



FQO-03

EVALUACIÓN SISTEMÁTICA DE NIVELES DE TEORÍA PARA LA PREDICCIÓN DE RMN Y APLICACIÓN EN DP4+

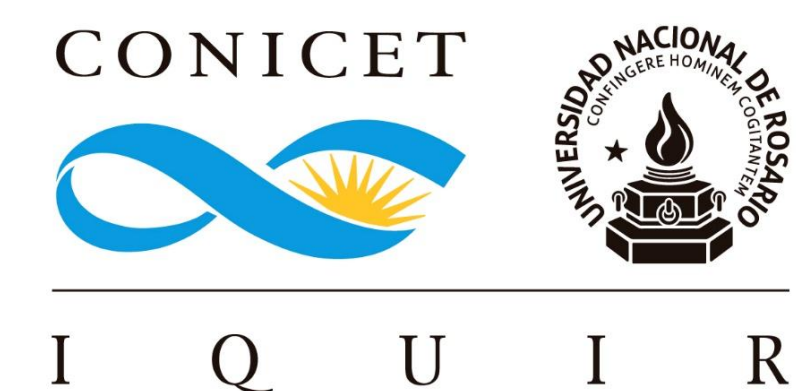


Ezequiel R. Luciano¹, Martín G. Armas Argentino¹, María B. Comba¹, Lucas Passaglia^{1,2}, Bruno A. Franco¹, Ignacio Macia¹, Milagros D. Amichetti^{1,2}, María M. Zanardi* y Ariel M. Sarotti^{2*}

¹ Instituto de Investigaciones en Ingeniería Ambiental, Química y Biotecnología Aplicada (INGEBIO), Facultad de Química e Ingeniería del Rosario, Pontificia Universidad Católica Argentina (UCA), Rosario, S2002QE0, Argentina. *E-mail: zanardi@inv.rosario-conicet.gov.ar

² Instituto de Química Rosario (QUIR-CONICET), Universidad Nacional de Rosario (UNR), Rosario, S2002LRK, Argentina.

*E-mail: sarotti@quir-conicet.gov.ar

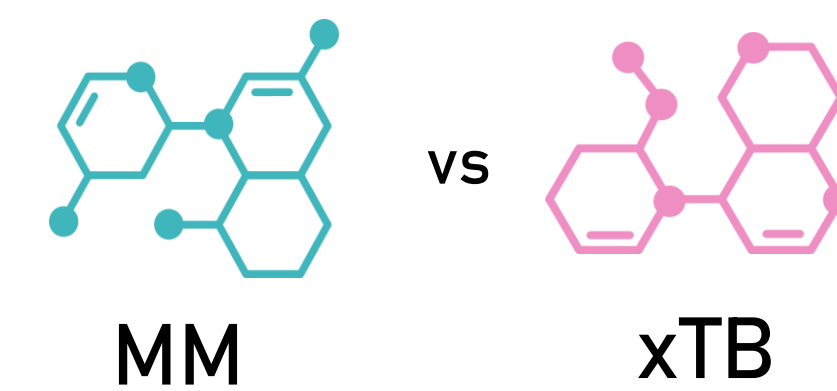


Introducción

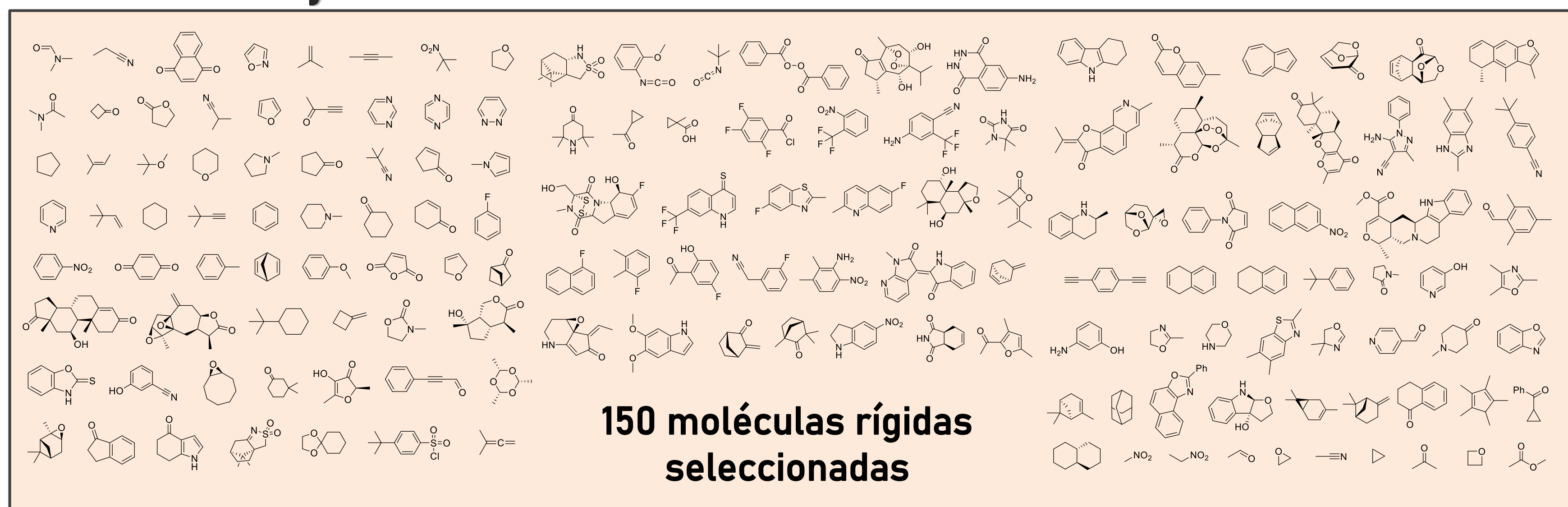
La predicción teórica de desplazamientos químicos de RMN es una herramienta clave en la elucidación estructural de compuestos orgánicos. Sin embargo, el equilibrio entre precisión y costo computacional sigue siendo un desafío, especialmente en grandes moléculas o sistemas flexibles, en donde la optimización de geometría a nivel cuántico sigue siendo limitante. Los métodos semiempíricos ocupan una posición intermedia, permitiendo obtener resultados de buena calidad a con bajos tiempos de cálculo. El método xTB logra un equilibrio entre velocidad y precisión, permitiendo estudiar moléculas grandes sin necesidad de utilizar tratamientos específicos para cada tipo de átomo. Su versión más usada, GFN2-xTB,^[1] incluye efectos de dispersión y electrostática anisotrópica, ofreciendo una descripción precisa de geometrías, frecuencias e interacciones no covalentes.

Objetivos

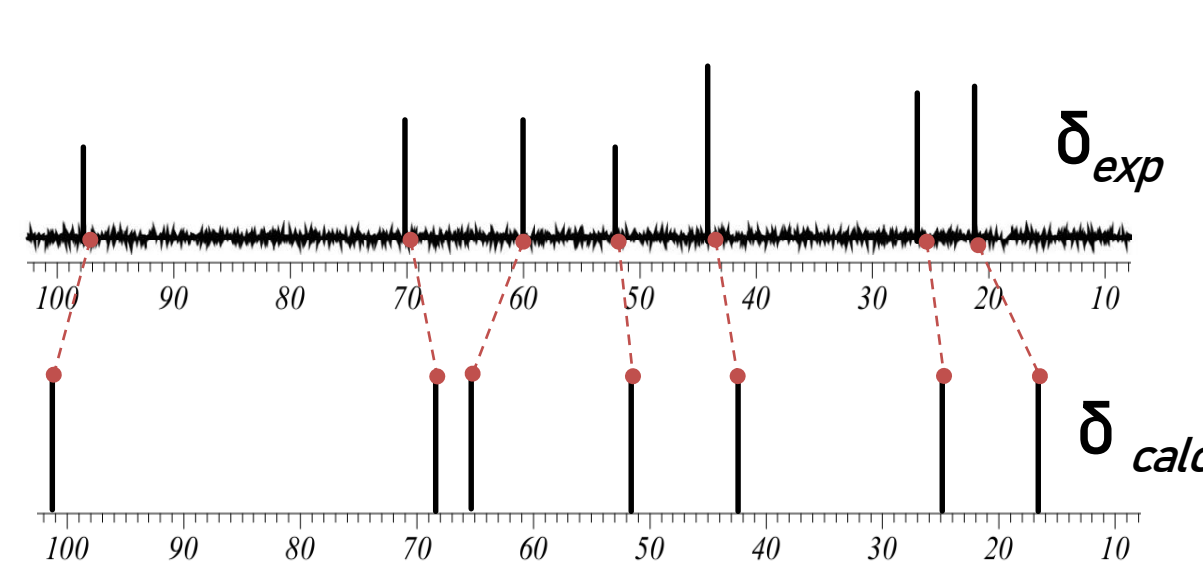
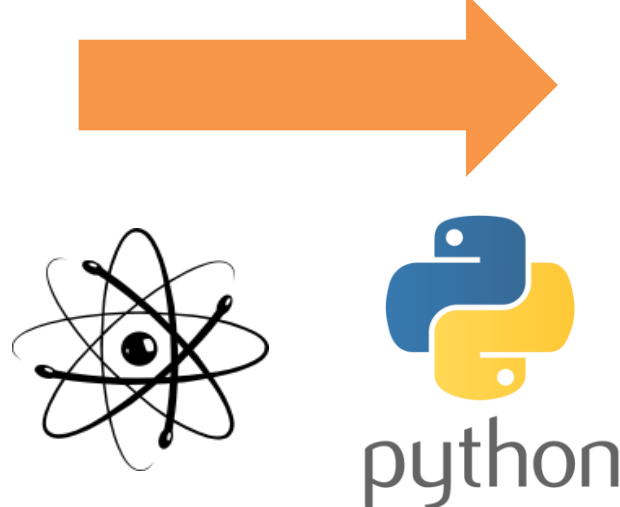
- Evaluar diferentes niveles de teoría para la predicción de desplazamientos químicos de ¹³C y ¹H en 150 moléculas con baja flexibilidad conformational.
- Identificar el nivel de teoría óptimo con buena predicción y bajo costo computacional.
- Analizar el desempeño del método DP4+^[2] bajo distintas calidades geométricas (MM, ^[3] xTB, DFT) y su impacto en la asignación estructural por RMN.



Resultados y Discusión



Modeladas con RDKit de Python, y optimizadas a MMFF

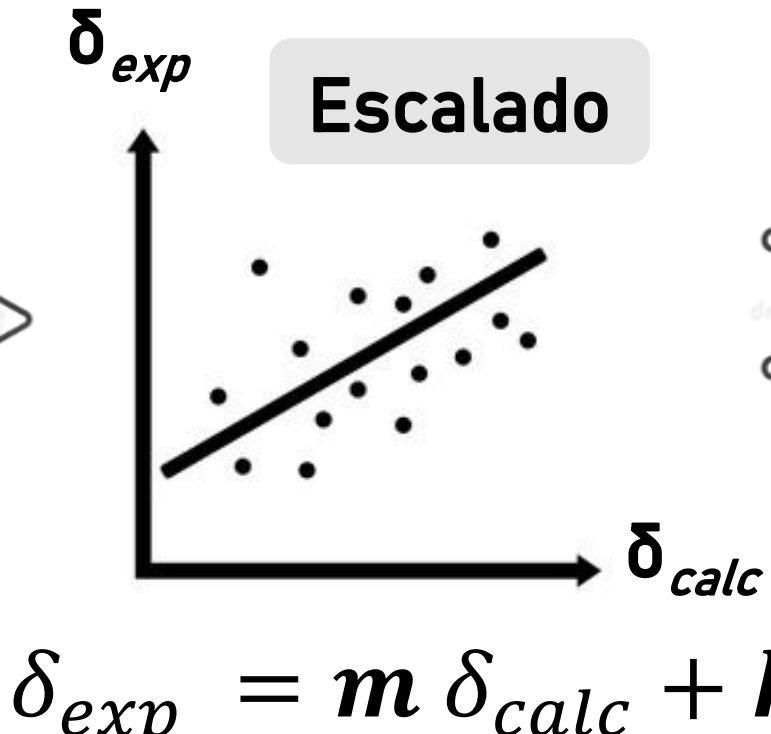


Cálculo de RMN a DFT

Error no escalado

$$e_{uns} = \delta_{calc} - \delta_{exp}$$

Escalado



$$\delta_{exp} = m \delta_{calc} + h$$

Error escalado

$$(\delta_{exp} - m)/h = \delta_{sca}$$
$$e_{sca} = \delta_{sca} - \delta_{exp}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_{uns}|$$
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_{uns})^2}$$
$$CMAE = \frac{1}{n} \sum_{i=1}^n |e_{sca}|$$
$$CRMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_{sca})^2}$$

Métricas de error

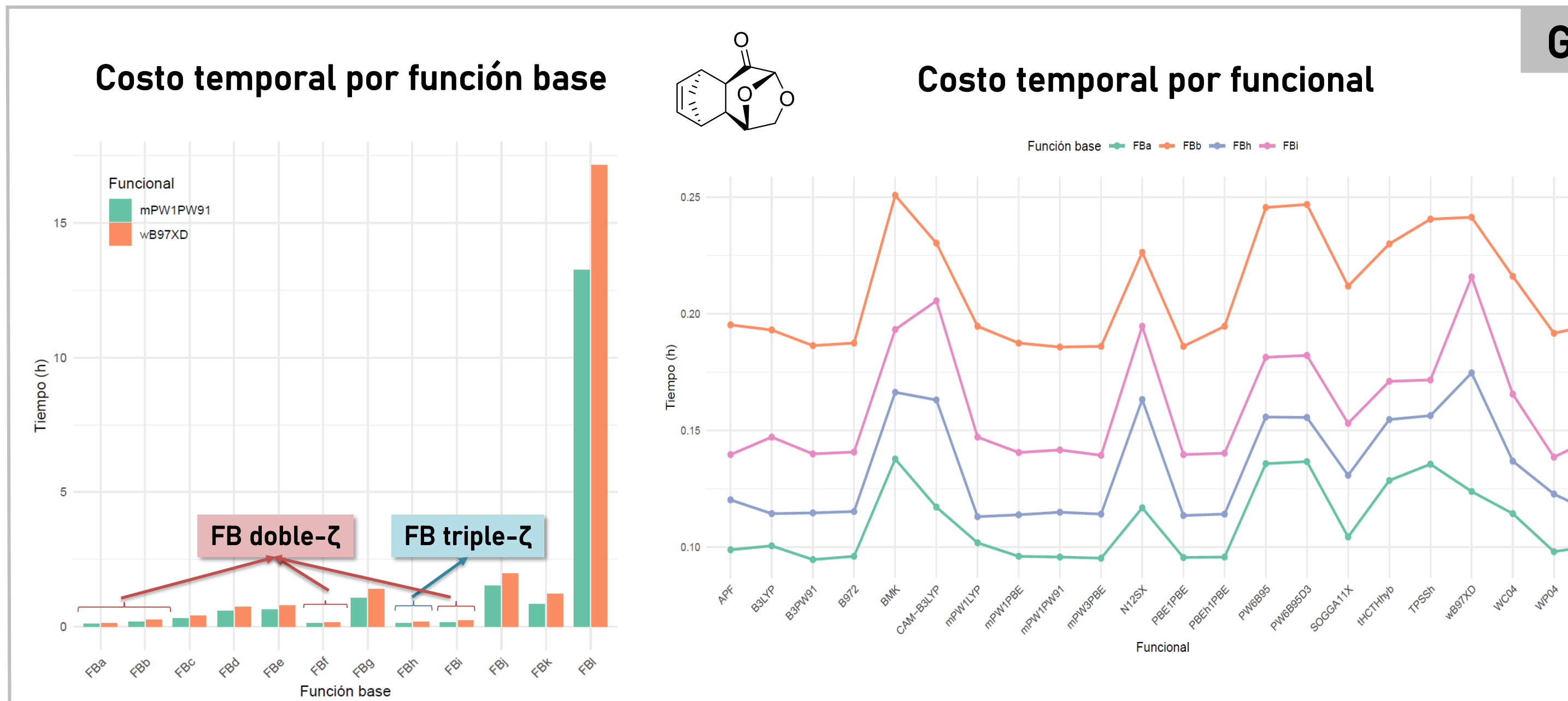
Funciones base doble- ζ :

FBa, FBb, FBf, Fbi, FBk

Funciones base triple- ζ :

FBc, FBd, FBe, FBg, FBh, FBj, FBl

Solvatación SMD



Geometrías optimizadas a MMFF

Factor de costo temporal

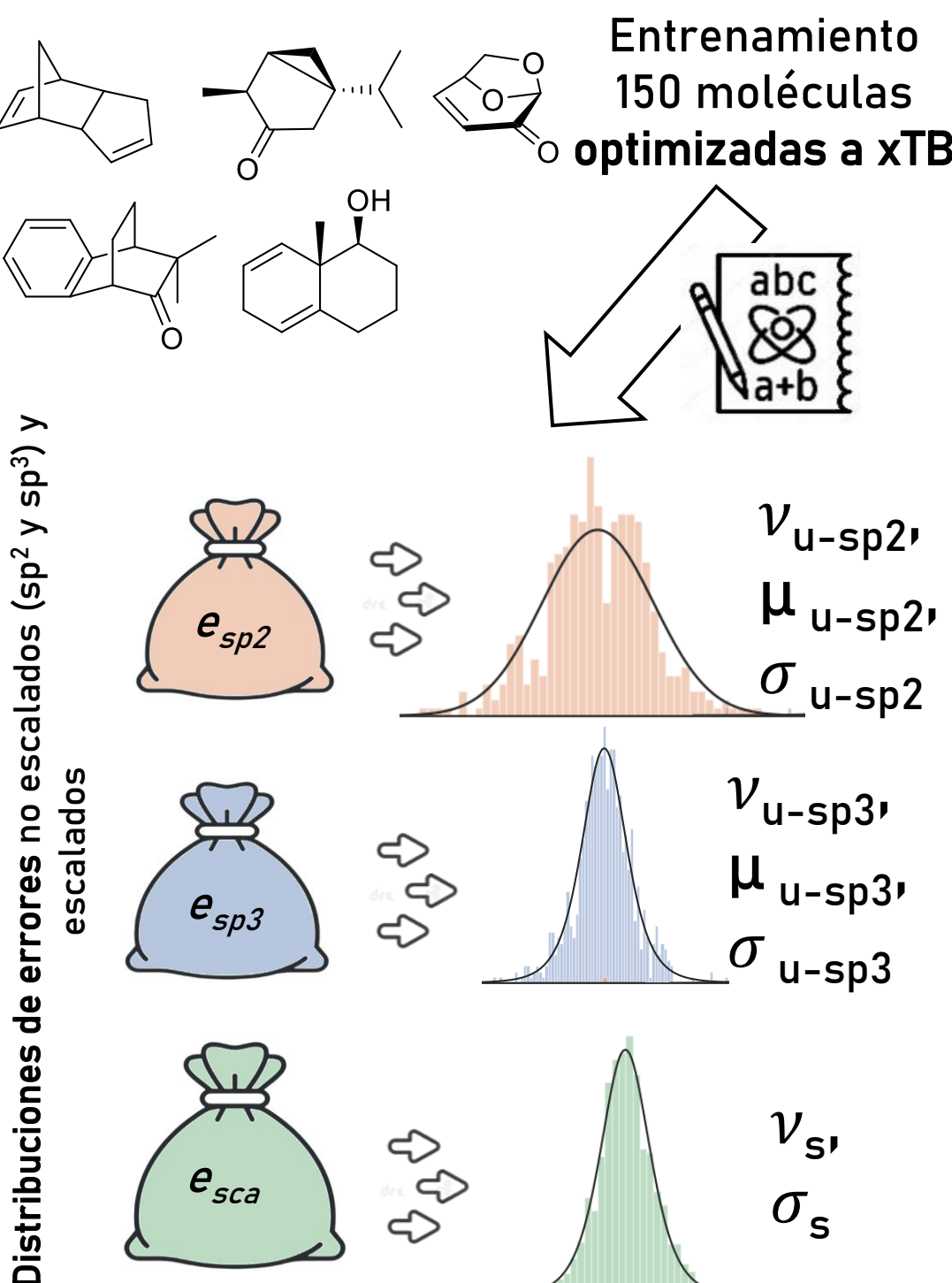
Funcional	X3LYP	WC04	wB97XD	TPSSH	hTCHThyb	SOGGA11X	PW6B95D3	PW6B95	PBE1PBE	PBEh1PBE	CAM-B3LYP	B3LYP	B3PW91	APF
Función base	1.035	2.042	3.448	6.24	6.811	1.295	11.904	1.361	1.709	1.724	0.714	1.014	2	3.376
6-31G**	1.014	2	3.376	6.24	6.811	1.295	11.904	1.361	1.709	1.724	0.714	1.014	2	3.376
6-31+G**	1.17	2.309	3.897	7.054	7.699	1.464	13.543	1.538	1.932	1.947	0.810	1.107	2.164	3.697
6-31++G(2d,p)	1.376	2.716	4.585	8.288	8.957	1.722	15.688	1.809	2.273	2.287	0.913	1.251	2.525	4.275
6-31++G(2d,p)	1.283	2.632	4.275	7.736	8.445	1.605	14.638	1.687	2.119	2.133	0.810	1.107	2.164	3.697
def2-SVP	1.259	2.477	4.180	7.586	8.281	1.57	14.631	1.65	2.075	2.089	0.714	1.014	2	3.376
def2-TZVP	1.107	2.164	3.697	6.872	7.283	1.385	12.828	1.455	1.825	1.839	0.610	0.810	1.014	2
def2-TZV	1.3	2.566	4.331	7.636	8.556	1.627	14.828	1.709	2.147	2.161	0.810	1.107	2.164	3.697
cc-PVDZ	1.298	2.561	4.323	7.824	8.541	1.624	14.802	1.709	2.143	2.157	0.810	1.107	2.164	3.697
cc-PVTZ	1.006	1.984	3.35	6.063	6.618	1.258	11.488	1.322	1.661	1.675	0.610	0.810	1.014	2
aug-cc-PVDZ	0.997	1.967	3.32	6.008	6.558	1.247	11.368	1.31	1.646	1.66	0.610	0.810	1.014	2
aug-cc-PVTZ	1.298	2.561	4.323	7.824	8.541	1.624	14.802	1.709	2.143	2.157	0.810	1.107	2.164	3.697
6-31G**	1.298	2.561	4.323	7.824	8.541	1.624	14.802	1.709	2.143	2.157	0.810	1.107	2.164	3.697
6-31+G**	0.990	1.97	3.325	6.02	6.571	1.249	11.361	1.312	1.648	1.662	0.610	0.810	1.014	2
6-31++G(2d,p)	1.021	2.012	3.396	6.146	6.708	1.275	11.624	1.34	1.683	1.697	0.610	0.810	1.014	2
6-31++G(2d,p)	1.305	2.575	4.346	7.866	8.556	1.632	14.81	1.715	2.155	2.169	0.810	1.107	2.164	3.697
def2-SVP	1.356	2.676	4.516	8.177	8.925	1.697	15.688	1.783	2.24	2.254	0.913	1.251	2.525	4.275
def2-TZVP	0.997	1.968	3.321	6.011	6.562	1.247	11.371	1.311	1.647	1.661	0.610	0.810	1.014	2
def2-TZV	0.990	1.97	3.325	6.02	6.571	1.249	11.361	1.312	1.648	1.662	0.610	0.810	1.014	2
cc-PVDZ	1.021	2.012	3.396	6.146	6.708	1.275	11.624	1.34	1.683	1.697	0.610	0.810	1.014	2
cc-PVTZ	1.026	2.028	3.426	6.2	6.762	1.286	11.728	1.352	1.698	1.712	0.610	0.810	1.014	2
aug-cc-PVDZ	1.026	2.028	3.426	6.2	6.762	1.286	11.728	1.352	1.698	1.712	0.610	0.810	1.014	2
aug-cc-PVTZ	1.026	2.028	3.426	6.2	6.762	1.286	11.728	1.352	1.698	1.712	0.610	0.810	1.014	2

El impacto del funcional sobre el tiempo de cálculo es considerablemente menor que el efecto de la función base
Nivel más rápido de cálculo de RMN (f = 0,997): SMD/PBE1PBE/6-31G**

Funcional	X3LYP	WC04	wB97XD	TPSSH	hTCHThyb	SOGGA11X	PW6B95D3	PW6B95	PBE1PBE	PBEh1PBE	CAM-B3LYP	B3LYP	B3PW91	APF
Función base	2.28	2.47	2.26	2.25	2.19	2.28	2.62	2.35	2.18	2.74	2.22	2.06	2.11	2.03
6-31G**	2.06	2.11	2.03	2.02	2.06	2.12	2.55	2.17	2.04	2.49	2.06	2.08	2.14	2.06
6-31+G**	2.30	2.45	2.19	2.23	2.18	2.27	2.65	2.36	2.17	2.69	2.19	2.11	2.17	2.10
6-31++G(2d,p)	2.11	2.18	2.20	2.18	2.17	2.17	2.10	2.56	2.14	2.10	2.38	2.13	2.11	2.10
6-31++G(2d,p)	2.11	2.17	2.10	2.11	2.19	2.08	2.07	2.55	2.15	2.03	2.59	2.04	2.11	2.10
def2-SVP	2.11	2.17	2.10	2.11	2.10	2.08	2.07	2.55	2.15	2.03	2.59	2.04	2.11	2.10
def2-TZVP	2.07	2.15	2.06	2.08	2.07	2.03	2.06	2.51	2.11	2.01	2.39	2.02	2.07	2.14
def2-TZV	2.07	2.14	2.06	2.08	2.07	2.03	2.05	2.51	2.10	2.01	2.38	2.02	2.07	2.14
cc-PVDZ	2.08	2.13	2.07	2.06	2.06	2.04	2.03	2.51	2.10	1.99	2.37	2.01	2.08	2.15
cc-PVTZ	2.15	2.25	2.09	2.12	2.12	2.06	2.12	2.55	2.17	2.04	2.49	2.06	2.11	2.03
aug-cc-PVDZ	2.08	2.16	2.06	2.08	2.07	2.04	2.05	2.52	2.11	2.00	2.41	2.02	2.07	2.14
aug-cc-PVTZ	2.07	2.14	2.06	2.07	2.06	2.03	2.05	2.51	2.09	2.00	2.38	2.01	2.07	2.14
6-31G**	2.25	2.46	2.21	2.26	2.25	2.18	2.28	2.63	2.33	2.19	2.74	2.23	2.20	2.15
6-31+G**	2.06	2.15	2.05	2.06	2.05	2.04	2.46	2.09	2.00	2.46	2.03	2.08	2.15	2.05
6-31++G(2d,p)	2.33	2.31	2.29	2.22	2.22	2.36	2.21	2.61	2.47	2.15	2.53	2.23	2.20	2.15
6-31++G(2d,p)	2.15	2.25	2.09	2.13	2.12	2.08	2.14	2.54	2.19	2.06	2.48	2.07	2.15	2.25
def2-SVP	2.15	2.25	2.09	2.13	2.12	2.07	2.10	2.54	2.18	2.03	2.48	2.05	2.15	2.25
def2-TZVP	2.31	2.30	2.21	2.26	2.25	2.20	2.29	2.63	2.37	2.19	2.76	2.23	2.20	2.31
def2-TZV	2.10	2.18	2.07	2.09	2.08	2.04	2.07	2.52	2.13	2.01	2.41	2.03	2.10	2.18
cc-PVDZ	2.10	2.18	2.07	2.09	2.08	2.04	2.07	2.52	2.13	2.01	2.41	2.03	2.10	2.18
cc-PVTZ	2.10	2.18	2.07	2.09	2.08	2.04	2.07	2.52	2.13	2.01	2.41	2.03	2.10	2.18
aug-cc-PVDZ	2.10	2.18	2.07	2.09	2.08	2.04	2.07	2.52	2.13	2.01	2.41	2.03	2.10	2.18
aug-cc-PVTZ	2.10	2.18	2.07	2.09	2.08	2.04	2.07	2.52	2.13	2.01	2.41	2.03	2.10	2.18

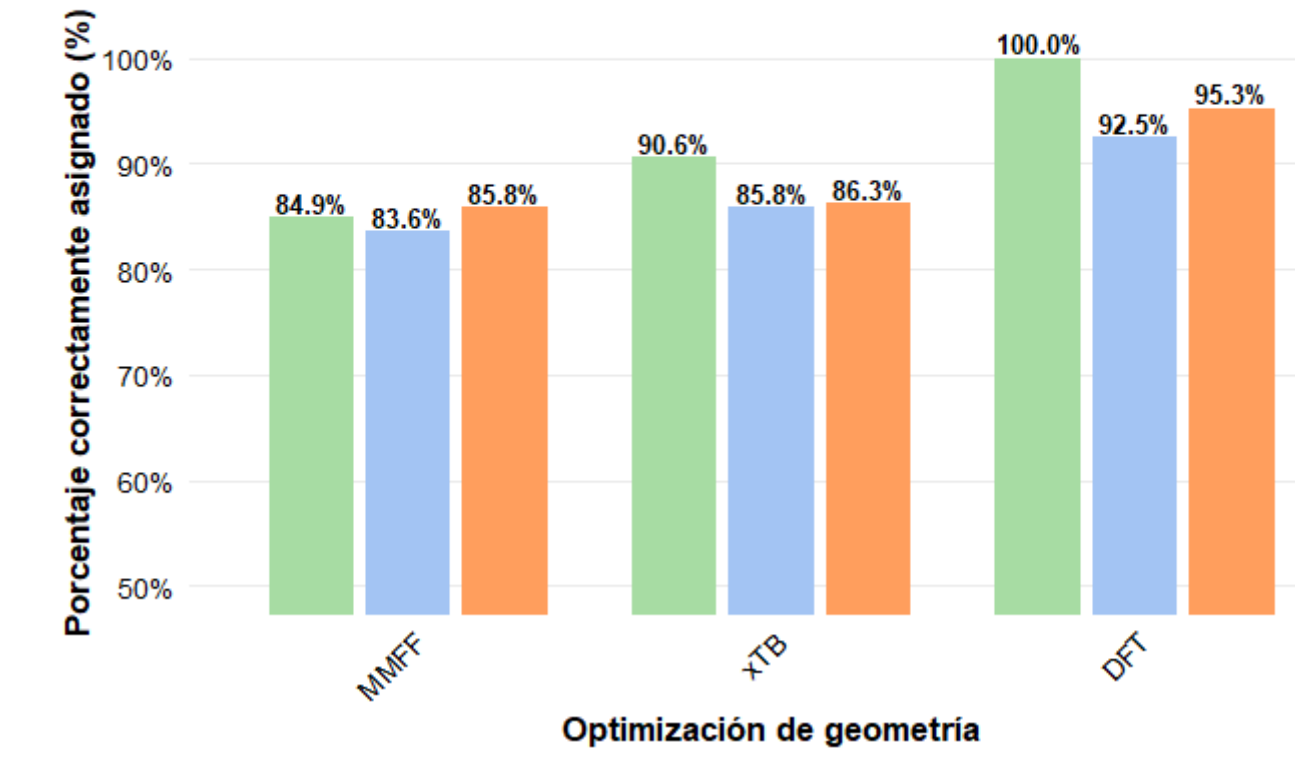
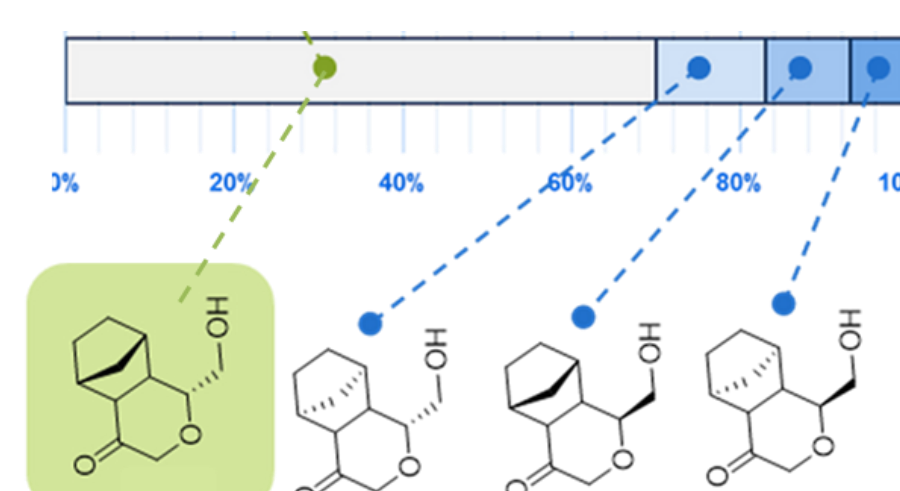
Valores muy similares entre niveles
Mala predicción: funcional WC04 y funciones base FBh, FBk, FBf, FBl

Impacto de xTB en DP4+



Validación con 104 casos de diversidad estructural a los niveles:

- mPW1PW91/6-31+G** Nivel de RMN recomendado para DP4+ (geometrías a DFT)
- wB97XD/6-31+G** Nivel de RMN recomendado para MM-DP4+ (geometrías a MMFF)
- PBE1PBE/6-31G** Nivel de RMN encontrado como óptimo para la predicción.



Conclusiones

- Las funciones base doble- ζ : 6-31G** (FBa), 6-31+G** (FBb), def2-SVP (FBf), cc-PVDZ (Fbi) mostraron bajo error y alta eficiencia para cálculo de RMN.
- La función base utilizada para el cálculo de RMN es el factor predominante en el tiempo de cálculo.
- Las geometrías optimizadas con xTB mejoran sistemáticamente la predicción de RMN.
- Nivel óptimo: Geometría: xTB, Cálculo de RMN: PBE1PBE/6-31G** (FBa) Excelente compromiso entre exactitud y tiempo de cómputo.
- La optimización de geometría a xTB mostró una performance de DP4+ intermedia entre MMFF y DFT, siendo el mejor nivel de RMN mPW1PW91/6-31+G** (90.6%)

Referencias

- [1] Bannwarth, C; Ehlert, S; Grimme, S. *J. Chem. Theory Comput.* **2019**, *15*, 3, 1652–1671
- [2] Grimblat N.; Zanardi, M. M.; Sarotti, A. M. *J. Org. Chem.* **2015**, *80*, 12526–12534
- [3] Franco, B. A.; Luciano, E. R.; Sarotti, A. M.; Zanardi, M. M. *J. Nat. Prod.* **2023**, *86*, 10, 2360–2367

Agradecimientos

ANPCyT INGENIO IQUIR
CONICET UCA UNR

