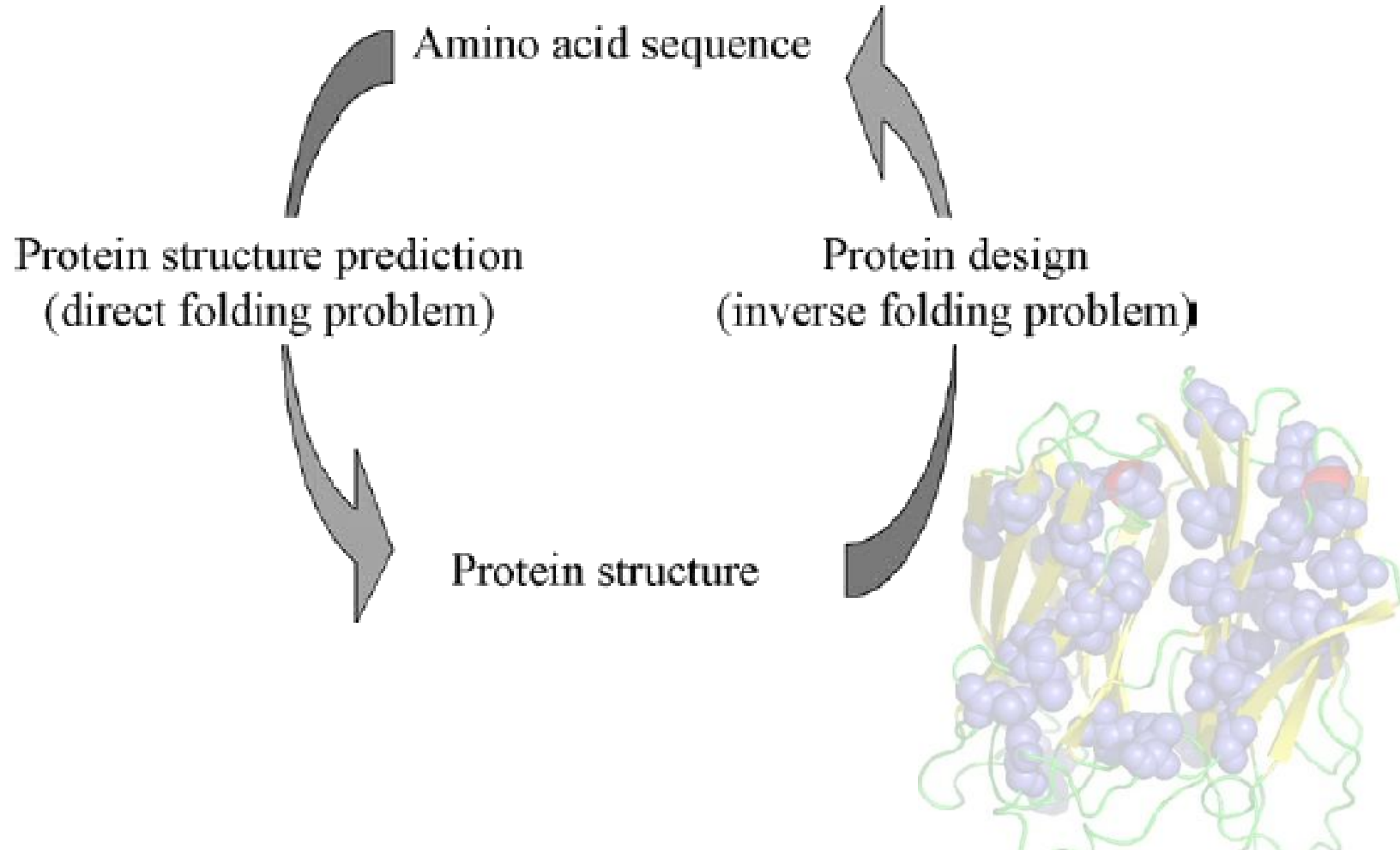


# Agenda

- Problema del Plegamiento Directo e Inverso.
- Evolución estructural de las proteínas.
- Dificultades.
- Interacciones en la proteínas.
- Arquitecturas usadas en el diseño de proteínas.
- Protocolo utilizado con éxito para el diseño de proteínas.
- Consideraciones Finales.

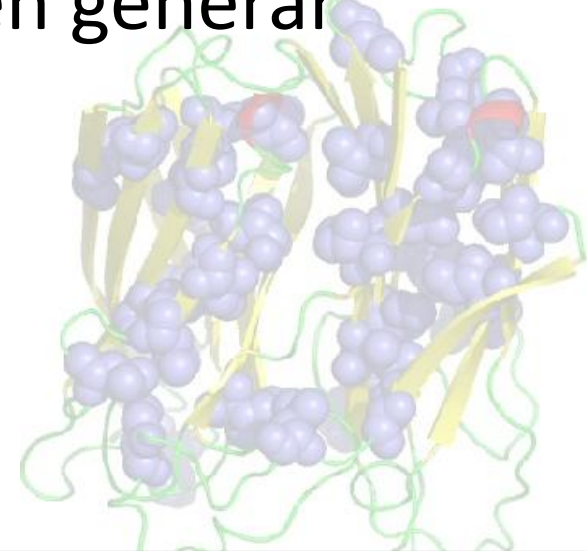


# Problema del Plegamiento Directo e Inverso



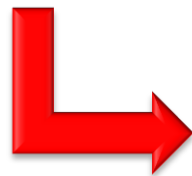
# Evolución estructural de las proteínas

- Se ha observado que a través de la evolución las proteínas que tienen relación han preservado su estructura.
- Por otra parte, también se ha observado que proteínas no relacionadas pueden generar una estructura similar.



# Evolución estructural de las proteínas

- Se ha observado que a través de la evolución las proteínas que tienen relación han preservado su estructura.
- Por otra parte, también se ha observado que proteínas no relacionadas pueden generar una estructura similar.

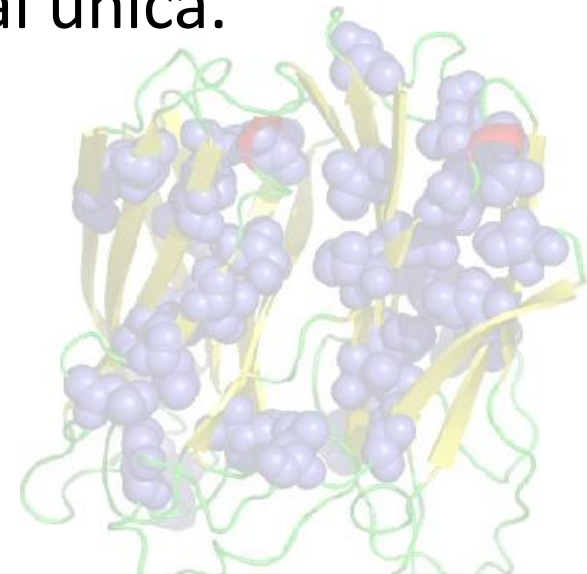


**TEORÍA SECUENCIA –ESTRUCTURA EN RIESGO!**



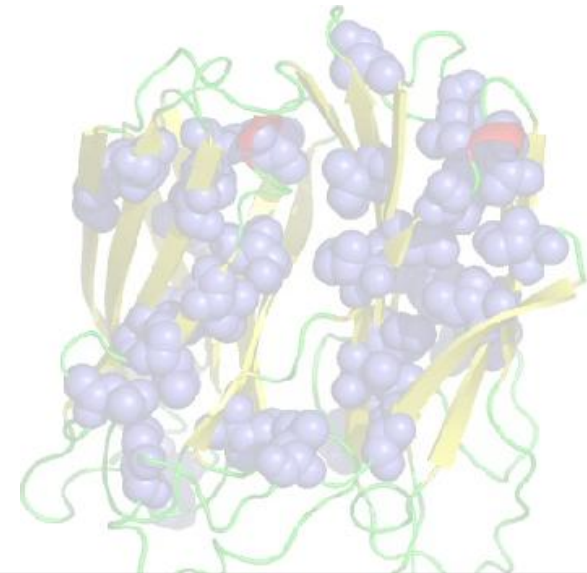
# Oportunidad de seleccionar una secuencia que se pliegue y forme una determinada estructura

- Para una secuencia de 100 aminoácidos.
  - $20^{100}$  secuencias posibles.
  - 1 de  $10^4$  secuencias es capaz de plegarse teniendo una energía mínima global única.



# Más dificultades....

1. Calcular la energía para una determinada conformación.
2. Evaluación de los resultados.



# Algunas interacciones en la proteínas

- Hidrofóbicas
- Hidrofílicas
- Enlaces de Hidrogeno
- Electroestáticas

Un simple patrón de residuos hidrofóbicos e hidrofílicos pueden ser especificados Para la conformación de una hélice  $\alpha$ .



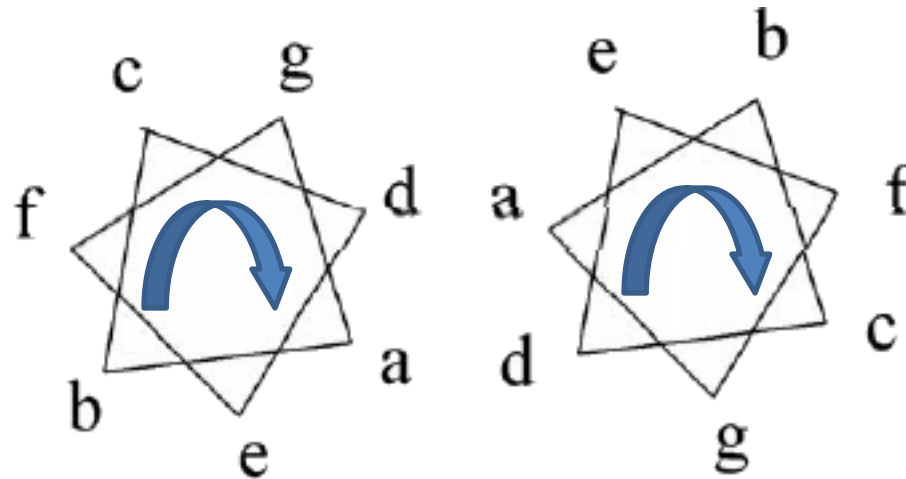
# Arquitecturas Usadas en el Diseño de Proteínas

- Parallel/Antiparallel Helix Bundle
- TIM(Triosephosphate isomerase) Barrel.

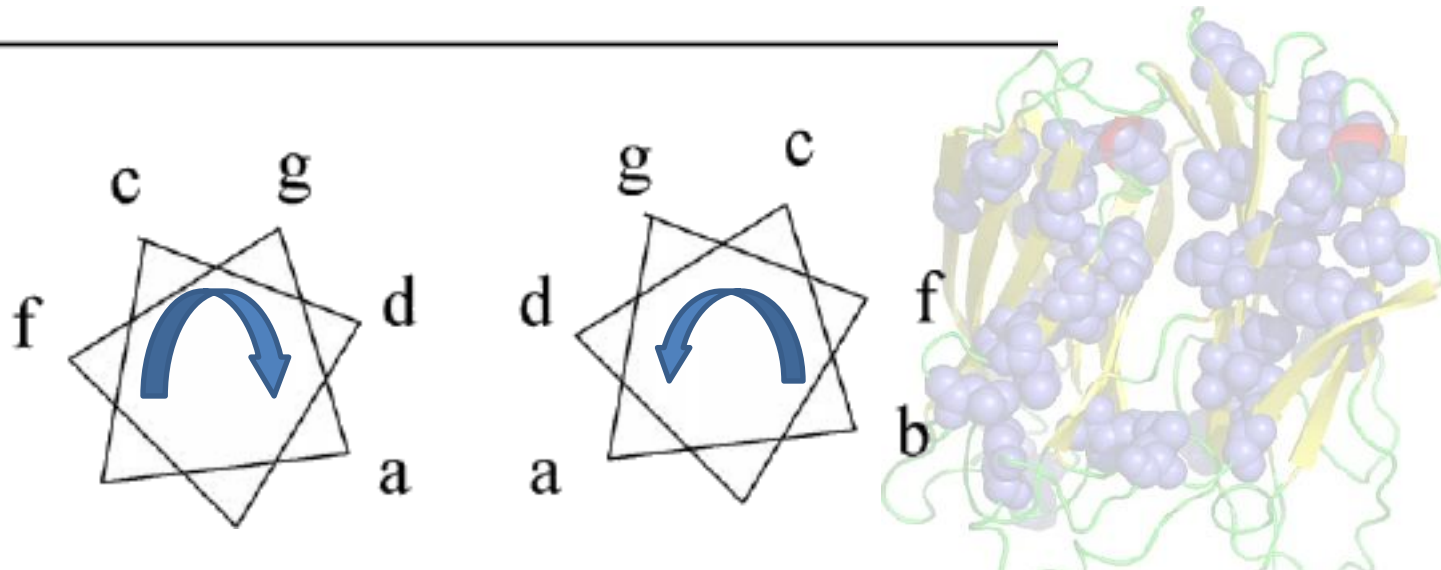
En 1986 Chris Sander organizó un curso de diseño de proteínas. Científicos reconocidos se dividieron en grupos y cada grupo trabajó en distintos proyectos de diseño de proteínas. Esto se repitió en 1990 y 1994.

# Parallel and Antiparallel Helices

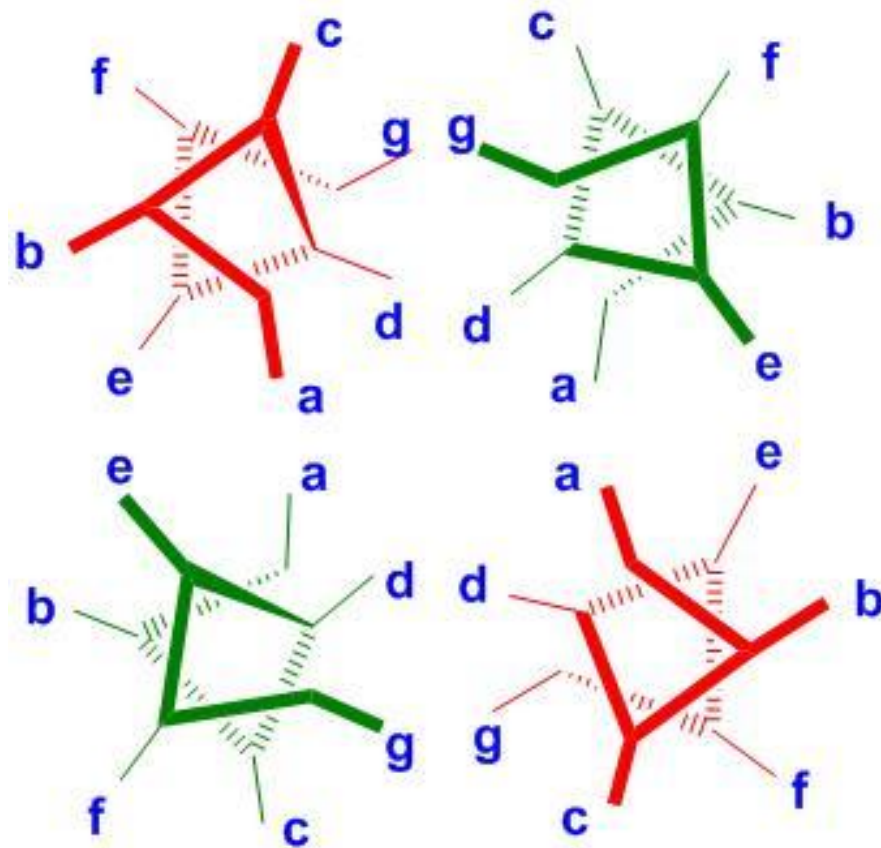
Paralelas



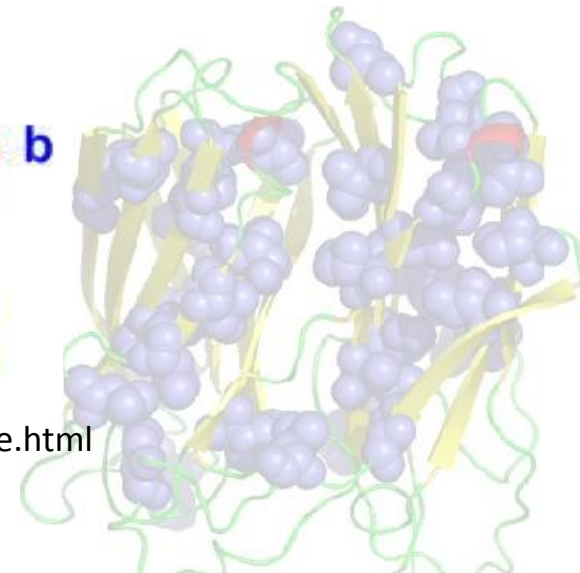
Anti paralelas



# Four-Helix Antiparallel Bundle(1)



<http://www.umich.edu/~marshlab/FluorousProteinsFrame.html>

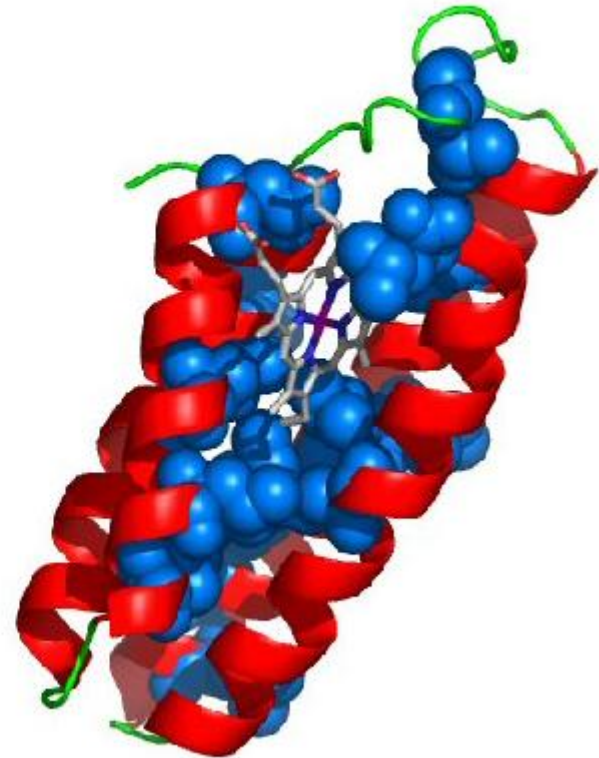


# Four-Helix Antiparallel Bundle(2)

Four-Helix bundle in Human Growth Hormone (1hgu)



Four-Helix bundle in Cytochrome b (1lm3)

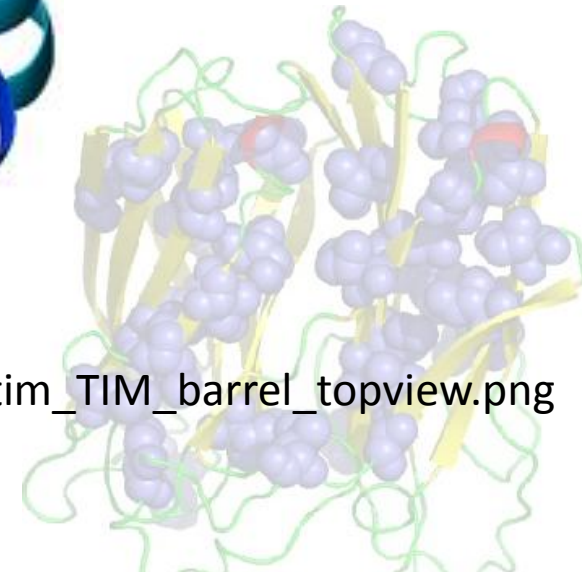


<http://chemistry.umeche.maine.edu/CHY431/Proteins9.html>

# TIM Barrel

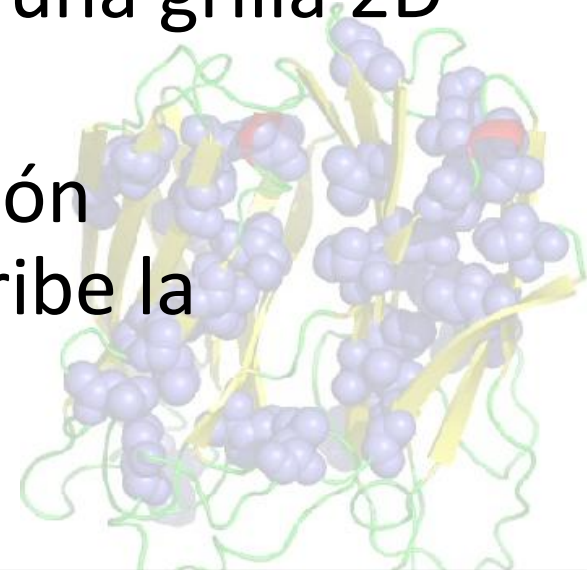


[http://en.wikipedia.org/wiki/File:8tim\\_TIM\\_barrel\\_topview.png](http://en.wikipedia.org/wiki/File:8tim_TIM_barrel_topview.png)



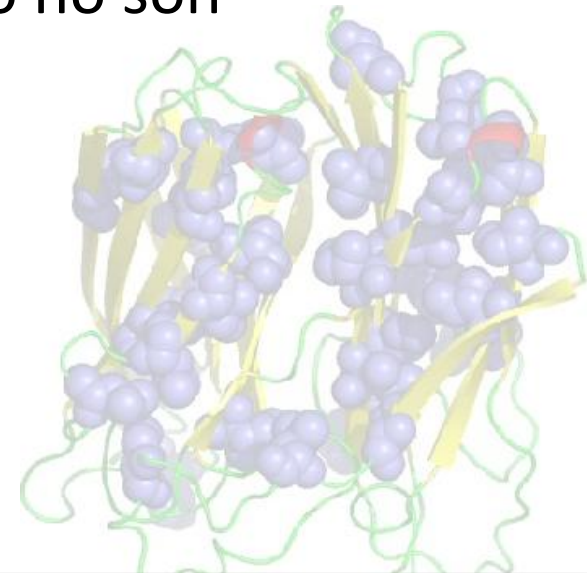
# Lattice Model

- Método extremadamente simplificado y automático para crear modelos no realistas de proteínas.
- Relaciones secuencia-estructura.
- La estructura es representada en una grilla 2D o 3D.
- El objetivo es optimizar una función simplificada de energía que describe la interacción entre símbolos.

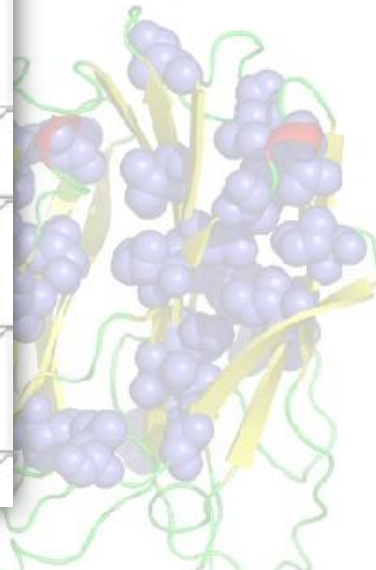
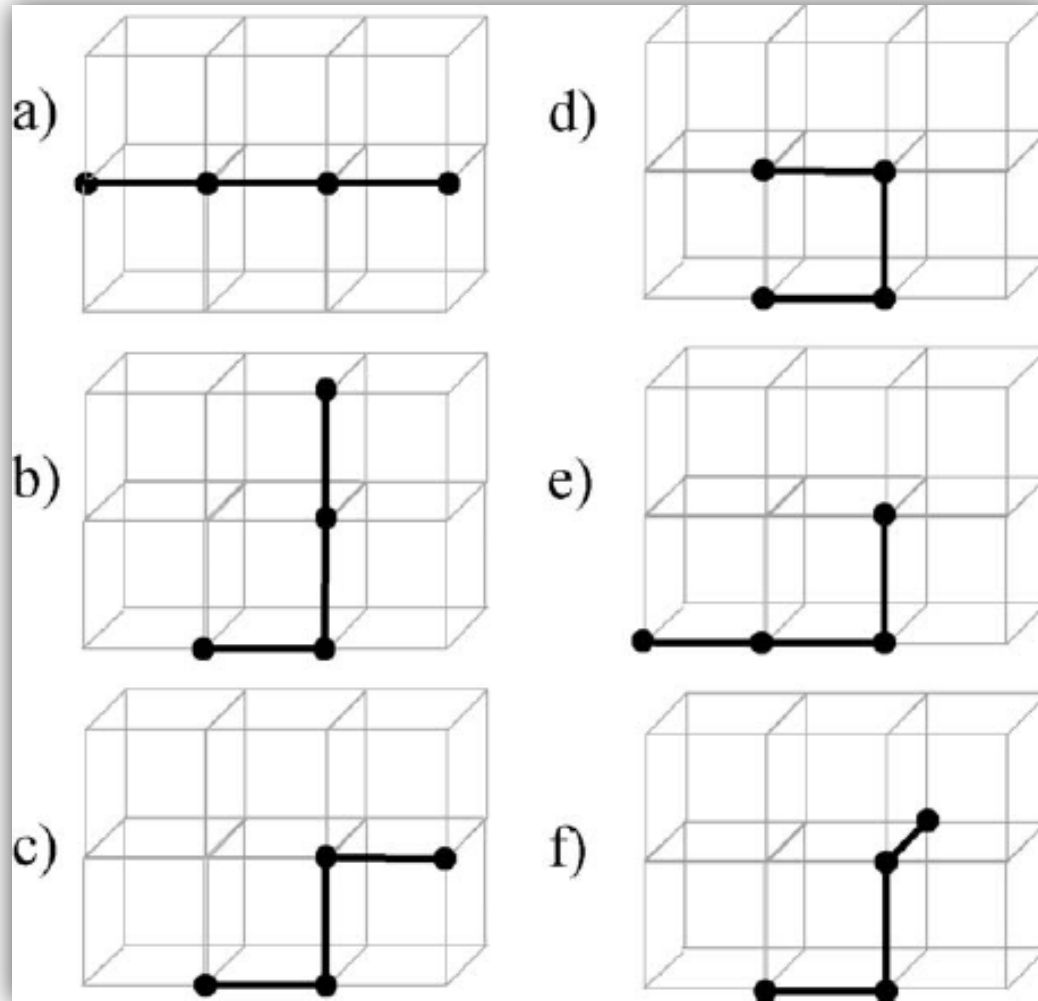


# Ejemplo: Lattice Model(1)

- Tenemos dos tipos de aminoácidos:
  - P : Polares
  - H: Hidrofóbicos
- Función de energía:
  - Dos H interaccionan favorablemente si tienen posiciones adyacentes en la grilla pero no son consecutivos en la secuencia.
    - H-H Favorable = -1
    - H-H No Favorable = 0
  - La interacción H-P y P-P es neutral.
    - H - P = 0
    - P - P = 0



# Ejemplo: Lattice Model(2)

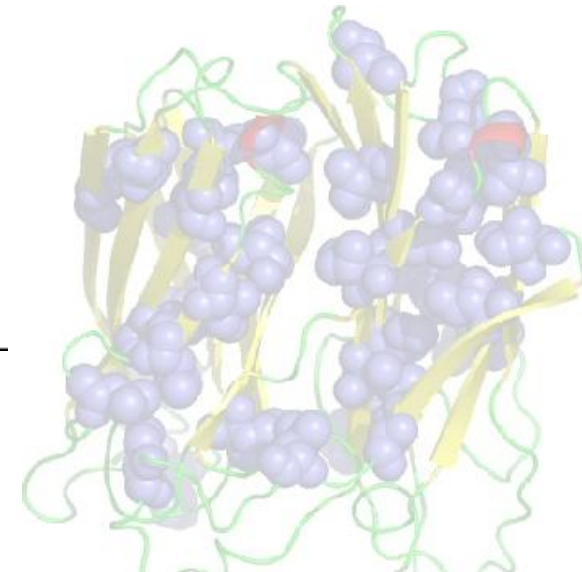
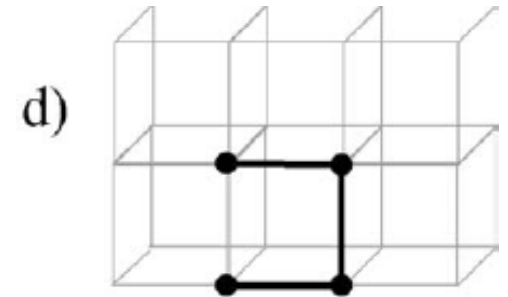


H-H Favorable = -1  
H-H No Favorable = 0  
H - P = 0  
P - P = 0



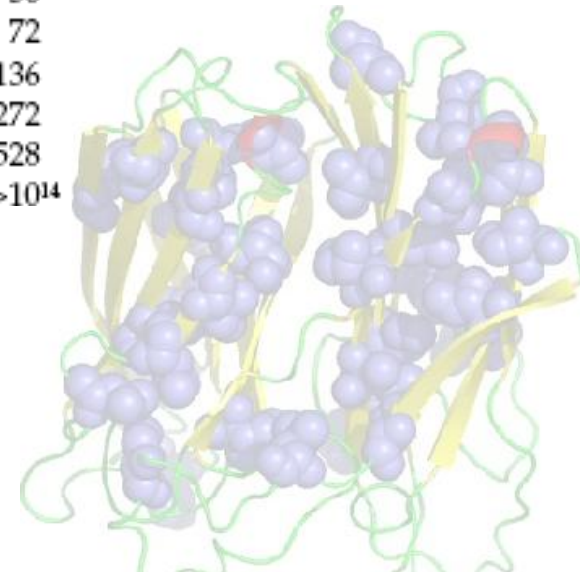
# Ejemplo: Lattice Model(3)

Position 1	Position 2	Position 3	Position 4	Energy in Conformation $e$
P	P	P	P	0
P	P	P	h	0
P	P	h	p	0
P	P	h	h	0
P	h	p	h	0
P	h	h	p	0
P	h	h	h	0
h	p	p	h	-1
h	p	h	h	-1
h	h	h	h	-1
p	h	p	p	0
h	p	p	p	0
h	p	h	p	0
h	h	p	p	0
h	h	p	h	-1
h	h	h	p	0



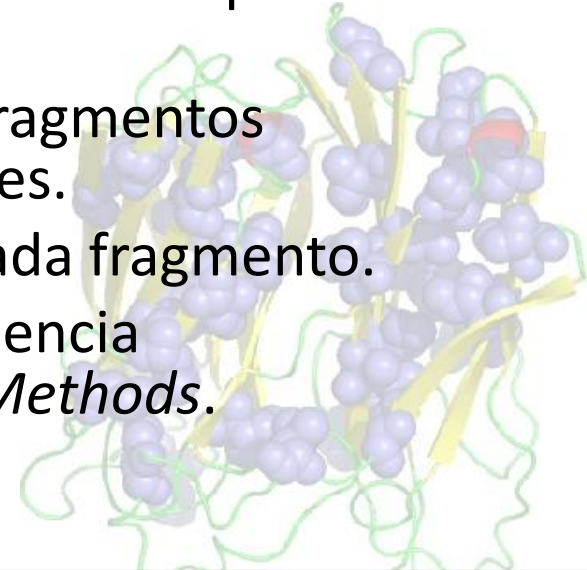
# Posibles conformaciones de Secuencias de H y P en un espacio 3D para diferentes tamaños de secuencia

Number of Elements	Number of Sequences ( $2^N$ )	Number of Possible Paths in a 3D Grid
4	16	10
5	32	20
6	64	36
7	128	72
8	256	136
9	512	272
10	1024	528
50	$>10^{15}$	$>10^{14}$



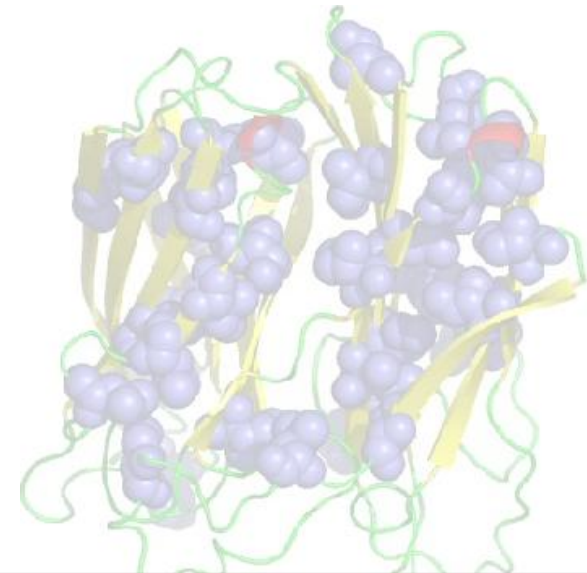
# Protocolo utilizado con éxito para el diseño de proteínas

- El siguiente protocolo ha dado muy buenos resultados y se repite cíclicamente siguiendo los siguientes pasos:
  1. Seleccionar una topología inicial de una estructura objetivo.
  2. Definir un conjunto de restricciones espaciales que define el plegamiento deseado.
  3. Generar conformaciones iniciales por fragmentos basándose en las restricciones espaciales.
  4. Aplicar técnicas de Monte Carlo para cada fragmento.
  5. Predicción de la estructura de una secuencia seleccionada usando *fragment-based Methods*.



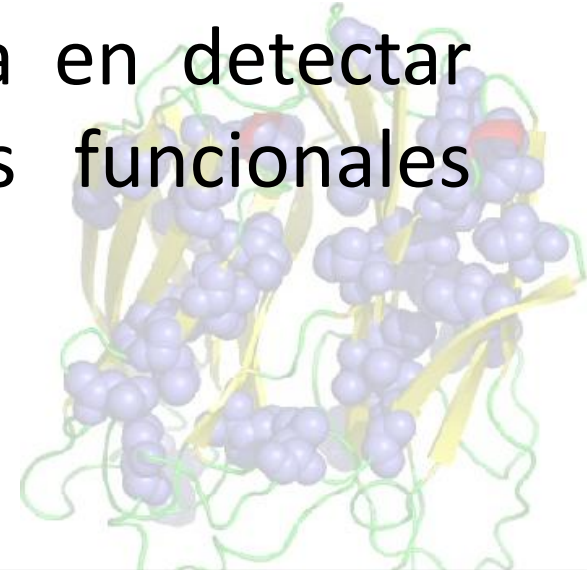
# Consideraciones Finales(1)

- Los métodos de Lattice nos ofrecen una vista general del problema, pero ayudan a entender reglas básicas del espacio de las proteínas.
- Debido a que un objetivo es la minimización de la función de energía, se han utilizado técnicas de optimización como:
  - Técnicas de Monte Carlo
  - Algoritmos Genéticos
  - *Dead-End Elimination* (DEE)



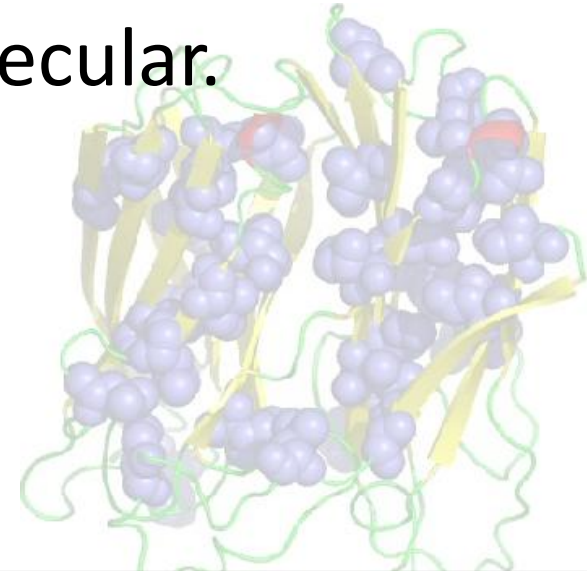
# Consideraciones Finales(2)

- El aspecto más difícil a tener en cuenta, es la desestabilización de plegamiento alternativos.
- El entendimiento de la evolución de proteínas es una herramienta que contribuye en asignar estructura y función de proteínas desconocidas. Pero el problema en detectar regiones conservadas de sitios funcionales aun persiste.



# Consideraciones Finales(3)

- Para el avance en diseño de proteínas se necesita un trabajo colaborativo entre teoría y experimentación.
- Los casos de éxito en diseño de proteínas se evidencian en la interacción entre la biología computacional, estructural y molecular.



Gracias!

