Nanoscale

An Insight into Embryogenesis Interruption by Carbon Nitride Dots: Can They Be Nucleobase Analogs?

Journal:	*Nanoscale*
Manuscript ID	NR-ART-08-2022-004778.R1
Article Type:	Paper
Date Submitted by the Author:	04-Nov-2022

Complete List of Authors:	Zhou, Yiqun; University of Miami, Chemistry Chen, Jiuyan; University of Miami - Coral Gables Campus Kirbas, Emel; University of Miami, Department of Chemistry Zhang, Wei; University of Miami - Coral Gables Campus Gonzalez, Lemay; Miami Dade College Perez, Samuel; Miami Dade College Davila, Arjuna; Miami Dade College Brejcha, Nicholas; Miami Dade College, Gu, Jun; University of Miami, Shi, Wenquan; Yunnan University Domena, Justin; University of Miami, Department of Chemistry Ferreira, Braulio; University of Miami - Coral Gables Campus Zhang, Fuwu; University of Miami, Chemistry Vallejo, Frederic ; University of Miami Miller School of Medicine, Department of Neurosurgery Toledo, Daniela; University of Miami Liyanage, Piumi; University of Miami Graham, Regina; University of Miami School of Medicine, Department of Neurological Surgery Dallman, Julia; University of Miami School of Medicine, Biology Peng, Zhili; Yunnan University, School of Materials and Energy Agatemor, Christian; University of Miami, Chemistry Catenazzi, Alessandro; Florida International University, Leblanc, Roger; University of Miami, Department of Chemistry

An Insight into Embryogenesis Interruption by Carbon Nitride Dots: Can They Be Nucleobase Analogs?

Yiqun Zhou,[1,2,3,§] Jiuyan Chen,[1,§] Emel Kirbas Cilingir,[1] Wei Zhang,[1] Lemay Gonzalez,[4] Samuel Perez,[4] Arjuna Davila,[4] Nicholas Brejcha,[4] Jun Gu,[1] Wenquan Shi,[5] Justin B. Domena,[1] Braulio C.L.B. Ferreira,[1] Fuwu Zhang,[1] Frederic A. Vallejo,[6] Daniela Toledo,[1] Piumi Y. Liyanage,[7] Regina M. Graham,[6] Julia Dallman,[8] Zhili Peng,[5] Christian Agatemor,[1,9] Alessandro Catenazzi,[3,*] Roger M. Leblanc[1,*]

[1] Department of Chemistry, University of Miami, Coral Gables, FL 33146, USA
[2] C-Dots, LLC, Miami, FL 33136, USA
[3] Department of Biological Sciences, Florida International University, Miami, FL 33199, USA
[4] Miami Dade College North Campus, Miami, FL 33167, USA
[5] National Center for International Research on Photoelectric and Energy Materials, School of Materials and Energy, Yunnan University, Kunming, Yunnan 650091, People's Republic of China
[6] Department of Neurological Surgery, Miller School of Medicine, University of Miami, Miami, FL 33136, USA
[7] Indium Corporation of America, Clinton, NY 13323, USA
[8] Department of Biology, University of Miami, Coral Gables, FL 33146, USA
[9] Sylvester Comprehensive Cancer Center, University of Miami, Miami, FL 33136, USA

[§] Both authors contributed equally to this work
* Corresponding authors: acatenazzi@gmail.com (AC), rml@miami.edu (RML)

ABSTRACT: Carbon nitride dot (CND) is an emerging carbon-based nanomaterial. It possesses rich surface functional moieties and a carbon nitride core. Spectroscopic data have demonstrated the analogy between CNDs and cytosine/uracil. Recently, it was found that CNDs could interrupt the normal embryogenesis of zebrafish. Modifying CNDs with various nucleobases, especially cytosine, further decreased embryo viability and increased deformities. Physicochemical property characterizations demonstrated that adenine- and cytosine-incorporated CNDs are similar but different from guanine-, thymine- and uracil-incorporated CNDs in many properties, morphology, and structure. To investigate the embryogenesis interruption at the cellular level, bare and different nucleobase-incorporated CNDs were applied to normal and cancerous cell lines. A dose-dependent decline was observed in the viability of normal and cancerous cells incubated with cytosine-incorporated CNDs, which matched results from the zebrafish embryogenesis experiment. In addition, nucleobase-incorporated CNDs were observed to enter cell nuclei, demonstrating a possibility of CND-DNA interactions. CNDs modified by complementary nucleobases could bind each other via hydrogen bonds, which suggests nucleobase-incorporated CNDs can potentially bind the complementary nucleobases in DNA double helix. Nonetheless, neither bare nor nucleobase-incorporated CNDs were observed to intervene in the amplification of the zebrafish polymerase-alpha 1 gene in quantitative polymerase chain reactions. Thus, in conclusion, the embryogenesis interruption by bare and nucleobase-incorporated CNDs might not be a consequence of CND-DNA interactions during DNA replication. Instead, CND-Ca^{2+} interactions offer a plausible mechanism that hindered cell proliferation and zebrafish embryogenesis originated from disturbed Ca^{2+} homeostasis by CNDs. Eventually, the hypothesis that raw or nucleobase-incorporated CNDs can be nucleobase analogs proved to be invalid.

INTRODUCTION

Nucleobase analogs, molecules that can substitute normal nucleobases in nucleic acids, may lead to altered base pairings and subsequently affect DNA replication and gene transcription,[1] which disrupts cell growth and biological development, and may promote cancer.[2] Meanwhile, nucleobase analogs are important antimetabolites in cancer treatment.[3] The most famous nucleobase analogs include 5-bromouracil, 2-aminopurine, 6-mercaptopurine, and acycloguanosine. Among them, 5-bromouracil is mutagenic because it has a similar structure to thymine so that it can pair with adenine. But it can also change shape and form base pairs with guanine.[1] During DNA replication or gene transcription, the corresponding nucleoside, 5-bromodeoxyuridine, once incorporated in DNA, can cause transitions from AT to GC.[4] Similarly, 2-aminopurine is an adenine analog that usually pairs with thymine, but it can also pair with cytosine, resulting in both AT-to-GC and GC-to-AT transitions.[5] In addition, 2-aminopurine possesses a high fluorescence quantum yield (68%) in an aqueous solution that can be significantly reduced when incorporated into nucleic acids.[6] This sensitivity to the surroundings can be utilized in the studies of structure and dynamics within DNA and RNA, and electron transfer within DNA. Furthermore, 6-mercaptopurine is a hypoxanthine analog, and it can inhibit purine nucleotide synthesis.[1] Acycloguanosine is a guanine analog, and it is one of the most commonly used antiviral drugs. Some nucleobase analogs such as 5-fluorouracil can also inhibit DNA synthesis by inhibiting certain enzymes.[1] Overall, the development of nucleobase analogs is becoming increasingly important in molecular biology, biochemistry, medicine especially in oncology, and DNA nanotechnology.

However, an outstanding question is whether nanoparticles (NPs) can serve as nucleobase analogs. Carbon dots (CDs) are a group of recently discovered, carbon-based spherical NPs with a mean diameter smaller than 10 nm.[7] Many CDs are smaller than 5 nm with a broad size distribution among the CDs produced by Leblanc's group,[8-10] and some of them are comparable to that of a nucleotide unit 3.3 Å (0.33 nm) in the 22-26 Å (2.2-2.6 nm) wide DNA chain.[11] Regarding structure, CDs are often compared to traditional quantum dots (QDs) and hypothesized to have a core-shell nanostructure whose studies are actively carried out.[12] Depending on the precursors and dopants, the CD shells are composed of a variety of functional moieties,[7] such as carboxyl, carbonyl, hydroxyl, and amine groups that endow CDs with the potential to mimic nucleobases. In addition, CDs are well characterized for high photostability, water solubility, biocompatibility, nontoxicity, and photoluminescence (PL).[8, 13] And the highest fluorescence quantum yield (QY) of CDs according to extensive literature studies has reached up to 93.3%.[14] Due to these excellent properties, CDs are versatile and their applications are rapidly rising in various nanotechnology fields,[9, 15-17] which results in more profound discoveries such as penetration of the blood-brain barrier (BBB) and cell membranes,[18, 19] localization in cell nuclei and mitochondria,[20, 21] and inhibition of tumors, fungi, bacteria and viruses.[22, 23] Considering the nanoscale size of CDs, it is not surprising that CDs can penetrate membranes of cells and organelles but it is unknown what CDs do inside cell nuclei and how they interfere with cellular activities and biological development. We can expect embryogenesis when DNA replication and gene expression are actively taking place to be especially sensitive to the interference by CDs.

Relatively little is known about the interactions between CDs and DNA/RNA. Because DNAs bound to positively charged CDs, Feng et al. speculated that hydrophobic interactions, electrostatic interactions, and van der Waals forces, could take place between DNA nucleobases or backbone, and CDs.[24] The electrostatic interaction between CDs and DNA proved to be significant for DNA conformation or structural transition. Later, Khakbaz and Mahani observed single-stranded DNA was adsorbed on CDs by π-π interaction to quench the fluorescence of CDs.[25] Meanwhile, Godavarthi et al. assumed that lone pairs of electrons from oxygen and nitrogen in passivated CDs bound with the nucleobases by means of hydrogen bonds which helped in visualizing the DNA.[26] Subsequently, in 2018, Cheng et al. reported that CDs derived

from m-phenylenediamine showed a unique RNA affinity that could be attributed to the isoquinoline moieties formed during the CD synthesis and amines on the surface of CDs, which could bind RNA through π-π stacking and electrostatic interaction, respectively.[27] In addition, Khan et al. hypothesized that CDs could bind with 2-hydroxyl groups of RNA through chemical bonding.[28] Furthermore, it was presented by Jhonsi et al. that CDs with antimicrobial activity derived from tamarind and calf thymus DNA could interact with DNA via intercalation.[29] Jhonsi et al. also summarized the two modes of how DNA binds with antimicrobial agents: intercalation, which is stabilized electronically in the helix by π-π stacking and dipole-dipole interactions, or groove binding, which is stabilized by electrostatic, hydrophobic and hydrogen-bonding interactions. Thus, the predominant interactions between CDs and DNA/RNA include π-π stacking, electrostatic interaction, and hydrogen bonding.

Although the CD-DNA/RNA interaction studies have gained much attention lately, such studies are incomplete because: (1) the structures of most CDs are still unknown, (2) the studied CDs were mostly positively charged and the predominant interaction between positively-charged CDs and negatively-charged DNA/RNA backbone is electrostatic force,[30, 31] (3) previous studies emphasized the superficial CD-DNA binding without considering the role of CDs in DNA replication and gene transcription following CD-DNA binding, when the DNA structures are relatively unstable and vulnerable to external mutagens, and (4) nucleobases are basic constituents of DNA/RNA and have been used to synthesize CDs. However, such CDs are more often developed for the purpose of sensing,[32, 33] but their advantages over other CD species have not been fully revealed especially in nucleic acid analogy. Recently, Kokic et al. reported that as the only FDA-approved drug for the treatment of COVID-19 patients, the active form of remdesivir acts as a nucleoside analog and inhibits the RNA-dependent RNA polymerase (RdRp) of coronaviruses including SARS-CoV-2.[34] According to Kokic et al., remdesivir is incorporated by the RdRp into the growing RNA product and allows for the addition of three more nucleotides before RNA synthesis stalls.[34] Additionally, many FDA-approved commercial anticancer agents such as gemcitabine, doxorubicin, epirubicin, idarubicin, and temozolomide were all reported with a DNA-associated mechanism of action.[35-39] And it is worth noting that gemcitabine is also a deoxycytidine nucleoside analog.[40] Thus, we are wondering if CDs with or without surface modifications could serve as nucleobase or nucleoside analogs similar to remdesivir and gemcitabine? In addition, to fill the gap in understanding the CD-DNA interactions, this study is novel in the following aspects: (1) the structure of our focal CD species is well characterized; (2) the CDs are negatively charged, in contrast to most studies investigating positively charged CDs; (3) this study focuses on CD-DNA interactions during DNA replication and gene transcription; (4) nucleobases are used during CD synthesis but not for sensing purposes.

In addition, if CDs are nucleobase analogs or can intervene in DNA replication and gene transcription during susceptible periods of embryogenesis, contact with CDs might damage DNA, proteins, and cell membranes, and thus interfere with cell division, proliferation, differentiation, and embryonic development. In terms of cell type, damages to cancerous cells will be more significant than to normal cells due to the enhanced permeability and retention effect. In order to study the effect of CDs on embryogenesis, we used a zebrafish model that is genetically similar to humans.[41] Therefore, it has become one of the most popular animal models for studying developmental processes.[42, 43] Zebrafish was used as an *in vivo* model in this study considering several advantages over the mouse model: (1) zebrafish do not require large space and also they are more cost-effective to keep than mice colonies;[44] (2) adult zebrafish breed approximately every 10 days and can lay 50 to 300 eggs at a time.[45] In contrast, mice in general produce a litter of up to 10 pups but can only bear around five litters for the whole lifetime. Therefore, the zebrafish model is more helpful than the mouse model for this study because zebrafish can produce a large

number of offspring to help repeat experiments with multiple replicates to obtain robust results; (3) zebrafish embryos and larvae are nearly transparent, which allows researchers to observe the real-time development of tissues and any fluorescently-labeled activity in zebrafish body.[46] In contrast, mouse embryos are not translucent and naturally develop inside the mother, so it is impossible to observe live embryo development and fluorescently-labeled activities as seen in zebrafish.

In this study, CDs were studied for their capacity to mimic nucleobases by observing zebrafish embryonic development, cell division and proliferation, and DNA replication in presence of CDs. Carbon nitride dots (CNDs) were used as a CD model because they feature a clearer core structure and were synthesized by following a well-established protocol.[47] Also, CNDs have exhibited a great application potential in drug delivery across the blood-brain barrier and brain tumor-specific targeted therapy. In addition, CNDs were modified by different nucleobases, namely adenine (A), cytosine (C), guanine (G), thymine (T), and uracil (U), to better mimic nucleobases and common surface modification practice since CDs, in general, are often conjugated to nucleobase/nucleoside analog drugs such as the aforementioned gemcitabine.[48] Various spectroscopic and microscopic characterizations were performed to confirm the success of different nucleobase modifications. To investigate if CNDs could mimic nucleobases and intervene in the routine cell growth and biological development, natural and nucleobase-incorporated CNDs were first applied to zebrafish embryo media to study their effects on embryogenesis. Of particular interest were their effects on embryo viability, morphology, and post-embryonic development. Subsequently, *in vitro* studies including cell viability and cellular uptake were conducted on cancerous and normal cell lines to understand their impacts on cell proliferation and cell nuclear targeting capability, respectively. Later, in order to investigate whether bare and nucleobase-incorporated CNDs interfere with DNA replication, quantitative polymerase chain reactions (qPCR) were conducted to compare the amplification rates of the zebrafish polymerase-alpha 1 gene.

EXPERIMENTAL SECTION

Materials. Adenine (≥ 99%), cytosine (≥ 99%), guanine (98%), thymine (≥ 99%), uracil (98%), citric acid (≥99.5%) and urea (≥99.5%) were purchased from Sigma-Aldrich (St. Louis, MO); Deionized (DI) water with 18.2 MΩ·cm resistivity and surface tension of 72.6 $mN \cdot m^{-1}$ was obtained from a Modulab 2020 water purification system (San Antonio, TX) at room temperature (20.0 ± 0.5 °C). All reagents were used without modification.

Characterization Methodologies and Instruments. UV/vis absorption spectroscopy was performed using a Cary series UV-Vis Spectrophotometer (Agilent Technologies, USA) with a 1 cm quartz cuvette (Starna Cells). Fluorescence emission spectroscopy was conducted in aqueous medium (1-cm pathlength quartz cuvette) with a Fluorolog-3 (Horiba Jobin Yvon, USA) with a slit width of 5 nm for both excitation and emission. Fourier-transform infrared (FTIR) spectra were obtained with air as background using a FITR spectrometer (Frontier, PerkinElmer, USA) equipped with attenuated total reflectance accessories. Raman spectra were recorded using an InVia (Renishawin) equipment with a 785 nm laser source. X-ray photoelectron spectroscopy (XPS) was applied with a spectrometer (Thermo Fisher Scientific, USA). The operational details include X-ray source Mono Al Kα, with a $h\nu$=1486.6 eV photon energy operated at 6 mA×12 kV. Thermogravimetric analysis (TGA) and derivative thermogravimetry (DTG) of bare and nucleobase-incorporated CNDs were analyzed using a thermo-microbalance (TG 209 F3 Tarsus, Netzsch, USA) while heating under the protection of a flow of nitrogen gas from 40 to 1000 °C at

a rate of 10 °C/min. A nano series Malvern Zetasizer was utilized to perform the zeta potential measurements. Atomic force microscopy (AFM) was applied using an Agilent 5420 atomic force microscope (5420, Agilent Technologies, USA) with one drop of each CND solution (1 mg/mL) placed on a clean silica mica slide and imaged with tapping mode. Transmission electron microscopy (TEM) images were obtained using a JEOL 1200X TEM with a drop of each CND solution sonicated for 5 min, deposited on a carbon-coated copper grid, and dried with air before the TEM screening. It is noteworthy that the 5-min ultrasonication applied prior to both AFM and TEM screening was to avoid self-aggregation of CND particles. A Malvern Zetasizer (nano series) was applied for the zeta potential measurements.

Syntheses of Nucleobase-incorporated Carbon Nitride Dots. To synthesize nucleobase-incorporated CNDs, we followed similar procedures as of the synthesis of bare CNDs.[9] In detail, 0.5 g of citric acid, 0.5 g of urea, and 0.5 g of each distinct nucleobase were mixed, dissolved in 25 mL of PBS (pH: 7.4) solution and stirred overnight. Following the stirring procedure, each mixed sample was microwaved for seven minutes at 700 watts. Once cooled down, each sample was named as A-CND, C-CND, G-CND, T-CND, U-CND, and added with 25 mL of deionized (DI) water and ultrasonicated for 1-2 min, which was followed by ultracentrifugation twice at 3,000 RPM for 30 minutes. The supernatant solutions were collected and filtered using a 5-mL syringe and a 25-mm syringe filter with 0.2 μm cellulose acetate membrane. Then each sample was concentrated by a rotary evaporator and dialyzed using 1 kDa dialysis membrane. The dialysis process lasted for three days with DI water changed every 12-24 hours. Eventually, five samples were lyophilized 2 days into solid state.

Zebrafish Incubation. Wild-type zebrafish embryos were obtained from the Zebrafish Core Facility at University of Miami. Aqueous solutions of bare and nucleobase-incorporated CNDs at 1 and 100 μg/mL were prepared to incubate zebrafish embryos since 0 hpf; a control group without any CNDs was also implemented. For the zebrafish embryo viability study, 10 embryos were observed in each group and the viability test was repeated three more times. The incubated zebrafish embryos were subjected for observation using a ZEISS fluorescence microscope and a Leica SP5 confocal microscope under white light. Animal care protocols for all procedures involving animals in this study were approved by the University of Miami Animal Care and Use Committee. Protocols are also in compliance with the guidelines of the National Science Foundation.

CND-Ca^{2+} Interaction Study. To study the interactions between different CNDs and Ca^{2+}, 0.2 and 1.1 mg of $CaCl_2$ (>96%) bought from Sigma-Aldrich (St. Louis, MO) were independently added into 5 mL of each CND solution (bare CNDs, A-CNDs, C-CNDs, G-CNDs, T-CNDs, and U-CNDs) at 1 μg/mL to achieve different concentrations of Ca^{2+} (0.3 and 2.0 mM). Then, fluorescence spectroscopy was performed on all the samples. PL emission intensities were recorded at 456 nm upon 375 nm excitation.

Cell Viability Study. Human mesenchymal stem cells (MSCs) were obtained from Thermo Fisher Scientific (Waltham, MA). Pediatric GBM cell line SJGBM2 was obtained from Children's Oncology Group. MSCs were cultured in MEM alpha supplemented with heat-inactivated 20% FBS and 1% penicillin/streptomycin. SJGBM2 cells were cultured in RPMI-1640 supplemented with heat-inactivated 10% FBS and 1% penicillin/streptomycin. Cell lines were grown in a humidified incubator kept at 37 °C and 5% CO_2 and routinely tested for mycoplasma using LookOut mycoplasma PCR detection kit (Sigma). Viability was determined using CellTiter 96 AQueous One Solution Cell Proliferation Assay, a colorimetric MTS assay kit (Promega). MSC

and SJBGM2 cells were seeded into 96-well plates using media containing 5% FBS at a density of 5000 and 10,000 cells per well, respectively, and 24 h later drug treated. After 72 h, media was aspirated and 100 µL of a 1:5 solution of MTS to cell culture media was added to each well and incubated for 1-4 h and absorbance was measured at 490 nm using BioTek Synergy HT plate reader. Experiments were repeated at least 3 times. Cell viability data were analyzed using Student's t-tests for all pairwise comparisons of the different treatments that were tested.

Human prostate cancer cells PC-3 and mouse breast cancer cell 4T1 were plated in 96-well plates ($5x10^3$ cells/well) in complete RPMI-1640 medium (10% fetal bovine serum and 1% penicillin/streptomycin), respectively. Cells were incubated at 37 °C in a humidified atmosphere containing 5% CO_2 atmosphere. The medium was replaced with a fresh medium 24 h after seeding. Culture medium was replaced with serial dilutions of different CNDs (100 µL) in fresh medium and incubated for 24 h. The medium was then replaced with 100 µL medium containing 0.5 mg/mL of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium (MTT). After 3 h incubation, the medium containing unreacted MTT was carefully removed, and the obtained blue formazan crystals were dissolved in 100 µL DMSO. The absorbance at 570 nm was measured by SpectraMax i3x. The cell viability was calculated based on the relative absorbance to untreated cells (100% cell viability) and control medium (no cells, 0% cell viability).

For the purpose of bioimaging, HPNE-KRAS and ASPC-1 cells were cultured in serum free DMEM media supplemented with 10% fetal bovine (Thermofisher Scientific, Waltham, MA). All cell lines were routinely tested for mycoplasma using LookOut mycoplasma PCR detection kit (Sigma Aldrich, St. Louis, MO) according to the manufacturer's instructions and were maintained at 37 °C in a humidified 5% CO_2 incubator. Then, the HPNE-KRAS and ASPC-1 cells were plated in 24 well plates at a density of $1X10^5$ cells/well. After the incubation for 24 h at 37 °C, cells were treated with A, C, G, T, U-incorporated CNDs at 1 mg/mL for 1 hour. It is worth noting that the concentration applied was higher than those used for viability measurements to enhance bioimaging. After the treatment, cells were washed with PBS and treated with 1 µg/mL Hoechst for 15 min. Then, cells were washed with PBS 2 more times and fixed with 4% formaldehyde for 20 min. Once the fixation was completed, 100 µL of citiflour antifading agent was added on top of the cells for 1 hour. Finally, cells were washed with PBS 3 times to wash off the antifading agent and examined by fluorescence microscopy with an Invitrogen EVOS M7000 Imaging System (Thermo Fisher Scientific, Waltham, MA USA).

Quantitative polymerase chain reactions in presence of bare and nucleobase-incorporated CNDs. We compared amplification of the zebrafish polymerase-alpha 1 gene (Gene ID: 317740) in presence of bare and different types of nucleobase-incorporated CNDs by using a SYBR green qPCR assay. We extracted DNA from three pooled samples of 5-day old zebrafish embryos using a commercial kit (gMAX Mini Genomic DNA Kit, IBI Scientific, Iowa, USA), and then diluted the three DNA extracts to the same concentration of 2.80 ng/µL. We used the highest concentration, non-diluted DNA extract to generate serial dilutions from 3.82 to 0.000382 ng/µL. Our calculated quantities related to the quantity of polymerase alpha-1 gene originally present in this 3.82 ng/µL DNA extract. To amplify the polymerase-alpha 1 gene, we used the primers (5'→3'): forward: TTAAGGGCAAAGGTGACGAG and reverse: CCTTCAAAACCGCTACAAGC. 25 µL volume assays were prepared with 0.75 µL of each primer at 10 µM, 6 µL of water or CND-spiked water, and 5 µL of the DNA sample. We compared eight treatments at both low (10 µg/mL) and high concentrations (100 µg/mL): water (control), a solution containing bare CNDs, a solution containing all the five nucleobase-incorporated CNDs, and one solution each for A-CNDs, C-

CNDs, G-CNDs, T-CNDs, and U-CNDs. A quantification SYBR green with melt curve protocol was run on a QuantStudio 3 qPCR system (Applied Biosystems). Examination of melt curves showed the presence of a single amplicon at Tm = 82.3±0.2 °C for all the samples and serial dilutions. To analyze the effects of treatment and concentration, a GLM was performed with Gaussian family and identity link function, where quantity indicates treatment + CND concentration after omitting the two outliers (because there were no differences among the three samples, we did not consider sample as a random factor in a mixed GLM).

Statistical analysis. Variations in cell viability between CND treated groups and the control were analyzed using a one-way analysis of variance (ANOVA). The results are presented as the mean ± SEM (*p<0.05, **p<0.01, ***p<0.001). In addition, the zebrafish embryo viability results are cumulative.

RESULTS AND DISCUSSION

Can CNDs serve as nucleobase analogs? Based on our previous studies,[9, 49] bare CNDs possess an average particle size of 2 ± 1 nm that is close to a nucleotide unit measuring 3.3 Å (0.33 nm), and the width of the DNA chain measuring 22-26 Å (2.2-2.6 nm). Regarding the structure, the FTIR spectra of bare CNDs and different nucleobases (**Fig. S1**) are presented and compared. The nucleobases display characteristic functional groups in their structures, including carbonyl groups, sp^2 nitrogen and carbon corresponding to the purine and pyrimidine structure[1]. Similarly, most of these chemical bonds and functional groups can be observed in different CD species (**Fig. S1**). However, when the FTIR spectra of CDs that originate from different synthesis pathways are compared to that of each nucleobase, CNDs show a better overlap with different nucleobases in vibration/stretching bands than the other three CD species developed in our group called Y-CDs, G-CDs, and B-CDs. To the best of our knowledge, CNDs are commonly prepared from a cyclic polymerization of citric acid and urea, which can form pyrimidine and purine in the structure as different nucleobases.[50]

When comparing the solid-state ^{13}C NMR spectra of bare CNDs and different nucleobases (**Fig. S2),** we observed three peaks at 100, 163, and 180 ppm in the spectrum of bare CNDs, which can be assigned to a 2-hydroxypyrimidine structure that is similar to that of cytosine and uracil according to a table of typical chemical shifts in ^{13}C NMR spectra (Chemguide). Thus, it suggests 2-hydroxypyrimidine structures in bare CNDs, which is generally consistent with our previous study[51]. Peaks at 35, 43, and 75 ppm for bare CNDs represented $-CH_3$, $-CH_2-$ or -CH=, and C-O, respectively. And the minor chemical shifts could be ascribed to the deshielding of the surrounding electronegative atoms, such as O in C-O. Thus, FTIR and ^{13}C NMR spectroscopies preliminarily confirmed our hypothesis on the analogy between the bare CNDs and nucleobases, especially cytosine and uracil, which will be further studied in the future.

Syntheses and purification of nucleobase-incorporated CNDs. For the purpose of positive control for the bare CNDs, CNDs were modified with five nucleobase species. Also, CDs are generally applied as drug nanocarriers so this practice will be able to assess the risks of CND-based drug delivery systems *in vitro* and *in vivo*. The molecular weight of different nucleobases, namely adenine, cytosine, guanine, thymine, and uracil, are 135, 111, 151, 126, 112 g/mol, respectively. Considering the molecular weight of bare CNDs and various nucleobases are around 750 and 100-150 Da, respectively, after dialysis with a dialysis tubing (molecular-weight cut-off: 1000 Da), the substances that remained in the dialysis tubing after three days' dialysis are various nucleobase-incorporated CNDs. Considering low water solubilities, especially for guanine (insoluble), thymine and uracil, and the lack of active surface functional moieties, each type of

nucleobase was added as a third precursor during the synthesis instead of the post-synthesis surface modification process. Meanwhile, nucleobases were separately mixed with citric acid and urea that underwent the same treatment for the synthesis of CNDs, but no CNDs were acquired, which motivated us to follow the same synthetic approach while wrapping nucleobases into the matrix covalently or noncovalently. In addition, to increase the solubility of nucleobases, PBS was used instead of deionized water as the solvent during the CND syntheses.

Optical properties. In comparison to that of bare CNDs,[9] the UV/vis absorption spectra of various nucleobase-incorporated CNDs (**Fig. S3A**) do not differ at the four peaks at 242, 279, 338 and 399 nm that correspond to C=C, C=N, C=O and NO_2 moieties, respectively.[9, 18] Additionally, the excitation-dependent fluorescence spectra of nucleobase-incorporated CNDs (**Fig. S3B-F**) are similar to that of bare CNDs with similar maximum excitation and emission wavelengths of ~375 and ~450 nm, respectively.[49] And the similar excitation-dependent PL behaviors of different CNDs were likely to be caused by the similar surface states of different fractions. Thus, modifications with various nucleobases do not affect these optical properties of CNDs, and nucleobases are hypothesized to be simply incorporated in the structure of CNDs with noncovalent interactions. Nonetheless, the full width at half maximum (FWHM) of CNDs was reduced by about 15 nm by various nucleobase modifications (**Table S1**). FWHM, the width of an optical signal at half of its maximum intensity, is an important property in the determination of color representation after light is modulated by color filters and polarizing layers. Also, a decrease in FWHM could be due to a reduction in the vibrational energy levels in the valence band,[51] which might result from the presence of abundant intact and uniform nucleobase molecules in different nucleobase-incorporated CND structures.

Structural analyses. FTIR analyses show that the purine structure was found in A-CNDs and G-CNDs as in their precursors, and the pyrimidine rings were found in C-CNDs, T-CNDs, and U-CNDs (**Fig. 1**). Thus, as hypothesized previously, different nucleobase structures, especially the purine and pyrimidine rings, remained intact together with the s-triazine structures formed in the syntheses of different CNDs, which supports our claim that CNDs are effectively modified by the five nucleobases. The formation of triazine structures was confirmed with the peaks at around 1700 (C=N stretching) and 740 cm^{-1} (breathing mode of s-triazine rings) typically found in carbon nitride materials.[52, 53] According to the IR Spectrum Table & Chart (Sigma-Aldrich), the 3200 and 1770 cm^{-1} peaks correspond to amide N-H and C=O stretching, respectively. The peaks at 1185 and 1070 cm^{-1} indicate C-N stretching in most CNDs except G-CNDs. Furthermore, the 3400, 1590, 1340 and 1185 cm^{-1} bands are likely due to aliphatic primary amine stretching, amine N-H bending, C-N stretching of aromatic amines, and regular amine N-H stretching at the fingerprint region, respectively. And the peak at 2770 cm^{-1} may be ascribed to O-H stretching, which suggests the availability of -COOH groups in the CNDs' structures. In addition, the peak at around 3060 cm^{-1} suggests the existence of sp^2 carbons in the structure of various nucleobase-incorporated CNDs. At last, the band observed at 2390 cm^{-1} can only be attributed to a few isocyanates (-N=C=O stretching) present in different nucleobase-incorporated CNDs except G-CNDs. Therefore, FTIR analyses have confirmed the presence of five nucleobase species in the carbon nitride matrix. However, due to the poor water solubility of guanine, there are relatively fewer complex structures formed around guanine in G-CNDs compared to the other nucleobase-incorporated CNDs.

Due to the interference of strong autofluorescence, the original Raman spectra (**Fig. S4**) of different nucleobase-incorporated CNDs did not show clear D (~1350 cm^{-1}) and G (~1580 cm^{-1})

characteristic peaks seen in many carbon dot studies[54]. Even after baseline subtraction, all the Raman spectra exhibited weak D and G bands.

With A-CNDs as an example, the XPS full spectrum (**Fig. S5A**) shows that the sample is mainly composed of C (285 eV), N (400 eV) and O (531 eV), and a small amount of impurities (Na^+ and PO_3^-). According to the XPS periodic table (Thermo Fisher Scientific), the high-resolution C 1s spectrum (**Fig. S6A**) can be deconvoluted into three peaks at 284.6, 286.3 and 288.0 eV, corresponding to C-C, C-O-C and O=C-O, respectively. According to **Fig. S7A**, the N 1s peak center at 400 eV denotes N bound to a graphene lattice as C=N that has been confirmed by FTIR spectroscopy. The high-resolution O1s spectrum (**Fig. S8A**) also can be deconvoluted into three peaks at 530.8, 532.4 and 535.5 eV, corresponding to O=C-O/C=O, C-O-C/C-O-H, -OH, respectively. The XPS spectra of the other four nucleobase-incorporated CNDs are similar to that of A-CNDs. Although the nucleobases were different, there is no appreciable difference in the elemental composition of the five nucleobase-incorporated CNDs. Nonetheless, in comparison to bare CNDs,[51] the variations in XPS atomic percentages and binding energies (BE) are sufficient to confirm surface modifications (**Table S2-4**).

Furthermore, different functional moieties can be identified and quantified by the TGA based on their own decomposition temperatures. The TGA (**Fig. 2A**) shows a similar thermal decomposition trend among different nucleobase-incorporated CNDs. It is worth noting that there were some residues with certain mass percentages after decomposition of different nucleobase-incorporated CNDs. It is due to the use of PBS as a solvent during the syntheses of different CNDs and some Na^+ and PO_3^- salts were trapped in their structures, which has been demonstrated previously by the XPS full spectra (**Fig. S5**). In addition, it is observed that the remnants (Na^+ and PO_3^- salts) of C-CNDs and U-CNDs are less than the others, which indicates higher purities of C-CNDs and U-CNDs. Furthermore, the DTG (**Fig. 2B**) reveals that: (1) due to few phase transitions in the TGA, DTG have to be smoothened by the software, Origin; (2) the sharp peak at 508 °C in the DTG of C-CNDs suggests a rapid decomposition, which implies that a few cytosine molecules were trapped in the core of C-CNDs, which prevented cytosine from decomposition until the shell of C-CNDs were completely decomposed; (3) with the DTG of bare CNDs as a reference,[49] decompositions of different nucleobase-incorporated CNDs are categorized into 5 decomposition stages. To be specific, the peaks between 40-122 °C indicate loss of water molecules adsorbed on the surface of or trapped in different CND structures. The mass loss at 122-168 °C shows the loss of water molecules generated through dehydration condensation reactions with alcoholic -OH.[55] The stage between 168-338 °C can be further divided into two substages at 168-250 and 250-338 °C. The stage at 168-250 °C can be attributed to the decomposition of edge-plane oxygen-containing functional moieties such as epoxy, carboxyl (-COOH) and carbonyl (C=O) groups into phenol groups while releasing CO and CO_2.[55, 56] The stage at 250-338 °C is ascribed to the decompositions of some relatively stable oxygen-containing functional moieties and sublimation of a few carbon frameworks.[55] The stage at 338-448 °C demonstrates the decompositions of amines with release of ammonia gas.[57, 58] And the stage at 448-1000 °C is due to the decomposition of graphene and heptazine-based carbon nitride structures;[59-61] (4) with the mass of adsorbed water and remnant at the end of thermal decomposition taken into consideration, the mass loss at each stage of different CNDs was recalculated by dividing the mass loss recorded by TGA by the sum of actual mass loss without that of adsorbed water and remnant. After recalculation, the actual mass loss of each functional moiety on different CNDs is listed in **Table S5**. In comparison, the content of alcoholic -OH among different CNDs is arranged in the order of C-CNDs (3%)>bare CNDs (2%)=A-CNDs=T-CNDs>G-CNDs (1%)=U-CNDs; the

content of various amines are settled in the order of T-CNDs (53%)>U-CNDs (51%)>G-CNDs (49%)>A-CNDs (38%)>C-CNDs (31%)>bare CNDs (20%); the content of oxygen-containing functional moieties and connected carbon frameworks are arranged in the order of A-CNDs (35%)=T-CNDs>U-CNDs (29%)>C-CNDs (28%)>G-CNDs (26%)>bare CNDs (20%); the content of core that means graphene and heptazine-based carbon nitride structures is arranged in the order of bare CNDs (58%)>C-CNDs (38%)>A-CNDs (25%)>G-CNDs (24%)>U-CNDs (19%)>T-CNDs (10%). Thus, on the contrary, the content of shell in different CNDs is settled in the order of T-CNDs (90%)>U-CNDs (81%)>G-CNDs (76%)>A-CNDs (75%)>C-CNDs (62%)>bare CNDs (42%). Most importantly, the content of hydrophilic functional moieties including alcoholic -OH and amines on the shell of different CNDs was calculated, which is arranged in the order of G-CNDs (66%)>U-CNDs (64%)>T-CNDs (61%)>bare CNDs (57%)>C-CNDs (55%)>A-CNDs (53%). Oxygen-containing functional moieties are not included in the calculation of hydrophilic functional moiety surface content because of the difficulty in separating carbon framework from oxygen-containing functional moieties such as -COOH and C=O. According to the calculation, A-CNDs and C-CNDs possess lower amounts of hydrophilic functional moieties than bare CNDs, which demonstrates lower hydrophilicity of A-CNDs and C-CNDs than the other CNDs.

Morphological studies. The AFM images (**Fig. 3**) show the mean particle sizes of A-CNDs, C-CNDs, G-CNDs, T-CNDs, and U-CNDs along the z-axis are 3.5, 2.0, 2.0, 2.0, and 3.0 nm, respectively. However, due to a limited number of particles taken into account, the AFM results have to be confirmed with other morphology analytical techniques. The TEM images (**Fig. 4**) show the mean particle sizes of A-CNDs, C-CNDs, G-CNDs, T-CNDs, and U-CNDs along the x-y plane are 1.8 ± 0.6, 1.9 ± 0.5, 2.1 ± 0.5, 2.1 ± 0.5, and 2.1 ± 0.5 nm, respectively. Taken together, AFM and TEM data suggest that different nucleobase-incorporated CNDs are spherical with a diameter of 2-3 nm that is similar to that of bare CNDs.[51] However, our finding that particle sizes are larger than a nucleotide unit 3.3 Å (0.33 nm) complicates our hypothesis that CNDs are typical nucleobase analogs.

Zeta potential measurements. Zeta potential values of nucleobase-incorporated CNDs are similar to that of bare CNDs between -20 to -30 mV (**Table S6**).[49] However, among the five types of nucleobase-incorporated CNDs, A-CNDs and C-CNDs possess lower zeta potential absolute values, suggesting a tendency to aggregate in water. In other words, A-CNDs and C-CNDs are more amphiphilic than the other three CND species, which matches our findings from TGA. More significantly, amphiphilicity benefits the penetration of A-CNDs and C-CNDs through lipid layers such as cell membranes.

Above all, various physicochemical property characterizations demonstrate: (1) although differences are not clear across nucleobase-incorporated CNDs, A-CNDs and C-CNDs are more similar compared to the others in terms of fluorescence FWHM, hydrophilicity, particle size, and surface charge; (2) nucleobase molecules remain intact in the structures of nucleobase-incorporated CNDs and they were likely to be incorporated in the carbon nitride matrix noncovalently; (3) the poor water solubility of guanine hindered the formation of complicated structures in G-CNDs.

Effect of bare and nucleobase-incorporated CNDs on zebrafish embryogenesis. According to **Fig. 5**, when bare and various nucleobase-incorporated CNDs were applied to incubate zebrafish embryos at 0 hpf, bare CNDs at a lower concentration (1 μg/mL) showed the least effect on embryo viability within 48 hpf that was close to that of the control (~100%). In contrast, when a higher concentration (100 μg/mL) of bare CNDs was applied, the embryo viability was largely

declined to 70% at 48 hpf. The increased cytotoxicity could only be induced by the interactions between bare CNDs and diverse biomolecules in cells and not by oxidative damages since bare CNDs were unable to induce reactive oxygen species.[49] In addition, modifications with various nucleobases enhanced the toxicity of CNDs. It is noteworthy that compared to bare CNDs at 1 μg/mL, different nucleobase-incorporated CNDs at 1 μg/mL exhibited enhanced toxicity. Among them, the C-CNDs exhibited the highest toxicity (35%) closely followed by the U-CNDs (30%), which might be due to a higher purity of these two samples than the others confirmed by TGA. The toxicity was more pronounced when the concentration of nucleobase-incorporated CNDs was increased to 100 μg/mL. To be specific, in comparison to bare CNDs at 100 μg/mL with a toxicity of 30%, the C-CNDs showed a slightly enhanced toxicity of 35%, while the others were similar to bare CNDs. Moreover, the C-CNDs showed the highest toxicity at both concentrations (1 and 100 μg/mL) among different groups at the same concentrations.

In addition, a few deformities were observed during the zebrafish embryogenesis. For instance, at 48 hpf, more evident abnormal developments were observed in the zebrafish embryos incubated in a high concentration (100 μg/mL) of A-CNDs and C-CNDs, such as tail detachment (**Fig. S9**). For another example, cell adhesion reduction was observed with embryos (8 hpf) in most groups except the control and bare CNDs at 1 μg/mL (**Fig. 6**). And the ratios of embryos exposed to different CNDs showing reduced cell adhesion are summarized in **Table S7** according to **Fig. S10**. In comparison to the control, most treatment groups showed an estimated 10-30% odd except bare CNDs at 1 μg/mL, which again demonstrates the low toxicity and low pathogenicity of CNDs at low concentrations. However, both toxicity and pathogenicity can be amplified with high concentrations and nucleobase modifications. Additionally, Ca^{2+} homeostasis and cell adhesion proteins such as cadherins named for calcium-dependent adhesion play crucial roles in vertebrate development.[62, 63] Especially, cadherins, a class of type-1 transmembrane proteins, are important in the formation of adherens junctions allowing cells to adhere to each other and dependent on Ca^{2+} to function.[64] Hence, in this study, it is likely that Ca^{2+} uptake, cadherin formation and functions such as Ca^{2+} binding in the early stage of zebrafish development were obstructed by bare CNDs at a high concentration and all the nucleobase-incorporated CNDs regardless of the concentration. In order to verify this hypothesis, Ca^{2+} (0.3 and 2.0 mM) was added to different CND solutions at 1 μg/mL. The concentrations of Ca^{2+} applied were selected considering the fact that zebrafish continually face changes in environmental Ca^{2+} concentrations between 0.01–3 mM and the plasma Ca^{2+} level is maintained at around 2.5 mM in vertebrates.[63] A low concentration (1 μg/mL, one concentration to incubate zebrafish embryos) of different CND solutions was applied considering the maximum PL emission intensity limit. As a result, with the addition of Ca^{2+}, no peak shift was observed in the fluorescence emission spectra of different CND solutions (**Fig. S11**). Nonetheless, interestingly, different patterns were observed in the change of PL emission intensities of different CND solutions. As displayed in **Fig. 7**, the Y axis indicates $F_0/F-1$. It suggests that if $F_0/F-1>0$, the PL emission is reduced after the addition of Ca^{2+}. On the contrary, if the value is negative, the PL emission is enhanced. In comparison, it is observed that bare CNDs were not significantly affected by Ca^{2+} even at a high concentration of 2.0 mM so it was used as a standard for comparison. However, as for different nucleobase-incorporated CNDs, with the addition of Ca^{2+} at 0.3 mM, both A-CNDs and C-CNDs showed PL emission reduction while G-CNDs, T-CNDs and U-CNDs exhibited PL emission improvement, suggesting two major different CND-Ca^{2+} interaction modes for A-CNDs and C-CNDs (electron/energy transfer), and G-CNDs, T-CNDs, and U-CNDs (chelation).[65] These results are consistent with high contents of surface oxygen-containing functional moieties for A-CNDs, C-

CNDs and high contents of surface hydrophilic functional groups such as -OH and -NH_2 for the rest three nucleobase-incorporated CNDs. However, after the Ca^{2+} concentration was increased to 2 mM, the F_0/F-1 value for C-CNDs remained positive while it became negative for A-CNDs, which demonstrated a change in the A-CND-Ca^{2+} interaction from electron/energy transfer to chelation. In contrast, at a higher concentration of Ca^{2+}, all the G-CNDs, T-CNDs and U-CNDs demonstrated a gradual transition in the CND-Ca^{2+} interaction from chelation to electron/energy transfer although the F_0/F-1 of U-CNDs was still negative. In fact, it is not difficult to understand it from a viewpoint of the structure of different nucleobase modifiers. In brief, except for adenine, all the other four nucleobases, including cytosine, guanine, thymine, and uracil, have negatively charged carbonyl groups and all the nucleobases remain structurally intact in the modified CNDs according to the optical property studies and FTIR analyses. Thus, all the nucleobase-incorporated CNDs at 1 μg/mL can interact with Ca^{2+} at two major interaction modes, which could disturb the Ca^{2+} homeostasis, leading to an abnormality in embryogenesis.

Furthermore, the impacts of CNDs on zebrafish embryogenesis can be assessed by measuring the body length and eye size of zebrafish embryos (48 hpf) by following a previous study[66]. It is observed in **Fig. 8** that both body length and eye size of the zebrafish embryos treated with different CNDs are generally smaller than those of the control group. And, among different treatment groups, the group with U-CNDs displays the smallest values in body length and eye size, which demonstrates a significant impact of U-CNDs on post-embryonic development. In addition, the inconsistency between the embryonic and post-embryonic developments demonstrated a higher impact of CNDs on the early stages of embryogenesis.

Effect of bare and nucleobase-incorporated CNDs on single cells. To elucidate the effect of bare and nucleobase-incorporated CNDs on cell viability and hence embryogenesis, four cell lines were treated with increasing concentrations of bare and five different nucleobase-incorporated CNDs. In detail, human mesenchymal stem cells, MSCs, pediatric glioblastoma cells, SJGBM-2, human prostate cancer cells, PC-3, and mouse breast cancer cells, 4T1, were subjected to increasing concentrations of bare CNDs, A-CNDs, C-CNDs, G-CNDs, T-CNDS, and U-CNDs from 10-100 μg/mL. Since the cytotoxicity of bare CNDs has been studied with MSCs and SJGBM-2 cell lines in our previous studies,[9, 48] it is worth noting that relevant data are not presented in **Fig. 9**.

In terms of cell viability, differences in cytotoxicity were observed across the four cell lines. To be specific, according to **Fig. 9**, with the PC-3 cell line, A-CNDs at 100 μg/mL showed the highest toxicity. As for the 4T1 cell line, it was most affected by C-CNDs at 100 μg/mL. However, with the concentration increased from 100 to 500 μg/mL, both cancer cell lines were impacted the most by C-CNDs as demonstrated in **Fig. S12**. Interestingly, the SJGBM-2 cell line was not significantly impacted by any of the CND groups across all concentrations tested. In contrast, MSC cells experienced a significant loss in cell viability when exposed to C-CNDs, T-CNDs, and U-CNDs. This effect was most dramatic when it was exposed to 100 μg/mL of C-CND, inducing a ~20% reduction in cell viability. It was hypothesized that SJGBM2 did not undergo a similar drop in viability due to many well-studied mechanisms by which aggressive cancer cells are able to evade drug-induced cytotoxicity.[67] Malignant cells have been shown to upregulate DNA repair proteins, drug-efflux membrane protein pumps, and modify the expression of apoptotic proteins to facilitate survival.[67] Glioblastoma is known to be a particularly aggressive primary brain tumor, portending an extremely poor prognosis.[68] The characteristic resistance to chemo, radio, and non-conventional therapeutics evident in glioblastoma and other aggressive tumors offers an

explanation to why these cells could resist the possible DNA damages exerted by our nucleobase-incorporated CNDs. While MSC cells do exhibit many pro-proliferative qualities as stem cells, they largely lack the mutational burden, expression dysregulation, and protein aberrations accrued in cancer cells, which confer therapeutic resistance. These cells are analogous to non-cancerous proliferating cells like those in an embryo. Thus, these observed reductions in the viability of both cancerous and non-cancerous cells demonstrate that the surface-modification of CNDs with cytosine damages the cell development, which is accidentally consistent with the zebrafish embryo viability results, specifically the finding that the C-CNDs significantly disrupted zebrafish embryogenesis.

To assess the cellular uptake and organelle localizations of A, C, G, T, U-incorporated CNDs, we performed *in vitro* bioimaging studies. Comparing the cellular uptake rates of A, C, G, T, U-incorporated CNDS with HPNE KRAS and ASPC-1 cell lines, we observed that ASPC-1 cells (**Fig. 10A**) exhibited brighter green fluorescence, which came from different CNDs than the HPNE KRAS cells (**Fig. 10B**). In other words, cellular uptake of these CNDs was more favored in pancreatic cancer cells than in the mutated pancreas cells, probably due to the enhanced penetration and retention effect. Cancer cells may be prone to uptake nucleobases faster than normal cells.[69] Since these CNDs' main precursors are nucleobases, it is plausible that cancer cells can take up these CNDs faster than normal cells, which is significant for the future cancer treatments. The observation of a bright turquoise color by merging the green fluorescence (CNDs) and the blue fluorescence (Hoechst-nucleus) channels (**Fig. 10A)** suggests colocalization of various nucleobase-incorporated CNDs in the cell nuclei. Nucleobase-incorporated CNDs in the nuclei of normal cells once again raised our question of whether they can interact with DNA to interfere with cell proliferation and embryogenesis. Additionally, a previous study showed the bare CNDs could not penetrate cell nuclei to interact with DNA,[9] which is consistent with the results in zebrafish embryogenesis, cell viability and uptake. And the enhanced cell nuclear targeting by nucleobase modifications suggests that nucleobases can promote cell nuclear targeting. Also, since bare CNDs could not enter cell nuclei, their nucleobase analog nature remains to be determined.

Effect of bare and nucleobase-incorporated CNDs on DNA. Now that nucleobase-incorporated CNDs could target the cell nuclei of both cancerous and non-cancerous cells, they are likely to interact with the DNA in cell nuclei. Thus, in order to investigate the possibility of nucleobase-incorporated CNDs binding complementary nucleobases via hydrogen bonds as nucleobase analogs even in extreme conditions, CNDs modified by complementary bases were mixed and their interactions were studied. Modified by A and T separately, A-CNDs and T-CNDs were observed to bind to each other, which is confirmed by the increased particle sizes (4.5 nm) revealed by AFM images (**Fig. 11A**). Due to the presence of complementary A and T in A-CNDs and T-CNDs, respectively, the interactions between A-CNDs and T-CNDs are theoretically dominated by hydrogen bonding. Moreover, the weak hydrogen bonding between A-CNDs and T-CNDs was validated by a reduction in the particle size (3.5 nm) upon an ultrasonication treatment and proved by the AFM image (**Fig. 11B**). Thus, this study showed that nucleobase-incorporated CNDs could utilize hydrogen bonds to bind complementary nucleobases in DNA double helix.

In addition, in order to study the impacts of bare and modified CNDs on DNA replication, we performed qPCR of the zebrafish polymerase-alpha 1 gene using SYBR green chemistry (Applied Biosystems, California, USA). We expected to find differences in gene amplification based on the

previously observed effects of nucleobase-incorporated CNDs acting as nucleobase analogs. Specifically, we expected a reduction in amplification in the presence of C-CNDs considering their enhanced toxicity compared with the other modified and bare CNDs. However, we found no difference in amplification across the treatments of bare, individual modified CNDs, and a mix of five modified CNDs at low (10 μg/mL) and high concentrations (100 μg/mL) with respect to our water control lacking CNDs (**Fig. S13** and **Table S8**). Also, two out of three samples with G-CNDs at high concentrations generated large values (which we considered outliers for the generalized linear model (GLM)) corresponding to a hypothetical enhancement of amplification, which seemed to be confirmed by melt curves and Tm matching the pattern of all the other samples; however, the biological significance of these outliers is unclear. Therefore, CNDs did not affect the amplification of the zebrafish polymerase-alpha 1 gene, which makes the CND-DNA interactions during DNA replication less important. In comparison, the results of cell studies and zebrafish embryogenesis are more consistent with the idea that CND-Ca^{2+} interactions are responsible for CNDs interfering with cell processes.

CONCLUSION

Due to the significance of nucleobase and nucleoside analogs in medicinal chemistry and abundant surface functional groups of CDs, this study was first conducted to validate the hypothesis that CDs mimic nucleobases. To simplify the study, a well-characterized CND was selected as a CD model. The comparison in FTIR and solid-state ^{13}C NMR data basically revealed a structural similarity in 2-hydroxypyrimidine between CNDs and cytosine or uracil. To better mimic nucleobases and common surface modification practice, the CNDs were separately modified by five distinct nucleobases. The obtained nucleobase-incorporated CNDs together with bare CNDs were characterized by multiple spectroscopic and microscopic techniques. Many physicochemical property parameters were similar across the bare and different nucleobase-incorporated CNDs, but a deep comparison demonstrated that A-CNDs and C-CNDs possess similar properties and structures while the other three nucleobase-incorporated CNDs are similar in terms of these factors. And the differences between A-CNDs, C-CNDs, and the other three modified CNDs might originate from the different water solubilities of A, C, G, T, and U during the CND syntheses. Furthermore, interestingly, A-CNDs and C-CNDs similarly exhibited lower contents of surface hydrophilic functional groups, smaller particle sizes based on TEM, and lower zeta potential values than the other three modified CNDs. These traits all benefit their cell membrane penetrations. Later, it was observed that the zebrafish embryogenesis was significantly affected by nucleobase-incorporated CNDs, especially the C-CNDs. And the effects included increased toxicity over time, cell adhesion reduction at 8 hpf, abnormal embryogenesis such as tail detachment at 48 hpf and inhibited post-embryonic development. It was hypothesized that the reduced cell adhesion was associated with the disturbance of Ca^{2+} homeostasis by different CND-Ca^{2+} interactions. Furthermore, different CNDs were applied to incubate both normal and cancerous cells. It turned out that normal cells and breast and prostate cancer cells were most vulnerable to the C-CNDs, which was consistent with the zebrafish embryogenesis results. Additionally, different nucleobase-incorporated CNDs were observed to target cell nuclei, especially of cancerous cells, due to the enhanced penetration and retention effect, which makes it viable to interact with DNA in cell nuclei. To further study the interactions between CNDs and DNA, CNDs modified by complementary nucleobases were mixed and it was observed that modifying with complementary nucleobases enabled CNDs to bind each other via hydrogen bonds, which suggests that nucleobase-incorporated CNDs can potentially bind nucleobases in DNA double helix. However, in the end, q-PCR tests revealed that nucleobase-incorporated

CNDs might not intervene in DNA replication. In other words, raw and nucleobase-incorporated CNDs are not nucleobase analogs. Instead, CND-Ca^{2+} interactions are more consistent with the results of zebrafish embryogenesis and cell studies as a plausible mechanism. Further studies are required to understand how CDs affect cell proliferation and embryogenesis. Most importantly, this study warns the scientific community that surface modifications of CDs may result in genotoxicity. However, the outcomes of this work also lay foundations for possible applications in cancer treatment by inhibiting cell proliferation. Considering the widespread presence of CDs, this study is of great significance to improve our understanding of the interactions between CDs and DNA during DNA replication and gene expression, and their impacts on embryogenesis, phenotype, and biological evolution of living organisms.

ACKNOWLEDGMENT

Professor Roger M. Leblanc thanks the support from National Science Foundation under the grants 1809060 and 2041413.

REFERENCE

1. N. V. Bhagavan and C.-E. Ha, in *Essentials of Medical Biochemistry (Second Edition)*, eds. N. V. Bhagavan and C.-E. Ha, Academic Press, San Diego, 2015, pp. 381-400.
2. O. Modén and B. Mannervik, in *Advances in Cancer Research*, eds. D. M. Townsend and K. D. Tew, Academic Press, 2014, vol. 122, pp. 199-244.
3. N. Tsesmetzis, C. B. J. Paulin, S. G. Rudd and N. Herold, *Cancers (Basel)*, 2018, **10**, 240-277.
4. T. R. Skopek, in *Encyclopedia of Genetics*, eds. S. Brenner and J. H. Miller, Academic Press, New York, 2001, pp. 1264-1268.
5. Y. Y. Polosina and C. G. Cupples, in *Brenner's Encyclopedia of Genetics (Second Edition)*, eds. S. Maloy and K. Hughes, Academic Press, San Diego, 2013, pp. 525-527.
6. A. C. Jones and R. K. Neely, *Q. Rev. Biophys.*, 2015, **48**, 244-279.
7. Y. Zhou, S. K. Sharma, Z. Peng and R. M. Leblanc, *Polymers*, 2017, **9**, 67-86.
8. Y. Zhou, E. M. Zahran, B. A. Quiroga, J. Perez, K. J. Mintz, Z. Peng, P. Y. Liyanage, R. R. Pandey, C. C. Chusuei and R. M. Leblanc, *Appl. Catal. B*, 2019, **248**, 157-166.
9. P. Y. Liyanage, R. M. Graham, R. R. Pandey, C. C. Chusuei, K. J. Mintz, Y. Zhou, J. K. Harper, W. Wu, A. H. Wikramanayake, S. Vanni and R. M. Leblanc, *Bioconjug. Chem.*, 2019, **30**, 111-123.
10. E. S. Seven, S. K. Sharma, D. Meziane, Y. Zhou, K. J. Mintz, R. R. Pandey, C. C. Chusuei and R. M. Leblanc, *Langmuir*, 2019, **35**, 6708-6718.
11. M. Mandelkern, J. G. Elias, D. Eden and D. M. Crothers, *J. Mol. Biol.*, 1981, **152**, 153-161.
12. K. J. Mintz, Y. Zhou and R. M. Leblanc, *Nanoscale*, 2019, **11**, 4634-4652.
13. Y. Zhou, K. J. Mintz, S. K. Sharma and R. M. Leblanc, *Langmuir*, 2019, **35**, 9115-9132.
14. C. Zheng, X. An and J. Gong, *RSC Adv.*, 2015, **5**, 32319-32322.
15. T. Kong, L. Hao, Y. Wei, X. Cai and B. Zhu, *Cell Prolif.*, 2018, **51**, e12488-e12496.
16. J. Li, K. Tang, J. Yu, H. Wang, M. Tu and X. Wang, *R. Soc. Open Sci.*, 2019, **6**, 181557-181566.
17. S. Cailotto, R. Mazzaro, F. Enrichi, A. Vomiero, M. Selva, E. Cattaruzza, D. Cristofori, E. Amadio and A. Perosa, *ACS Appl. Mater. Interfaces*, 2018, **10**, 40560-40567.
18. Y. Zhou, P. Y. Liyanage, D. Devadoss, L. R. Rios Guevara, L. Cheng, R. M. Graham, H. S. Chand, A. O. Al-Youbi, A. S. Bashammakh, M. S. El-Shahawi and R. M. Leblanc, *Nanoscale*, 2019, **11**, 22387-22397.

19. E. S. Seven, Y. B. Seven, Y. Zhou, S. Poudel-Sharma, J. J. Diaz-Rucco, E. Kirbas Cilingir, G. S. Mitchell, J. D. Van Dyken and R. M. Leblanc, *Nanoscale Adv.*, 2021, **3**, 3942-3953.
20. E. Kirbas Cilingir, E. S. Seven, Y. Zhou, B. M. Walters, K. J. Mintz, R. R. Pandey, A. H. Wikramanayake, C. C. Chusuei, S. Vanni, R. M. Graham and R. M. Leblanc, *J. Colloid Interface Sci.*, 2021, **592**, 485-497.
21. Y. K. Jung, E. Shin and B.-S. Kim, *Sci. Rep.*, 2015, **5**, 18807-18815.
22. J. Xia, Y. Kawamura, T. Suehiro, Y. Chen and K. Sato, *Drug Discov. Ther.*, 2019, **13**, 114-117.
23. X. Dong, W. Liang, M. J. Meziani, Y.-P. Sun and L. Yang, *Theranostics*, 2020, **10**, 671-686.
24. L. Feng, A. Zhao, J. Ren and X. Qu, *Nucleic Acids Res.*, 2013, **41**, 7987-7996.
25. F. Khakbaz and M. Mahani, *Anal. Biochem.*, 2017, **523**, 32-38.
26. S. Godavarthi, K. Mohan Kumar, E. Vázquez Vélez, A. Hernandez-Eligio, M. Mahendhiran, N. Hernandez-Como, M. Aleman and L. Martinez Gomez, *J. Photochem. Photobiol. B*, 2017, **172**, 36-41.
27. Y. Cheng, C. Li, R. Mu, Y. Li, T. Xing, B. Chen and C. Huang, *Anal. Chem.*, 2018, **90**, 11358-11365.
28. S. Khan, N. C. Verma, Chethana and C. K. Nandi, *ACS Appl. Nano Mater.*, 2018, **1**, 483-487.
29. M. A. Jhonsi, D. A. Ananth, G. Nambirajan, T. Sivasudha, R. Yamini, S. Bera and A. Kathiravan, *Spectrochim. Acta A*, 2018, **196**, 295-302.
30. B. Wang, J. Huang, M. Zhang, Y. Wang, H. Wang, Y. Ma, X. Zhao, X. Wang, C. Liu, H. Huang, Y. Liu, F. Lu, H. Yu, M. Shao and Z. Kang, *ACS Appl. Bio Mater.*, 2020, **3**, 8857-8864.
31. S. H. Schwartz, B. Hendrix, P. Hoffer, R. A. Sanders and W. Zheng, *Plant Physiol.*, 2020, **184**, 647-657.
32. Y. Hou, X. Liu, X. Tang, T. Li, Q. Wu, Y. Jiang, J. Yi and G. Zhang, *Talanta*, 2017, **173**, 107-112.
33. Q. Li, Z. Bai, X. Xi, Z. Guo, C. Liu, X. Liu, X. Zhao, Z. Li, Y. Cheng and Y. Wei, *Spectrochim. Acta A*, 2021, **248**, 119208-119216.
34. G. Kokic, H. S. Hillen, D. Tegunov, C. Dienemann, F. Seitz, J. Schmitzova, L. Farnung, A. Siewert, C. Höbartner and P. Cramer, *Nat. Commun.*, 2021, **12**, 279-285.
35. L. de Sousa Cavalcante and G. Monteiro, *Eur. J. Pharmacol.*, 2014, **741**, 8-16.
36. H. Taymaz-Nikerel, M. E. Karabekmez, S. Eraslan and B. Kırdar, *Sci. Rep.*, 2018, **8**, 13672-13685.
37. L. T. Ferreira, J. Rodrigues, G. C. Cassiano, T. A. Tavella, K. C. P. Tomaz, D. C. Baia-da-Silva, M. F. Souza, M. N. d. N. Lima, M. Mottin, L. D. Almeida, J. Calit, M. C. S. d. B. Puça, G. C. Melo, D. Y. Bargieri, S. C. P. Lopes, M. V. G. Lacerda, E. Bilsland, P. Sunnerhagen, B. J. Neves, C. H. Andrade, P. V. L. Cravo and F. T. M. Costa, *Antimicrob. Agents Chemother.*, 2020, **64**, e02041-02019.
38. T. Fukushima, T. Ueda, M. Uchida and T. Nakamura, *Int. J. Hematol.*, 1993, **57**, 121-130.
39. J. R. Wesolowski, P. Rajdev and S. K. Mukherji, *Am. J. Neuroradiol.*, 2010, **31**, 1383-1384.
40. D. S. Gesto, N. M. Cerqueira, P. A. Fernandes and M. J. Ramos, *Curr. Med. Chem.*, 2012, **19**, 1076-1087.
41. U. Langheinrich, *BioEssays*, 2003, **25**, 904-912.
42. L. A. Rudner, K. H. Brown, K. P. Dobrinski, D. F. Bradley, M. I. Garcia, A. C. H. Smith, J. M. Downie, N. D. Meeker, A. T. Look, J. R. Downing, A. Gutierrez, C. G. Mullighan, J. D. Schiffman, C. Lee, N. S. Trede and J. K. Frazer, *Oncogene*, 2011, **30**, 4289-4296.
43. P. Outtandy, C. Russell, R. Kleta and D. J. P. N. Bockenhauer, *Pediatr. Nephrol.,* 2019, **34**, 751-762.

44. S. Li, Z. Peng, J. Dallman, J. Baker, A. M. Othman, P. L. Blackwelder and R. M. Leblanc, *Colloids Surf. B*, 2016, **145**, 251-256.
45. A. Avdesh, M. Chen, M. T. Martin-Iverson, A. Mondal, D. Ong, S. Rainey-Smith, K. Taddei, M. Lardelli, D. M. Groth, G. Verdile and R. N. Martins, *J. Vis. Exp.*, 2012, **69**, e4196-e4203.
46. J. R. Meyers, *Curr. Protoc.*, 2018, **16**, e19-e44.
47. Y. Zhou, A. E. ElMetwally, J. Chen, W. Shi, E. K. Cilingir, B. Walters, K. J. Mintz, C. Martin, B. C. L. B. Ferreira, W. Zhang, S. D. Hettiarachchi, L. F. Serafim, P. L. Blackwelder, A. H. Wikramanayake, Z. Peng and R. M. Leblanc, *J. Colloid Interface Sci.*, 2021, **599**, 519-532.
48. P. Y. Liyanage, Y. Zhou, A. O. Al-Youbi, A. S. Bashammakh, M. S. El-Shahawi, S. Vanni, R. M. Graham and R. M. Leblanc, *Nanoscale*, 2020, **12**, 7927-7938.
49. Y. Zhou, N. Kandel, M. Bartoli, L. F. Serafim, A. E. ElMetwally, S. M. Falkenberg, X. E. Paredes, C. J. Nelson, N. Smith, E. Padovano, W. Zhang, K. J. Mintz, B. C. L. B. Ferreira, E. K. Cilingir, J. Chen, S. K. Shah, R. Prabhakar, A. Tagliaferro, C. Wang and R. M. Leblanc, *Carbon*, 2022, **193**, 1-16.
50. J. A. Milner, *J. Nutr.*, 1985, **115**, 516-523.
51. K. J. Mintz, M. Bartoli, M. Rovere, Y. Zhou, S. D. Hettiarachchi, S. Paudyal, J. Chen, J. B. Domena, P. Y. Liyanage, R. Sampson, D. Khadka, R. R. Pandey, S. Huang, C. C. Chusuei, A. Tagliaferro and R. M. Leblanc, *Carbon*, 2021, **173**, 433-447.
52. C. Fan, J. Miao, G. Xu, J. Liu, J. Lv and Y. Wu, *RSC Adv.*, 2017, **7**, 37185-37193.
53. A. Ben-Refael, I. Benisti and Y. Paz, *Catal. Today*, 2020, **340**, 97-105.
54. H.-y. Wang, L. Zhou, H.-m. Yu, X.-d. Tang, C. Xing, G. Nie, H. Akafzade, S.-y. Wang and W. Chen, *Adv. Opt. Mater.*, 2022, **10**, 2200678-2200688.
55. G. Zhang, M. Wen, S. Wang, J. Chen and J. Wang, *RSC Adv.*, 2018, **8**, 567-579.
56. J. Li, D. Liu, B. Li, J. Wang, S. Han, L. Liu and H. Wei, *CrystEngComm*, 2015, **17**, 520-525.
57. C. Li, Q. Ran, R. Zhu and Y. Gu, *RSC Adv.*, 2015, **5**, 22593-22600.
58. N. Ravi, S. A. Anuar, N. Y. M. Yusuf, W. N. R. W. Isahak and M. S. Masdar, *Mater. Res. Express*, 2018, **5**, 055501-055514.
59. H. Dai, X. Gao, E. Liu, Y. Yang, W. Hou, L. Kang, J. Fan and X. Hu, *Diam. Relat. Mater.*, 2013, **38**, 109-117.
60. L.-Y. Chen and W.-D. Zhang, *Appl. Surf. Sci.*, 2014, **301**, 428-435.
61. Z.-l. Wang, D. Xu, Y. Huang, Z. Wu, L.-m. Wang and X.-b. Zhang, *ChemComm*, 2012, **48**, 976-978.
62. Q. Liu, J. A. Marrs, R. L. Londraville and A. L. Wilson, *Cell Tissue Res.*, 2008, **334**, 37-45.
63. C. H. Lin and P. P. Hwang, *Int. J. Mol. Sci.*, 2016, **17**, 1783-1798..
64. H. Harjunpää, M. Llort Asens, C. Guenther and S. C. Fagerholm, *Front. Immunol.*, 2019, **10**, 1078-1101.
65. A. Q. Hassan, R. K. Barzani, K. M. Omer, B. R. Al-Hashimi, S. Mohammadi and A. Salimi, *Arab. J. Chem.*, 2021, **14**, 103452-103462.
66. M. Zoupa and K. Machera, *Int. J. Mol. Sci.*, 2017, **18**, 817-842.
67. B. Mansoori, A. Mohammadi, S. Davudian, S. Shirjang and B. Baradaran, *Adv. Pharm. Bull.*, 2017, **7**, 339-348.
68. R. Stupp, W. P. Mason, M. J. van den Bent, M. Weller, B. Fisher, M. J. Taphoorn, K. Belanger, A. A. Brandes, C. Marosi, U. Bogdahn, J. Curschmann, R. C. Janzer, S. K. Ludwin, T. Gorlia, A. Allgeier, D. Lacombe, J. G. Cairncross, E. Eisenhauer, R. O. Mirimanoff, R. European Organisation for, T. Treatment of Cancer Brain, G. Radiotherapy and G. National Cancer Institute of Canada Clinical Trials, *N. Engl. J. Med.*, 2005, **352**, 987-996.
69. C. M. Galmarini, J. R. Mackey and C. Dumontet, *Lancet Oncol.*, 2002, **3**, 415-424.

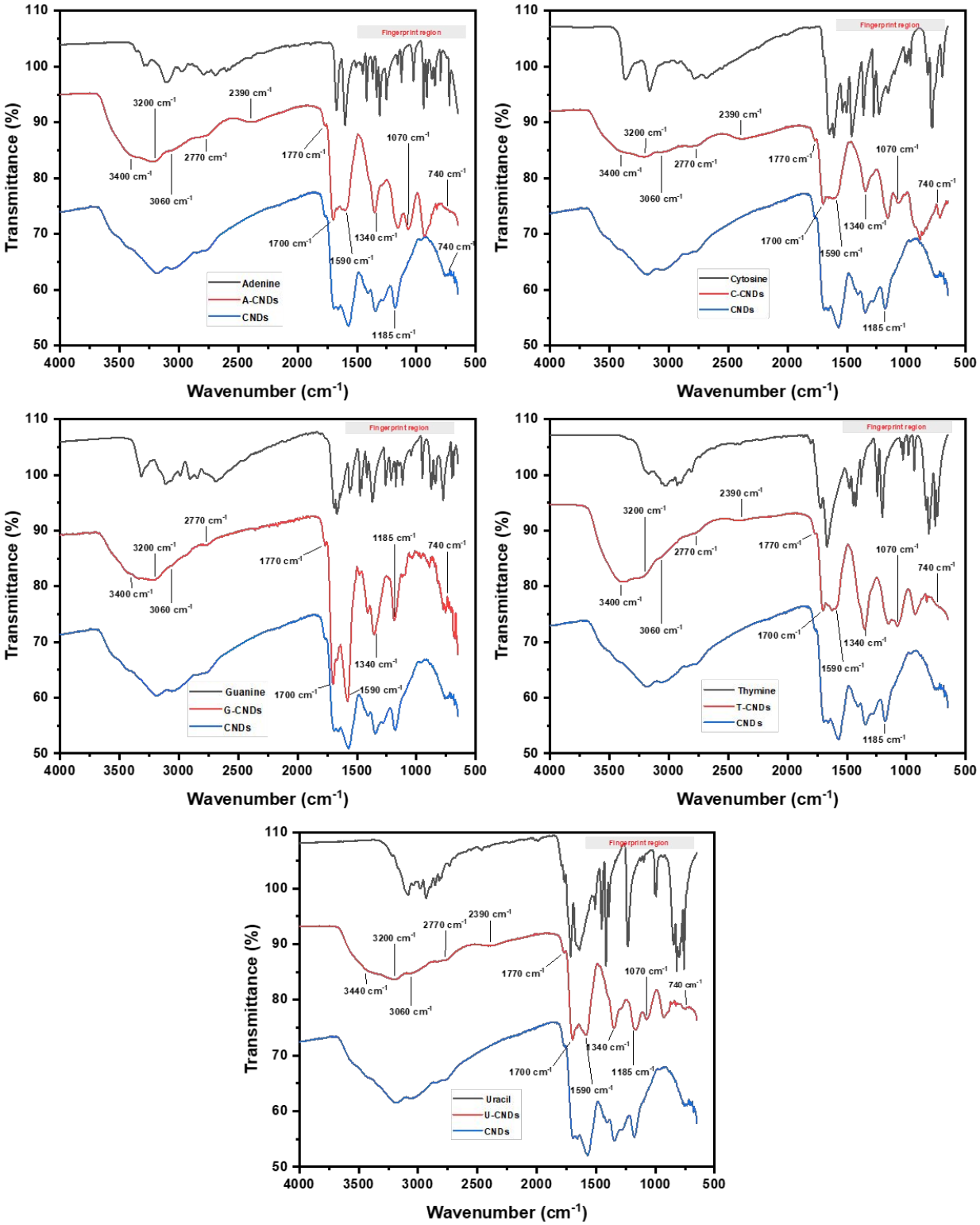

Fig. 1. FTIR spectra of nucleobase-incorporated CNDs in powder in contrast to their corresponding nucleobase precursors and bare CNDs in powder.

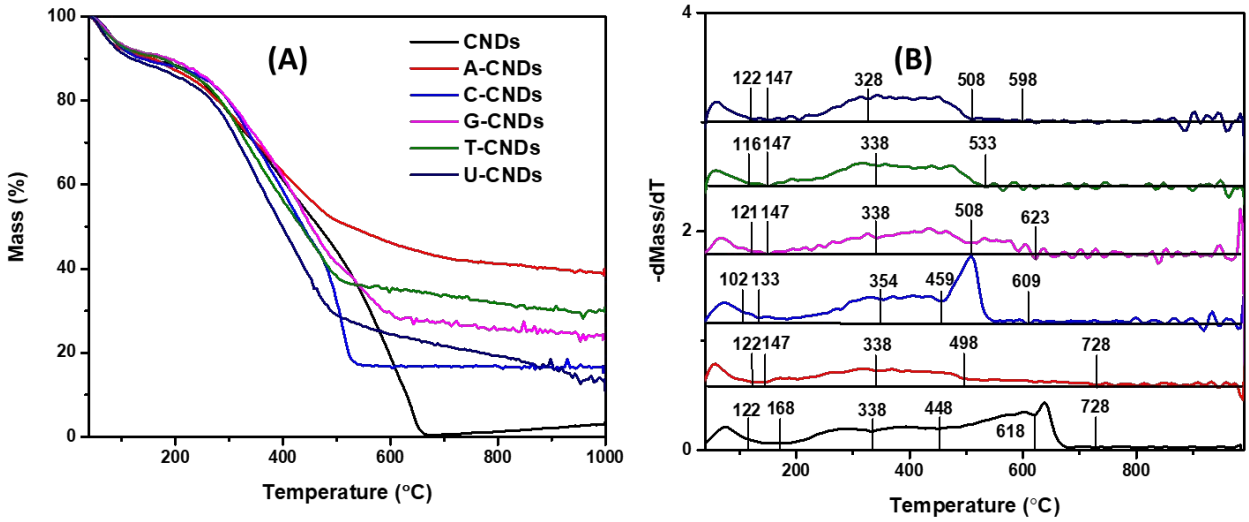

Fig. 2. TGA (A) and DTG (B) of different nucleobase-incorporated CNDs.

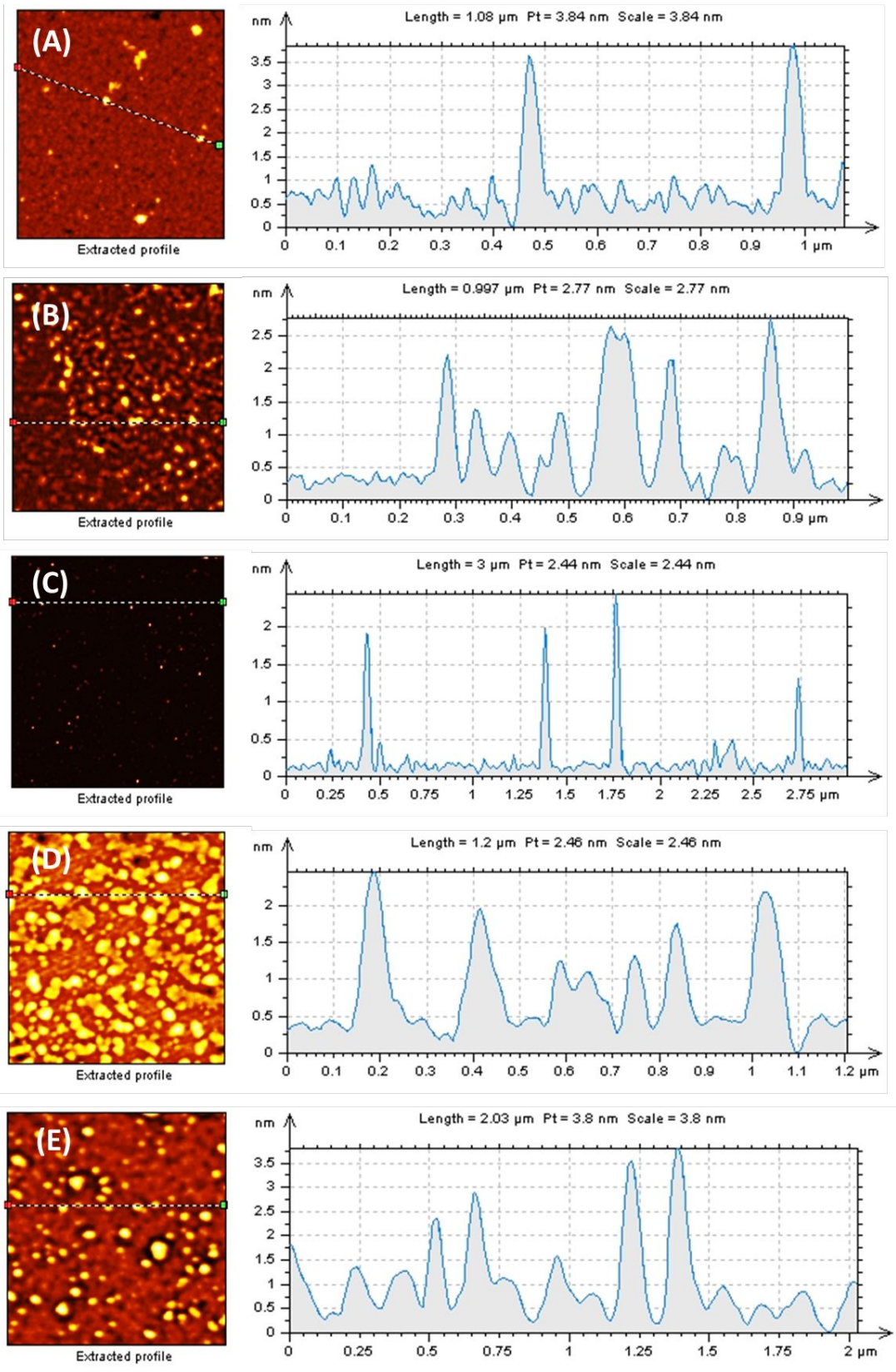

Fig. 3. AFM images of various nucleobase-incorporated CNDs (left) and corresponding size distribution histograms (right). (**A-E**, A-CNDs, C-CNDs, G-CNDs, T-CNDs, U-CNDs, respectively)

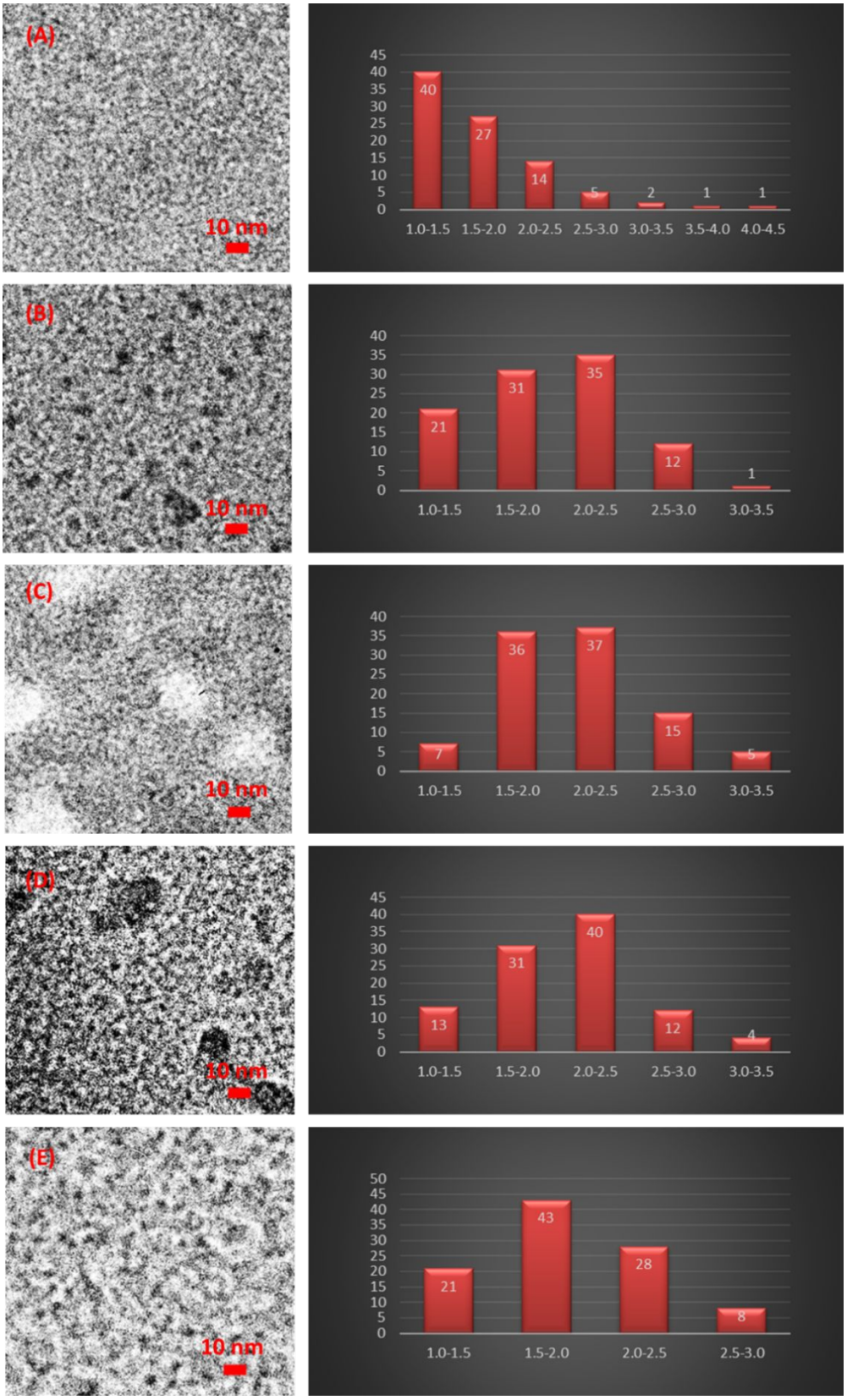

Fig. 4. TEM images of nucleobase-incorporated CNDs (left) and the corresponding size distribution histograms (right). (**A-E**, A-CNDs, C-CNDs, G-CNDs, T-CNDs, U-CNDs, respectively)

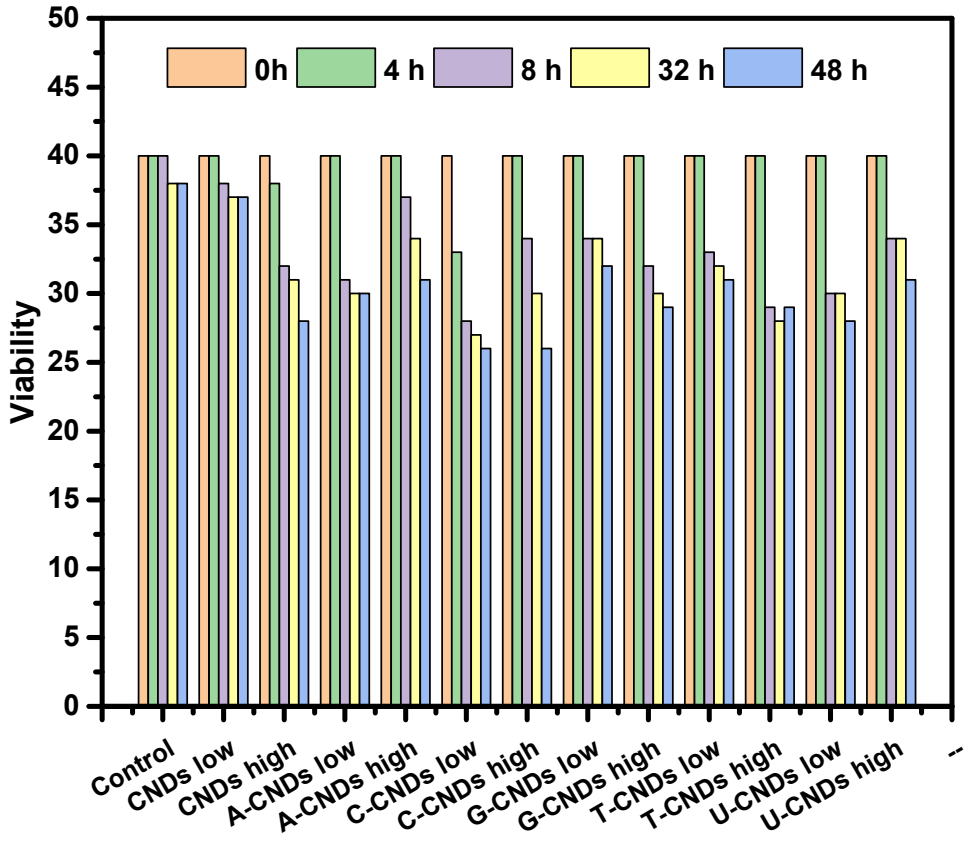

Fig. 5. Zebrafish embryo viability in presence of different CNDs at low (1 μg/mL) and high (100 μg/mL) concentrations.

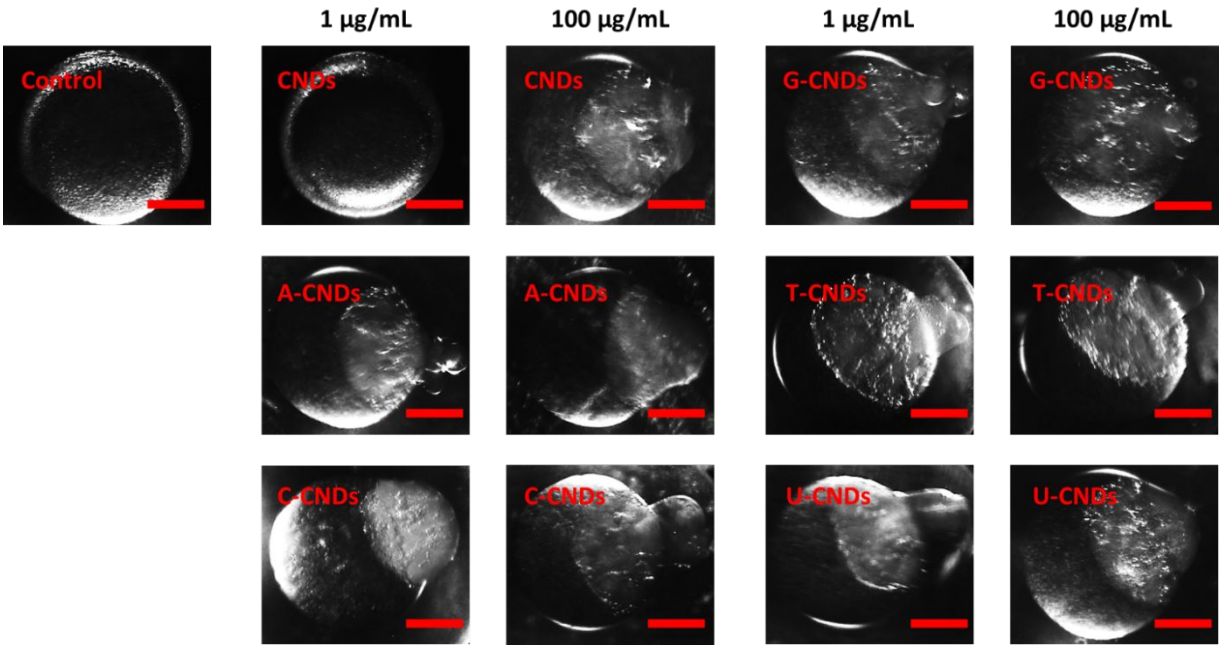

Fig. 6. Cell adhesion reduction observed in the zebrafish embryos (8 hpf) in presence of different concentrations (1 and 100 μg/mL) of bare and various nucleobase-incorporated CNDs. The control was conducted in absence of any CNDs. The scale bars represent 160 pixels.

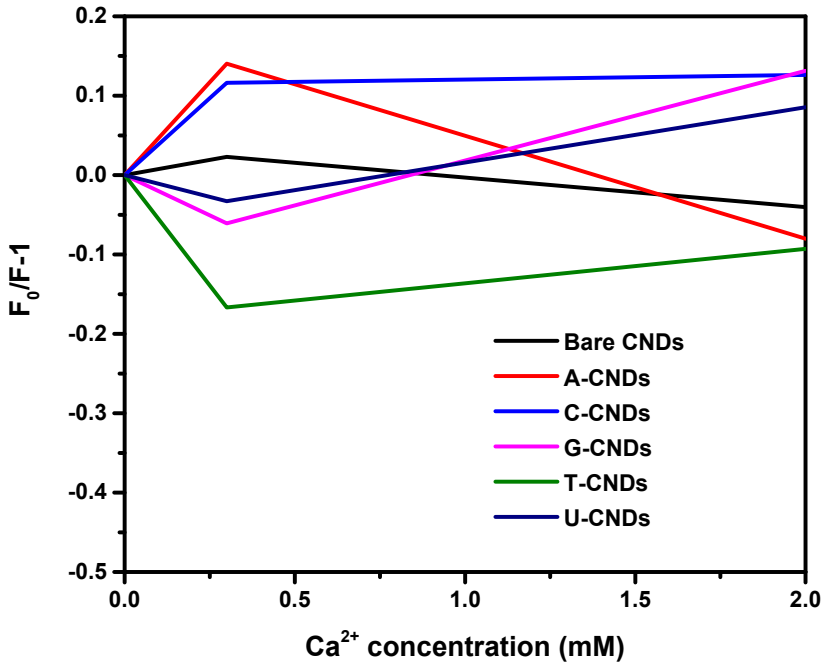

Fig. 7. Effect of Ca^{2+} (0, 0.3 and 2.0 mM) on the fluorescence intensity of different CNDs.

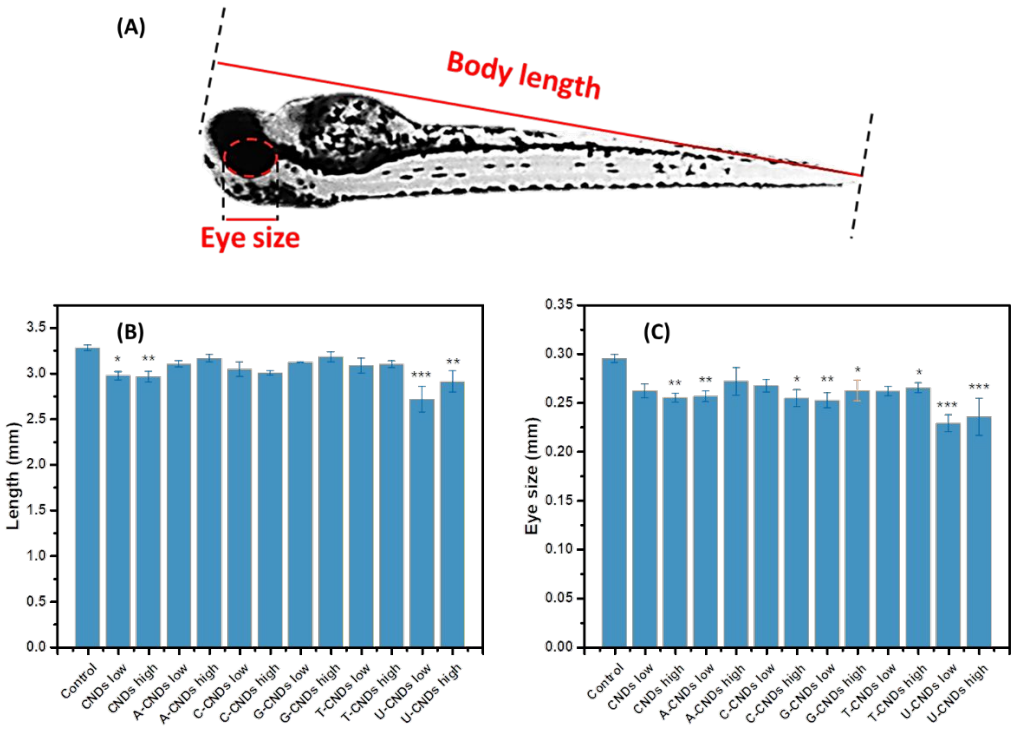

Fig. 8. Measurements of zebrafish embryo body length and eye size **(A).** Zebrafish embryo length **(B)** and eye size **(C)** are affected by different CNDs at two different concentrations (low: 1 μg/mL; high: 100 μg/mL). *p<0.05, **p<0.01, ***p<0.001 compared to non-treated controls.

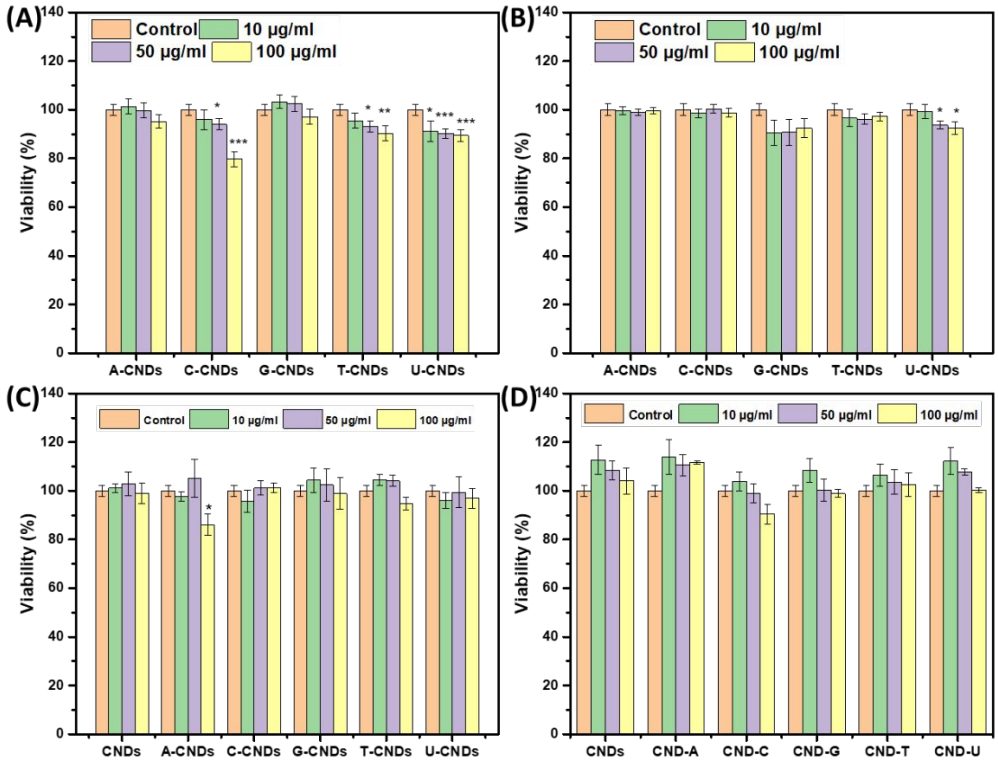

Fig. 9. Viability of MSCs **(A)**, SJGBM-2 **(B)**, PC-3 **(C)** and 4T1 **(D)** cells treated with different concentrations (0, 10, 50 and 100 μg/mL) of bare and various nucleobase-incorporated CNDs for 72 hours. *p<0.05, **p<0.01, ***p<0.001 compared to non-treated controls.

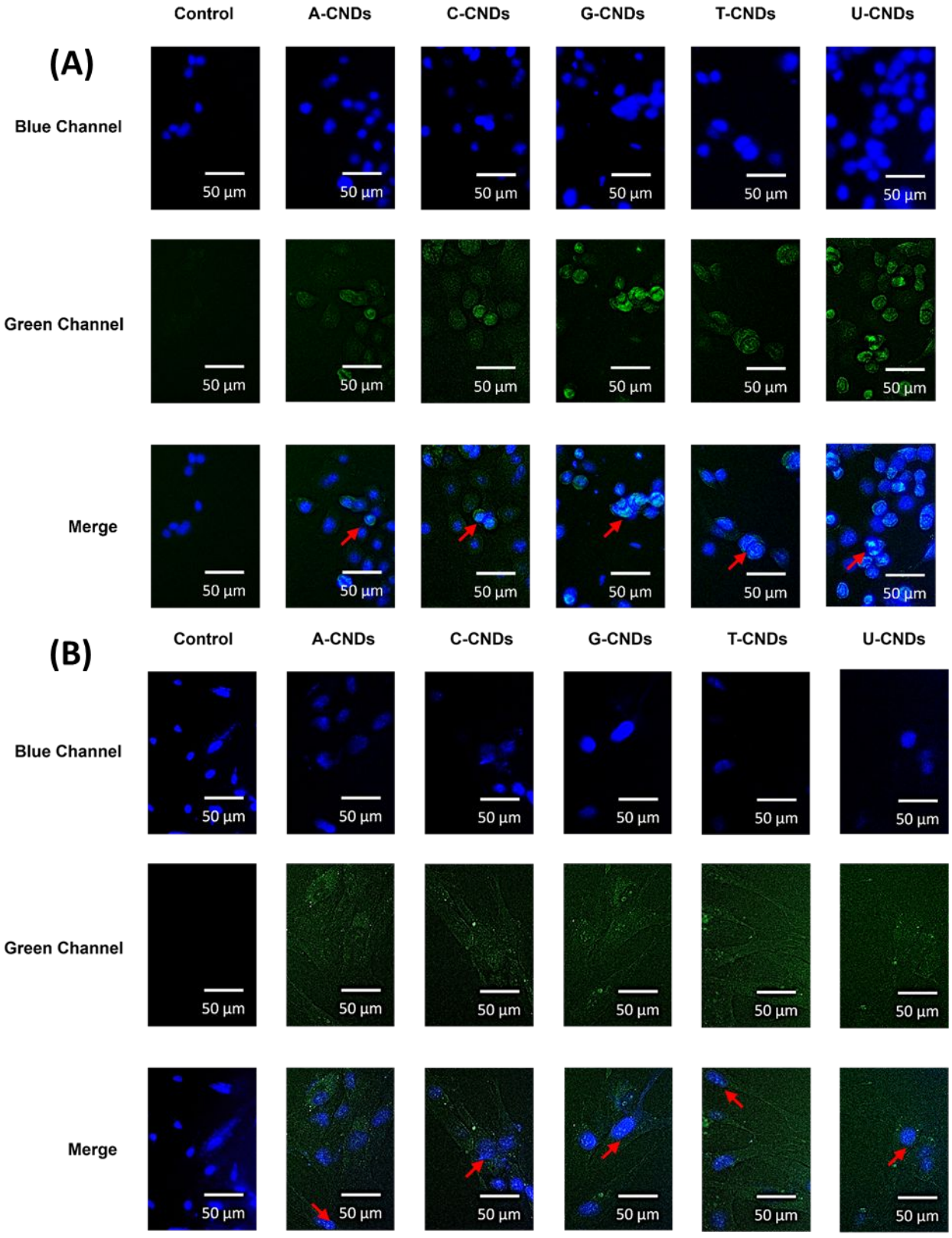

Fig. 10. Fluorescence microscopy images of **(A)** ASPC-1 pancreatic cancer cells, and **(B)** HPNE KRAS mutated pancreas cells incubated with A, C, G, T, U-incorporated CNDs. Both cell lines were treated with 1 mg/mL concentration of A, C, G, T, U-incorporated CNDs for 1 h. Excitation wavelengths: Blue, 358 nm; Green, 488 nm.

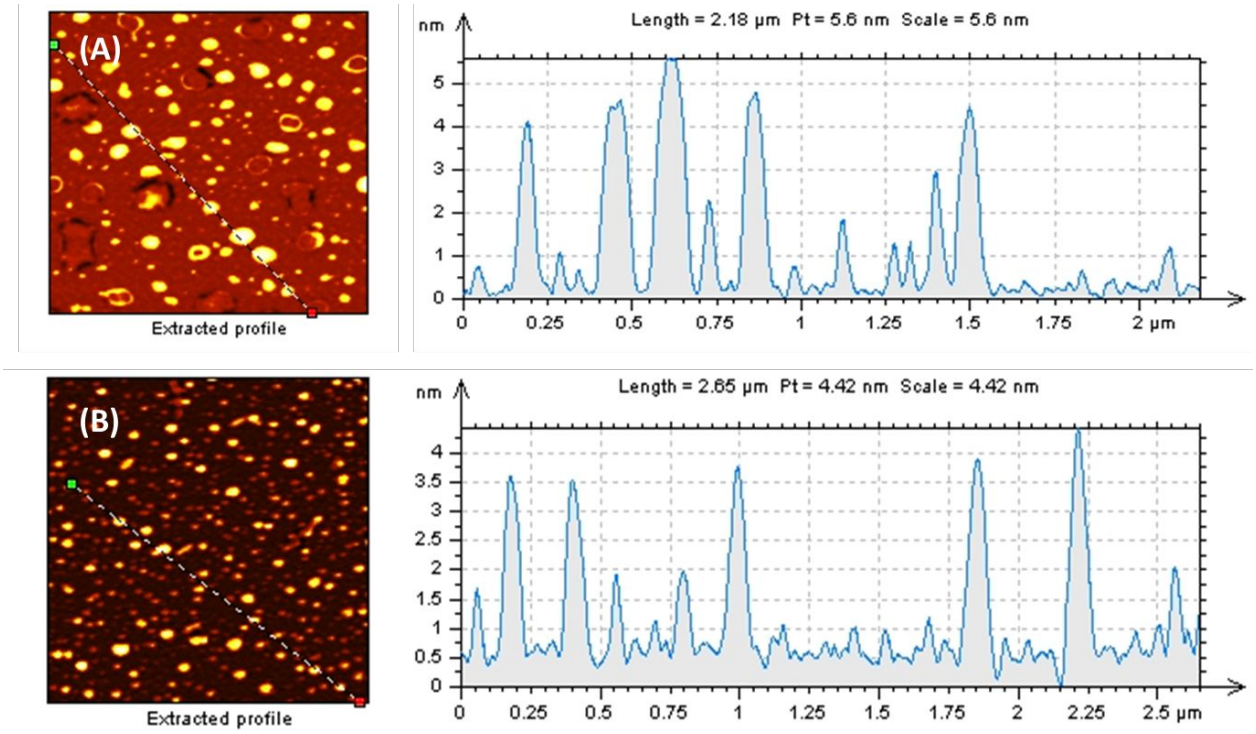

Fig. 11. AFM images of a mixture of A-CNDs and T-CNDs that underwent **(A)** no sonication and **(B)** sonication.