ChemComm

COMMUNICATION



View Article Online View Journal | View Issue

Check for updates

Cite this: Chem. Commun., 2018, 54, 3715

Received 8th February 2018, Accepted 19th March 2018

DOI: 10.1039/c8cc01134f

rsc.li/chemcomm

Real-space evidence of Watson–Crick and Hoogsteen adenine–uracil base pairs on Au(111)†

Yuanqi Ding, Lei Xie, Xinyi Yao and Wei Xu 🗅 *

From the interplay of high-resolution scanning tunneling microscopy imaging and density functional theory calculations, we show the real-space evidence of the formation of Watson–Crick and Hoogsteen adenine–uracil base pairs on an Au(111) surface with the employment of base derivatives, and further investigate the relative stability of the two types of adenine–uracil base pairs.

DNA and RNA are both essential building blocks for life, which play pivotal roles in biological processes with their abilities to store and reproduce genetic information. The general doublehelix structure of DNA discovered by Watson and Crick¹ and the duplex structure of RNA²⁻⁶ in vivo have shown that both of them are built up on the principle of complementary base-pairing sets of guanine(G)/cytosine(C), adenine(A)/thymine(T) and adenine(A)/uracil(U). These specific bindings rely on the basis of molecular recognition processes where genetic codes are carried and are moreover responsible for the interstrand binding and stabilization of DNA and RNA.1,7,8 Thus, the principle of complementary base-pairing is believed to be the foundation of genetic information. A series of previous studies have been reported to investigate specific molecular recognitions of base pairs by scanning tunneling microscopy (STM), such as adeninethymine quartets,9 guanine-uracil wobble base pair,10 and guaninecytosine Watson-Crick base pair.11,12 As the peculiar base pair in RNA, the adenine-uracil (AU) pair has attracted tremendous interest over the past decades with theoretical investigations¹³⁻¹⁷ and spectroscopy studies¹⁸⁻²¹ performed in solution systems. However, the realspace evidence of the AU base pair has not been demonstrated to the best of our knowledge. Thus, it is of general interest to construct a system under well-controlled ultrahigh vacuum (UHV) conditions with the aim of forming different types of AU base pairs (*e.g.* Watson–Crick and Hoogsteen) on the surface and resolving the atomic-scale structures by STM imaging.

According to our previous studies, molecular recognitions are more efficient between complementary base derivatives than archetypal nucleobases on surfaces.^{11,12} In this study, firstly, we also try to deposit prototypal uracil and adenine molecules on the Au(111) surface, and find out that the uracil and adenine molecules keep their own self-assembled phases (close-packed structure^{22–25} and honeycomb network,^{26,27} respectively) which are separated on the surface, as shown in Fig. 1. Annealing the sample step by step still results in phase separation up to 390 K (the desorption temperature of uracil molecules).

Herein, we choose modified bases, *i.e.*, 9-methyladenine (9mA) molecule and 1-ethyluracil (1eU) molecule together with the noble Au(111) surface as a model system to explore the possibility of molecular recognitions between complementary bases. Both base derivatives are modified at the same sites as those of the natural nucleosides (*cf.* Scheme 1). The Au(111) surface is employed as a template to ensure that the molecules adopt flat adsorption geometries to facilitate the potential base pairing.^{9–12} From the interplay of high-resolution STM imaging and density functional theory (DFT) calculations, we show that (i) the formation of Watson–Crick and Hoogsteen AU base pairs (*cf.* Scheme 1) is feasible on Au(111); (ii) by regulating the



Fig. 1 (a) Large-scale and (b) close-up STM images showing the coexistence of uracil and adenine molecules on Au(111).

Interdisciplinary Materials Research Center, Tongji-Aarhus Joint Research Center for Nanostructures and Functional Nanomaterials, College of Materials Science and Engineering, Tongji University, Shanghai 201804, P. R. China.

E-mail: xuwei@tongji.edu.cn

 $[\]dagger$ Electronic supplementary information (ESI) available. See DOI: 10.1039/ <code>c8cc01134f</code>



Scheme 1 Schematic illustration of Watson-Crick and Hoogsteen AU base pairs. 9-Methyladenine (9mA) molecule and 1-ethyluracil (1eU) molecule are indicated by A and U notations, respectively.

deposition sequence and the ratio of 1eU/9mA to ~1:1 results in the formation of a network structure composed of pure Hoogsteen AU base pairs; (iii) DFT calculations demonstrate that the Hoogsteen AU base pair (0.80 eV per pair) is slightly thermodynamically more stable than the Watson–Crick AU base pair (0.75 eV per pair). Such a study may provide a model system to unravel some biologically relevant issues in RNA duplex, *e.g.*, the relative stability of different kinds of AU base pairs.

After deposition of the 1eU and 9mA molecules on Au(111) at room temperature (RT), respectively, we observed the formation of the corresponding self-assembled structures as shown in Fig. 2. The large-scale STM image (Fig. 2a) shows a well-ordered selfassembled structure of the 1eU molecules on Au(111). From the close-up STM image (Fig. 2b), we identify a single 1eU molecule as a relatively dark ellipse part (*i.e.* the uracil moiety) with a bright round protrusion (*i.e.* tilted ethyl group). On the basis of our



Fig. 2 Self-assembled structures of 1eU and 9mA molecules on Au(111), respectively. (a and c) Large-scale STM images showing the self-assembled structures after deposition of 1eU and 9mA molecules on Au(111) held at RT, respectively. (b and d) Close-up STM images showing more details of the two self-assembled structures where the corresponding DFT-optimized gas phase models are superimposed and the hydrogen bonds are depicted by blue dashed lines. The unit cell of the 1eU self-assembled structures is indicated by white rhomboid, and the 9mA ribbons are separated by dashed wavy lines. H: white; C: gray; N: blue; O: red.

experimental observations, DFT calculations are performed and the most stable DFT-optimized model in the gas phase is superimposed on the STM image (Fig. 2b). As seen in the model, we identify that the 1eU molecules bind to each other via four NH···O hydrogen bonds within the unit cell (as indicated by a white rhomboid), and the molecules interact with each other via van der Waals forces from the tilted ethyl groups in between the unit cells. As for the 9mA self-assembled structure, the large-scale STM image (Fig. 2c) reveals that the 9mA islands are formed by laterally connected zigzag ribbons. From the DFT-optimized gas phase model superimposed on the close-up STM image (Fig. 2d), we identify that, along the ribbon, each molecule binds to the other two via double NH···N hydrogen bonds, and between the ribbons, the molecules interact with each other via van der Waals forces from the methyl groups. It is noted that the desorption of the 1eU and 9mA molecules occurs at 370 K and 400 K on Au(111), respectively.

In order to obtain and investigate AU base pairs, we then introduce the 1eU and 9mA molecules onto the Au(111) surface together. We first deposit the 9mA molecules and then the 1eU molecules, followed by annealing the sample to 330 K resulting in the formation of intercalated structures as shown in Fig. 3. As shown in the large-scale STM image (Fig. 3a), the 1eU molecular strips are inserted into the 9mA islands at the interior and the edge. The close-up STM image (Fig. 3b) allows us to distinguish the Watson-Crick AU base pair and Hoogsteen AU base pair at the 1eU/9mA interface. Upon inserting into the interior of the 9mA islands, the 1eU molecules recognize and match the 9mA molecules laterally on both sides of strips by forming Watson-Crick AU base pairs; in the other case, the 1eU molecules match the 9mA molecules on one side forming Hoogsteen AU base pairs at the edge of the 9mA islands. The corresponding Watson-Crick AU base pairs and Hoogsteen AU base pairs are depicted by white and blue contours, respectively, as shown in Fig. 3b. On the basis of these experimental observations, we build up structural models containing both AU base pairs, and relax by DFT calculations. The most stable DFT-optimized gas phase model is overlaid on the top of the high-resolution STM image as shown in Fig. 3c. Hydrogen bonds are depicted by dashed lines which demonstrate the different types of AU base pairs. The structures remain unchanged upon further annealing up to 370 K (the desorption temperature of 1eU molecules).

Interestingly, when changing the deposition sequence, *i.e.*, we first deposited the 1eU molecules and then the 9mA molecules, followed by annealing the sample to 330 K resulted in the coexistence of the two phases composed of the mixed Watson–Crick and Hoogsteen AU pairs (outlined by white rectangles in Fig. 4a) and the pure Hoogsteen AU pairs forming well-ordered islands as shown in Fig. 4a. Note that the codeposition of the 1eU and 9mA molecules at a ratio of 1:1 to the Au(111) surface leads to similar results. The close-up STM images (Fig. 4b and c) with sub-molecular resolution allow us to identify the alternating 1eU and 9mA molecules from the morphologies. In comparison with the well-analysed structure in Fig. 3, the Watson–Crick and Hoogsteen AU base pairs can be distinguished as depicted and notated in Fig. 4b. The DFT-optimized

ChemComm



Fig. 3 STM images and DFT-optimized structural models of the 1mA selfassembled island with 1eU strips inserted. (a) A large-scale STM image showing the 1mA self-assembled islands with the 1eU strips inserted at the interior and edge (highlighted by black outlines and pointed out by white arrows). (b) A close-up STM image with sub-molecular resolution to distinguish the Watson–Crick AU pairs and Hoogsteen AU pairs (depicted by white and blue contours, and indicated by W–C and H notations, respectively). (c) Highresolution STM image superimposed with the DFT-optimized gas phase models, where the hydrogen bonds are depicted by dashed lines. 1eU strips are highlighted by black dash lines. H: white; C: gray; N: blue; O: red.

gas phase model is superimposed on the corresponding STM image (Fig. 4c). The hydrogen bonds depicted in Fig. 4c allow identification of the two different types of AU base pairs in the structure.

Noticeably, a well-ordered 2D nanostructure is more attractive as it is homogeneous in a relatively large area on the terrace as shown in Fig. 4d. The close-up STM image (Fig. 4e) allows identifying that the structure is composed of pure Hoogsteen AU pairs. Based on the experimental observations, DFT calculations are performed in the gas phase, and the optimized structural model together with the STM simulation (the gray part) are partially superimposed on the STM image (Fig. 4e). The two phases remain coexisting on the surface after further annealing up to the desorption temperature of 1eU molecules at 370 K.

These various structures formed by the Watson–Crick or/and Hoogsteen AU base pairs show the comparable stability of the AU base pairs on the Au(111) surface. To further quantify the relative stabilities of different AU pairs, DFT calculations are performed in the gas phase and the binding energy (E_b) is computed to characterize their relative stabilities. As it is known that the inert Au(111) surface has little influence on



Fig. 4 Adenine–uracil base pairs on Au(111). (a) Large-scale STM image showing the coexistence of the two phases formed by the AU base pairs. One phase is mixed with Watson–Crick and Hoogsteen AU pairs that are outlined by white rectangles. The other phase is composed of pure Hoogsteen AU pairs. (b) A close-up STM image showing the phase composed of the Watson–Crick and Hoogsteen AU pairs (depicted by white and blue contours, and indicated by W–C and H notations, respectively). (c) High-resolution STM image super-imposed with the DFT-optimized gas phase model. (d) A large-scale STM image showing the phase composed by the pure Hoogsteen AU pairs. (e) A close-up STM image partially superimposed with the DFT-optimized gas phase model and the STM simulation (the gray part) at a bias voltage of 1.2 V. One of the Hoogsteen AU pairs is depicted by a blue contour and indicated by the H notation. Hydrogen bonds are depicted by dashed lines. H: white; C: gray; N: blue; O: red.

the hydrogen bonded molecular structures,^{9–12} the interactions between the adsorbed molecules and the substrate are mainly van der Waals forces, which justifies the gas-phase calculation. $E_{\rm b}$ is defined as the total energy of the whole relaxed system minus the individual energies of the relaxed components. That is, $E_{\rm b} = E_{\rm sys} - E_{\rm com}$, where $E_{\rm sys}$ represents the total energy of the AU base pair, and $E_{\rm com}$ represents the individual energies of the 9mA and 1eU molecules. The $E_{\rm b}$ of the Watson–Crick AU base pair is calculated to be 0.75 eV per pair, and the $E_{\rm b}$ of the Hoogsteen AU base pair is 0.80 eV per pair. From the above, it is seen that the Hoogsteen AU pair is slightly more stable than the Watson–Crick AU pair, which is in accordance with the theoretical studies of the archetypal adenine–uracil base pair.^{14,16}

From the interplay of sub-molecularly resolved STM imaging and DFT calculations, we have presented the real-space experimental evidence of Watson–Crick and Hoogsteen adenine–uracil base pairs on Au(111) with the employment of base derivatives, and further revealed their comparable stabilities. Such a surface science study on the essential building blocks of RNA duplex may help to increase our fundamental understanding of basepairing principles in RNA.

The authors acknowledge financial support from the National Natural Science Foundation of China (21473123, 21622307, 21790351).

Conflicts of interest

There are no conflicts to declare.

Notes and references

- 1 J. D. Watson and F. H. C. Crick, Nature, 1953, 171, 737.
- 2 S. Rouskin, M. Zubradt, S. Washietl, M. Kellis and J. S. Weissman, *Nature*, 2014, **505**, 701.
- 3 Y. Ding, Y. Tang, C. K. Kwok, Y. Zhang, P. C. Bevilacqua and S. M. Assmann, *Nature*, 2014, **505**, 696.
- 4 Y. Wan, K. Qu, Q. C. Zhang, R. A. Flynn, O. Manor, Z. Ouyang, J. Zhang, R. C. Spitale, M. P. Snyder, E. Segal and H. Y. Chang, *Nature*, 2014, 505, 706.

- 5 M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna and E. Charpentier, *Science*, 2012, **337**, 816.
- 6 Z. Lu, Q. C. Zhang, B. Lee, R. A. Flynn, M. A. Smith, J. T. Robinson, C. Davidovich, A. R. Gooding, K. J. Goodrich, J. S. Mattick, J. P. Mesirov, T. R. Cech and H. Y. Chang, *Cell*, 2016, 165, 1267.
- 7 G. F. Joyce, Nature, 1989, 338, 217.
- 8 L. E. Orgel, Nature, 1992, 358, 203.
- 9 W. Mamdouh, M. Dong, S. Xu, E. Rauls and F. Besenbacher, J. Am. Chem. Soc., 2006, **128**, 13305.
- 10 W. Mamdouh, R. E. A. Kelly, M. Dong, L. N. Kantorovich and F. Besenbacher, J. Am. Chem. Soc., 2008, 130, 695.
- 11 R. Otero, W. Xu, M. Lukas, R. E. A. Kelly, E. Lægsgaard, I. Stensgaard, J. Kjems, L. N. Kantorovich and F. Besenbacher, *Angew. Chem., Int.* Ed., 2008, 47, 9673.
- 12 W. Xu, J. G. Wang, M. F. Jacobsen, M. Mura, M. Yu, R. E. A. Kelly, Q. Meng, E. Lægsgaard, I. Stensgaard, T. R. Linderoth, J. Kjems, L. N. Kantorovich, K. V. Gothelf and F. Besenbacher, *Angew. Chem.*, *Int. Ed.*, 2010, **49**, 9373.
- 13 P. Hobza, J. Šponer, E. Cubero, M. Orozco and F. J. Luque, *J. Phys. Chem. B*, 2000, **104**, 6286.
- 14 R. E. A. Kelly and L. N. Kantorovich, J. Phys. Chem. C, 2007, 111, 3883.
- 15 B. Herrera and A. Toro-labbé, J. Phys. Chem. A, 2007, 111, 5921.
- 16 S. J. Sowerby, P. A. Stockwell, W. M. Heckl and G. B. Petersen, *Origins Life Evol. Biospheres*, 2000, **30**, 81.
- 17 M. K. Shukla and J. Leszczynski, J. Phys. Chem. A, 2002, 106, 1011.
- 18 A. Müller, F. Talbot and S. Leutwyler, J. Am. Chem. Soc., 2002, 124, 14486.
- 19 Y. Kyogoku, R. C. Lord and A. Rich, J. Am. Chem. Soc., 1967, 89, 496.
- 20 M. C. Chen and R. C. Lord, *Biochim. Biophys. Acta*, 1974, **340**, 90.
- 21 M. Rozenberg, G. Shoham, I. Reva and R. Fausto, Spectrochim. Acta, Part A, 2005, 62, 233.
- 22 S. J. Sowerby and G. B. Petersen, J. Electroanal. Chem., 1997, 433, 85.
- 23 T. Dretschkow, A. S. Dakkouri and T. Wandlowski, *Langmuir*, 1997, 13, 2843.
- 24 A. C. Papageorgiou, S. Fischer, J. Reichert, K. Diller, F. Blobner, F. Klappenberger, F. Allegretti, A. P. Seitsonen and J. V. Barth, ACS Nano, 2012, 6, 2477.
- 25 S. Irrera, G. Portalone and N. H. De Leeuw, Surf. Sci., 2013, 614, 20.
- 26 R. E. A. Kelly, W. Xu, M. Lukas, R. Otero, M. Mura, Y. J. Lee, E. Lægsgaard, I. Stensgaard, L. N. Kantorovich and F. Besenbacher, *Small*, 2008, 4, 1494.
- 27 M. Lukas, R. E. A. Kelly, L. N. Kantorovich, R. Otero, W. Xu, E. Lægsgaard, I. Stensgaard and F. Besenbacher, *J. Chem. Phys.*, 2009, **130**, 024705.