

THE “LEGO BRICKS” APPROACH: A LOW-COST STRATEGY TO PREDICT ACCURATE ROTATIONAL SPECTRA

F. TONOLO, A. MELLI, *Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy. Dipartimento di Chimica “Giacomo Ciamician”, Università di Bologna, Via F. Selmi 2, I-40126 Bologna, Italy.* ; **V. BARONE**, *Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy.* ; **C. PUZZARINI**, *Dipartimento di Chimica “Giacomo Ciamician”, Università di Bologna, Via F. Selmi 2, I-40126 Bologna, Italy.*

When looking to the first prediction and interpretation of rotational spectra, computational equilibrium structures determinations require an accuracy that demands a very high-level quantum chemical treatment, *e.g.* exploiting the good performances of composite schemes rooted in the coupled cluster theory. However, for medium- to large- sized molecules, the computational cost to achieve high accuracy might be unaffordable. In these cases, cheaper yet accurate approaches need to be defined.

In Melli *et al.* (2021)^{1,2}, a benchmark study, based on the availability of accurate semi-experimental equilibrium rotational constants for 21 molecular species, tested and validated a new approach which relies on a “LEGO bricks” model. This approach is rooted in the idea that a molecular system can be seen as formed by different fragments (*i.e.* the “LEGO bricks”), whose accurate semi-experimental (SE) equilibrium geometries are available³. To join together the different “LEGO bricks”, this approach relied on: (i) the template molecule approach (TM) in order to account, at the rev-DSD-PBEP86-D3/jun-cc-pVTZ (revDSD) level, for the changes occurring when moving from the smaller fragment to the larger molecular system; (ii) the linear regression model (LR) to correct the inter-fragment parameters⁴. The TM-SE_LR scheme has thus been validated by comparison with their semi-experimental equilibrium structures (when available) and rotational constants (which straightforwardly depend on the equilibrium geometry). The main outcome of this study is the reliability, robustness and accuracy of this novel approach, showing structural differences of about 0.001 Å for bond lengths and 0.1° for angles and an average relative deviation with respect to the accurate semi-experimental equilibrium rotational constants of ~ 0.2%.

The application of this approach gave a contribution to several recent results. For example, the successful performance of this methodology has allowed for the spectroscopic assignment of three isomeric forms of cyanoethynylbenzene, thus ensuring an

¹A. Melli and F. Tonolo equally contributed to this work.

²[doi:10.1021/acs.jpca.1c07828](https://doi.org/10.1021/acs.jpca.1c07828) A. Melli, F. Tonolo, V. Barone and C. Puzzarini, *J. Phys. Chem. A*, **125** (2021) 9904.

³[doi:10.1021/jp511432m](https://doi.org/10.1021/jp511432m) M. Piccardo, E. Penocchio, C. Puzzarini, M. Biczysko and V. Barone, *J. Phys. Chem. A*, **119** (2015) 2058.

⁴[10.1021/acs.jctc.1c00788](https://doi.org/10.1021/acs.jctc.1c00788) G. Ceselin, N. Tasinato and V. Barone, *JCTC*, **17** (2021) 7290.

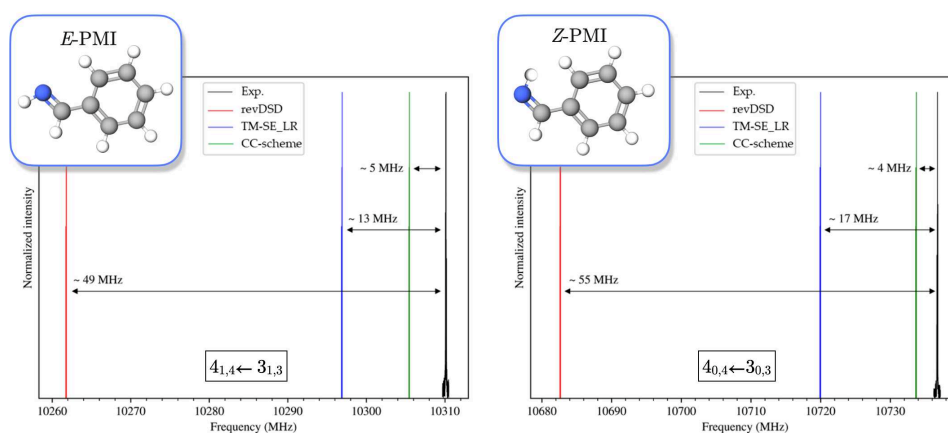


Figure 1: Comparison between the experimental spectrum (Exp.) and different simulations (revDSD, TM-SE_LR and a computationally expensive composite scheme based on coupled-cluster theory (CC)) for a transition of both isomers of Phenylmethanimine (PMI). This example shows how the TM-SE_LR approach is able to provide a remarkable accuracy in rotational spectral predictions, without increasing the computational effort with respect to revDSD.

analysis both accurate and competitive with the computational and spectral acquisition times⁵. Another example is provided by the recent application of this approach to a set of small polycyclic aromatic hydrocarbons (PAHs) and their derivatives, where a comparison with experimental data demonstrated a relative accuracy for rotational constant predictions below 0.1%⁶. Therefore, we expect that the “LEGO bricks” model will represent a valid approach for the spectroscopic characterization of large molecular systems that were previously inaccessible for accurate computational studies.

⁵doi: [10.1039/D2CP04825F](https://doi.org/10.1039/D2CP04825F) J. T. Spaniol, K. L. K. Lee, O. Pirali, C. Puzzarini and M. A. Martin-Drumel, *PCCP*, **25** (2023) 6397.

⁶doi: [10.1039/D2CP04706C](https://doi.org/10.1039/D2CP04706C) C. Puzzarini, J. F. Stanton, *PCCP*, **25** (2023) 1421; doi: [10.1039/D2CP03294E](https://doi.org/10.1039/D2CP03294E) H. Ye, S. Alessandrini, M. Melosso and C. Puzzarini, *PCCP*, **24** (2024) 23254.