Protonation, Tautomerism, and Base Pairing of the Antiviral Favipiravir (T-705)

Gabriel da Silva*

Department of Chemical Engineering, The University of Melbourne, Victoria 3010 Australia

*gdasilva@unimelb.edu.au

ABSTRACT

Favipiravir (T-705) is an antiviral medication used to treat influenza. T-705 is also currently being trialled as a repurposed COVID-19 treatment. To help accelerate these efforts, this study provides important solution-phase properties of T-705 determined via computational chemistry. Density functional theory (DFT) calculations combined with the SMD continuum solvation model demonstrate that T-705 prefers the aromatic enol form in solution over the ketone tautomer. Deprotonation constants for the conjugate acids of T-705 (pK_as) are then evaluated, by combining the DFT/SMD calculations with accurate G4 gas-phase basicities. These calculations indicate that T-705 is a weak base that should not significantly protonate at physiological pH. The preferential site for protonation is at the ring nitrogen *ortho* to the alcohol functional group ($pK_a \sim 7.4$), followed by protonation of the oxygen on the amide side-chain at more acidic conditions ($pK_a \sim -9.8$). Significantly, protonation of the ring nitrogen produces an acid that can deprotonate to the enol form ($pK_a \sim -5.1$), providing a pathway for their interconversion. Finally, base-pairing of the active ribose-bound form of T-705 to cytidine and uridine is also examined. These calculations indicate that both base pairs have large binding free energies of around 4 – 5 kcal/mol, supporting previous findings that T-705 can bind with both nucleobases, leading to mis-incorporation of these pairs into viral RNA.

Manuscript

Favipiravir (T-705, brand name Avigan) is a broad antiviral medication used to treat novel strains of influenza in Japan. As an antiviral, T-705 has been considered under drug repurposing efforts as a possible COVID-19 treatment.[1] T-705 is particularly attractive as it is thought to inhibit the RNA-dependent RNA polymerase enzyme,[2-4] a key protein target in the SARS-CoV-2 virus.[5] Indeed, results of preliminary trials have shown that T-705 appears to be somewhat effective when used to treat COVID-19 patients.[6]

T-705 has both enol and ketone tautomers, as shown in Figure 1. The enol form has an aromatic phenol-like structure, whereas the ketone form has a series of four conjugated double bonds, with a hydrogen on the ring nitrogen *ortho* to the carbonyl group. Recent calculations show that the enol form is considerably more stable than the ketone form.[7] However, there is little information available on the protomers of the T-705 tautomers and their relative stability.

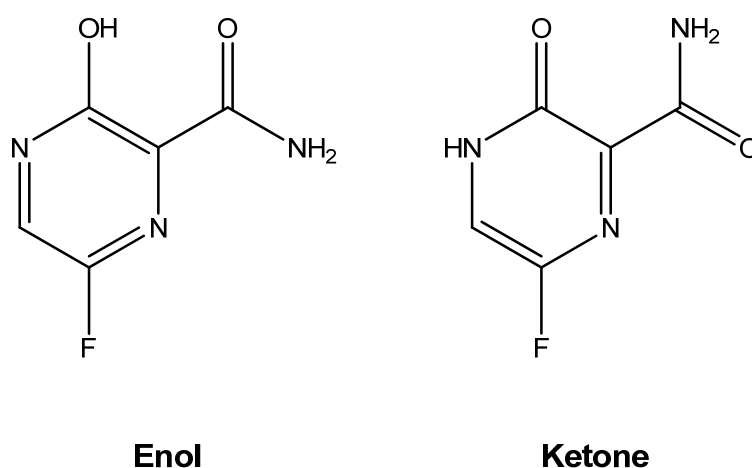
Enol   **Ketone**

Figure 1. Enol and ketone tautomers of T-705.

T-705 is a prodrug, which is metabolised *in vivo* to the active form – a ribosyl triphosphate (T-705RTP, Figure 2). The active form is then believed to act as an analogue of the guanosine and adenosine nucleotides, producing mis-incorporated base pairs with cytidine and uridine.[8] Interestingly, because T-705RTP incorporates the ribosyl triphosphate at the *ortho* ring nitrogen it does not have an enol tautomer available to it.
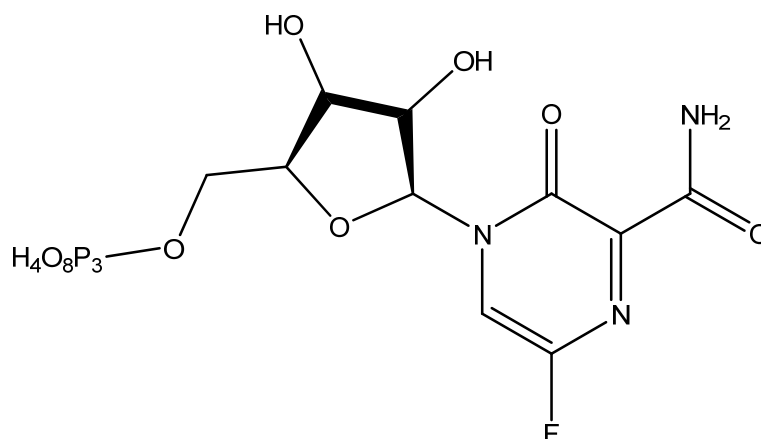
Figure 2. T-705RTP, the active form of T-705.

There is at present an intensive world-wide research effort into discovering treatments for COVID-19, which includes work directed at potentially repurposing antivirals such as T-705.[1,9] It is therefore critical that we understand the properties of these drugs and their modes of action. To aid in these activities this study reports calculated properties of the T-705 tautomers and their protomers, particularly the site-specific pK_a values[10] for deprotonation of the conjugate acids. Additionally, the base-pairing of T-705RTP with cytidine and uridine is also investigated.

As reported recently, the enol tautomer of T-705 is more stable than the keto form (which is subsequently referred to as T-705K).[7] At the M05-2X/cc-pVTZ@SMD level of theory this study predicts that Gibbs energy of the enol form is 5.7 kcal/mol lower than that of the keto form at 298 K. This corresponds to an equilibrium constant of *ca.* 6×10^{-5}, and a correspondingly small equilibrium concentration of the keto tautomer is expected in solution.

Figure 3 illustrates the network of protonation reactions available to T-705, which includes the step-wise process for T-705K formation via T-705+-1. Calculated pK_a values for each step are listed in Table 1, with the average value of the three thermodynamic cycles also included in Figure 3. The mean uncertainty of these calculations is expected to be around 1 pK_a unit, and we see from Table 1 that across the three thermodynamic cycles the standard deviation is 0.4 pK_a units.

The most favorable protonation site on T-705 is at the *ortho* ring nitrogen, with pK_a ~ -7.4, leading to T-705+-1. Given this value, T-705 should not protonate to a significant extent at physiological pH. This reaction leads to protomer T-705+-1, which can subsequently deprotonate to T-705K (pK_a ~ 5.1). Accordingly, this provides a mechanism for equilibration of the keto and enol tautomers of T-705. A second significant protonation reaction is available to T-705, being protonation of the O atom on the amide functional group, with pK_a ~ -9.8. Beyond this no other protomer is predicted to be competitive. For T-705K other than

protonation to T-705[+] 1 as mentioned above, protonation of the amide oxygen to give T-705[+]-K2 has a similar pK_a (-6.5) and may also take place.
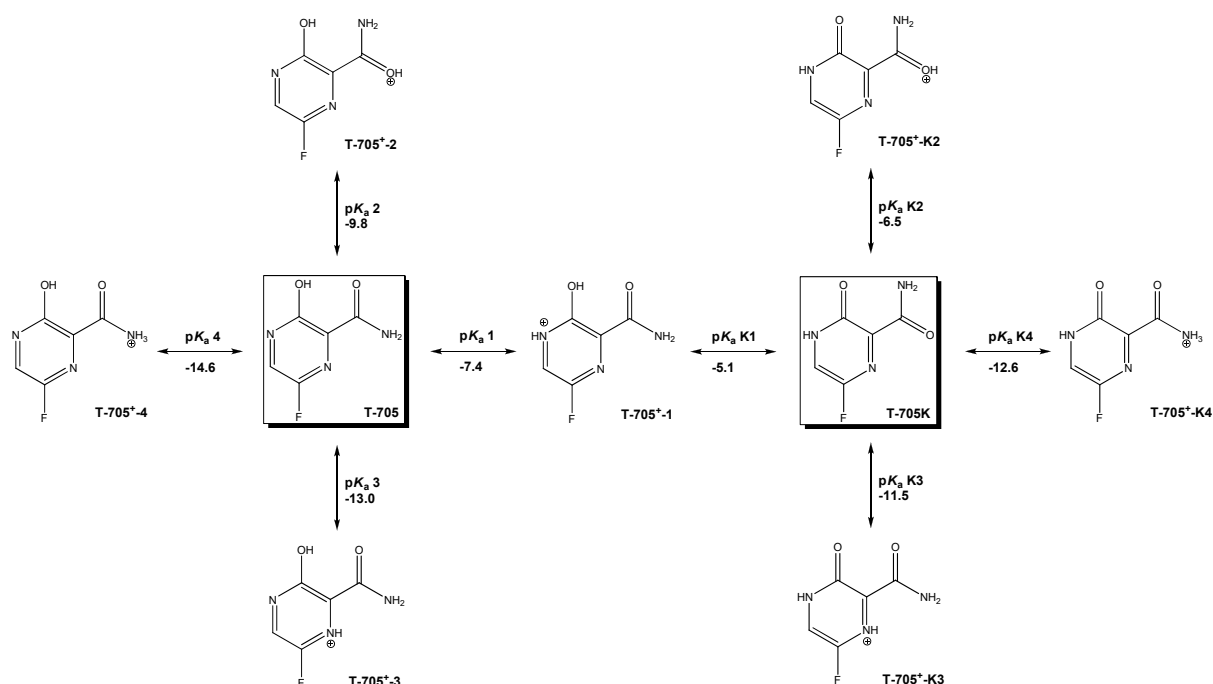

Figure 3. Network of protonation reactions in T-705 and T-705K with calculated pK_a values.

Table 1. Calculated pK_a values for the conjugate acids of T-705 (pK_a 1 – 4) and T-705K (pK_a K1 – K4), determined using the direct, modified direct, and proton exchange thermodynamic cycles. Corresponding protomers identified in Figure 3.

	pK_a 1	pK_a 2	pK_a 3	pK_a 4	pK_a K1	pK_a K2	pK_a K3	pK_a K4
Direct	-7.5	-9.8	-13.1	-14.7	-5.2	-6.6	-11.6	-12.7
Modified Direct	-7.0	-9.3	-12.6	-14.2	-4.7	-6.9	-11.9	-13.0
Proton Exchange	-7.8	-10.1	-13.4	-15.0	-5.5	-6.1	-11.1	-12.2
Average	-7.4	-9.8	-13.0	-14.6	-5.1	-6.5	-11.5	-12.6

Finally, calculations have been carried out to examine the pairing of T-705RTP with the uridine and cytidine nucleotides. Calculations were carried out at the M05-2X/cc-pVTZ@SMD level of theory on structures optimized with the 6-31G(d) basis set, with a truncated T-705RTP model in which the triphosphate group is removed. The optimized structures for these two pairs are shown in Figure 4. T-705RTP is able to both donate and accept H bonds to uridine and cytidine, via the amide side-chain.

The T-705RTP-cytidine dimer has a predicted binding free energy of 5.4 kcal/mol, whereas that of T-705RTP-uridine is 4.6 kcal/mol. These are large binding energies for aqueous base

4

pairs,[11] and support the mechanism of action in which T-705 binds to these viral nucleotides to produce mis-incorporated base pairs. The NH2 group donates a H bond into lone pairs on a carbonyl moiety in uridine (with interatomic H···O distance of 1.6 Å) and a ring N atom in cytidine (H···O distance of 2.0 Å). The amide O atom is the hydrogen bond acceptor in both pairs, with uridine donating from a ring NH group (O···H distance of 2.0 Å) and cytidine donating from NH2 (O···H distance of 1.9 Å).
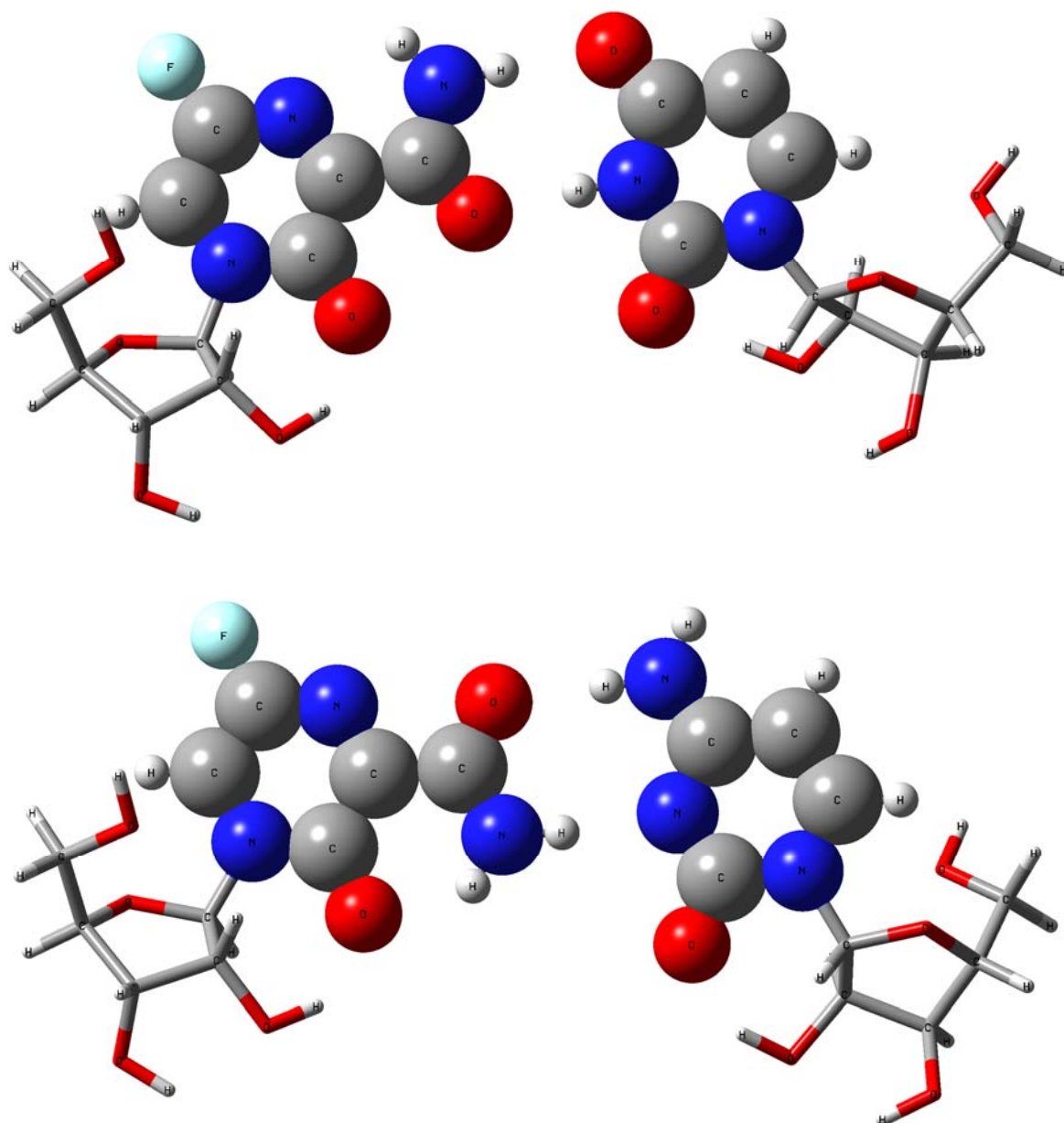

Figure 4. Optimized structures of the T-705RTP model base pairs with uridine (top) and cytidine (bottom), calculated at the M05-2X/6-31G(d)@SMD level of theory. For clarity the furanosyl moieties are depicted as tube structures.

Methods

Electronic structure theory calculations were carried out using the Gaussian 16 code.[12] All solution phase calculations are carried out for water using the SMD[13] continuum solvation model, in conjunction with the M05-2X functional (which was used in the SMD model parametrization). The SMD solvation model is parameterized to provide Gibbs energies at 298 K. Typically, the cc-pvTZ basis set was used, with the smaller 6-31G(d) basis set employed for the larger base pairing calculations. For pK_a calculations, the protocol of Sutton was used,[14] which combines M05-2X/cc-pVTZ@SMD solvation Gibbs energies with (i) high-level gas phase basicities, (2) a thermodynamic cycle which connects to the gas phase, and (3) requisite experimental reaction energies. In this study the G4[15] composite model chemistry was used for single-point energies on M05-2X/cc-pVTZ optimized structures. The direct (D), modified direct (MD), and proton exchange (PX) thermodynamic cycles were used, which incorporate the experimental solvation Gibbs energy of H^+ (D, MD) or the gas-phase basicity of water (PX). The exact nature of the thermodynamic cycles and detailed description of the calculation protocol is given by Sutton et al. Calculations reported here all refer to the most stable conformation, and all optimized structures possess zero imaginary frequencies. Coordinates for all optimized species are provided as Supporting Information.

Supporting Information Available: Optimized geometries for all species.

References

1. Du, Y.-X.; Chen, X.-P. Favipiravir: Pharmacokinetics and concerns about clinical trials for 2019-nCoV infection. *Clinical Pharmacology & Therapeutics* **2020**, doi: 10.1002/cpt.1844.
2. Furuta, Y.; Takahashi, K.; Fukuda, Y.; Kuno, M.; Kamiyama, T.; Kozaki, K.; Nomura, N.; Egawa, H.; Minami, S.; Watanabe, Y.; Narita, H.; Shiraki, K. In vitro and in vivo activities of anti-influenza virus compound T-705. *Antimicrobial Agents and Chemotherapy* **2002**, *46*, 977-981.
3. Furuta, Y.; Takahashi, K.; Kuno-Maekawa, M.; Sangawa, H.; Uehara, S.; Kozaki, K.; Nomura, N.; Egawa, H.; Shiraki, K. Mechanism of action of T-705 against influenze virus. *Antimicrobial Agents and Chemotherapy* **2005**, *49*, 981-986.
4. Gao, Y.; Yan, L.; Huang, Y.; Liu, F.; Zhao, Y.; Cao, L.; Wang, T.; Sun, Q.; Ming, Z.; Zhang, L.; Ge, J.; Zheng, L.; Zhang, Y.; Wang, H.; Zhu, Y.; Zhu, C.; Hu, T.; Hua, T.; Zhang, B.; Yang, X.; Li, J.; Yang, H.; Liu, Z.; Xu, W.; Guddat, L. W.; Wang, Q.; Lou, Z.; Rao, Z. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **2020**, doi: 10.1126/science.abb7498.
5. Furuta, Y.; Gowen, B. B.; Takahashi, K.; Shiraki, K.; Smee, D. F.; Barnard, D. L. Favipiravir (T-705), a novel viral RNA polymerase inhibitor. *Antiviral Research* **2013**, *100*, 446-454.
6. Cai, Q.; Yang, M.; Liu, D.; Chen, J.; Shu, D.; Xia, J.; Liao, X.; Gu, Y.; Cai, Q.; Yang, Y.; Shen, C.; Li, X.; Peng, L.; Huang, D.; Zhang, J.; Zhang, S.; Wang, F.; Liu, J.; Chen, L.; Chen, S.; Wang, Z.; Zhang, Z.; Cao, R.; Zhong, W.; Liu, Y.; Liu, L. Experimental treatment with favipiravir for COVID-19: An open-label control study. *Engineering* **2020**, doi: 10.1016/j.eng.2020.03.007.
7. Antonov, L. Favipiravir tautomerism: A short theoretical report. *ChemRxiv* **2020**, doi: 10.26434/chemrxiv.12115620.

8. Jin, Z.; Smith, L. K.; Rajwanshi, V. K.; Kim, B.; Deval, J. The ambiguous base-pairing and high substrate efficiency of T-705 (favipiravir) ribofuranosyl 5'-triphosphate towards influenza A virus polymerase. *PLOS One* **2013**, 8, e68347.

9. Harrison, C. Coronavirus puts drug repurposing on the fast track. *Nature Biotechnology* **2020**, 38, 379-391.

10. Manallack, D. T. The pK_a distribution of drugs: Application to drug discovery. *Perspectives in Medicinal Chemistry* **2007**, 1, 25-38.

11. Freier, S. M.; Sugimoto, N.; Sinclair, A.; Alkema, D.; Neilson, T.; Kierzek, R.; Caruthers, M. H.; Turner, D. H. Stability of GGCGCp, GCGCYp, and XGCGCTp helixes: An empirical estimate of the energetics of hydrogen bonds in nucleic acids. *Biochemistry* **1986**, *25*, 3214-3219.

12. Frisch, M. J. et al. *Gaussian 16 Rev. B.01*, Gaussian Inc., Wallingford CT, 2016.

13. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *Journal of Physical Chemistry B* **2009**, *113*, 6378-6396.

14. Sutton, C. C. R.; Franks, G. V.; da Silva, G. First principles pKa calculations on carboxylic acids using the SMD solvation model: Effect of thermodynamic cycle, model chemistry, and explicit solvent molecules. *Journal of Physical Chemistry B* **2012**, *116*, 11999-12006.

15. Curtiss, L. A.; Redfern, P. C.; Raghavachari, K., Gaussian-4 theory. *The Journal of Chemical Physics* **2007,** *126*, 084108.