

## MATHEMATICAL METHOD FOR SUBMOLECULAR RESOLUTION OF HELICENE-BASED MACROCYCLES BY ATOMIC FORCE MICROSCOPY IN AIR

<sup>1,2</sup>Egor UKRAINTSEV, <sup>3,4</sup>Václav HOUSKA, <sup>3</sup>Jaroslav VACEK, <sup>3</sup>Ivo STARÝ, <sup>3</sup>Irena G. STARÁ,  
<sup>1</sup>Bohuslav REZEK

<sup>1</sup>*Faculty of Electrical Engineering, Czech Technical University, Prague, Czech Republic, EU*

<sup>2</sup>*Institute of Physics, Academy of Sciences of the Czech Republic, Czech Republic, EU*

<sup>3</sup>*Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Prague, Czech Republic, EU*

<sup>4</sup>*Department of Physical and Macromolecular Chemistry, Faculty of Science, Charles University, Prague, Czech Republic, EU*

<https://doi.org/10.37904/nanocon.2019.8492>

### Abstract

We introduce a straightforward mathematical method for improving the AFM image resolution, applied to image analysis of helicene-based macrocycles adsorbed on HOPG. The method reveals structural details from insufficiently resolved AFM images and attributes them to internal structure and ordering of the macrocycles. Our findings are also corroborated by molecular mechanics simulations, validating that the structure provided by the method has lower potential energy compared to other tested macrocycle arrangements.

**Keywords:** Atomic force microscope, mathematical modeling, submolecular resolution, macrocycles, molecular mechanics simulations

### 1. INTRODUCTION

Mathematical data analysis plays a key role in many areas of science. Modeling and simulations are often used to improve the image resolution, sometimes obtained on the instrumentation limit. Recently, the first image of a black hole has been obtained using a mathematical approach [1]. For many years, astronomers have been using multiple images of stars and planets to combine information only from the part of images with best resolution and thus obtain an image with higher resolution than original images [2]. Some methods use a precisely moved low resolution camera to get slightly shifted images which are then merged into a single higher quality image [3,4]. Similar approach can be used for tiny movement or rotation of the object [5]. If the imaging process is too long, the compressive imaging method can be applied [6]. If the measurements are undersampled and must be done too quickly, reconstruction of data can be used [7].

In atomic force microscopy (AFM) or scanning tunneling microscopy (STM) studies, computer models and simulations are also widely employed. For example, a numerical STM/AFM model can take into account the probe relaxation due to the tip-sample interaction. This way, not only the experimentally observed intra- and intermolecular contrasts but also their evolution upon tip approach are described [8]. Computer models can be used to analyze nanoparticle size distribution and estimate the uncertainties of nanoparticle size on rough substrates or non-isolated nanoparticles [9]. Analysis of large datasets was also utilized to evaluate friction and energy dissipation between MoS<sub>2</sub> sample and silicon cantilever [10]. Computer simulations together with optical data can be used to study cell behavior on patterned samples [11].

High resolution AFM/STM microscopy has become an extremely powerful tool for the study of individual molecules. Images with submolecular resolution can be routinely obtained today [12]. However, drawbacks of this method are a very expensive UHV AFM/STM apparatus, long experimental times and difficulties in sample

preparation, especially in the case of large molecules. For these reasons, methods increasing the resolution of ambient AFM/STM images are highly desirable and could circumvent the aforementioned challenges.

Herein, we introduce a straightforward and relatively simple mathematical method for obtaining sub-molecular structural details from ambient AFM images of molecules. The method is demonstrated on images of 4 nm large helicene-based macrocycle molecules adsorbed on highly ordered pyrolytic graphite (HOPG) surface. The initial stage of the proposed method is extraction of an average surface cell, summarizing the prominent structural features and arrangement of the adsorbed molecules. The structural details of the system are highlighted while noise and imperfections in the arrangement are inherently suppressed at this stage. The obtained average unit cell is subsequently analyzed by comparison with a simple 3D surface model obtained from an atomic molecular structure, providing a valuable information about the morphology and orientation of the molecules on the surface.

## 2. EXPERIMENTAL DETAILS

### 2.1. AFM measurement

The synthesis procedure of the helicene-based macrocycles will be presented elsewhere. Its structure was inferred from spectroscopic data (MALDI MS, NMR, IR). For AFM measurements, 20  $\mu$ l of the macrocycle ( $c = 10^{-7}$ g/ml) solution in dichloromethane (VWR, for HPLC 99.8 %, filtered through a microfilter) was drop casted on a freshly cleaved piece of HOPG substrate. Then the sample was used for AFM studies on Dimension Icon AFM (Bruker). The samples were scanned in ambient conditions in Peak Force QNM mode using ultrasharp SLN-B cantilevers (Bruker) with tip radius  $r = 2$  nm. 100 x 100 nm images were obtained with resolution 2000 x 2000 px (20 px/nm) at scanning speed 0.1 Hz. Applied load was 20 pN with peak force amplitude of 30 nm. The obtained images were processed by 1D FFT filtering to remove clear periodical AFM noise and blurring to reduce the overall noise level.

### 2.2. Computer analysis

The computer analysis of AFM experimental data was performed using Matlab software on HP Pro 3400 computer with Intel® Core™ i5-2400 processor and on cluster Ferret5 (CIIRC CTU in Prague) with 12 cores. The geometry of the macrocycle molecule was obtained from a DFT (B3LYP/6-31G/GD3) calculation using Gaussian [13]. Mutual potential energy of two macrocycle molecules placed on HOPG surface was investigated using QuantumATK P-2019 software [14] with ReaxFF CHONi 2015 force field [15].

## 3. RESULTS AND DISCUSSION

### 3.1. Model system

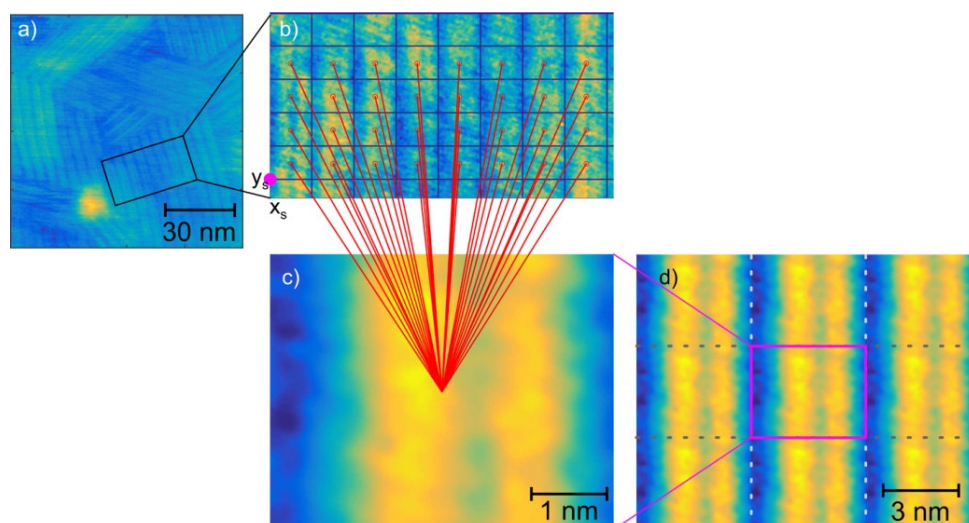
We demonstrate our method on an ambient AFM image of a thin layer of the helicene-based macrocycles adsorbed on HOPG surface (**Figure 1a**). Based on the AFM results the measured height of helicene molecule is 0.4 nm, suggesting that the AFM tip slightly presses the molecule even at quite a low applied force of 20 pN. The layer of molecules consists of parallel stripes organized into domains of 120° relative orientation. It is obvious from the magnified internal structure of the domains (**Figure 1b**) that the stripes are divided into two substripes but no other structural details can be visually distinguished in the image. This can be attributed to the instrument's resolution limit under ambient conditions, considering the RMS roughness in the whole image of 80 pm and noise level of 70 pm.

The basic principle of the method is presented in **Figures 1b to 1c**. It comprises of four consequent steps as described in detail in the following chapters. First, a region of interest is chosen in the original AFM image. This region is divided into a matrix of cells. The dimensions of the unit cell are automatically varied during the

optimization step. Each pixel of specific coordinates from these cells is summarized in the average unit cell and this process is repeated for every single pixel of the average unit cell. The obtained average cell is analyzed by comparison with the geometry model of the studied molecule.

### 3.2. Complete model cell formation

We selected a 37x20 nm<sup>2</sup> rectangular region in the original 100 x 100 nm<sup>2</sup> image (black rectangle in **Figure 1a** and **1b**) with partially visible structural details. It was divided into  $v \times w$  closely packed rectangular cells (8 x 4 = 32 in this case) with a variable cell size of  $m \times n$  pixels (91 x 73 px in this case). The offsets of the cell in the left bottom corner (marked with magenta dot) were  $x_s$  and  $y_s$  (0 px and 40 px in this case). Real AFM height values from  $v \times w$  cells were then summarized (as represented by the red lines) in one point of an average unit cell using Matlab software. The incomplete cells on the edges of AFM images were not taken into account. This procedure was repeated for each point (91 x 73 px = 6643 points) in a model cell to create the average unit cell  $H(x,y)$ , shown in **Figure 1c**. Note that only experimental values were used to create the average unit cell. Subsequently, we copied the average unit cell (purple rectangle in **Figure 1d**) three times in vertical direction and three times in horizontal direction to create a complete model cell (**Figure 1d**).

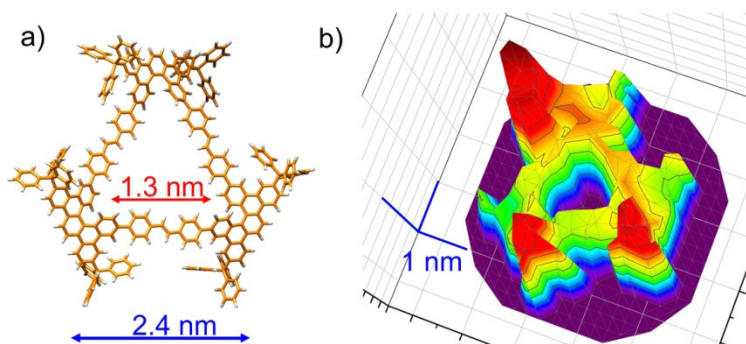


**Figure 1** a) AFM topography image of thin film of the helicene-based macrocycle molecules on HOPG substrate. Z scale=0.4 nm. b) 37 x 20 nm<sup>2</sup> AFM image with 32 cells 91 x 73 px. c) Averaged unit cell, size 91 x 73 px, 4.55 x 3.65 nm. d) Complete model cell, size 273 x 219 px, 13.65 x 10.95 nm. Yellow areas are the highest.

### 3.3. Helicene-based macrocycle and its molecular model

The structure of the helicene-based macrocycle molecule is shown in **Figure 2a**. It is a non-planar molecule with 2 nm height and ~4 nm diameter. For the purpose of the image analysis, we created its molecular model. A 20 x 20 points (5 x 5 nm) mesh was set up and all z values of the mesh were changed to 0 as a background. Therefore, this mesh had 400 points with  $(x_i, y_i, 0)$  coordinates at the beginning. Then we removed all points which were outside a circle of 2.5 nm radius, placed in the middle of the mesh. The macrocycle molecule comprising a total of 408 atoms (carbon and hydrogen) was placed to the center of the mesh. For each atom, its lateral position in the mesh was found and changed from the  $x_i, y_i, 0$  to  $x_j, y_j, z_j$  value, where  $z_j$  was obtained from the DFT optimized structure. This procedure was then repeated for each atom in the molecule. When several atoms were placed in the same lateral position in the mesh, the  $x_j, y_j, z_j$  position of the atom with higher  $z_j$  was taken. As a result, this mesh had 305 points, 184 were points with zero height, while 121 were points with the position of the highest atom, as depicted in **Figure 2b**. Such mesh can be understood as an AFM

image of an individual molecule placed on a flat surface obtained with a very sharp tip (diameter  $\sim 0.25$  nm, cone angle  $0^\circ$ ). This model was introduced in order to keep zero height values not only outside the molecule but also inside it, and to consider the fact that many interior atoms cannot be reached by the AFM tip. **Figure 2b** also shows that for the case of  $20 \times 20$  point mesh there are no holes inside the macrocycle helicenes and trityl groups, so there is no need to take into account molecular orbitals, the knowledge of atoms position is enough.



**Figure 2** a) The top view of the macrocycle molecule structure. b) The top view of the 3D macrocycle molecule modelled by a mesh.

### 3.4. Procedure for complete model cell analysis

To analyze the possible positions and orientations of the molecules in the image, the following automatic data analysis was performed employing the Matlab software. We introduced a sum of differences  $S$ , quantifying the agreement between the atom heights in the AFM complete model cell, characterized by  $H(x,y)$  topography surface, and the molecular model defined by the mesh  $z_j(x_j, y_j)$ .

$$S = \sum_{j=1}^{305} (rz_j(x_j, y_j) - H(x_j, y_j))^2 \quad (1)$$

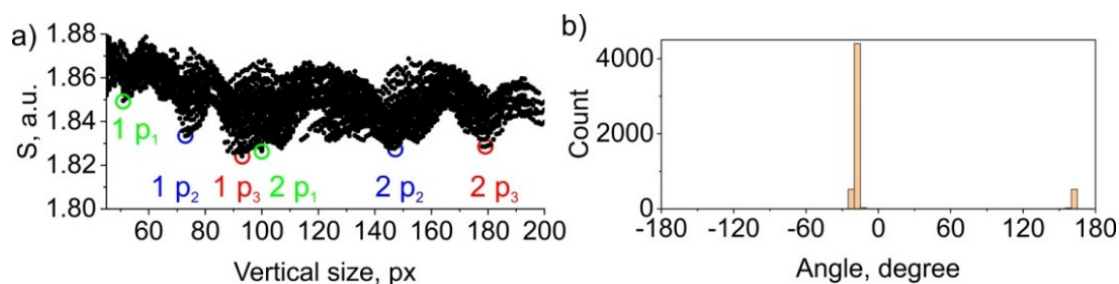
First, the  $z$  height of the macrocycle was linearly adjusted by defining the parameter  $r$  which reflects how the AFM tip compresses the molecule. For our molecules, we used value  $r = 0.0984$ . Next, the molecule was automatically moved across the complete model cell (1681 different positions for each molecule across the central part of a complete model cell, all area inside purple rectangle in **Figure 1d** to cover all possible positions) and rotated (360 different angles with  $1^\circ$  step to cover all possible orientations) to find the minimum  $S$  value. Subsequently, the size of the complete model cell (156 different sizes, 45-200 px, with 1 px step) and the position of the lower left corner of the first cell (39 different positions, 0-10 px in horizontal direction and 0-60 px in vertical direction with 5 px step) were changed in order to find out which complete model cell gives the lowest  $S$  value. Altogether, the  $S$  value was calculated approx.  $3 \times 10^9$  times.

### 3.5. Optimization analysis of the complete model cells for a $37 \times 20$ nm<sup>2</sup> image

Values  $S$  for all the model cells (horizontal size 91 px, vertical sizes 45-200 px, horizontal shift 0-10 px and vertical shifts 0-60 px) are shown in **Figure 3a**. We defined the horizontal size of the complete model cell  $m = 91$  px from the profile taken across AFM image shown in **Figure 1b**. We also performed calculations for horizontal sizes 89 px, 90 px, and 92 px with quite similar results, except a slightly worse alignment of both substripes. At least 16 cells were combined to create the complete model cell.

We identified three groups of vertical sizes with low  $S$  values (local minima). The first group had vertical sizes  $1p_1$  and  $2p_1$ , second group had vertical sizes  $1p_2$  and  $2p_2$ , third group had vertical sizes  $1p_3$  and  $2p_3$ .  $n_{p1} = 51$  px,  $n_{p2} = 73$  px,  $n_{p3} = 93$  px. Sizes  $1p$  correspond to the placement of one molecule into the averaged unit cell. Sizes  $2p$  can be called “second harmonic” or “doubled sized cells”, because these cells are large and two molecules can be placed in such averaged unit cell.

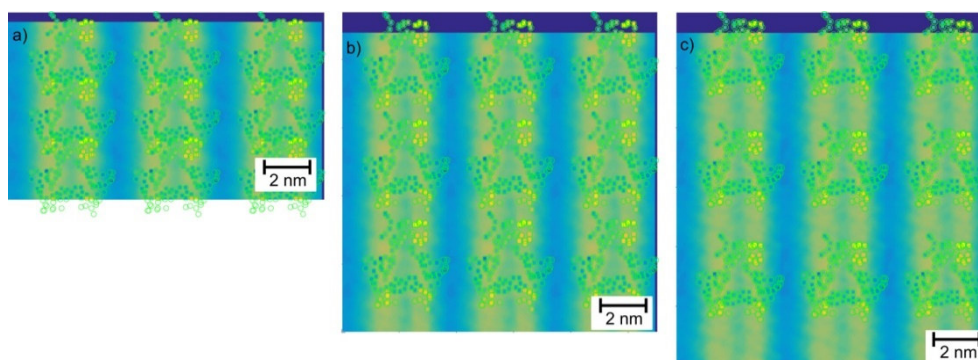




**Figure 3** a) Values  $S$  for complete model cells (horizontal size  $m=91$  px, vertical sizes  $n=45-200$  px, horizontal shift  $x_s=0-10$  px, vertical shift  $y_s=0-60$  px).  
b) Angular distribution of the macrocycles with the lowest  $S$  value.

To get our results looking closer to obtained AFM data, we placed one macrocycle to each of the 8 outside parts of the complete model cell at the positions analogous to the central macrocycle, altogether 9 molecules. In this and all other presented images, only the positions of the highest atoms in the mesh are shown. Points with zero  $z_j$  values are not shown either. Individual atoms are marked with green circles while the inside color scale indicates the height of the atom (yellow corresponds to the highest atom). We choose the same color scale for the molecule and for the complete model cell on purpose. In this case, the worse color visibility of the molecule indicates better height alignment of the molecule.

**Figure 4a** shows that the vertical size  $1p_1$  probably corresponds to the packing of the macrocycle molecules far too dense for the rectangular arrangement ( $180^\circ$  or  $90^\circ$  symmetry),  $n_{p1} = 51$  px. We consider such arrangement as practically impossible. **Figure 4b** shows that the vertical size  $1p_2$  could correspond to the densest packing of the macrocycle molecules with a quite narrow free space,  $n_{p2} = 73$  px = 3.65 nm. We believe that this is one of the most probable arrangements of the macrocycles on HOPG surface with monolayer coverage. **Figure 4c** shows that the vertical size  $1p_3$  (global minimum) corresponds to the packing of the macrocycle molecules with a lot of free space which does not seem to be a realistic configuration either,  $n_{p3} = 93$  px = 4.65 nm.



**Figure 4** a) The complete model cell  $273 \times 153$  px with 9 macrocycles at the same positions as the central macrocycle. b) The complete model cell  $273 \times 219$  px with 9 macrocycles. c) The complete model cell  $273 \times 279$  px with 9 macrocycles

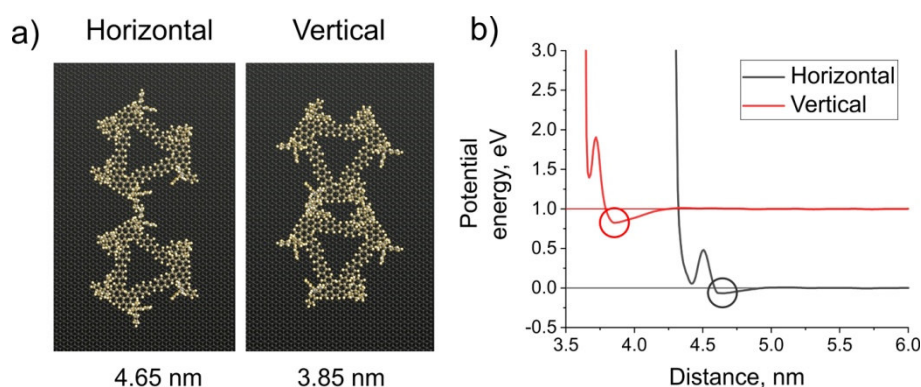
One can see the importance to copy the averaged unit cell for making larger model cell. Otherwise, the migration of the macrocycle outside the averaged unit cell caused the decrease of the  $S$  value and the cell with the lowest  $S$  had the lowest number of the mesh points inside the averaged unit cell.

For each studied model cell (with different  $m$ ,  $n$ ,  $x_s$  and  $y_s$  values) we found one angle of the macrocycle rotation, corresponding to the lowest  $S$  value (between different positions of the macrocycle relative to the

complete model cell and the rotation of the macrocycle), analyzed the angle of rotation for different complete model cells and found two clusters with slightly different orientations of the macrocycle molecules in each cluster (**Figure 3b**). Two clusters ( $-19.7 \pm 1.2^\circ$ ,  $159.9 \pm 0.6^\circ$ ) are separated by  $180^\circ$ , which is basically the same macrocycle orientation. This result is a good indication for only one possible orientation of macrocycles on HOPG surface with possible  $90^\circ$  symmetry.

### 3.6. Molecular mechanics simulations

To corroborate the above results, we investigated relative potential energy of the molecules on a graphite surface by molecular mechanics. **Figure 5** shows two studied arrangements of the macrocycles (horizontal and vertical) and relative energy with respect to the mutual separation between the molecules (center to center). The attraction between the molecules manifests itself in global energy minima before the steep energy rise as the molecules start to collide. The arrangements corresponding to the global minima are shown in **Figure 5a**. In accordance with the Van der Waals character of the interaction, the depth of the energy minima depends mainly on the number of closely interacting atoms of the two molecules. While the horizontal orientation profile shows only a very shallow minimum (one pair of trityl groups is colliding), the vertical orientation is markedly lower in energy (two pairs of trityl groups are colliding).



**Figure 5** a) Horizontal and vertical arrangements studied by molecular mechanics simulations. The corresponding separation distances for those images are written below the images. b) The potential energy vs macrocycle separation. Positions of minima are marked with circles.

### 3.7. Comparison between experimental and theoretical results.

The separation of molecules in the global minimum of the horizontal arrangement (4.65 nm) is similar to the experimental value for distance between the observed stripes (4.55 nm), indicating good agreement between theoretical and experimental results. This agreement also indicates that molecules indeed can form so tightly packed stripes. Similarly, the separation of molecules in the global minimum of the vertical arrangement (3.85 nm) is also in a good agreement with experimentally observed value for distance of the molecules within the stripes (3.65 nm), indicating that molecules can indeed be so tightly packed within the stripes in vertical direction, too. Therefore, the arrangement shown in **Figure 4b** can be considered as an energetically favorable monolayer with the densest packing of the macrocycles on the HOPG surface with  $90^\circ$  symmetry.

## 4. CONCLUSIONS

In this work we present straightforward mathematical method for analyzing ordinary AFM images obtained in ambient conditions on the samples with self-assembled patterns of helicene-based macrocycle molecules. We were able to reveal sub-molecular features from AFM images and we found possible arrangement of macrocycles on HOPG surface. Molecular mechanics simulations found corresponding arrangement with

minimum in potential energy. The presented method can be applied for any molecule of known atomic xyz coordinates and to any height image with periodic pattern and rectangular symmetry. The approach is sufficiently general so that it can be applied also to other properties, not only structural but for instance electronic.

## ACKNOWLEDGEMENTS

***This research was supported by the project “Centre of Advanced Applied Sciences”, No. CZ.02.1.01/0.0/0.0/16\_019/0000778 and the project GAUK No. 818319. The work occurred in frame of the LNSM infrastructure.***

## SUPPORTING INFORMATION

The Matlab script which can perform the presented calculation is available free of charge for non-commercial use. Please contact the corresponding author ([ukraiego@fel.cvut.cz](mailto:ukraiego@fel.cvut.cz)).

## REFERENCES

- [1] JOHNSON, M.D., BOUMAN, K.L., BLACKBURN, L., CHAEL, A.A., ROSEN, J., SHIOKAWA, H., ROELOFS, F., AKIYAMA, K., FISH, V.L. and DOELEMEN, S.S., Dynamical imaging with interferometry, arXiv:1711.01286v1.
- [2] LIM, A.C., JAENISCH, H., HANDLEY, J., FILIPOVIC, M., WHITE, G., HONS, A., BERREVOETS, C., DERAGOPHAN, G., PAYNE, J., SCHNEIDER, M. and EDWARDS, M., Image resolution and performance analysis of webcams for ground-based astronomy, *Proceedings Volume 5489, Ground-based Telescopes*. 2004, <https://doi.org/10.1117/12.549297>.
- [3] PELEG, S., KEREN, D. and SCHWEITZER, L., Improving image resolution using subpixel motion, *Pattern Recognition Letters*. 1987. vol. 5, pp. 223-226.
- [4] KHADEMI, S., DARUDI, A. and ABBASI, Z., A Sub Pixel Resolution Method, *World Academy of Science, Engineering and Technology*. 2010. vol. 70.
- [5] YUE, L., SHEN, H., LI, J., YUAN, Q., ZHANG, H. and ZHANG, L., Image super-resolution: The techniques, applications, and future, *Signal Processing*. 2016. vol.128. pp. 389-408.
- [6] ŽIDEK, K., DENK, O. and HLUBICEK, J. Lensless Photoluminescence Hyperspectral Camera Employing Random Speckle Patterns. *Scientific Reports*. 2017. vol.7, pp. 15309, <https://doi.org/10.1038/s41598-017-14443-4>.
- [7] HAN, G., NIU, Y., ZOU, Y. and LIN, B., Reconstruction of undersampled atomic force microscope images using block-based compressive sensing, *Appl. Surf. Sci.* 2019. vol. 484, pp. 797-807.
- [8] HAPALA, P., KICHIN, G., WAGNER, C., TAUTZ, F.S., TEMIROV, R. and JELINEK, P., Mechanism of high-resolution STM/AFM imaging with functionalized tips, *Phys.Rev.B*. 2014. vol. 90, p. 085421.
- [9] KLAPETEK, P., VALTR, M., NECAS, D., SALYK, O. and DZIK, P., Atomic force microscopy analysis of nanoparticles in non-ideal conditions, *Nanoscale Research Letters*. 2011. vol. 6, p.514.
- [10] CAMMARATA, A., NICOLINI, P., SIMONOVIC, K., UKRAINTSEV, E. and POLCAR, T., Atomic-scale design of friction and energy dissipation, *Phys.Rev.B*. 2019. vol.99, p.094309.
- [11] UKRAINTSEV, E., BROZ, A., HUBALEK KALBACOVA, M., KROMKA, A. and REZEK, B., Stochastic model explains formation of cell arrays on H/O-diamond patterns, *Biointerphases*. 2015. vol. 10, p.041006.
- [12] CHENG, L., PENG, X., ZHANG, S., DENG, K., SHU, L., WAN, J. and ZENG, Q., Selective and competitive adsorption behaviors of guest molecules in donor-acceptor conjugated macrocycles networks on liquid/solid interface, *Appl.Surf.Sci.* 2018. vol. 462, pp. 1036-1043.
- [13] FRISH, M.J., TRUCKS, G.W. at all. Inc., Wallingford CT, 2009. Gaussian 09, Revision E.01.
- [14] SCHNEIDER, J., HAMAEEKERS, J., CHILL, S.T., SMIDSTRUP, S., BULIN, J., THESEN, R., BLOM, A. and STOKBRO, K., ATK-ForceField: a new generation molecular dynamics software package, *IOP Publishing, Modelling Simul. Mater. Sci. Eng.* 2017. vol. 25, p. 085007.
- [15] TAVAZZA, F., SENFTLE, T.P., ZOU, C., BECKER, C.A. and DUIN, A.C., Molecular Dynamics Investigation of the Effects of Tip-Substrate Interactions during Nanoindentation, *J. Phys. Chem. C*. 2015. vol. 119, p.13580.