Supporting Information for

Identification of DNA base-pairing via tunnel-current decay

Jin He[†], Lisha Lin[†,‡], Peiming Zhang[†1] and Stuart Lindsay[†,‡,§]

Arizona State University, Tempe, AZ 85287, USA

AUTHOR EMAIL ADDRESS Stuart.Lindsay@asu.edu

[†]Biodesign Institute, Tempe, AZ 85287-5601

[‡]Department of Chemistry and Biochemistry

[§]Department of Physics

General. All reagents were used as received. TLC plates were from EMD Chemical Inc (Gibbstown, NJ). NMR spectra were recorded on the Varian Inova 500 MHz instrument in the Magnetic Resonance Research Center at ASU. Matrix assisted laser desorption ionization time-of-flight (MALDITOF) mass spectra were recorded on a VG TofSpec spectrometer, Proteomics and Protein Chemistry Lab at ASU. The ozone cleaner is an UV Clean #135500 (Boekel Inc.). The thickness measurement was performed on LSE Stokes Ellipsometer (Gaertner Scientific Corp.). FTIR spectra were collected at a grazing angle of 80 °on the Smart SAGA (Specular Apertured Grazing Angle) accessory coupled into a Thermo-Nicolet 6700 FTIR spectrometer that is equipped with a mercury cadmium tellurium (MCT) detector. The entire FTIR system was protected by nitrogen generated from a flow-controlled liquid nitrogen tank.

1. Synthesis and Characterization of thiolated nucleosides

Synthesis of 5'-S-acetyl-*5'-thiodeoxycytidine*[1]

Diethyl azodicarboxylate (0.4ml,) and thiolacetic acid (0.218ml,) was added successively to a solution of deoxycytidine (0.5g, 2.4 mmol) and PPh3 (0.785g, 3mmol)in anhydrous THF (25 ml) under nitrogen in ice bath. The reaction mixture was stirred at room temperature overnight and then concentrated under diminished pressure to produce a residue. The residue was purified on silica gel by column chromatography (eluent: methylene chloride/methanol = 9:1). The product was isolated as a white solid (yield: 15%). ^{1}H NMR (CD_3OD, TMS) δ: 2.08 (s, 3H), 2.09(2.20(m, 1H), 3.65(m, 2H), 3.89(q, 1H), 4.26(q, 1H), 6.13(t, 1H), 7.30(d, 1H), 8.34(d, 1H). MALDI MS m/z: 308.06 ($M+Na^+$) (calcd for 308.31)

Synthesis of 5'-acetyl-5'-thiothymidine[2]

5'-Iodothymidine (250mg, 0.71mmol) was added to a solution of potassium thioacetate (121.4mg, 1.0 mmole) in anhydrous DMF (20 ml). The mixture was stirred overnight at 80 ˚C. The solvent was evaporated under diminished pressure. Crude product was separated by column

chromatography on silica gel with 100% ethyl acetate, providing a white solid (yield: 30%). ^{1}H NMR ($CDCl_3$, TMS) δ: 1.93(s, 3H), 2.089(s, 3H), 2.24(m, 1H), 2.38(s, 3H), 3.28(q, 2H), 4.15(q, 1H), 5.05(t, 1H), 6.23(t, 1H), 7.24(s, 1H), MALDI MS m/z: 365.075($M+Na^+$) (calcd for 365.36)

2. Preparation of functionalized probes STM tips were made from 0.25 mm diameter gold wire (purity 99.999% from Alfa Aesar) by either AC (30V, 4.2 KHz) or DC (2V) electrochemical etching in mixed solutions of concentrated HCl and ethanol (50:50, v/v). The gold tips were immersed in Piranha Solution (3:1 H_2SO_4:H_2O_2 - 30% by volume - *use caution as this solution is extremely caustic and will explode on contact with organics*) for 30 s, rinsed with DI water and dried in an N_2 stream prior to use or modification. For recognition experiments, gold tips were immersed in 0.5 mM thiol derivative of guanine in DMF, 2mM thiol derivative of cytosine in methanol or 1mM thiophenol in methanol for 2 hours or overnight. The tips were rinsed with DMF and ethanol (or methanol alone for the cytosine and thiophenol modification) and blown dry in an N_2 stream before use. Tips for measurements in water were coated with high density polyethylene as described elsewhere.[3] About half of the 8-mercaptoguanine modified tips fail to give recognition signals, the curves appearing identical to those obtained from unmodified probes, indicative of a lack of functionalization at the end of the tip. There is no analytical probe that will address this problem independently, but, in the case of the polyethylene insulated probes, SEM images showed that at least some of the failures were due to excess insulation. Importantly, false positives were never observed with bare or thiophenol modified tips. We conclude therefore that the positive signals are a robust indicator of the chemical composition of the target. We used only those tips that yielded sharp molecular images (c.f. Figure S2) as these were most likely to yield data from single molecules (i-z curves from blunt probes were clearly more complex).

3. Preparation and Characterization of Nucleoside SAMs The thiothymidine monolayer on the gold substrate was prepared as follows: A 1 mM methanolic solution of 5'-S-acetyl-5'-thiothymidine (2 ml) was first treated with Pyrrolidine (0.2 ml) for one hour to remove the thiol protecting group. Then, a freshly hydrogen-flame annealed gold substrate[4] (Agilent, Tempe, AZ) was immersed in this solution

for another hour. The sample was rinsed thoroughly to remove physisorbed molecules. The thiodeoxycytidine monolayer was prepared in the same way in dichloromethane (DCM) or methanol. For the mixed SAM, 1 mM 5'-S-acetyl-5'-thiodeoxycytidine and 1 mM 5'-S-acetyl-5'-thiothymidine solutions were mixed in equal volumes and the procedure described above followed.

The thickness of the nucleoside SAM was measured by ellipsometry at a wavelength of 632.8 nm with an incident angle of 70 degrees. The optical constants of the bare gold substrate were measured before deposition of molecules, yielding n=0.2 and k=-3.53, values slightly different from sample to sample and from bulk gold. The nucleoside SAM optical constants were chosen to be $n_f = 1.5$ and $k_f = 0$. After one hour of deposition, ellipsometry gave a thickness of 0.38±0.1 nm for the 5'-thiodeoxycytidine SAM, 0.50±0.08nm for the 5'-thiothymidine SAM, and 0.38±0.07 nm for the C/ T mixed SAM. These data imply the formation of sub-monolayer coverage on the gold surface.

FTIR clearly demonstrated the presence of nucleosides on gold substrates. FTIR spectra of these nucleoside SAMs are presented in Figure S1, and are divided into two regions: the 700 – 2000 cm^{-1} region and the 2500 – 3600 cm^{-1} region for the sake of clarity. Both the thiodooxycytidine SAM and thiothymidine SAM show characteristic IR bands. The tentative band assignments are listed in Table S1. Without further experiments and theoretical calculations, it is not possible to give definitive assignment to all the vibrations of these nucleosides involved in SAM formation. The FTIR spectra of the SAM generated from the mixed solution of thiothymidine and thiodeoxycytidine are different from those of SAMs composed of individual nucleosides, which indicates formation of a mixed SAM.

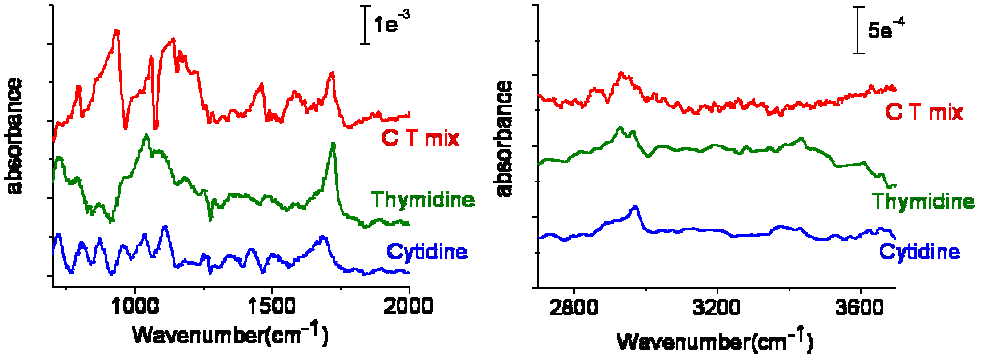

Figure S1. FTIR spectra of nucleoside SAM on the gold substrate

Table S1. Bands observed in FTIR spectra of nucleoside SAMs. (ν = stretching mode, δ = bending mode, s =strong, w =weak)

Thiodeoxycytidine SAM			Thiothymidine SAM		
ν (cm^{-1})	Intensity	Assignment	ν (cm^{-1})	Intensity	Assignment
3383	w	ν(N-H)	3427	w	ν(N-H)
2976	s	ν(C-H)	2963, 2972	s	ν(C-H)
1683	s	ν(C=O),δ(N-H)	1722	s	ν(C=O)
1520	w	Pyrimidine ring	1476	w	Pyrimidine ring
1419	w	Sugar ring	1403	w	Sugar ring

4. STM Imaging.

STM imaging (using a PicoSTM system from Agilent, Tempe) was conducted at a sample bias of 0.5V and a tunneling current of 20~30pA. STM images of Au(111) surfaces modified with thiodeoxycytidine and thiothymidine are shown in Figures S2a and S2b. Figure S2c shows an image of a mixed C/T monolayer taken with a thio-phenol modified Au tip. The contrast enhancement obtained with an 8-mercaptoguanine modified Au tip is shown in Figure S2d. Figures S2c and S2d are adjusted to the same height scale (300 pm) and selected from areas that appeared to be free from topographical features like steps or large pits. Presumably the apparent height is enhanced over deoxycytidine-rich regions. The distribution of

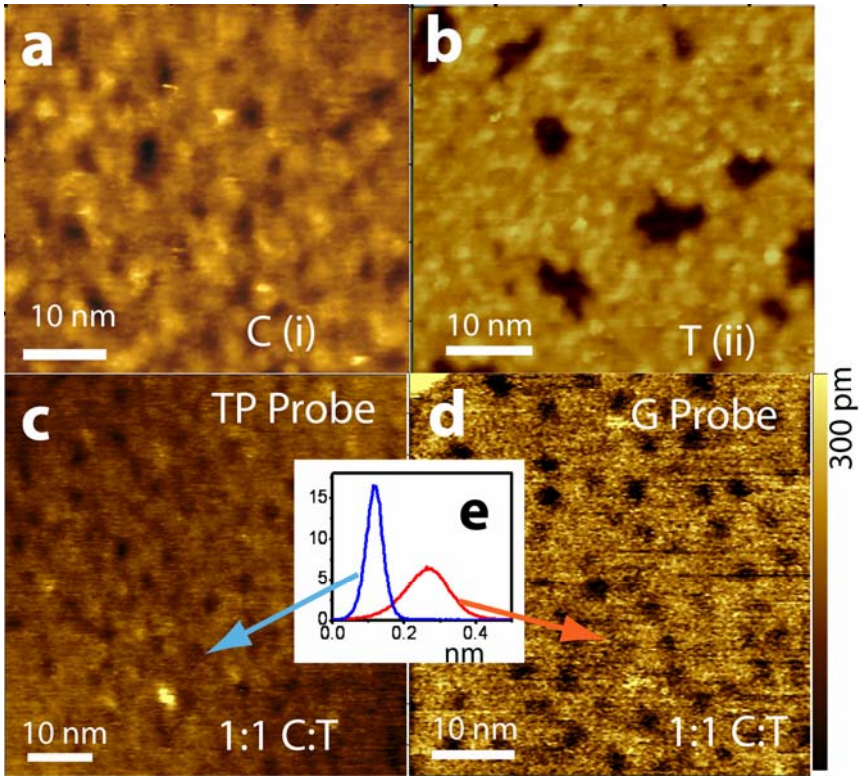

Figure S2. STM images of 5'-thio-cytidine (a) and 5'-thio-thymidine (b) on Au(111) taken in tricholorbenzene (current 0.3 nA, bias 0.5V). Mixed film imaged with a thio-benzene modified Au tip (c) and an 8-mercaptoguanine modified tip (d). Height scale for d and e is shown on right. Distribution of pixel heights is shown in e (red line for image d, blue line for image c).

heights in each image is quantified in the histogram of pixel heights (Figure S2e) and it shows that the brightest features appear to be about 2x higher with the 8-mercaptoguanine (hydrogen-bonding) tip than with the thio-phenol (non-hydrogen bonding) tip. These data confirm the phenomenon reported by Ohshiro and Umezawa,[5] extending the result to these nucleosides.

5. STM i-z measurements

1,2,4-trichlorobenzene was used as the working solution for STM measurements in organic

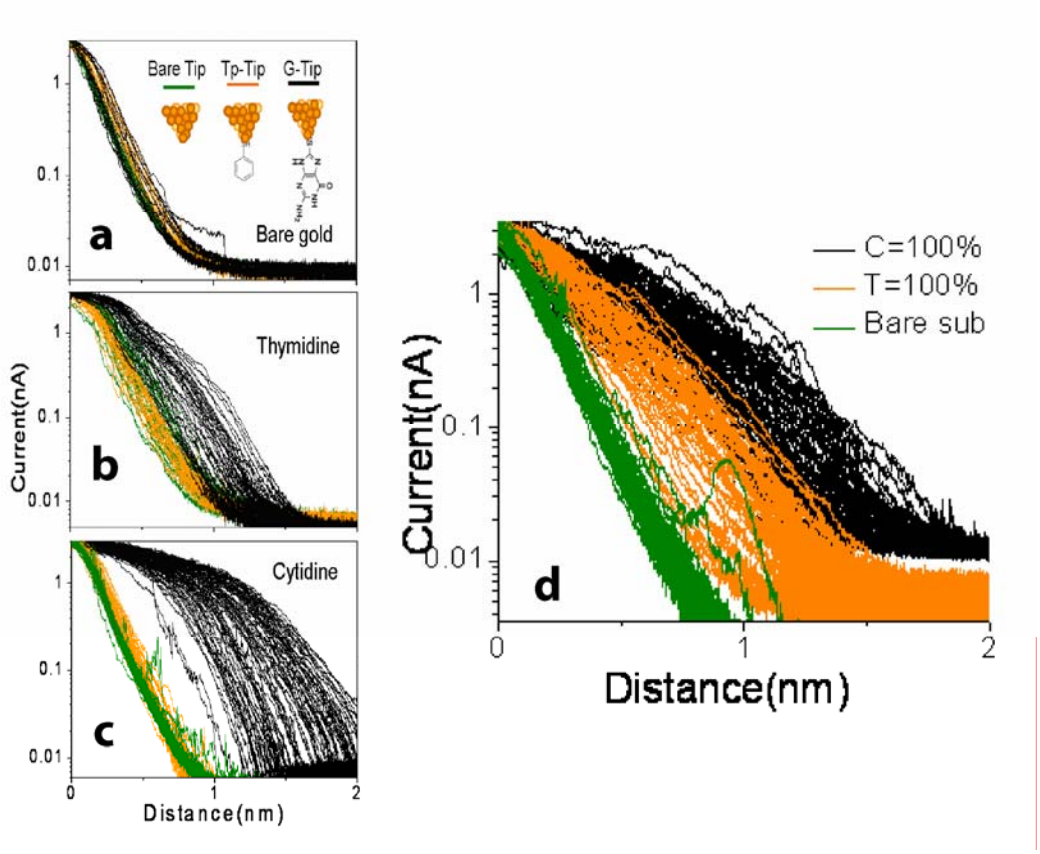

Figure S3: i-z curves on log-linear scales for trichlorobenzene (a – bare gold, b – thymidine SAM and c cytindine SAM). Curves in water are shown in d. Color coding for a,b,c inset in a, for d, inset in d. The baseline currents in (d) are a consequence of leakage currents.

solvents. For i-z measurements, the tip was positioned using a set point of 3nA at a bias of 0.5V (sample positive relative to the tip). The tip was withdrawn 2nm in the z direction over a period 15mS and current data collected. We waited several seconds between repeated sweeps and changed tip location frequently. We found no polarity dependence in experiments in which the potential was reversed. Exmples of i-z curves plotted on a logarithmic current scale are given in Figure S3.

6. Analysis of decay curves

Decay curves in tricholorbenzene (a, b, c) and water (d) are plotted on logarithmic current scales in Figure S3. Taken together with the linear plots in Figures 1 and 4 (main text) these show that the decay consists of an almost linear region followed by an exponential decay. The decay constant analysis of each curve (equations 1a and 1b) was carried out using NI Labview 8.0. Histograms of derived values for β_1 in trichlorobenzene are shown in Figure S4 and for β_2 in Figure S5. Most of the data for β_1 are clustered around 0.1 $Å^{-1}$ (green shaded region) but thymidine samples show a substantial number of slopes that were much steeper (pink shaded region). These are consistent with a ca. 35% failure rate for forming G-T basepairs (compared to an essentially 100% success rate for G-C basepairs, though not all of the G-C basepairs are fully bonded). For β_2, average decays are all similar, though G-C pairs show a number of events with somewhat smaller decay lengths.

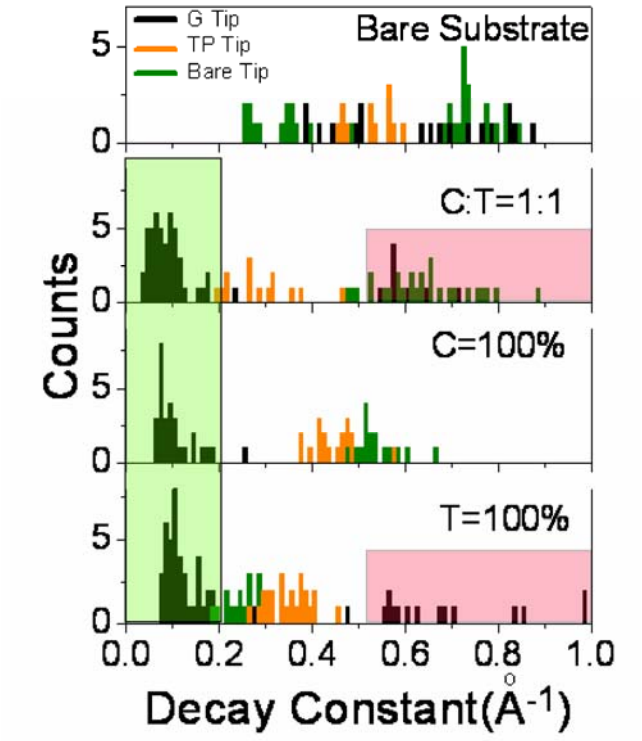

Figure S4: Histogram of fitted values of β_1 for (top to bottom) a bare substrate, an equimolar mix of C and T, 100% C, and 100% T. Tip modification is indicated by the color coding shown in the top panel.

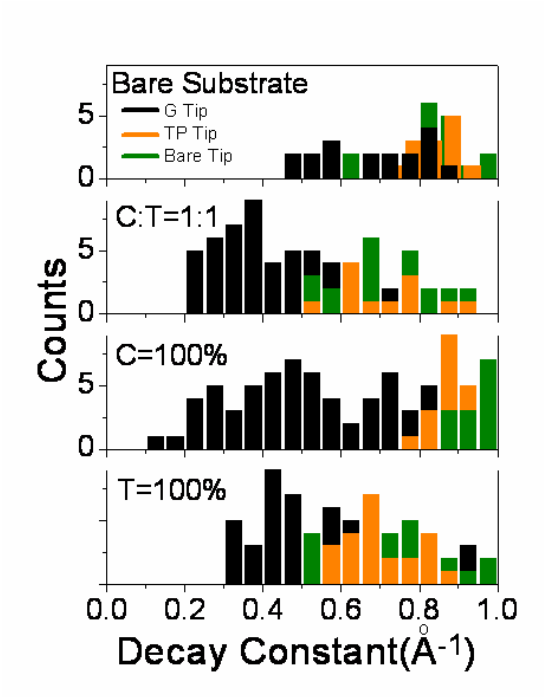

Figure S5: Histogram of fitted values of β_2 for (top to bottom) a bare substrate, an equimolar mix of C and T, 100% C, and 100% T. Tip modification is indicated by the color coding shown in the top panel.

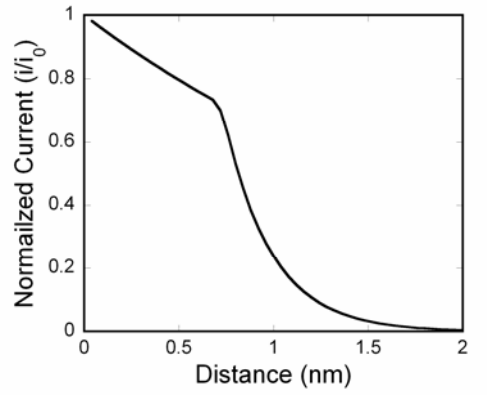

Figure S6: Simulated decay curve for z_c = 0.7nm, β = 0.4Å^{-1} and κ_H/κ_1 = 8.

6. Elastic model of decay curves

We have arbitrarily chosen β = 0.4 Å^{-1} for both the hydrogen bond stretching and the decay in solvent, and taken z_c to be 0.7nm (c.f., Figure 3, main text). A good likeness of the experimental curves is obtained with $\kappa_H = 8\kappa_1$ (Figure S6). If, in fact, the intrinsic value of β for hydrogen bonds is as high as 2Å^{-1} (O.F. Sankey, personal communication) this would imply a spring constant ratio of 40:1. Note that the first decay appears to be linear because of the 'amplification' (equation 2 in the main text) of what is actually a very small motion. Note also that the slope in this region is relatively insensitive to the intrinsic value of β, so the small difference between G:C and G:T basepairs (β_1 in Table 1) may mask a significant electronic difference between the types of basepair.

References for the Supporting Information

1 Henningfeld, K. A.; Arslan, T.; Hecht, S. M., Alteration of DNA Primary Structure by DNA Topoisomerase I. Isolation of the Covalent Topoisomerase I-DNA Binary Complex in Enzymatically Competent Form. *J. Am. Chem. Soc.* **1996,** 118, (47), 11701-11714.

2 Zheng, T.-C.; Burkart, M.; Richardson, D. E., A General and Mild Synthesis of Thioesters and Thiols from Halides. *Tetrahedron Letters* **1999,** 40, 603-606.

3 Visoly-Fisher, I.; Daie, K.; Terazono, Y.; Herrero, C.; Fungo, F.; Otero, L.; Durantini, E.; Silber, J. J.; Sereno, L.; Gust, D.; Moore, T. A.; Moore, A. L.; Lindsay, S. M., Conductance of a biomolecular wire. *Proc. Nat. Acad. Sci.* **2006,** 103, 8686–8690.

4 DeRose, J. A.; Thundat, T.; Nagahara, L. A.; Lindsay, S. M., Gold Grown Epitaxially on Mica: Conditions for Large Area Flat Faces. *Surf. Sci.* **1991,** 256 102-108.

5 Ohshiro, T.; Umezawa, Y., Complementary base-pair-facilitated electron tunneling for electrically pinpointing complementary nucleobases. *Proc. Nat. Acad. Sci.* **2006,** 103, 10-14.