

Predicción de Estructuras Cristalinas de Interés Farmacológico vía Algoritmos Genéticos

DDF 2015

6 de julio de 2015

Gabriel I. Pagola

Departamento de física, FCEN,
UBA.

gpagola@df.uba.ar



Departamento
de física.
FCEN-UBA



Utah, US



Grupo de trabajo

Predicción de estructuras cristalinas (Crystal Structure Prediction ,CSP)

Gabriel Pagola, Marta Ferraro Depto de Física, FCEN, UBA

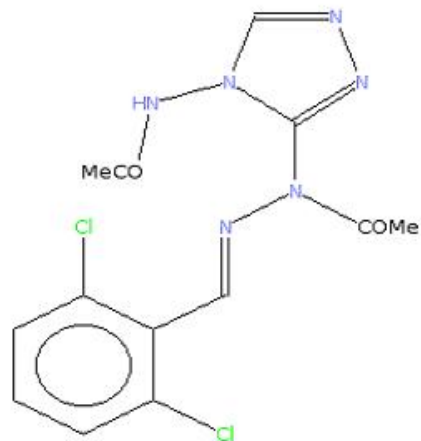
Albert Lund, CHPC and Department of Chemistry, UTAH

Anita Orendt, CHPC, UTAH

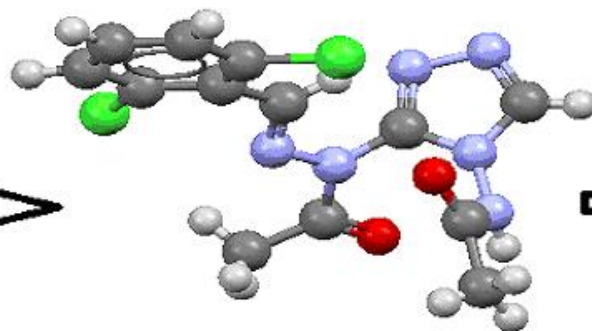
Julio Facelli, Department of Biomedical Informatics, UTAH

Financiamiento: CONICET y

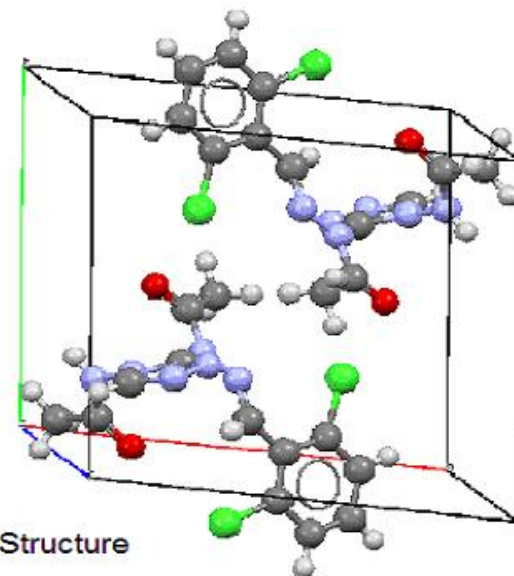
Universidad de Buenos Aires



Chemical Diagram



Chemical Structure



Crystal Structure

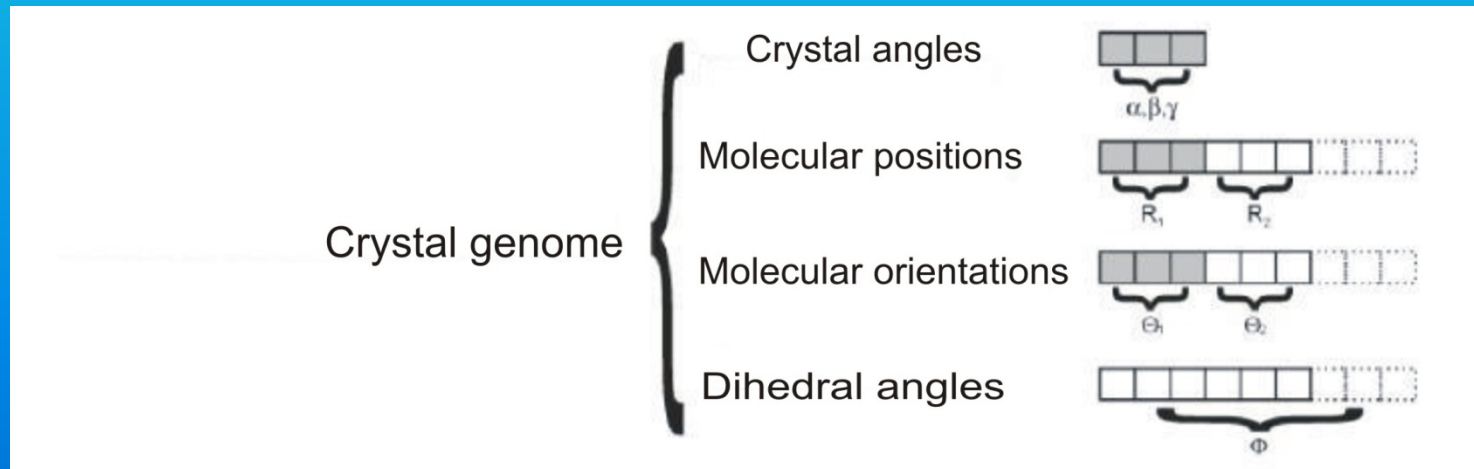
*Proceso de Predicción de Estructuras
Cristalinas*

Por qué es importante la predicción de estructuras cristalinas

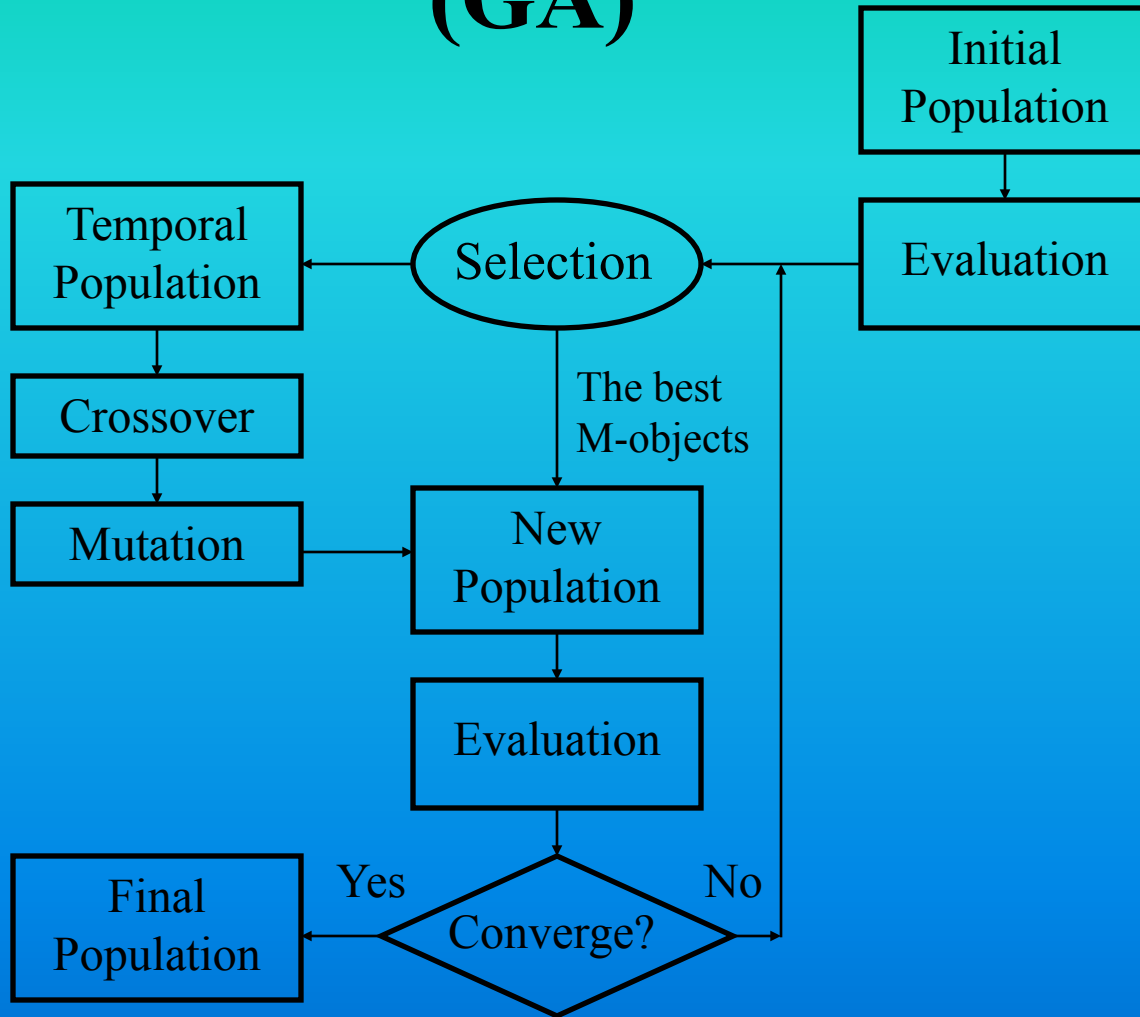
- Polimorfismo en drogas
- Diseño de Cristales
- Estudio de co-cristales
- No hay una manera “**intuitiva**” de *predecir* estructuras

MGAC (Modified Genetic Algorithms for Crystals)

- Algoritmo elitista
- Criterio de selección de las mejores estructuras cristalinas: energía del cristal candidato
- Se optimizan un conjunto de parámetros específicos (genoma)



Genetic Algorithms (GA)



MGAC

Optimización local y evaluación de la energía de los candidatos

- **CHARMM** (potenciales semiempíricos denominados GAFF(1.27))
- **Quantum Espresso, QE.** (método **DFT-D**, Dispersion – Corrected Density Functional Theory)

Búsqueda de los mejores candidatos

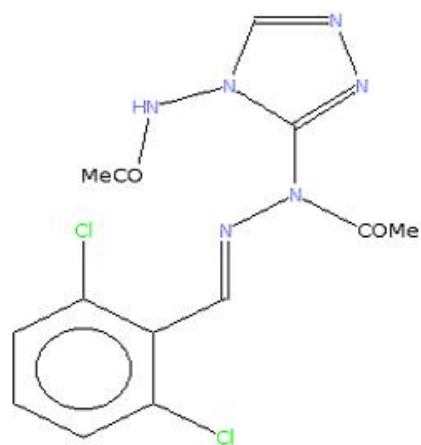
- **14 grupos cristalográficos más comunes** en moléculas orgánicas:
 - P1, P-1, P21, C2, Pc, Cc, P21/c, C2/c, P212121, Pca21, Pna21, Pbcn, Pbca, Pnma
 - **Poblaciones de 30** estructuras y **3** corridas independientes de **150 generations**

Selección

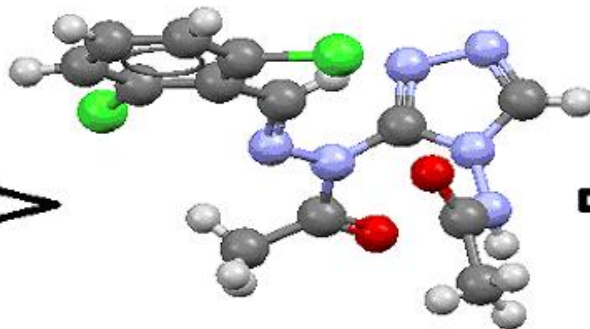
- Se *filtran* los resultados para obtener una lista con las **300** estructuras con energía más baja.

6th Blind Test (Cambridge Crystallographic Data Centre)

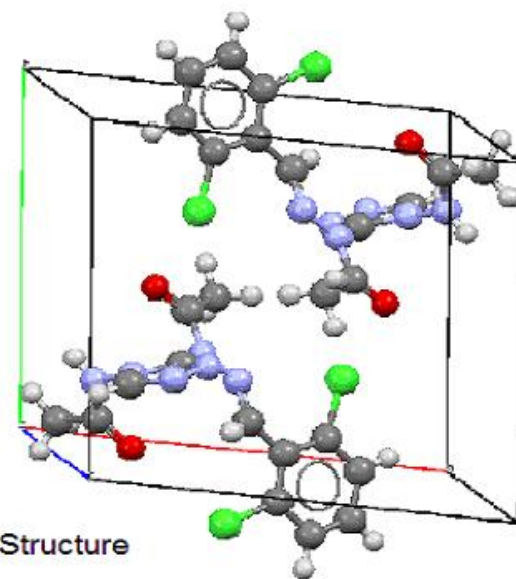
Combinación de MGAC- CHARMM con re-optimización QE posterior.



Chemical Diagram



Chemical Structure

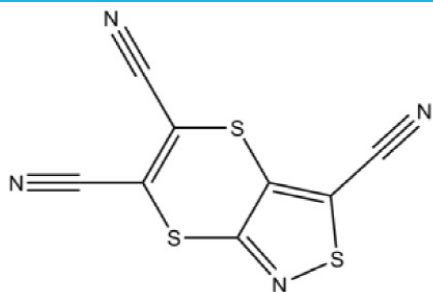


Crystal Structure

Lista de 100 candidatos (en la tabla superior, se muestran sólo los primeros 7)

# Charmm	ΔE (kJ/mol)	vol.(\AA^3)	Simetría	# átom.
3	0,00	463,6	'P b c a'	344
7	0,69	467,6	'P -1'	86
18	0,80	452,7	'P -1'	86
24	1,48	444,9	'P 1 21 1'	86
2	1,75	435,1	'P -1'	86
44	2,59	432,2	'P -1'	86
108	3,57	433,9	'P 1 21/c 1'	172

XXII

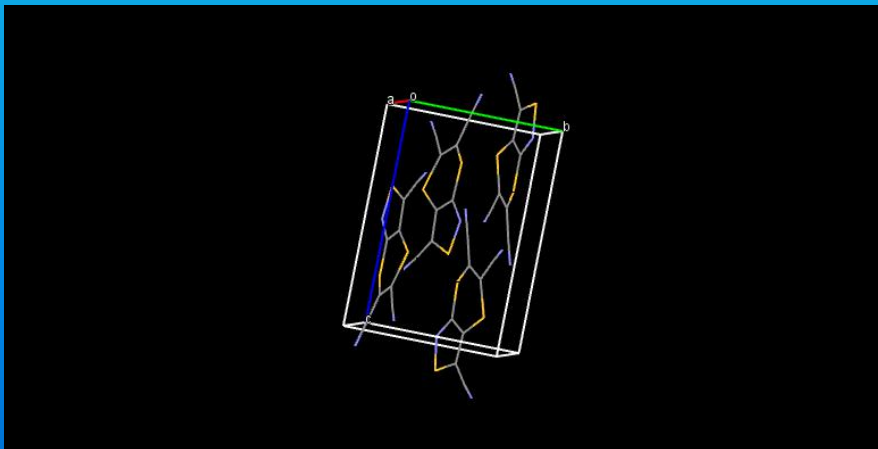
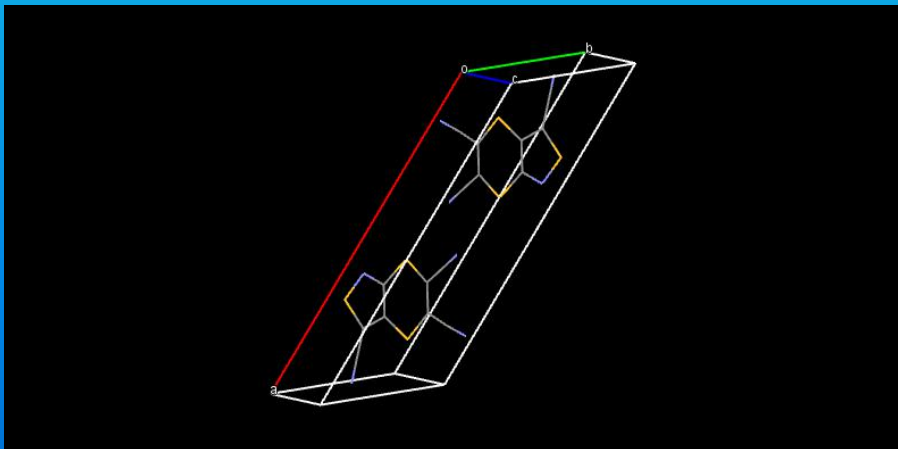
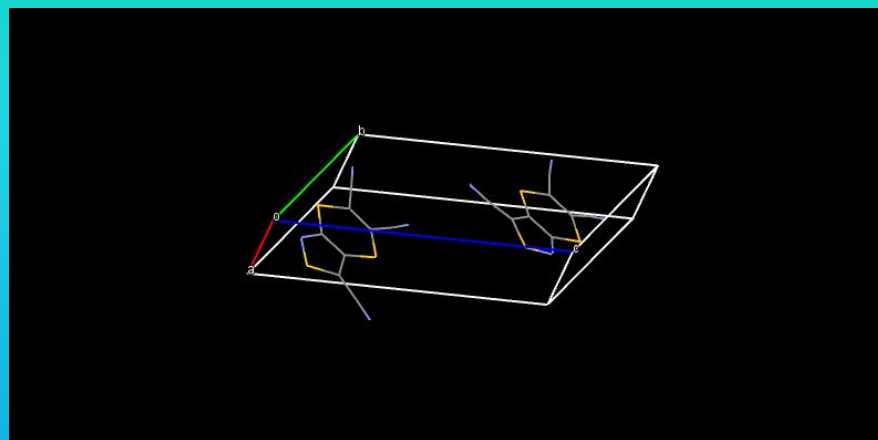
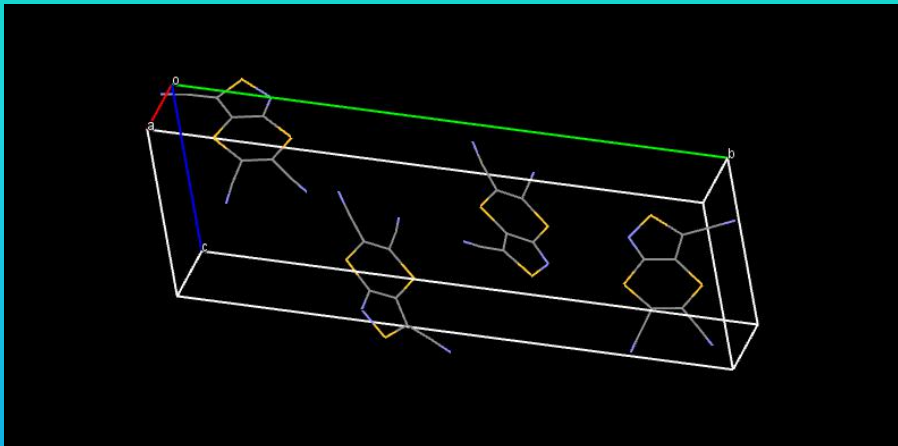


Resultados a presentar en el 6th Blind Test

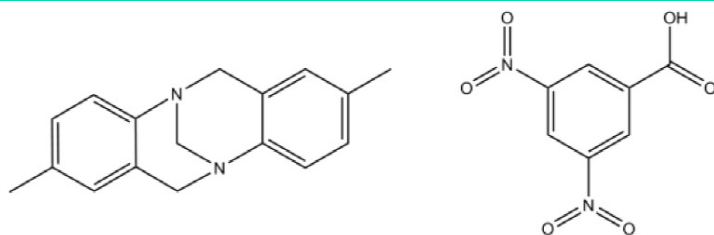
Energy QE (kJ/mol) vs specific volume (\AA^3)



mol-xxii: primeras 4 estructuras cristalinas candidato

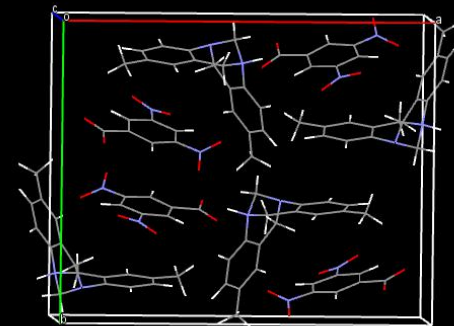
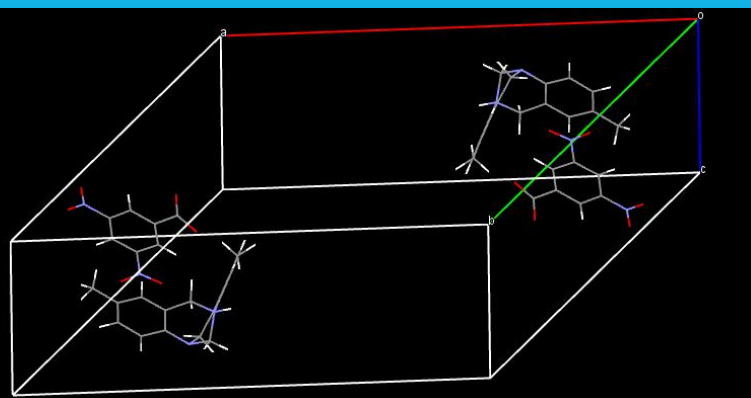
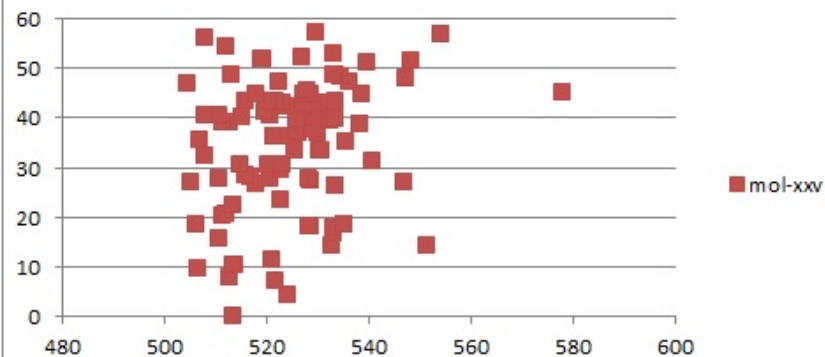


XXV

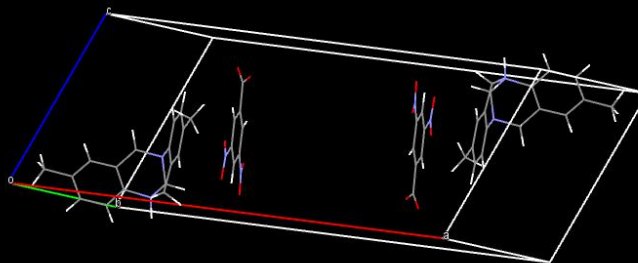


**Resultados a
presentar en el 6th
Blind Test**

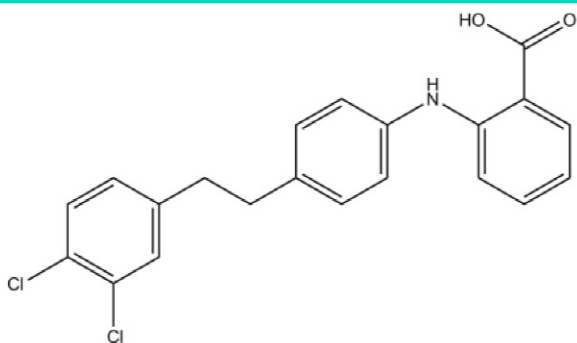
**Energy QE (kJ/mol) vs specific volume
(\AA^3)**



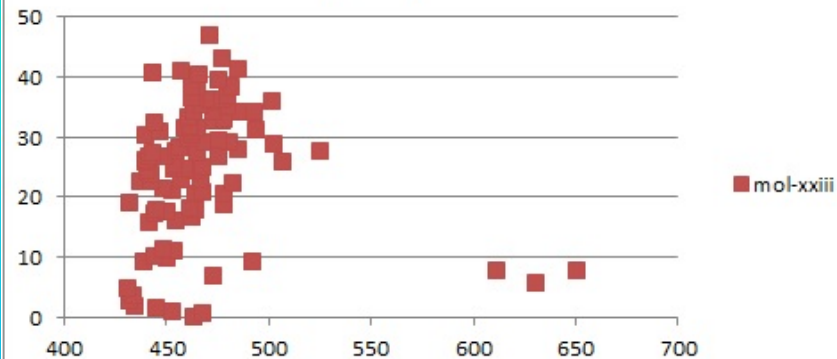
**mol-xxv: primeras 3
estructuras cristalinas
candidato**



XXIII

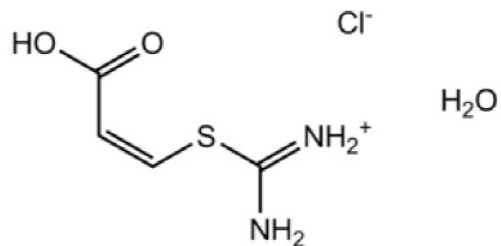


Energy QE (kJ/mol) vs specific volume (Å³)

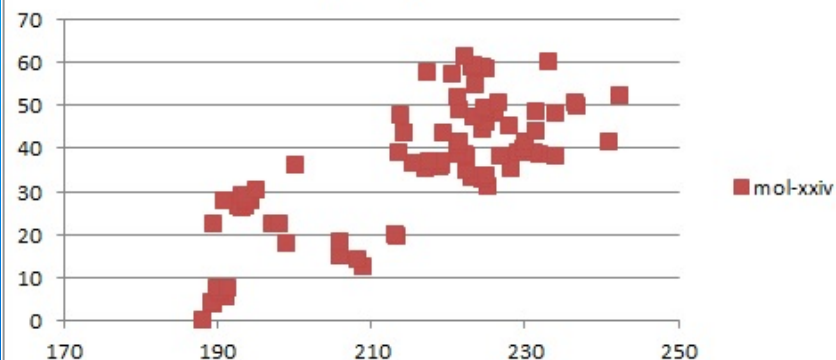


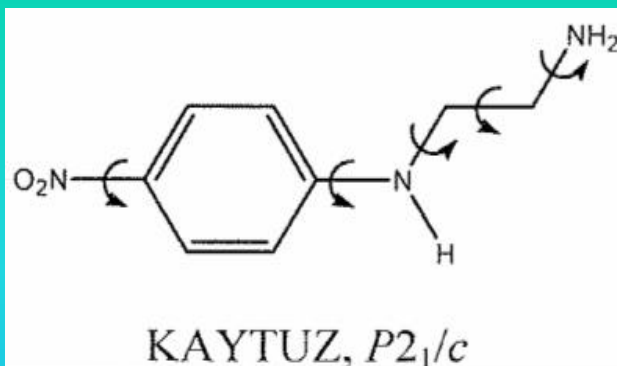
Resultados a presentar en el 6th Blind Test

XXIV

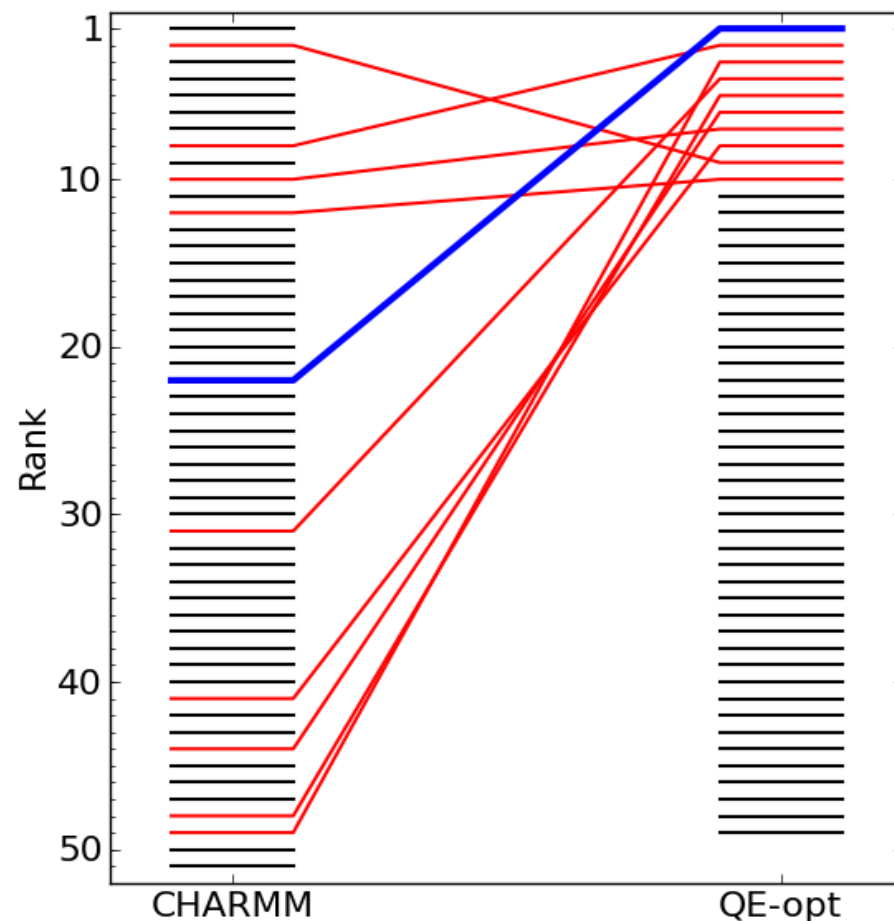


Energy QE (kJ/mol) vs specific volume (Å³)





KAYTUZ: columna izquierda, ranking de energías obtenida con MGAC/CHARMM (potenciales semi-empíricos)



columna derecha, se muestra el reordenamiento que se produce QE al re-optimizar cada estructura con QE (DFT-D)

Lund AM, Orendt AM, Pagola GI, Ferraro MB, Facelli JC (2013). *Cryst Growth Des*, 13(5), 2181-2189

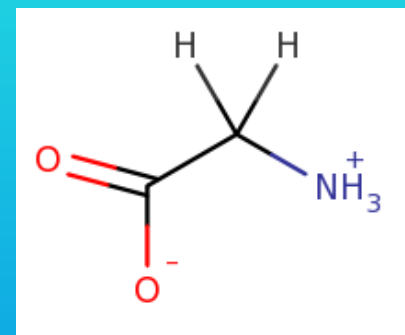
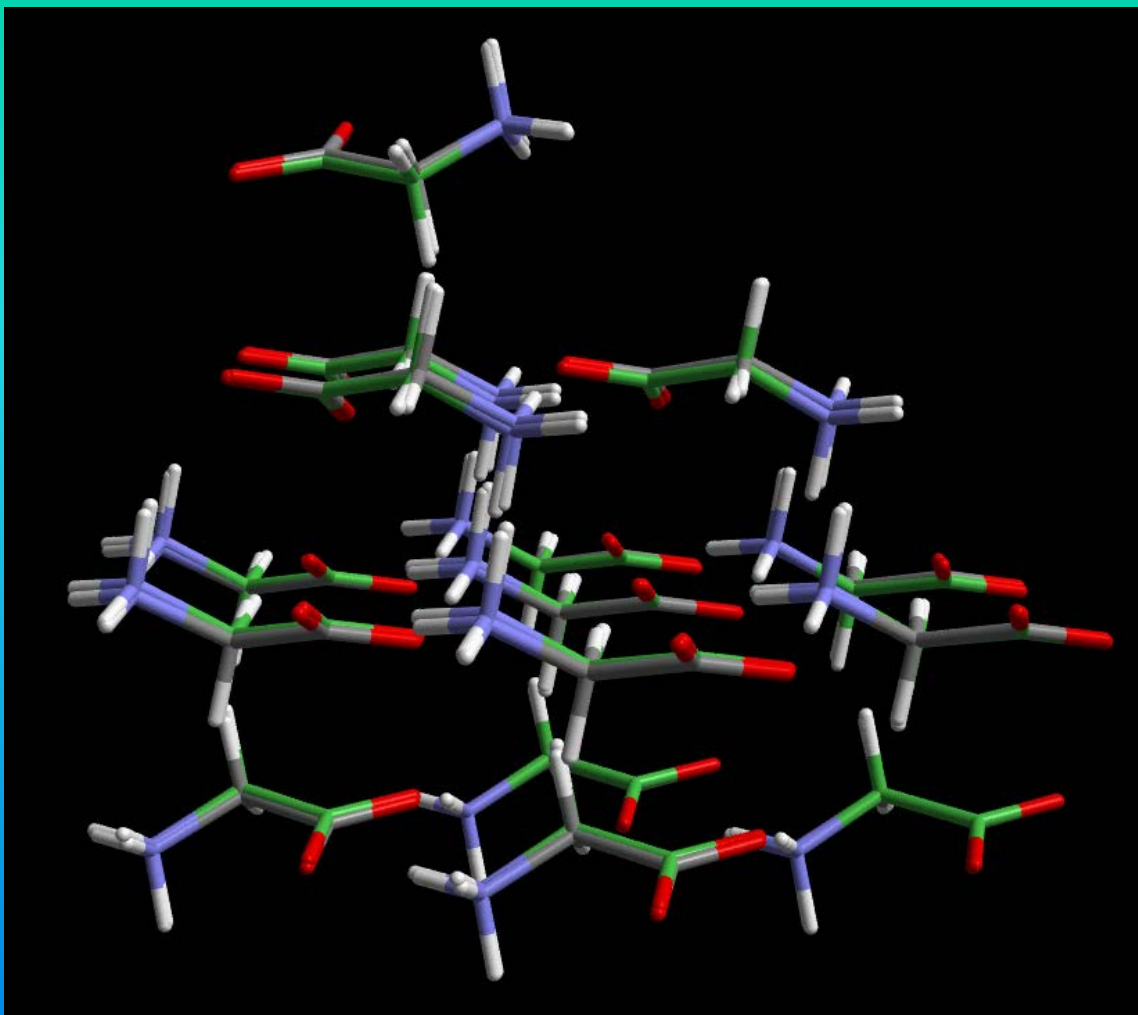
MGAC-QE

resultados para glicina

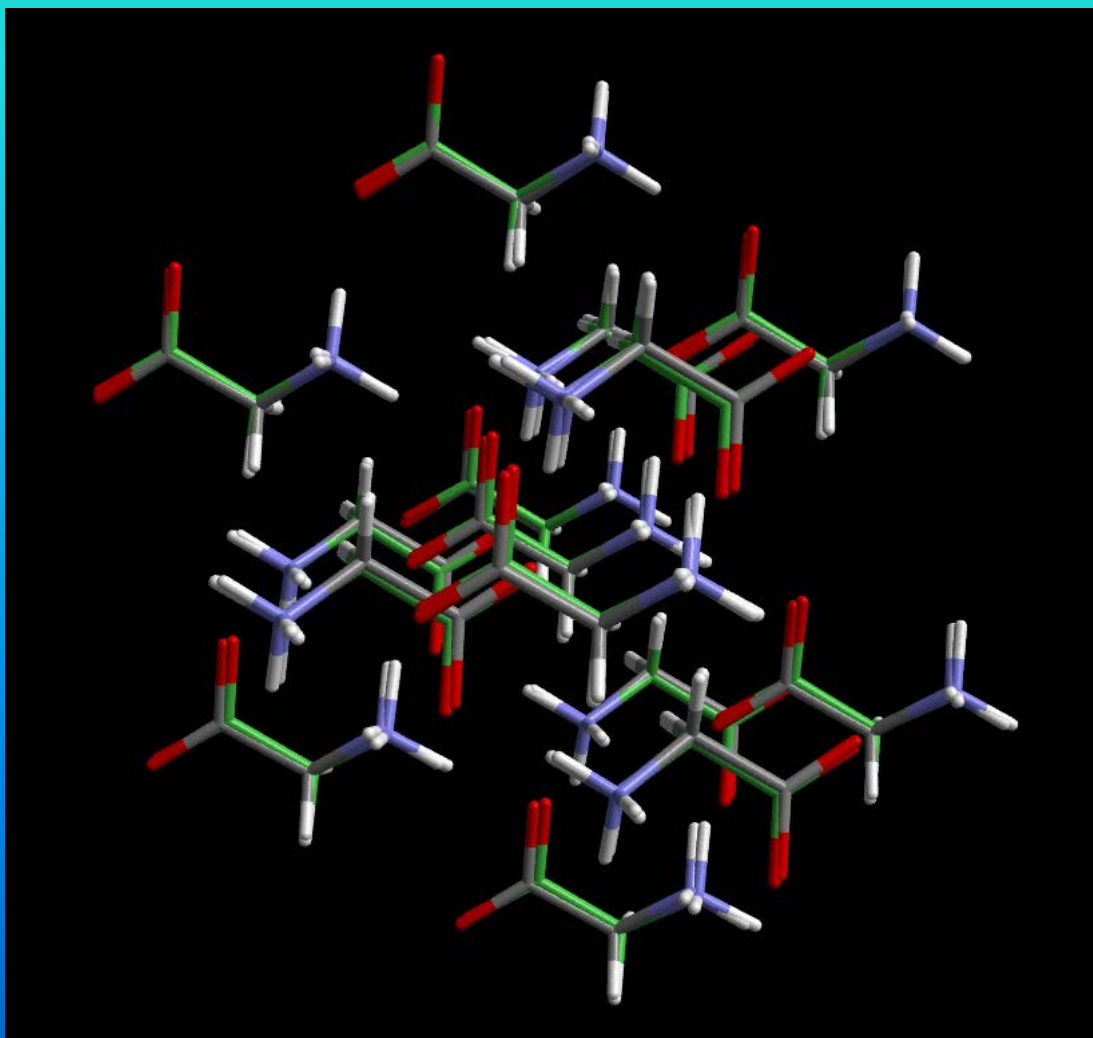
- MGAC con evaluación de energía (y optimización local) con QE
- alto costo en tiempos computacionales (inaplicable por ahora para el Blind Test):

CHARMM: horas QE: 1 mes (predicción completa)

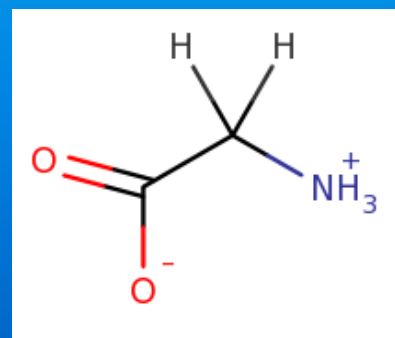
“Crystal structure prediction from first principles: The crystal structures of glycine”; A.M. Lund, G.I. Pagola, A.M Orendt, M.B. Ferraro, J.C. Facelli; Chemical Physics Letters **626**, (2015), pp 20-24



- **ALPHA glycine**, MGAC-QE. Generation 12, Rank 1, Rms = 0.094, Energy = -35292.11 kcal/mol (expected -35292.09). Generation count is low because there are four molecules in the unit cell. Experimental structure is displayed in **gray**.



BETA glycine, MGAC5-QE **Generation 51**, rank 1, rms = **0.118**, Energy = -35291.52 kcal/mol (expected -35291.54), the experimental structure is the **gray** one.



FIN