Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Electronic Supplementary Information

Fast Prediction of Oxygen Reduction Reaction Activity on Carbon Nanotubes with a Localized Geometric Descriptor

Kunran Yang[#], Jeremie Zaffran[#] and Bo Yang*

School of Physical Science and Technology, ShanghaiTech University, 393 Middle Huaxia Road, Shanghai 201210, China

*Corresponding author e-mail: yangbo1@shanghaitech.edu.cn

[#]These authors contribute equally to this manuscript

Content

1. Structures of carbon nanototubes

Carbon nanotubes (CNTs) are obtained from graphene sheets. Graphene is rolled up according to a specific direction vector denoted $m\vec{a}_1 + n\vec{a}_2$, where $\vec{a}_1$ and $\vec{a}_2$ are the lattice vectors of the graphene sheet (see Figure S1). CNTs are denominated after this rolling vector (*n, m*). We mainly distinguish two rolling kinds, armchair and zigzag (see Figure S1). In the armchair type, *m*=*n*, hence armchair CNTs are denoted (*n, m*). However, in the zigzag type, *m*=*0*, and thus zigzag CNTs are denoted *(n, 0)*. CNT diameter *d* can be obtained from the following equation:

$$d = \frac{a}{\pi}\sqrt{n^2 + nm + m^2}, \quad (1)$$

Where *a* is the graphene lattice parameter (0.246 nm)

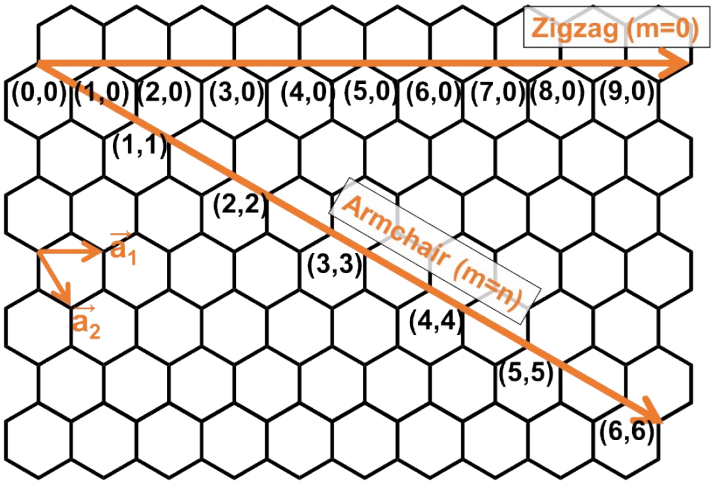

Figure S1. Armchair and zigzag rolling types and CNT (n ,m) denomination. $\vec{a}_1$ *and* $\vec{a}_2$ *are the graphene lattice vectors*

2. Generalization of the Haddon pyramidalization angle

In the pi-orbital axis vector (POAV) theory, it is necessary to identify an axis going through a central C atom π orbital. However, when the system is doped this axis is deviated in comparison with that of the undoped surface, since all the bonds are no more equal to each other after doping (see Figure S2).

For undoped CNT surfaces, the pi axis direction vector as defined by Haddon[1,2] (denoted as $\pi^{Ha\vec{d}don}$) is defined as follows:

$$\pi^{Ha\vec{d}don} = (C_0\vec{C}_1 - C_0\vec{C}_2) \times (C_0\vec{C}_1 - C_0\vec{C}_3) \quad (2)$$

Where C_0 is the C central atom, and C_1, C_2 and C_3 are the three C atoms surrounding the C_0 central atom.

For doped CNT surfaces, the pi axis direction vector (denoted as $\vec{\pi}$) is defined as follows:

$$\vec{\pi} = (C_0\vec{C}_1 - C_0\vec{C}_2) \times (C_0\vec{C}_1 - C_0\vec{X}) \quad (3)$$

Where X is the doping hetero-atom.

Hence, the deviation correction ($\Delta\theta$) between the pi orbital axis in undoped and doped CNT surfaces can be computed as follows:

$$\Delta\theta = arccos\left(\frac{\pi^{Ha\vec{d}don} \cdot \vec{\pi}}{\left|\pi^{Ha\vec{d}don}\right| \cdot |\vec{\pi}|}\right) \quad (4)$$

Where, $\left|\pi^{Ha\vec{d}don}\right|$ and $|\vec{\pi}|$ represent the modulus of the pi orbital axis direction vectors $\pi^{Ha\vec{d}don}$ and $\vec{\pi}$, respectively.

Finally, the pyramidalization angle in doped CNT surfaces (θ_p) is obtained as:

$$\theta_p = \theta_p^{Haddon} - \Delta\theta \quad (5)$$

[1] R. C. Haddon, Chem. Phys. Lett., 1986, 125, 231

[2] R. C. Haddon, Science, 1993, 261, 1545

where θ_p^{Haddon} is the pyramidalization angle in undoped CNTs in its orginial Haddon formulation.[1,2]

Let us highlight that our version of the pyramidalization angle is only a generalization of the original formulation, and not essentially different. That is why. our θ_p angle can be applied either to undoped or doped CNT surfaces.

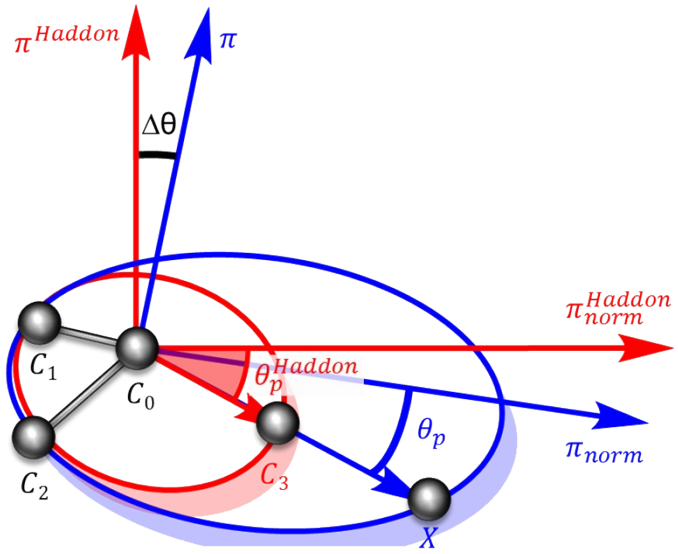

Figure S2. Illustration of the π-orbital axis deviation in undoped (in red) and doped (in blue) CNT surfaces. π^{Haddon} *and* π *denote the pi orbial axes in undoped and doped CNT surfaces, respectively.* π_{norm}^{Haddon} *and* π_{norm} *denote the normal axes to* π^{Haddon} *and* π*, respectively.* θ_p^{Haddon} *and* θ_p *represent the pyramidalization angle in undoped and doped surfaces, respectively, while* $\Delta\theta$ *is the pi orbial axis deviation.* C_0, C_1, C_2 *and* C_3 *represent the C atoms, whereas X denotes the heteroatom dopant.*

3. Surface species preferable adsorption sites

On pure CNT surface, all the C atoms are equivalent, and thus no preferential adsorption site can be identified. However, on N- and B- doped CNT surfaces, the position of the site of adsorption regarding the dopant atom determines the stability of the surface species. Indeed, for B-doped surfaces, we found the most stable configuration is obtained when OH (or OOH) adsorbate is binded to the dopant. To the opposite, on N-doped CNTs surface species preferentially adsorb in ortho position of the N atom (denoted "o" in Figure S4).

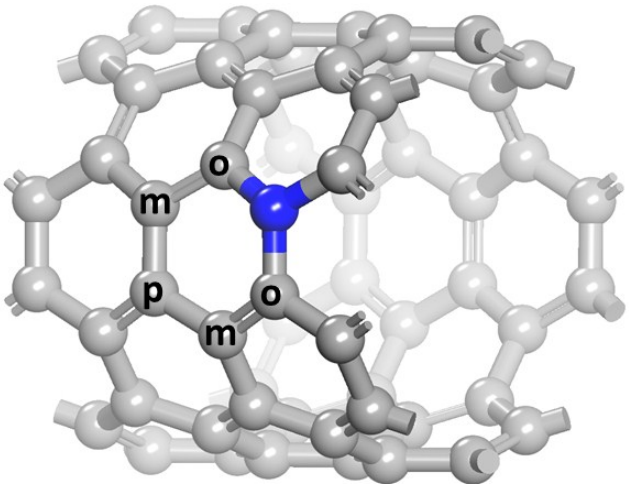

Figure S3. Preferential adsorption sites of ORR species on N- and B- doped CNT surfaces. Blue and grey balls represent dopant and C adsorption sites, respectively. Ortho, meta and para positions regarding the dopant are denoted "o", "m" and "p" respectively.

4. The CNT diameter vs. the pyramidalization angle descriptor

4.1 Influence of the pyramidalization angle on the nanotube diameter

We plotted in Figure S3 the diameters of various CNT structures according to the pyramidalization angle, calculated at the preferential ORR species adsorption site. We found three different trends for undoped, N- and B-doped CNTs. For each kind of those three surfaces, the global tendency is that the nanotube diameter decreases when the pyramidalization angle increases.

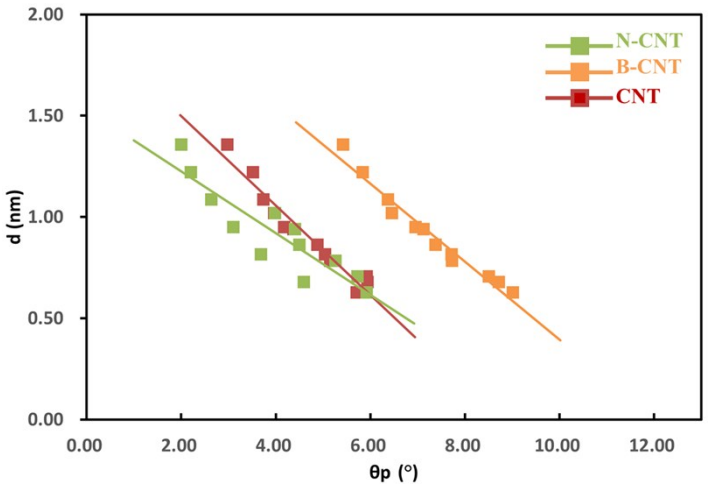

Figure S4. Trends in the evolution of CNT diameters (d) with the pyramidalization angle (θ_p). Color: green, N-doped CNT; orange, B-doped CNT; red, pure CNT

4.2 The CNT diameter as a descriptor of ORR intermediate adsorption energy

We reported in the following adsorption energy of OH and OOH species plotted against the CNT diameter.

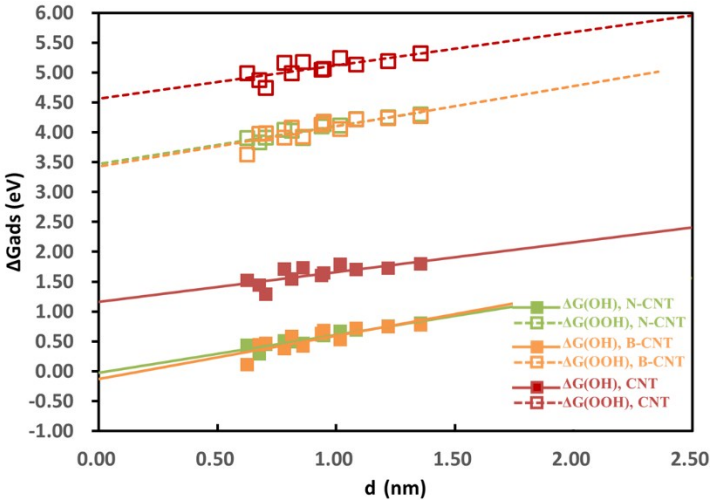

Figure S5. : Adsorption free energies of OH (solid line) and OOH (dashed line) species, plotted versus the CNT diameter (d). Color: green, N-doped CNT; orange, B-doped CNT; red, pure CNT

The statistical quality of Figure S5 plot is presented in Table S1.

Table S1. Statistical parameters given for OH and OOH adsorption energies (ΔGOH and ΔGOOH) vs. the CNT diameter (d) for N-, B- and undoped CNTs. R, MAE and MAX, represent respectively the coefficient of correlation, the mean absolute error (in eV) and the maximal absolute error (in eV).

	ΔG_{OH}			ΔG_{OOH}		
	R	*MAE*	*MAX*	*R*	*MAE*	*MAX*
N-CNT	0.94	0.04	0.12	0.93	0.04	0.12
B-CNT	0.84	0.09	0.22	0.83	0.08	0.22
CNT	0.73	0.08	0.23	0.76	0.08	0.22

5. Basics of statistics

Let us consider two set of variables $Y \in \{y_1, y_2 \ldots y_n\}$ (the "predicted values") and $X \in \{x_1, x_2 \ldots x_3\}$ (the "predicting value"), that are correlated together. It is possible to assess the quality of the predicting model using several quantitative parameters. Readers are referred to the statistic textbook for more information.[3]

- ***The coefficient of correlation, R***

The coefficient of correlation (R) is defined as the square root of the coefficient of determination R^2, with:

$$R^2 = 1 - \frac{\Sigma_i (y_i - \hat{y}_i)^2}{\Sigma_i (y_i - \bar{y}_i)^2} \quad (6)$$

where y_i and $\hat{y}_i$ respective denotes the effective value and the model estimated value of the variable Y.

- ***The mean absolute error, MAE***

The mean absolute error (MAE), is defined as follows:

$$MAE = \frac{\Sigma_i |y_i - \hat{y}_i|}{N} \quad (7)$$

where N is the total number of points in the sample.

- ***The maximal absolute error, MAX***

The maximal absolute error (MAX) is defined as follows:

$$MAX = \max_i (|y_i - \hat{y}_i|) \quad (8)$$

[3] McClave, J. T.; Sincich, T. Statistics; 10 ed.; Prentice-Hall: NJ, USA, 2006.

6. Table of data

Table S2. Calculated data for all of the considered CNTs. d is diameter of a CNT; θ_P is the pyramidalization angle at the active site; ΔG_{OH} and ΔG_{OOH} represent OH and OOH adsorption energies, respectively; ΔU_1 and ΔU_4 represent the reduction potential of ORR step 1 and 4 respectively.

	(n,m)	***d* /Å**	**θ_P /°**	**ΔG_{OH} /eV**	**ΔG_{OOH} /eV**	**ΔU_4 /V**	**ΔU_1 /V**
Undoped CNT	5,5	6.78	5.93	1.44	4.87	1.44	0.05
	6,6	8.14	5.04	1.54	4.98	1.54	-0.06
	7,7	9.49	4.17	1.64	5.06	1.64	-0.14
	8,8	10.85	3.73	1.70	5.13	1.70	-0.21
	9,9	12.20	3.51	1.72	5.18	1.72	-0.26
	10,10	13.56	2.97	1.79	5.32	1.79	-0.40
	8,0	6.26	6.48	1.52	4.98	1.52	-0.06
	9,0	7.05	5.61	1.29	4.74	1.29	0.18
	10,0	7.83	5.23	1.71	5.16	1.71	-0.24
	11,0	8.61	4.86	1.73	5.17	1.73	-0.25
	12,0	9.39	4.25	1.60	5.04	1.60	-0.12
	13,0	10.18	3.97	1.79	5.24	1.79	-0.32
N-doped CNT	5,5	6.78	4.58	0.28	3.83	0.28	1.09
	6,6	8.14	3.67	0.49	4.02	0.49	0.90
	7,7	9.49	3.09	0.59	4.13	0.59	0.79
	8,8	10.85	2.62	0.68	4.22	0.68	0.70
	9,9	12.20	2.19	0.75	4.24	0.75	0.68
	10,10	13.56	1.99	0.80	4.30	0.80	0.62
	8,0	6.26	5.81	0.43	3.90	0.43	1.02
	9,0	7.05	5.70	0.45	3.90	0.45	1.02
	10,0	7.83	5.17	0.51	4.04	0.51	0.88
	11,0	8.61	4.44	0.46	3.90	0.46	1.02
	12,0	9.39	4.34	0.61	4.09	0.61	0.83
	13,0	10.18	3.94	0.66	4.11	0.66	0.81

B-doped CNT	5,5	6.78	8.61	0.44	3.97	0.44	0.95
	6,6	8.14	7.64	0.58	4.07	0.58	0.85
	7,7	9.49	6.90	0.68	4.17	0.68	0.75
	8,8	10.85	6.32	0.71	4.21	0.71	0.71
	9,9	12.20	5.80	0.75	4.23	0.75	0.69
	10,10	13.56	5.39	0.78	4.27	0.78	0.65
	8,0	6.26	8.92	0.10	3.62	0.10	1.30
	9,0	7.05	8.39	0.46	3.98	0.46	0.94
	10,0	7.83	7.62	0.38	3.91	0.38	1.01
	11,0	8.61	7.32	0.42	3.92	0.42	1.00
	12,0	9.39	7.02	0.63	4.13	0.63	0.79
	13,0	10.18	6.40	0.53	4.05	0.53	0.87