

## Introduction

The main problem hindering the application of accurate methods to large species arises from a high scaling of required computational resources. The fragmentation methods are based on an idea of the fragmentation of the molecule into small overlapping fragments and the representation of the energy and other molecular properties as a sum of fragments' energies with positive or negative weights [1,2]. An important consideration in fragmentation-based methods, and in composite methods in general, is how to satisfy the valences of a broken bond in the calculation of a model subsystem. In general, a hydrogen atom is placed along the vector of the broken bond, but this procedure works only when a single bond is broken. Because of this, most fragmentation methods start by identifying each individual link in the system and then deciding where to cut single bonds to form fragments. In polycyclic aromatic hydrocarbons (PAHs), each carbon-carbon bond has a bond order greater than one, making it impossible to avoid cutting a "multiple bond" [2].

Previous approaches for fragmentation of large PAHs have been implemented in the different frameworks (molecular tailoring approach (MTA), kernel energy method (KEM), molecular fractionation with conjugate caps (MFCC)) and rely on generating large fragments to circumvent the effects of disturbing aromaticity [3]. However, neither of these is a systematic approach which is able to be applied on a large class of PAHs and usually they produce really big fragments, therefore they are not applicable on fullerene structures [4].

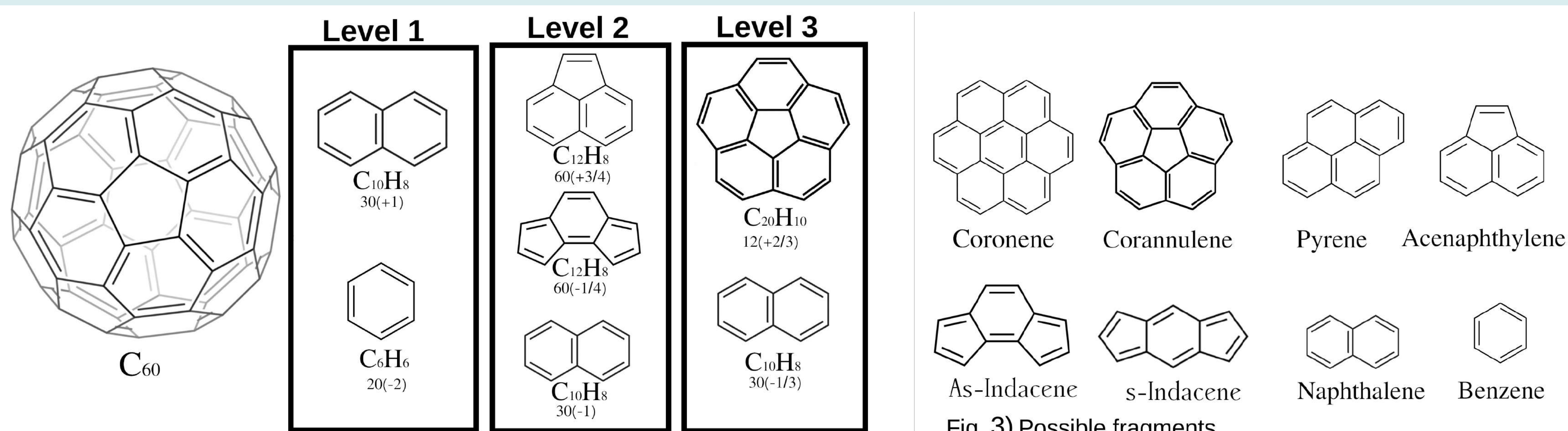


Fig. 1) Symmetric Systematic Molecular Fragmentation of  $C_{60}$

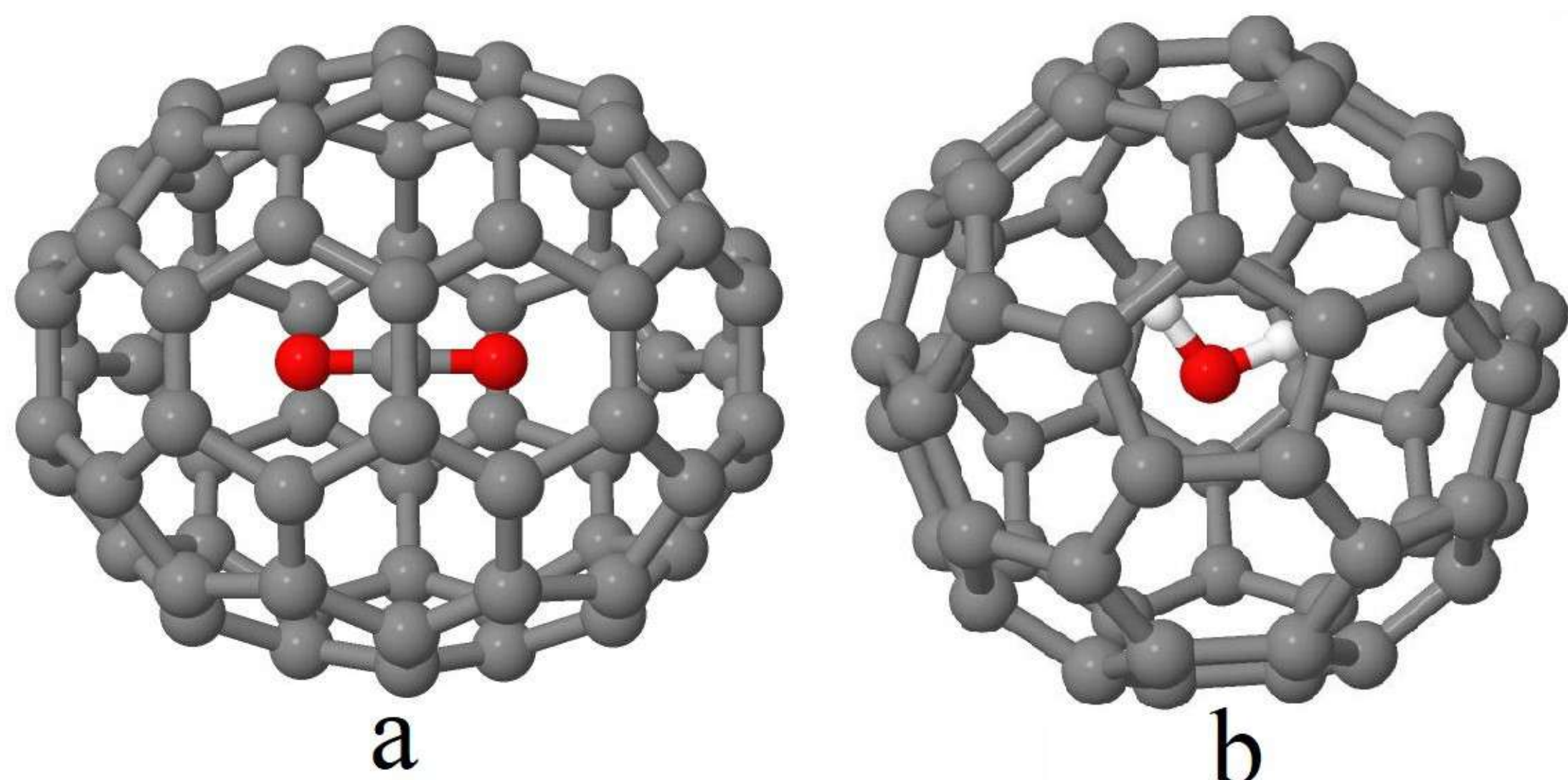


Fig. 2) a- $CO_2$  inside  $C_{70}$ , b- $H_2O$  inside  $C_{60}$

## Method

- 1) We present a novel algorithm which starts from the geometry of the PAH compound, from which all possible fully aromatic segments like naphthalene, pyrene, corannulene, coronene and etc. are extracted (the list of allowed fragments is predefined)
- 2) At the end, weights of each fragment are calculated according to the symmetry and stoichiometry of the PAH molecule. For instance, when applying this algorithm to  $C_{60}$ , twelve corannulene molecules with weight equal to  $2/3$  and thirty naphthalene molecules with weight equal to  $-1/3$  are produced at third level of fragmentation. The same algorithm can be applied to carbon nanotubes and graphene sheets.

## Results and discussion

- We calculated the interaction energy of several small molecules inside  $C_{60}$  and  $C_{70}$  (Fig. 2).
  - With symmetry-adapted perturbation theory (SAPT), where monomers are treated at the Hartree-Fock (HF) level (the so-called SAPT0), it turns out that the scheme reproduces the second-order dispersion contribution very accurately, with relative errors less than 2%.
  - On the other hand, the SAPT0 components which asymptotically correspond to the supermolecular HF interaction energy are reproduced with somewhat larger errors, e.g. the second-order induction gives errors of up to 10%.
- Overall, the error resulting from the HF interaction energy amounts to up to 20%.

- The HF part of the interaction energy can be easily obtained even for large molecules, without any fragmentation.
- We propose a hybrid approach, which consists of a calculation of the HF interaction energy for the whole complex and of making use of the fragmentation in order to obtain the electron correlation energy only.

	$E_{\text{elst}}^{(10)}$	$E_{\text{exch}}^{(10)}$	$E_{\text{ind}}^{(20)}$	$E_{\text{exch-ind}}^{(20)}$	$E_{\text{disp}}^{(20)}$	$E_{\text{exch-disp}}^{(20)}$	$E_{\text{int}}^{\text{HF}}$	
$CO@C_{60}$	No Fragmentation	-8.63	19.38	-9.25	7.71	-35.53	3.97	8.74
	Level 1	-7.80	13.57	-14.53	13.42	-35.17	3.84	3.56
	Level 2	-10.32	22.44	-13.29	11.33	-35.98	4.18	9.28
	Level 3	-7.03	15.37	-5.44	3.61	-35.57	3.80	6.31
$N_2@C_{60}$	No Fragmentation	-5.90	15.09	-4.31	3.62	-28.63	2.44	7.91
	Level 1	-6.02	12.29	-9.40	9.02	-28.29	2.49	4.73
	Level 2	-5.63	14.03	-4.09	3.38	-25.14	2.15	7.30
	Level 3	-5.05	12.81	-2.18	1.27	-28.55	2.32	6.47
$C_2H_2@C_{60}$	No Fragmentation	-63.31	160.89	-77.45	63.72	-129.85	23.62	77.61
	Level 1	-58.05	109.33	-84.01	69.97	-130.37	22.83	28.65
	Level 2	-57.71	165.39	-84.29	70.90	-134.05	25.03	74.10
	Level 3	-60.03	150.16	-73.02	59.72	-130.78	23.13	70.89
$CO_2@C_{60}$	No Fragmentation	-59.93	155.49	-53.56	49.85	-114.93	18.44	86.31
	Level 1	-50.76	109.30	-59.40	55.15	-113.96	16.43	49.61
	Level 2	-71.87	150.95	-60.37	54.47	-117.50	18.27	82.26
	Level 3	-54.90	142.59	-41.56	38.00	-115.34	17.55	78.26
$H_2O@C_{60}$	No Fragmentation	-3.86	10.25	-5.15	2.26	-28.68	2.30	2.91
	Level 1	-3.96	2.60	-8.08	5.19	-28.40	2.30	-5.12
	Level 2	-4.97	10.38	-7.35	4.24	-28.88	2.51	1.56
	Level 3	-3.16	8.92	-3.86	0.67	-28.75	2.14	2.22
$CH_3OH@C_{60}$	No Fragmentation	-95.30	252.80	-114.15	95.02	-138.12	27.81	137.52
	Level 1	-75.23	206.18	-107.15	91.51	-136.96	26.15	115.26
	Level 2	-99.38	271.85	-127.34	110.19	-142.68	29.44	146.53
	Level 3	-90.62	240.51	-106.23	87.63	-138.48	27.18	130.94
$CH_4@C_{60}$	No Fragmentation	-24.41	59.69	-20.42	18.67	-75.33	10.98	31.87
	Level 1	-23.09	38.85	-26.81	26.11	-74.43	11.08	12.23
	Level 2	-24.70	57.80	-24.04	22.40	-73.74	10.96	30.70
	Level 3	-22.68	55.90	-18.74	17.21	-75.47	10.76	30.58
$H_2S@C_{60}$	No Fragmentation	-25.88	56.24	-30.74	27.06	-86.31	13.02	24.63
	Level 1	-23.48	27.80	-42.77	39.82	-86.11	13.07	-2.79
	Level 2	-28.87	55.22	-38.65	34.10	-86.51	13.61	18.26
	Level 3	-23.59	51.66	-24.44	20.89	-86.91	12.88	22.80
$HCN@C_{60}$	No Fragmentation	-32.97	80.30	-34.66	27.25	-86.68	13.00	36.32
	Level 1	-31.77	62.74	-44.11	37.40	-85.98	13.12	21.10
	Level 2	-31.43	69.16	-37.89	30.80	-86.33	12.91	27.76
	Level 3	-31.09	75.58	-31.67	24.20	-86.68	12.69	34.43

## References

- [1] M. A. Collins, R. P. Bettens, *Chemical Reviews*, **2015**, *115*, 5607–5642.
- [2] E. Masoumifeshani, T. Korona, *Computational and Theoretical Chemistry*, **2021**, *1202*, 113303.
- [3] B. W. Noffke, D. Beckett, L. Li, K. Raghavachari, *Journal of Chemical Theory and Computation*, **2020**, *16*, 2160–2171.
- [4] D. Zhang, *Scientific Reports*, **2017**, *7*, 44645.

## Acknowledgments

This research was partially supported by PL-Grid Infrastructure and by PCSS infrastructure and PRACE-GIP 823767, MNISW DIR/WK/2016/18. The support from the National Science Centre of Poland through grant 2017/27/B/ST4/02699 is gratefully acknowledged.