

Tautomerism provides a molecular explanation for the mutagenic properties of the anti-HIV nucleoside 5-aza-5,6-dihydro-2'-deoxycytidine

Deyu Li^{a,b,c,1}, Bogdan I. Fedeles^{a,b,c,1}, Vipender Singh^{a,b,c,1}, Chunte Sam Peng^{a,1,2}, Katherine J. Silvestre^a, Allison K. Simi^b, Jeffrey H. Simpson^a, Andrei Tokmakoff^{a,2,3}, and John M. Essigmann^{a,b,c,3}

Departments of ^aChemistry and ^bBiological Engineering, and ^cCenter for Environmental Health Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139

Edited by Vern L. Schramm, Albert Einstein College of Medicine of Yeshiva University, Bronx, NY, and approved June 27, 2014 (received for review March 26, 2014)

Viral lethal mutagenesis is a strategy whereby the innate immune system or mutagenic pool nucleotides increase the error rate of viral replication above the error catastrophe limit. Lethal mutagenesis has been proposed as a mechanism for several antiviral compounds, including the drug candidate 5-aza-5,6-dihydro-2'-deoxycytidine (KP1212), which causes A-to-G and G-to-A mutations in the HIV genome, both in tissue culture and in HIV positive patients undergoing KP1212 monotherapy. This work explored the molecular mechanism(s) underlying the mutagenicity of KP1212, and specifically whether tautomerism, a previously proposed hypothesis, could explain the biological consequences of this nucleoside analog. Establishing tautomerism of nucleic acid bases under physiological conditions has been challenging because of the lack of sensitive methods. This study investigated tautomerism using an array of spectroscopic, theoretical, and chemical biology approaches. Variable temperature NMR and 2D infrared spectroscopic methods demonstrated that KP1212 existed as a broad ensemble of interconverting tautomers, among which enolic forms dominated. The mutagenic properties of KP1212 were determined empirically by *in vitro* and *in vivo* replication of a single-stranded vector containing a single KP1212. It was found that KP1212 paired with both A (10%) and G (90%), which is in accord with clinical observations. Moreover, this mutation frequency is sufficient for pushing a viral population over its error catastrophe limit, as observed before in cell culture studies. Finally, a model is proposed that correlates the mutagenicity of KP1212 with its tautomeric distribution in solution.

viral decay acceleration | KP1461 | viral extinction |
spectral deconvolution | site-specific mutagenesis

Many viruses exhibit a high mutation rate when replicating their genomes, enabling quick adaptation to both changing cellular environments and therapeutics (1–5). Mammalian innate immune systems have developed a mechanism to exploit this high mutation rate against the virus; in a phenomenon termed “lethal mutagenesis,” (6–14) the immune system employs nucleic acid-modifying enzymes (e.g., APOBEC and ADAR) to increase the viral mutation rate sharply, stressing the functional gene product repertoire of the virus to the point that the viral population collapses (15–17). Several antiviral agents are proposed to work at least in part by a chemical version of lethal mutagenesis [e.g., ribavirin against hepatitis C virus (18–22), 5-hydroxy-2'-deoxycytidine against HIV (7), and T-705 against influenza viruses (23)]. When a sufficient number of these mutagenic nucleoside analogs is incorporated into viral genomes, the analogs increase the viral mutation rate above the error catastrophe limit, the rate above which no viable progeny are produced (6, 24–27). This work aimed to understand the molecular basis underlying the biological phenomenon of lethal mutagenesis induced by mutagenic nucleotides.

The nucleoside analog 5-aza-5,6-dihydro-2'-deoxycytidine (KP1212) (Fig. 1A) is specifically designed to induce lethal mu-

tagenesis in HIV (28–30). KP1212, the only anti-HIV drug candidate in clinical trials to use this mechanism, has been shown to increase the mutation rate of HIV both in cell culture and in isolates from humans undergoing monotherapy (28, 29). The mutagenic properties of KP1212 in cell culture reveal that it is likely to base pair promiscuously with A and G, and that the progressive acquisition of mutations (primarily A-to-G and G-to-A transitions) precedes population collapse (Fig. 1B) (29). These data are supported by biochemical experiments performed using purified polymerases that establish the ability of KP1212 to pair with either A or G, both when the modified base enters DNA from the nucleotide pool and when it acts as a template base (30). Understanding the chemical and structural basis of mutagenesis of this drug candidate is critical for both its future clinical progress and the development of new therapeutic agents that work by the principle of lethal mutagenesis.

Tautomerism of KP1212 leading to viral mutagenesis has been proposed by us and others (29, 30) to be the basis for the clinical activity of this drug candidate. There are, however, no direct data to support that view. Tautomerism as the basis of mutagenesis of

Significance

Unlike conventional antiviral therapy, lethal mutagenesis is a therapeutic strategy that exploits the high mutation rates of certain viruses. It works by intentionally increasing the viral mutation rate, causing excessive error accumulation and viral population collapse. The mutagenic nucleoside analog 5-aza-5,6-dihydro-2'-deoxycytidine (KP1212) is specifically designed to use lethal mutagenesis against HIV. The mechanism of KP1212 mutagenesis was proposed to involve tautomerism—the repositioning of active protons on the nucleic acid base on a fast time scale. Using a multifaceted approach, we demonstrate that KP1212 exists in multiple tautomeric forms, and that the tautomeric distribution correlates with the mutagenic properties of KP1212. This work also provides a toolset for studying tautomerism in nucleic acids and developing the next-generation antiviral lethal mutagens.

Author contributions: D.L., B.I.F., V.S., C.S.P., A.T., and J.M.E. designed research; D.L., B.I.F., V.S., C.S.P., K.J.S., A.K.S., and J.H.S. performed research; D.L., B.I.F., V.S., C.S.P., K.J.S., A.K.S., and J.H.S. contributed new reagents/analytic tools; D.L., B.I.F., V.S., C.S.P., J.H.S., A.T., and J.M.E. analyzed data; and D.L., B.I.F., V.S., C.S.P., A.T., and J.M.E. wrote the paper.

Conflict of interest statement: J.M.E. declares a competing financial interest as a co-founder and advisor for a pharmaceutical company interested in developing mutagenic inhibitors of HIV.

This article is a PNAS Direct Submission.

¹D.L., B.I.F., V.S., and C.S.P. contributed equally to this work.

²Present address: Department of Chemistry, University of Chicago, Chicago, IL 60637.

³To whom correspondence may be addressed. Email: jessig@mit.edu or tokmakoff@uchicago.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1405635111/-DCSupplemental.

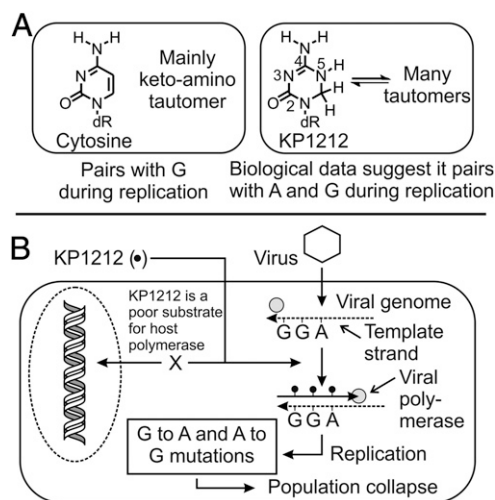


Fig. 1. Schematic presentation of KP1212's mutagenic effect on viruses. (A) KP1212 exists as an array of different tautomeric forms, whereas cytosine almost exclusively exists as one form, the canonical keto-amino tautomer. (B) The deoxynucleotide analog of KP1212 is incorporated by viral polymerases, causing G-to-A and A-to-G mutations during viral replication. KP1212 is a poor substrate for human polymerases, which provides selectivity in its action against the virus. The progressive acquisition of mutations in the viral genome leads to viral population collapse.

natural bases has long been proposed (31–35), and substantiated in part by experimental evidence of minor tautomeric forms of both canonical bases (36–38) and certain base analogs (e.g., 5-hydroxy-2'-deoxycytidine) (39). In a search for a chemical rationale to explain the ambiguous pairing of KP1212 during replication, the present study revealed that the compound readily adopts multiple tautomeric forms, some of which were unexpected. Previously, spectroscopic methods (e.g., UV, Raman, NMR) have been used to study tautomerism of nucleobases (37, 39, 40). In the current work, we also used a battery of spectroscopic tools (1D, 2D, and variable temperature NMR; FTIR; and 2D IR) (41) to quantify and structurally characterize the array of tautomers exhibited by KP1212. Tautomer interconversion equilibria deconvoluted from NMR spectra provided data on the relative levels of tautomers in solution. In parallel with the spectroscopic studies, the qualitative and quantitative features of KP1212 mutagenesis were directly determined by inserting the KP1212 base into a single-stranded viral vector and measuring the intrinsic mutagenic properties of the base, both in vitro and in vivo. Finally, a model is proposed that correlates the mutagenic and clinical properties of KP1212 with its ability to exist as multiple tautomers.

Results

DNA Polymerase Extension Assays Demonstrate the Mutagenic Properties of KP1212. Whereas the mutagenic properties of KP1212 have been demonstrated before (29, 30), it was desirable to determine the base-pairing preferences of the nucleoside analog in vitro and in living cells, under conditions that would allow direct comparison with other mutagenic base analogs and lesions that we and others have previously studied in a site-specific manner. Thus, using a well-established protocol (42), a DNA template containing a single KP1212 was constructed (Fig. 2A–C) and used to determine the amount and type of the opposing-strand base(s) placed by polymerases across from KP1212. In addition, the extent to which KP1212 present in the template strand acted as a replication block was determined.

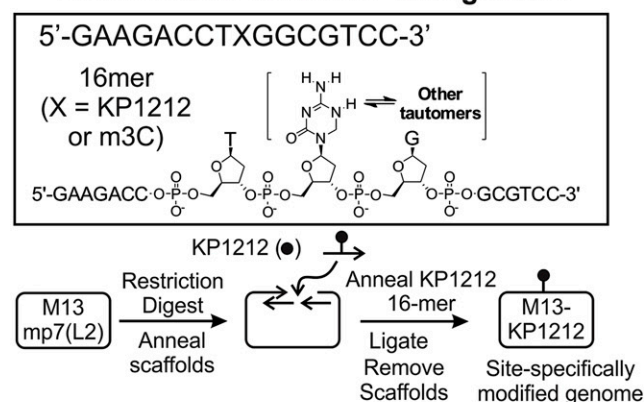
Mutagenicity and toxicity studies were performed by using methods established in our previous work, where a single nucleoside (in this case, KP1212) was incorporated at a defined site

in the single-stranded DNA circular genome of an M13 virus (42, 43). This viral genome was used as a template both for in vitro polymerase extension assays and, for the in vivo assays, as a vector that was replicated in *Escherichia coli* hosts. The experimental system was specifically designed to avoid certain DNA repair processes, because KP1212 was present in a single-stranded genome. Most glycosylases acting in the base excision repair pathway, as well as the enzymes of the nucleotide excision repair pathway, require double-stranded DNA substrates (44).

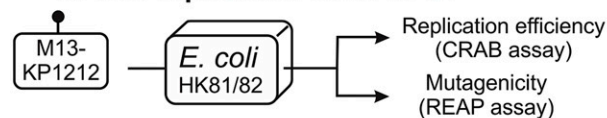
KP1212 was found not to be a replication block in *E. coli* (Fig. 2D). In this experiment 3-methylcytosine (m3C), an established toxic and mutagenic DNA lesion (43), was used as a positive control. As expected, the data showed that m3C was very toxic to replication in cells that lack α -ketoglutarate and iron(II) dependent dioxygenase AlkB (an enzyme that repairs m3C) and that toxicity was abolished in cells that expressed the repair enzyme for the lesion (43). By comparison with m3C, KP1212 was not a replication block, but it was mutagenic both in vitro and in vivo (Fig. 2C and *E* and *SI Appendix*, Tables S1 and S2). Once again, m3C was the positive control and its mutagenicity in repair-deficient cells was clearly evident (Fig. 2E). Consistent with previous reports, m3C was found to mispair with both A and T (43). By contrast with the deoxycytidine (dC) control, which only pairs with G, KP1212 was shown to pair in vivo with both G (90%) and A (10%) (Fig. 2E). Given that KP1212 is a cytosine analog, and therefore expected to pair with G, the mutation frequency of KP1212 can be defined as the frequency with which it paired with any non-G base, in this case A. Therefore, KP1212 has an in vivo mutation frequency of $\sim 10\%$. To rule out the indirect effects that the cellular environment might have on the mutagenicity of KP1212 (e.g., DNA repair or metabolism), the template containing KP1212 was also replicated in vitro, at 72 °C using PfuTurbo DNA polymerase. The reaction conditions (high temperature and a high-fidelity polymerase containing a powerful “proofreading” 3'-exonuclease) were specifically chosen to capture only the intrinsic base-pairing properties of KP1212. The results of the in vitro experiment indicated that G was placed opposite KP1212 90.5% of the time, whereas A was placed opposite of KP1212 9.5% of the time (Fig. 2C). These data were nearly identical to the results obtained in the in vivo experiment. Taken together, the polymerase extension assays established that: (i) KP1212 is mutagenic in living cells, by pairing with both A and G, and (ii) the mispairing properties of KP1212 are intrinsic to the base.

NMR Spectroscopy Demonstrates the Existence of Multiple Tautomers of KP1212. In an effort to understand the molecular basis of the observed mutagenic properties of KP1212, the nucleoside analog was studied using several spectroscopic techniques. Proton (^1H) variable-temperature (VT) NMR analysis was used to investigate the presence of tautomeric forms of KP1212. The analysis was performed at nine temperatures, which allowed resolution of multiple NMR peaks consistent with proton mobility within the KP1212 base as would occur during the process of tautomerization (Fig. 3). Fig. 4 depicts the structures of all five possible tautomers (1 to 5) of KP1212, among which the three active protons on the nucleobase shift among the O2, N3, N4, and N5 positions. The NMR experiments were carried out in a solution of KP1212 in heptadeuterated dimethylformamide (DMF- d_7); the solvent was chosen because it was aprotic, afforded good solubility of the compound, and had a low freezing point (-61°C). Before assigning the three active protons on the nucleobase moiety, all nonexchangeable protons in KP1212 and the two hydroxyl protons on the 3' and 5' positions were identified by carrying out 1D (^1H and ^{13}C) and 2D [COSY, heteronuclear single quantum coherence (HSQC), and heteronuclear multiple-bond correlation (HMBC)] NMR experiments. The assignments of individual protons and carbons are listed in *SI Appendix*, Tables S3

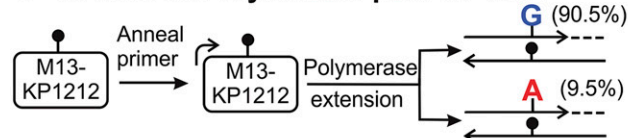
A Construction of M13-KP1212 genome



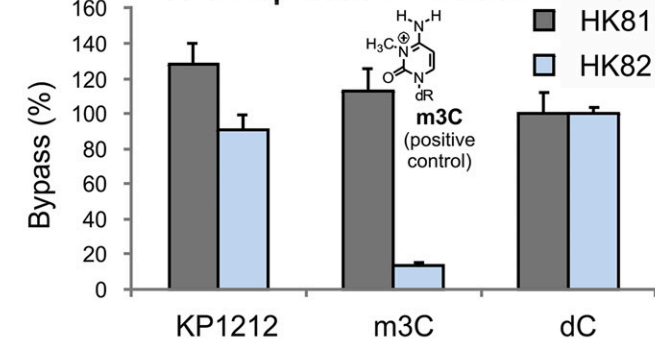
B In vivo replication of KP1212



C In vitro DNA synthesis past KP1212



D Lesion replication inhibition *in vivo*



E Lesion mutational specificity and frequency *in vivo*

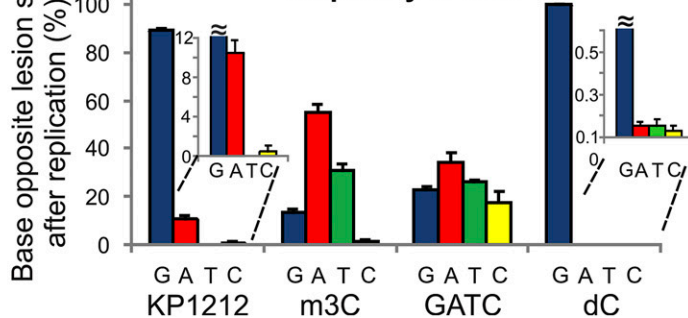


Fig. 2. In vivo and in vitro demonstration of the promiscuous base-pairing properties of KP1212. (A) A 16mer oligonucleotide containing KP1212 (or controls) at a specific site was chemically synthesized and ligated into the genome of an M13 bacteriophage. (B) The KP1212-containing M13 genomes were replicated within *E. coli* cells. Progeny were analyzed to characterize the amount and the type of nucleic acid base placed opposite the lesion during replication. The relative reduction in lesion vs. nonlesion competitor progeny formation was used as an estimate of the extent to which KP1212 and m3C inhibited DNA replication. (C) The KP1212-containing M13 genomes were used as templates for in vitro polymerase extension assays, carried out at 72 °C, using the high-fidelity DNA polymerase PfuTurbo. SD of the measurements is ~1% ($n = 3$). (D) Bypass efficiency (CRAB assay) of KP1212, m3C, C in HK81 (AlkB⁺), and HK82 (AlkB⁻) *E. coli* cells. m3C and undamaged C were used as controls. Genomes were made and normalized to one another before being combined with a competitor genome. Each mixture was transformed into the corresponding cell strain in triplicate, and bypass efficiency was calculated by using the undamaged C genome as 100% bypass, with error bars representing one SD ($n = 3$). (E) Mutagenesis (REAP assay) of KP1212, m3C, and C in HK82 (AlkB⁻) *E. coli* cells. m3C, undamaged C, and an approximately equimolar mixture of genomes carrying unmodified G/A/T/C bases at the site of inquiry (denoted as GATC) were used as controls. Genomes containing the lesions of interest were transfected into *E. coli* in triplicate. The percentage of G, A, T, and C incorporated opposite the lesion site reveals the base-pairing preference of the lesions, with error bars representing one SD ($n = 3$).

and S4 and in the NMR spectra included as *SI Appendix, Figs. S1–S5*. After these assignments were made, the NMR signals of the three active protons in the base portion of KP1212 were studied at different temperatures (Fig. 3). At 20 °C, the exchangeable protons on the KP1212 base (protons bound to either nitrogen or oxygen atoms on the base) displayed three broad peaks denoted i', ii', and iii' (Fig. 3). When the temperature was decreased to -50 °C, the three broad peaks resolved into six distinct proton resonances (Figs. 3 and 4B), denoted as peaks i–vi, which could be assigned to imino (designated blue in Fig. 4A–C), amido (purple), enolic (red), or amino (green) protons. Detailed assignments of these protons are listed in *SI Appendix, Supporting Note*. Because each tautomeric form has only a total of three exchangeable protons on the nucleobase, the observation of six proton resonances indicated the presence of a minimum of two tautomeric forms of KP1212. By contrast, the normal nucleoside deoxycytidine investigated under similar conditions displayed only one tautomeric form, the classical keto-amino tautomer (Fig. 1A) in DMF-d₇ over the same temperature interval (*SI Appendix, Fig. S6*).

In the NMR spectrum of KP1212 in DMF-d₇ solution at -50 °C, the chemical shifts for the imino/amido, enolic, and amino protons could be clearly differentiated. However, the imino proton signals corresponding to the Z and E geometric isomers of the

imino tautomers (i.e., isomer 2Z versus 2E, or isomer 5Z versus 5E), were not separated by VT-NMR under the above conditions (Fig. 4A). Therefore, the subsequent quantification of the relative amounts of each tautomer from NMR data treated the imino

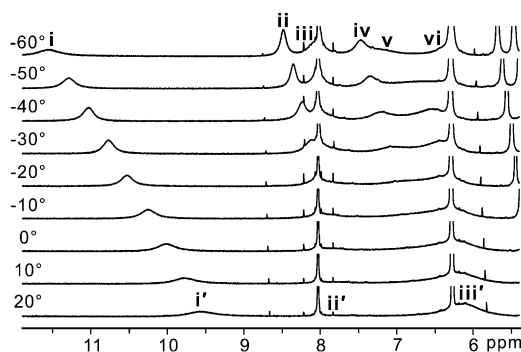


Fig. 3. Variable temperature ¹H NMR spectra of KP1212 in DMF-d₇ (20 °C to -60 °C). The peaks corresponding to the active protons on the base portion of KP1212 are labeled as i–vi at -60 °C and i–iii' at 20 °C.

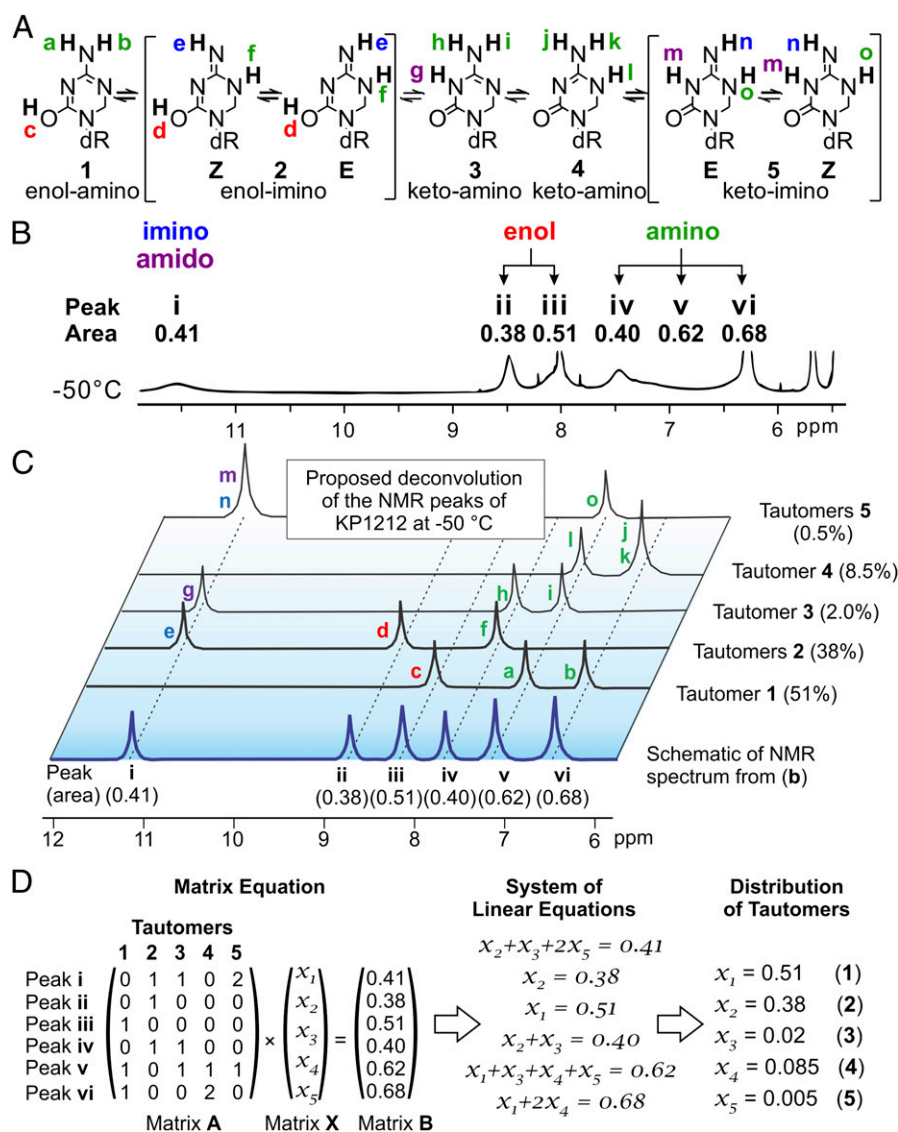


Fig. 4. NMR studies demonstrate the existence of different tautomeric forms of KP1212. (A) Structures of the five possible tautomeric forms of KP1212. The active protons (a–o) on the nucleobase portion of the molecule are designated with different colors to indicate their chemical environment (type): blue (imino), purple (amido), red (enol), and green (amino). (B) The ^1H NMR spectrum of KP1212 in DMF-d_7 at -50°C (5.5–12.0 ppm), from Fig. 3. The peaks from the active protons on the nucleobase portion are labeled as i–vi and their corresponding areas are indicated. According to their chemical shifts, the type of the active protons on the KP1212 nucleobase that contribute to each peak is indicated. (C) Schematic of the deconvolution process of the ^1H NMR spectrum of KP1212 at -50°C depicting how the active proton peaks corresponding to each tautomer contribute to the overall spectrum. Each of the six peaks identified in the NMR spectrum in B, denoted i–vi, is schematically represented as the bottom trace (in blue). To indicate each tautomer's respective contributions to the six peaks, schematic representations of the NMR signals of the active protons of each tautomer are shown (black traces). Each peak is labeled with a color coded letter (a–o), which corresponds to the active protons labeled in section A. (D) Mathematical analysis of NMR spectrum using matrix algebra to calculate relative distribution of tautomers. The elements of matrix A represent the number of active protons from each tautomer (columns) that contribute to each of the six NMR peaks (rows). The matrix X elements are the unknown variables, which represent the relative amounts of each tautomer. Matrix B contains the areas corresponding to each peak in the NMR spectrum at -50°C . Linear equations were generated from the matrix equation $A \times X = B$. Solving the system of linear equations yielded values for the unknowns, which provided the relative distribution of individual tautomers of KP1212.

geometric isomers together (i.e., **2Z** and **2E** together and **5Z** and **5E** together).

The NMR data also allowed the determination of the relative ratios of the tautomeric species present in the DMF solution at -50°C (Fig. 4 B and C). Several peaks in the NMR spectrum partially overlapped (e.g., iv and v), so the raw data were reconstructed using Varian NMR simulation software (*SI Appendix*, Fig. S8). The reconstructed peaks from the simulation were integrated, yielding an estimate of the relative amounts of each of the proton resonances in the spectrum. The deconvolution of the NMR spectrum was done in two steps: (i) Each tautomer was analyzed for the

types of protons (imino/enolic/amido/amino) it contained on the nucleobase; this analysis revealed which of the six NMR peaks observed contained which of the proton signals from a given tautomer. (ii) Based on the peak assignments from i, a system of linear equations was set up using the relative amounts of tautomers as unknown variables (Fig. 4D). The matrix form of this system is also shown in Fig. 4D. Each NMR peak provided a linear equation whereby the peak area should be equal to the sum of all tautomers (as relative amounts) containing protons that contribute to that peak. The solutions of these equations yielded the relative ratios of the tautomeric species present in DMF at -50°C (Fig. 4D).

The unique solution of the system of equations presented above also served as a validation for the correct NMR assignment of different tautomer resonances. An exhaustive analysis was performed where other possible assignments were explored (*SI Appendix, Supporting Note*). However, in all other cases, the resulting system of linear equations either had no solution or did not produce chemically sensible results (e.g., negative amounts of a given tautomer). Moreover, we also investigated whether a smaller number of tautomers could explain the spectroscopic data. A comprehensive analysis showed that only when five tautomeric species are considered, the assignment becomes possible.

It has been reported that nucleosides and nucleobases can exist as self-dimers (45, 46) and self-oligomers (47–49) under certain conditions. To exclude the contribution of any putative dimers or oligomers of KP1212 that could be forming in DMF solution at low temperatures, the NMR studies were repeated using 1:3.2 and 1:10 dilutions of the 10 mg/mL KP1212 solution. The NMR spectra of the diluted KP1212 solutions (*SI Appendix, Figs. S7 A and B*) were very similar to the spectra of the original, undiluted solution (Fig. 3), indicating that formation of dimers or higher order oligomers did not occur to any significant extent in the experiments reported.

The relative proportions of the tautomeric species of KP1212 deconvoluted from the NMR spectrum highlighted some unique and nonobvious features of the molecule. Contrary to our expectations, the enol tautomers of KP1212 were found to be the dominant species at -50°C in DMF (51% for **1**, 38% for **2**, Fig. 4C), whereas the keto tautomers **3** to **5** collectively constituted only 11% of the total mixture. By contrast, for normal DNA bases (such as cytosine), the keto tautomers are the major ones, whereas enol tautomers are very rare (31, 38, 50). Taken together, the NMR observations demonstrated the existence of various tautomeric forms of KP1212 in solution, suggesting

a putative mechanistic basis for the observed mutagenic properties of KP1212 (Fig. 2).

IR Spectroscopy Establishes the Tautomeric Properties of KP1212 in Aqueous Solution. Whereas NMR spectroscopy demonstrated the presence of different tautomeric forms of KP1212, most of the measurements were carried out at low temperature, with the nucleoside dissolved in the nonaqueous solvent DMF. To probe whether KP1212 can also form multiple tautomers under more physiologically relevant conditions, variable temperature FTIR and 2D IR experiments were performed in D_2O solutions of KP1212, buffered with deuterated potassium phosphate (0.5 M, pH in D_2O , $\text{pD} = 7.9$). In the region of in-plane double bond vibrations for aromatic heterocycles (Fig. 5A), vibrational bands from different tautomers are expected to have distinct patterns (50). The vibrational mode at $1,666\text{ cm}^{-1}$ was assigned to the carbonyl stretch present in the KP1212 keto tautomers based on the characteristic vibrational frequency and the broad line shape. Deoxycytidine is known to display a similar keto-carbonyl stretch at $1,651\text{ cm}^{-1}$ (51). However, compared with deoxycytidine at a similar concentration, the intensity of the KP1212 carbonyl stretch was clearly reduced, indicating a possible reduction in the population of keto tautomers and the presence of a significant enol tautomer population.

To provide further evidence of the existence of multiple tautomers, 2D IR spectra of KP1212 were recorded. The 2D IR spectroscopy is analogous to 2D NMR: sequences of ultrafast IR pulses are used to excite molecular vibrations, and the energy flow from one vibration to others is then detected. The correlation of excitation and detection frequencies allows mixtures of tautomers to be separated through the cross-peaks that encode their intramolecular vibrational couplings (52, 53). The 2D IR spectrum of KP1212 (Fig. 5B) showed intense cross-peaks among the three lower frequency diagonal peaks. However, no cross-peaks were detected between the $\text{C}=\text{O}$ stretch peak and the other peaks, indicating that the $1,666\text{ cm}^{-1}$ mode and the lower frequency modes originated from two separate species, namely, keto and enol tautomers, respectively. The variable temperature IR data were also consistent with the 2D IR result. When the temperature was increased, the $\text{C}=\text{O}$ peak intensity increased, whereas some of the lower frequency modes decreased in intensity (Fig. 5A). This result indicated that both keto and enol tautomeric forms were present in solution, in a thermodynamic equilibrium, and the proportion of the keto tautomeric form increased at elevated temperatures. Taken together, the IR data helped establish the presence of multiple tautomeric forms of KP1212 in aqueous solution at physiological pH.

Discussion

Multiple tautomeric forms of canonical nucleic acid bases have long been thought to exist and to provide a basis for spontaneous mutations (31–38). Because minor tautomers (e.g., enol or imino) are typically rare, obtaining direct evidence to support their existence has been a challenge to experimentalists. The present study, by using a battery of complementary spectroscopic tools (41), has shown that KP1212, unlike canonical nucleic acid bases, exists as an ensemble of up to five tautomers in solution. The data further show that KP1212 displays a strong propensity to exist in an enol form, rather than the classical keto form characteristic of nucleic acid bases. It is tempting to speculate that one or more members of the structural ensemble of KP1212 isomers were responsible for the mutagenic properties of the base. Nevertheless, the explicit relationship between the minor tautomers observed by spectroscopy and the mutagenic properties of KP1212 seen in the biological experiments is not immediately obvious. In the next section, we propose a model that attempts to correlate the mutagenic properties of KP1212 to its ability to adopt multiple tautomeric forms.

One of the fundamental assumptions in developing a mutagenesis model is that each KP1212 tautomer has a distinct base-pairing preference. To establish these preferences, in a relatively

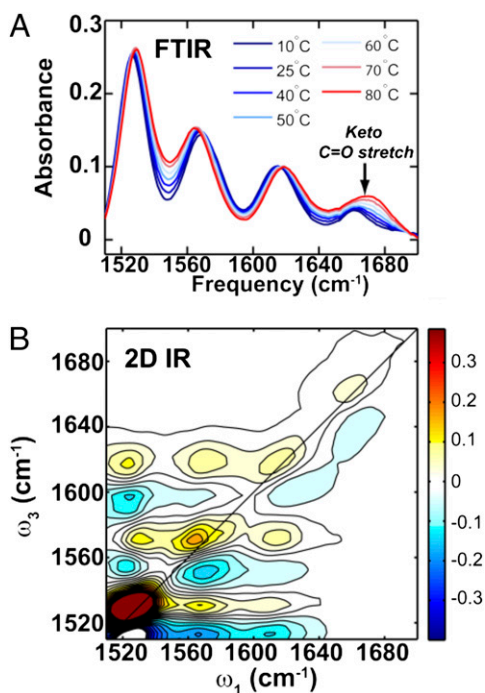


Fig. 5. IR studies demonstrate the existence of different tautomeric forms of KP1212. (A) Variable temperature FTIR spectra of KP1212 in deuterated 0.5 M phosphate buffer $\text{pD} = 7.9$ taken at various temperatures (10°C – 80°C). The black arrow indicates the temperature-dependent increase of the keto-carbonyl stretch. (B) 2D IR spectrum of KP1212 under the same solvent condition as FTIR at 20°C .

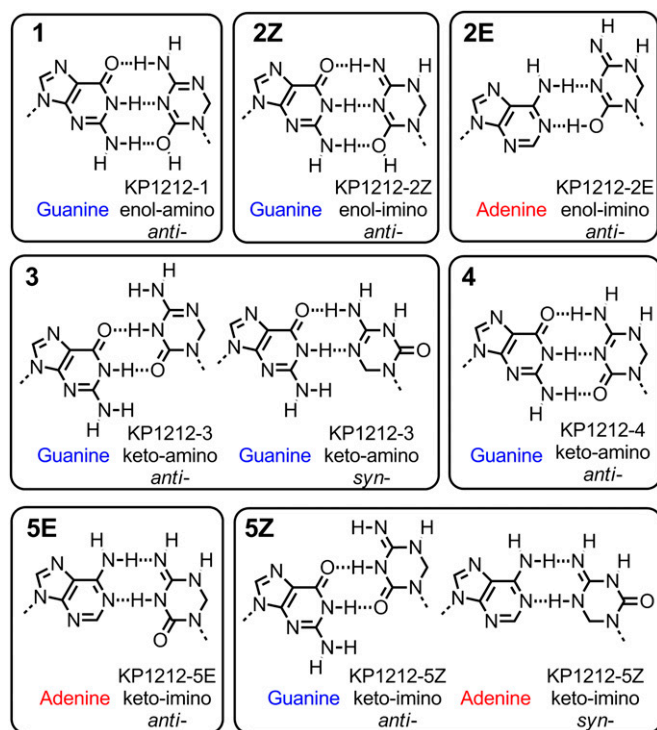


Fig. 6. A mechanistic model explaining the mutagenesis of KP1212 by the distinct base-pairing preferences of its different tautomers. KP1212 tautomers are paired with purine bases (G or A), by maximizing the number of possible hydrogen bonds between the bases. Tautomers **1**, **2Z**, **4**, and **5E** have a canonical Watson–Crick face and therefore are proposed to pair exclusively with either G or A. The remaining tautomers are proposed to pair either in wobble position (**2E**, **3**, and **5Z**) or involving a *syn*-conformer of KP1212 (**3** and **5Z**).

simple modeling exercise, each tautomer was analyzed for its most probable base-pairing partner on the basis of hydrogen bonding. The primary criterion for correct base pairing was formation of the maximum number of Watson–Crick type hydrogen bonds between the bases. Each type of hydrogen bond was considered to confer an equal stabilization energy. Configurations in which multiple base pairs had equal hydrogen bond numbers were considered equally probable. Because KP1212 is a cytosine (pyrimidine) analog, and the primary mutagenic events are A to G or G to A, our modeling exercise explored only purine–pyrimidine putative base pairs. Several KP1212 tautomers featured a canonical Watson–Crick face. Thus, tautomers enol-amino (**1**), enol-imino (**2Z**) and keto-amino (**4**) were predicted to pair with guanine, whereas keto-imino (**5E**) was predicted to pair with adenine (Fig. 6). The remaining tautomers could not adopt a canonical Watson–Crick face and therefore were hypothesized to pair in other geometries (i.e., wobble pairings, or *syn*-conformer pairings). Fig. 6 shows different possibilities for tautomers **2E**, **3**, and **5Z**, which can pair with guanine (**3**), adenine (**2E**), or both (**5Z**).

Using the aforementioned model, the tautomeric distribution deconvoluted from the NMR data (Fig. 4 *C* and *D*) allows predictions to be made on the type and amount of base pairing to be expected for KP1212. The model predicts that most of the time KP1212 would pair with G, which corresponded to the cumulative proportion of tautomers **1**, **2Z**, **3**, and **4**. Because KP1212 was also present as tautomers **2E** or **5E**, the model predicts it can also pair with A. Additionally, tautomer **5Z** can pair with either G or A. The relative distribution of tautomers deconvoluted from the NMR spectra did not provide quantitative information on the relative ratios between the *cis/trans* imine geometric isomers (i.e., **2Z** and **2E** or **5E** and **5Z**). Therefore, it was not possible to estimate pre-

cisely how often KP1212 pairs with G vs. A. However, if the additional assumption were made that the imine geometric isomers were present in equal proportions (i.e., their free energies were similar), then the model predicts that KP1212 would pair ~80% of the time with G and ~20% of the time with A. These numbers are reasonably in accord with the experimentally observed mutagenic properties of KP1212 (Fig. 2 *C* and *E*), correctly predicting that: (i) the dominant base-pairing partner of KP1212 is guanine, and (ii) KP1212 can also pair with adenine, at a lower frequency.

There are certain caveats to the proposed model. The tautomerization of KP1212 was established from our spectroscopic data on the free molecule in solution, which reflects the properties of the molecule present in the deoxynucleotide pool in living cells. Our genetic studies, by contrast, were carried out on a DNA oligonucleotide containing one KP1212 base.

One design feature of the KP1212 molecule was the lack of an aromatic system; as a consequence, it is not expected to be significantly stabilized by the cation– π interactions of normal DNA base stacking. Therefore, it seems reasonable to speculate that the electronic properties of KP1212, including its ability to exist as multiple tautomers, would be similar when the base is present as a free nucleoside and when it is present in an oligonucleotide, although the exact distribution of major and minor tautomeric forms would likely vary. Supporting this view is the fact that KP1212 is present as multiple tautomeric species both in DMF and in water, solvents of very different polarities. Nevertheless, it remains for further spectroscopic work to establish the differences between the tautomeric distributions of KP1212 as a free nucleoside and as part of a DNA strand. Further work will also address the issue of how the rate of tautomeric interconversion compares in the two cases and how it relates to the rate of nucleotide incorporation during nucleic acid synthesis.

The model presented above assumes that genetic information corruption induced by KP1212 is due to tautomerism and/or *syn*–*anti* rotamerism. However, other chemical properties, such as ionization and anomerization, could also contribute to the mutagenic properties of a nucleoside analog. The formation of ionized base pairs has been proposed to account for the mutagenicity of 5-fluoro-uracil (14). The telltale sign of the involvement of ionized base pairs is a pH dependence of mutagenicity (14). In preliminary studies using an *in vitro* polymerase extension reaction, the mispairing frequency of KP1212 did not vary significantly over a small pH range (pH 7.0 and 7.8). However, further work is needed to establish the extent (if any) to which ionized base pairs contribute to the mutagenicity of KP1212. The other mechanism that may contribute to mutagenesis involves the anomerization of the sugar portion of the nucleoside, generating an alpha-anomer (54). The presence of the alpha-anomer of the KP1212 nucleoside, however, was ruled out from ^1H and ^{13}C NMR studies (*SI Appendix*, Tables S3 and S4).

The genetic and spectroscopic results presented here complement earlier biochemical work by Anderson and coworkers who performed a detailed analysis of the kinetics of pairing of KP1212 by HIV RT, DNA polymerase γ , and DNA polymerase β (30). Using single turnover methods, they observed that the quantitative features of KP1212 pairing with opposing bases is very polymerase dependent. It is difficult to extrapolate their data to ours but some observations are noteworthy. They predict a mutation rate that is fivefold or more lower than what we observed (30), but this result seems reasonable given the differences in the systems used (e.g., they provided presteady-state and single turnover kinetic constants, whereas we looked at the end product under steady-state conditions in the presence of all four dNTPs; moreover, some of our studies were carried out in living cells, whereas their study was performed *in vitro*). Nevertheless, their data, as ours, predict pairing of KP1212 with both G and A, with pairing with G being the favored event (30).

A recent theoretical analysis of lethal mutagenesis applied to HIV predicts that a two- to sixfold enhancement in mutation rate will result in error catastrophe and population collapse (55). These data are in accord with more experimental predictions by Loeb et al. (7) and Harris et al. (29). In this context, how does the 10% mutation frequency of KP1212 measured in the current study relate to the viral population collapse observed in biological systems? One necessary assumption is that the 10% mutation frequency observed *in vitro* and *in vivo* in this study would also apply to a system in which HIV reverse transcriptase (RT) was the main replicative enzyme. When HIV RT replicates the HIV genome, it introduces about one error in 10^4 bases replicated, or one error per genome (the HIV genome is 9,749 bases long). The net increase in the viral mutation frequency due to KP1212 will depend not only on its intrinsic mutagenicity (10%), but also on how many KP1212 bases get incorporated every replication cycle. Given that each KP1212 has a 10% chance of miscoding, for every 10 KP1212s present in the HIV genome, on average one additional mutation per replication cycle will be introduced, which corresponds to a twofold increase in the viral mutation frequency. Therefore, to achieve a two to sixfold increase in mutation frequency, HIV RT needs to introduce one to five additional mutations, which can be achieved by incorporating 10–50 KP1212s every replication round.

The number of KP1212 bases incorporated depends on the concentration of the KP1212 triphosphate, relative to the concentration of dCTP (KP1212 is a deoxycytidine analog) in the dNTP pool, and the relative catalytic efficiency of the viral polymerase to incorporate KP1212 relative to dCTP. Kinetic data of Anderson and coworkers show that HIV RT is 15 times more efficient (k_{cat}/K_M) at incorporating dCTP over KP1212 triphosphate (30). Therefore, to incorporate 10–50 KP1212s per cycle into the HIV genome, which contains about 2,400 guanines (56), the concentration of the KP1212 triphosphate should be 6–30% of that of dCTP, which corresponds (assuming an intracellular concentration of dCTP of $\sim 10 \mu\text{M}$) (57) to 0.6–3.2 μM KP1212 triphosphate. Micromolar concentrations are achievable in clinical settings for nucleotide analogs, such as azidothymidine (AZT) (58). This calculation in fact is an overestimate of the level of KP1212 needed to cause HIV population collapse, because it only takes into account incorporation of KP1212 coming from the nucleotide pool opposite G. The observed A-to-G mutations both in cell culture (29) and clinical isolates (28) suggest that KP1212 is also incorporated opposite A. Because the HIV genome has an even higher percentage of A (35%) than G (24%) (56), the concentration of KP1212 triphosphate needed to achieve the two- to sixfold increase in the mutation rate of HIV should be even lower than the range estimated above (0.6–3.2 μM).

Whereas the 10% mutation frequency of KP1212 is sufficient to induce HIV population collapse in cell culture experiments (29), if one were to design a better lethal mutagen, what would its theoretically maximum mutation frequency be? This value will depend on the number of possible base-pairing partners of the candidate lethal mutagen. If a pool nucleotide can pair only with two partners (e.g., A and G, using KP1212 as an example), it would be maximally mutagenic when it can pair with A half of the time and with G half of the time. Its mutation frequency would be 50%. Similarly, if a nucleotide has three pairing possibilities, it could achieve a mutation frequency of 67%, and if it has four (the maximum) pairing partners, its maximal mutation frequency would be 75%. KP1212 pairs only with two partners and has a mutation frequency measured at 10%, so, by the criteria described above, it is still fivefold away from the theoretical maximum mutation rate. Thus, the experimental system developed here provides an approach whereby one can assess, in quantitative terms, the proximity of any future antiviral mutagenic candidate to the theoretically maximum mutation rate.

Stepping back, the lethally mutagenic power of a pool nucleotide derives from several factors (Fig. 1B): (i) the likelihood that its nucleoside precursor will appear in the bloodstream and penetrate target cells; (ii) its efficient conversion to the triphosphate form, which must be acceptable to the polymerase of the targeted system (e.g., HIV RT; and, of course the nucleotide ideally should be unacceptable to the host polymerases to avoid off-target effects); (iii) the likelihood that the molecule will diversify in solution into ambiguously pairing forms (e.g., enol tautomers, imine tautomers, syn-anti rotamers, etc.); (iv) the presence of ambiguously pairing forms of the lethal mutagen in the active site of a polymerase; (v) the actual formation of promiscuous base pairs in the active site; and (vi) avoidance of polymerase proofreading, enzymatic repair, and recombination to suppress mutations. Although complex, it is useful to break the overall problem down to discrete steps to evaluate each from the experimental standpoint. This work addressed the steps involving the diversification of the candidate lethal mutagen in solution and the ability of the diversified mutagen population to cause mutations *in vitro* and *in vivo*.

In summary, this study used in concert a slate of synthetic, spectroscopic, and genetic methods to investigate how a candidate antiviral agent achieved its mutagenic mechanism of action. The demonstration of multiple tautomeric forms of KP1212 in solution presents opportunities, through the synthesis of new analogs, for rationally reprogramming the genetic landscapes of replicating organisms. The toolset described herein could be applied to advance the development of other tautomerizable/rotamerizable nucleosides that could be used to elevate the inherent mutation rates of other fast-mutating viruses, such as those of the hepatitis C, influenza, and dengue families, above the error catastrophe limit.

Materials and Methods

Oligonucleotide Synthesis. KP1212 or the m3C control lesion were incorporated into 16mer oligonucleotides (5'-GAAGACCTXGGCGTCC-3', where X is the lesion or the cytosine control) by using phosphoramidite solid-phase methods described before (42, 43). The phosphoramidites were from Berry and Associates. The syntheses were done at the Keck Oligonucleotide Synthesis Facility, Yale School of Medicine, New Haven, CT.

Lesion Bypass and Mutagenesis Assays. For the *in vivo* toxicity and mutagenicity studies, the oligonucleotides were ligated into an M13mp7(L2) single-stranded viral genome by using reported methods (42, 43). The viral genomes were then electroporated into *E. coli* strains HK81 (as AB1157, but *nalA*) and HK82 (as AB1157, but *nalA alkB22*; *AlkB* deficient), or used as templates for *in vitro* polymerase extension studies. Lesion toxicity was measured using the competitive replication and bypass (CRAB) assay and mutational analysis, both for the *in vivo* and *in vitro* study was performed using the restriction endonuclease and postlabeling (REAP) assay (42). Further details are provided in *SI Appendix, SI Materials and Methods*.

NMR Analysis of KP1212. All NMR experiments were performed on Varian 500 MHz NMR spectrometers. The deuterated solvent DMF- d_7 was from Cambridge Isotope Laboratories. The 5-aza-5,6-dihydro-2'-deoxycytidine (KP1212) nucleoside was from Berry and Associates. The NMR sample was prepared by dissolving 1.0, 3.2, or 10 mg of KP1212 nucleoside in 1.0 mL DMF- d_7 (final concentration ~ 4.4 , 14.1, or 44 mM, respectively). The ^1H NMR spectra were reported in parts per million (ppm) and were referenced to the signals for DMF- d_7 (8.03 ppm). The ^{13}C NMR spectra were referenced to the signals for DMF- d_7 (163.15 ppm).

Infrared Spectroscopy. For both 1D FTIR and 2D IR experiments, the H/D exchanged KP1212 was dissolved at a concentration of 20 mg/mL (88 mM) in 0.5 M phosphate buffer pD (pH reading in D_2O) 7.9. About 25 μL of sample solution was sandwiched between two CaF_2 windows separated by a 50- μm Teflon spacer. Variable-temperature FTIR spectra were collected using Nicolet 380 FTIR spectrometer at 1.0 cm^{-1} resolution with 16 scans per spectrum. Spectra for both the sample and the D_2O were collected with the same procedure and the solvent spectra were subtracted from the sample spectra.

Absorptive 2D IR spectra were collected using a 2D IR spectrometer as described in detail previously (59). The relative polarizations of the pulses were set to be perpendicular (ZZYY). The waiting time (τ_2) between the first two pulses and the third pulse was fixed at 150 fs. The coherence time between the first and the second pulse was scanned in 4 fs steps from -60 fs to 2.8 ps and 2.0 ps for rephasing and nonrephasing spectra, respectively. The coherence time (τ_1) was Fourier transformed to obtain the first frequency axis ω_1 . The heterodyned signal was dispersed in a monochromator to obtain the ω_3 frequency dimension and collected using a 64×2 pixel mercury-cadmium-telluride (MCT) array detector. Linear absorption from the solvent and solute was divided out along both the ω_1 and ω_3 axes to remove spectral distortions (60).

- Frenkel LM, et al. (2003) Multiple viral genetic analyses detect low-level human immunodeficiency virus type 1 replication during effective highly active antiretroviral therapy. *J Virol* 77(10):5721–5730.
- Mullins JI, Jensen MA (2006) Evolutionary dynamics of HIV-1 and the control of AIDS. *Curr Top Microbiol Immunol* 299:171–192.
- Johnston R (2010) HIV cure: Controversy, consensus, and a consortium. *AIDS Res Hum Retroviruses* 26(9):943–946.
- Esté JA, Cihlar T (2010) Current status and challenges of antiretroviral research and therapy. *Antiviral Res* 85(1):25–33.
- Broder S (2010) The development of antiretroviral therapy and its impact on the HIV-1/AIDS pandemic. *Antiviral Res* 85(1):1–18.
- Eigen M (2002) Error catastrophe and antiviral strategy. *Proc Natl Acad Sci USA* 99(21):13374–13376.
- Loeb LA, et al. (1999) Lethal mutagenesis of HIV with mutagenic nucleoside analogs. *Proc Natl Acad Sci USA* 96(4):1492–1497.
- Smith RA, Loeb LA, Preston BD (2005) Lethal mutagenesis of HIV. *Virus Res* 107(2): 215–228.
- Clouser CL, Patterson SE, Mansky LM (2010) Exploiting drug repositioning for discovery of a novel HIV combination therapy. *J Virol* 84(18):9301–9309.
- Graci JD, Cameron CE (2008) Therapeutically targeting RNA viruses via lethal mutagenesis. *Future Virol* 3(6):553–566.
- Perales C, Martín V, Domingo E (2011) Lethal mutagenesis of viruses. *Curr Opin Virol* 1(5):419–422.
- Elena SF (2012) RNA virus genetic robustness: Possible causes and some consequences. *Curr Opin Virol* 2(5):525–530.
- de la Torre JC (2005) Arenavirus extinction through lethal mutagenesis. *Virus Res* 107(2):207–214.
- Bonnac LF, Mansky LM, Patterson SE (2013) Structure-activity relationships and design of viral mutagens and application to lethal mutagenesis. *J Med Chem* 56(23):9403–9414.
- Koito A, Ikeda T (2013) Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases. *Front Microbiol* 4:28.
- Jaszczur M, Bertram JG, Pham P, Scharff MD, Goodman MF (2013) AID and APOBEC3G haphazard deamination and mutational diversity. *Cell Mol Life Sci* 70(17):3089–3108.
- Smyth RP, Davenport MP, Mak J (2012) The origin of genetic diversity in HIV-1. *Virus Res* 169(2):415–429.
- Ortega-Prieto AM, et al. (2013) Extinction of hepatitis C virus by ribavirin in hepatoma cells involves lethal mutagenesis. *PLoS ONE* 8(8):e71039.
- Dietz J, et al. (2013) Deep sequencing reveals mutagenic effects of ribavirin during monotherapy of hepatitis C virus genotype 1-infected patients. *J Virol* 87(11): 6172–6181.
- Moreno H, Grande-Pérez A, Domingo E, Martín V (2012) Arenaviruses and lethal mutagenesis. Prospects for new ribavirin-based interventions. *Viruses* 4(11): 2786–2805.
- Graci JD, Cameron CE (2002) Quasispecies, error catastrophe, and the antiviral activity of ribavirin. *Virology* 298(2):175–180.
- Crotty S, et al. (2000) The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. *Nat Med* 6(12):1375–1379.
- Baranovich T, et al. (2013) T-705 (favipiravir) induces lethal mutagenesis in influenza A H1N1 viruses in vitro. *J Virol* 87(7):3741–3751.
- Manrubia SC, Domingo E, Lázaro E (2010) Pathways to extinction: Beyond the error threshold. *Philos Trans R Soc Lond B Biol Sci* 365(1548):1943–1952.
- Domingo E, Sheldon J, Perales C (2012) Viral quasispecies evolution. *Microbiol Mol Biol Rev* 76(2):159–216.
- Domingo E, Grande-Pérez A, Martín V (2008) Future prospects for the treatment of rapidly evolving viral pathogens: Insights from evolutionary biology. *Expert Opin Biol Ther* 8(10):1455–1460.
- Domingo E, et al. (2006) Viruses as quasispecies: Biological implications. *Curr Top Microbiol Immunol* 299:51–82.
- Mullins JI, et al. (2011) Mutation of HIV-1 genomes in a clinical population treated with the mutagenic nucleoside KP1461. *PLoS ONE* 6(1):e15135.
- Harris KS, Brabant W, Styrchak S, Gall A, Daifuku R (2005) KP-1212/1461, a nucleoside designed for the treatment of HIV by viral mutagenesis. *Antiviral Res* 67(1):1–9.
- Murakami E, Basavapathruni A, Bradley WD, Anderson KS (2005) Mechanism of action of a novel viral mutagenic covalent nucleotide: Molecular interactions with HIV-1 reverse transcriptase and host cell DNA polymerases. *Antiviral Res* 67(1):10–17.
- Watson JD, Crick FH (1953) Genetical implications of the structure of deoxyribonucleic acid. *Nature* 171(4361):964–967.
- Lowdin P-O (1963) Proton tunneling in DNA and its biological implications. *Rev Mod Phys* 35:724–732.
- Topal MD, Fresco JR (1976) Complementary base pairing and the origin of substitution mutations. *Nature* 263(5575):285–289.
- Goodman MF (1995) DNA models. Mutations caught in the act. *Nature* 378(6554): 237–238.
- Kunkel TA (2004) DNA replication fidelity. *J Biol Chem* 279(17):16895–16898.
- Rüterjans H, Kaun E, Hull WE, Limbach HH (1982) Evidence for tautomerism in nucleic acid base pairs. 1H NMR study of 15N labeled tRNA. *Nucleic Acids Res* 10(21):7027–7039.
- Purrello R, et al. (1993) Keto-iminol tautomerism of protonated cytidine monophosphate characterized by ultraviolet resonance Raman spectroscopy: implications of C+ iminol tautomer for base mispairing. *J Am Chem Soc* 115:760–767.
- Wang W, Hellinga HW, Beese LS (2011) Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. *Proc Natl Acad Sci USA* 108(43): 17644–17648.
- Suen W, Spiro TG, Sowers LC, Fresco JR (1999) Identification by UV resonance Raman spectroscopy of an imino tautomer of 5-hydroxy-2'-deoxycytidine, a powerful base analog transition mutagen with a much higher unfavored tautomer frequency than that of the natural residue 2'-deoxycytidine. *Proc Natl Acad Sci USA* 96(8):4500–4505.
- Sečková P, et al. (2004) Direct determination of tautomerism in purine derivatives by low-temperature NMR spectroscopy. *Tetrahedron Lett* 45:6259–6263.
- Singh V, et al. (2014) Direct observation of multiple tautomers of oxythiamine and their recognition by the thiamine pyrophosphate riboswitch. *ACS Chem Biol* 9(1):227–236.
- Delaney JC, Essigmann JM (2006) Assays for determining lesion bypass efficiency and mutagenicity of site-specific DNA lesions in vivo. *Methods Enzymol* 408:1–15.
- Delaney JC, Essigmann JM (2004) Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB *Escherichia coli*. *Proc Natl Acad Sci USA* 101(39):14051–14056.
- Friedberg EC (2006) *DNA Repair And Mutagenesis* (ASM Press, Washington, DC), 2nd Ed.
- Nir E, Kleinermanns K, de Vries MS (2000) Pairing of isolated nucleic-acid bases in the absence of the DNA backbone. *Nature* 408(6815):949–951.
- Armentano D, et al. (2004) Self-assembly of cytosine nucleoside into triply-bound dimers in acid media. A comprehensive evaluation of proton-bound pyrimidine nucleosides by electrospray tandem mass spectrometry, X-rays diffractometry, and theoretical calculations. *J Am Soc Mass Spectrom* 15(2):268–279.
- Giorgi T, et al. (2002) Gel-like lyomesophases formed in organic solvents by self-assembled guanine ribbons. *Chemistry* 8(9):2143–2152.
- Davis JT (2004) G-quartets 40 years later: From 5'-GMP to molecular biology and supramolecular chemistry. *Angew Chem Int Ed Engl* 43(6):668–698.
- Sivakova S, Rowan SJ (2005) Nucleobases as supramolecular motifs. *Chem Soc Rev* 34(1):9–21.
- Miles HT (1961) Tautomeric forms in a polynucleotide helix and their bearing on the structure of DNA. *Proc Natl Acad Sci USA* 47:791–802.
- Peng CS, Jones KC, Tokmakoff A (2011) Anharmonic vibrational modes of nucleic acid bases revealed by 2D IR spectroscopy. *J Am Chem Soc* 133(39):15650–15660.
- Peng CS, Tokmakoff A (2012) Identification of lactam–lactim tautomers of aromatic heterocycles in aqueous solution using 2D IR spectroscopy. *J Phys Chem Lett* 3(22): 3302–3306.
- Peng CS, Baiz CR, Tokmakoff A (2013) Direct observation of ground-state lactam–lactim tautomerization using temperature-jump transient 2D IR spectroscopy. *Proc Natl Acad Sci USA* 110(23):9243–9248.
- Matoušová M, et al. (2011) 2'-deoxy-5,6-dihydro-5-azacytidine - a less toxic alternative of 2'-deoxy-5-azacytidine: A comparative study of hypomethylating potential. *Epi-genetics* 6(6):769–776.
- Tripathi K, Balagam R, Vishnoi NK, Dixit NM (2012) Stochastic simulations suggest that HIV-1 survives close to its error threshold. *PLoS Comput Biol* 8(9):e1002684.
- van der Kuyl AC, Berkhout B (2012) The biased nucleotide composition of the HIV genome: A constant factor in a highly variable virus. *Retrovirology* 9:92.
- Gandhi VV, Samuels DC (2011) A review comparing deoxyribonucleoside triphosphate (dNTP) concentrations in the mitochondrial and cytoplasmic compartments of normal and transformed cells. *Nucleosides Nucleotides Nucleic Acids* 30(5):317–339.
- Rodman JH, et al. (1999) Systemic pharmacokinetics and cellular pharmacology of zidovudine in human immunodeficiency virus type 1-infected women and newborn infants. *J Infect Dis* 180(6):1844–1850.
- Chung HS, Khalil M, Smith AW, Tokmakoff A (2007) Transient two-dimensional IR spectrometer for probing nanosecond temperature-jump kinetics. *Rev Sci Instrum* 78(6):063101.
- Jones KC, Ganim Z, Peng CS, Tokmakoff A (2012) Transient two-dimensional spectroscopy with linear absorption corrections applied to temperature-jump two-dimensional infrared. *J Opt Soc Am B* 29:118–129.