Journal of Enzyme Inhibition and Medicinal Chemistry, October 2006; 21(5): 483–518

informa
healthcare

REVIEW ARTICLE

A kinetic, modeling and mechanistic re-analysis of thymidine phosphorylase and some related enzymes

PHILIP N. EDWARDS

School of Pharmacy & Pharmaceutical Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, UK

(Received 8 August 2005; in final form 24 February 2006)

Abstract
Thymidine phosphorylase (TP) is an important target enzyme for cancer chemotherapy but currently available inhibitors lack in vivo potency. Related enzymes also are therapeutic targets. A greater understanding of enzyme structure and mechanism may help in the design of improved drugs and this work assists in that regard. Also important is the correct identification of the ionization states and tautomeric forms of substrates and products when bound to the enzyme and during the course of the reaction. Approximate methods for estimating some ΔpK_as between aqueous and protein-bound substrates are exemplified for nucleobases and nucleosides. The estimates demonstrate that carbonyl-protonated thymidine and hydroxy tautomers of thymine are not involved in TP's actions. Other estimates indicate that purine nucleoside phosphorylase binds inosine and guanosine as zwitterionic tautomers and that phosphorolysis proceeds through these forms. Extensive molecular modeling based on an X-ray structure of human TP indicates that TP is likely to be mechanistically similar to all other natural members of the class in proceeding through a α-oxacarbenium-like transition state or states.

Keywords: *Thymidine phosphorylase, enzyme-bound substrate pK_as, anomalous tautomerism, purine nucleoside phosphorylase, isotope effects, arsenate esters, molecular modeling, kinetics, inhibition*

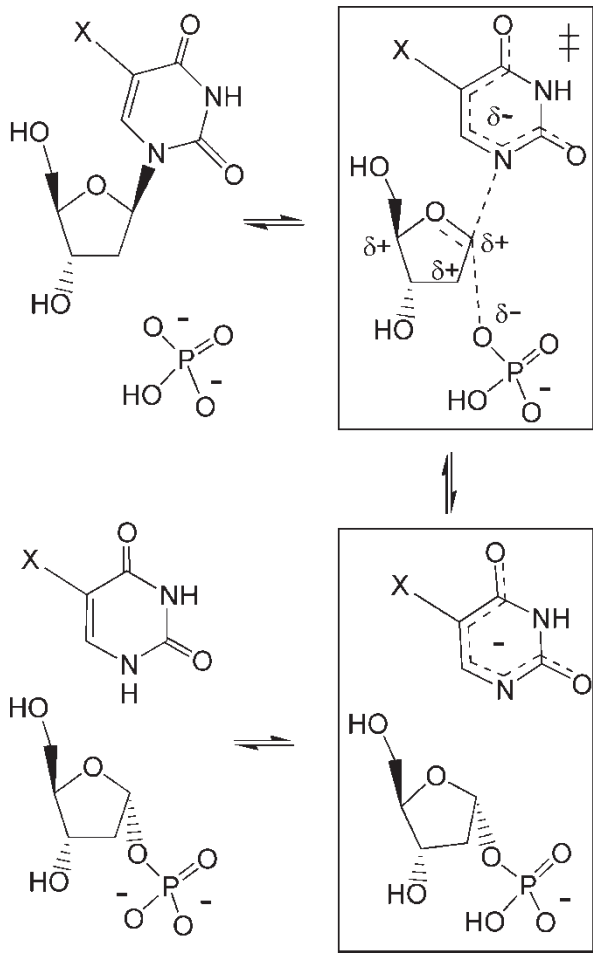
Introduction

Thymidine phosphorylase is a highly selective enzyme that in its normal intracellular environment catalyses the reactions depicted in Scheme 1 with X = CH_3. It catalyses analogous reactions with a variety of 5-substituted uracils and their β-D-2-deoxyribofuranosyl derivatives [1], but analogous purines are non-substrates. Corresponding 5-substituted uracil ribosides typically are very poor substrates due to large effects on k_{cat} with much smaller effects on K_m [2]. Ribose 1-phosphate is a competitive inhibitor of the normal 2-deoxy-substrate with similar binding affinity [3]. TP functions principally as a catabolic enzyme, generating 2-deoxyribose 1-phosphate (dR1P) during the process of maintaining adequate intra- and extra-cellular levels of thymine. After it is formed, dR1P mainly is hydrolysed to 2-deoxyribose, which then can be utilized as an oxygen-dependent energy source. Or, in hypoxic tumor tissue particularly, it can diffuse out of the cell and thereby function as an angiogenic factor (endothelial cell chemoattractant) [4, 5]. The widely used anticancer prodrug, 5-fluorouracil, is converted by this enzyme to 5-fluorodeoxyuridine, which, after phosphorylation at C5′-OH, is the form of the drug that inhibits thymidylate synthase [5].

A potent inhibitor of human TP (hTP), the N2′-protonated form of 5-chloro-6-[1-(2′-iminopyrrolidin-1′yl)methyl]uracil (TPI, see Figure 1.b), causes a modest but significant reduction in tumor growth rate when given every 12 hours for 22 days to nude mice

Correspondence: Dr. Philip Neil Edwards, School of Pharmacy & Pharmaceutical Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, UK. Tel: + 44(0)161 275 2403. Fax 44(0)161 275 2481. E-mail: Philip.N.Edwards@manchester.ac.uk

ISSN 1475-6366 print/ISSN 1475-6374 online © 2006 Informa UK Ltd
DOI: 10.1080/14756360600721075



Scheme 1. TP reaction pathways incorporating possible complications arising from use of arsenate as a substitute for phosphate.

with implanted experimental tumors [6]. In a separate study in which mice were dosed TPI by the intraperitoneal, intravenous or oral routes, drug clearance was so rapid that the duration of any pharmacodynamic effect was considered to be a potential problem. A more potent compound, or at least one with improved pharmacodynamics, is required if the therapeutic potential of TP inhibition in human disease is to be maximized. Alternatively, formation of a potent drug from an inactive or weakly potent prodrug, selectively within the target tissue, may improve therapeutic index and modify clearance and distribution [7].

Transition state (TS) analogues can prove to be potent enzyme inhibitors so information regarding the characteristics of the TP TS could be useful in inhibitor design. TP protein exists as a dimer but that fact seems to have no clear relevance to its function. Crystal structures, until recently, have revealed binding sites for phosphate/sulphate and thymine that are exposed to bulk water in an open cleft and are too far separated for reaction to occur. A large domain movement, closing the cleft and bringing the substrates together, was suggested as a necessary step towards an active complex [8]. A closed-cleft conformation of the suggested type was first observed in the X-ray-determined structure of pyrimidine nucleoside phosphorylase (PYNP) from *Bacillus stearothermophilus* [9]. This prokaryotic enzyme shares 40% sequence similarity and very significant structural similarity to human TP; the relative positions of phosphate and a uracil-like ligand, seen in the closed-cleft conformation, are analogous to those seen in other enzymes of this type. Consequently the mechanism is expected to involve an α-oxacarbenium-like transition state, or states. Ideally one would wish to have information pertinent to hTP, its substrates and their interactions at all critical points throughout the course of the reaction. Identification of the rate determining step(s) is an important first objective.

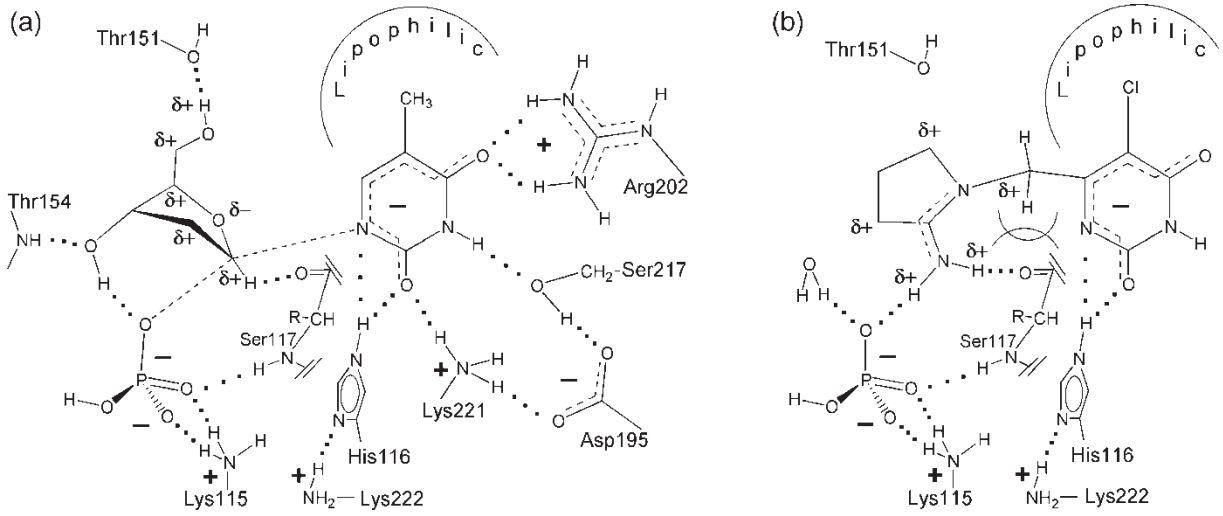


Figure 1. ***a***, Schematic drawing of hTP partial active-site containing a proposed TS-complex. ***b***, TP-bound TPI with HPO_4^{2-} and some binding partners.

Discussion

Protonation state of N1 in product 5-X-uracils can affect phosphorolysis rates of 5-X-deoxyuridines (5-X-dUrd)

The protonation states of TP-bound inorganic phosphate, dR1P and thymine/5-X-uracil (5XU) are unknown, but they are expected to vary with changes to the experimental pH which has been varied from ~4 to 9. In the following discussion the focus is on the closed form of the enzyme with both sets of substrates independently present at particular moments. In the synthesis direction, reaction of 5XU with dR1P is all but inconceivable if the 5XU N^1-hydrogen atom remains attached to N1. Therefore, for now, all 5XUs are assumed to be preferentially bound as their N^1-anions. The improbable intervention of an imino-hydroxy tautomer will be discussed later. With the N^1-anion assumption, microscopic reversibility requires the anion to be the initial product in the phosphorolysis direction. The possibility is now addressed that the extent of subsequent protonation of product N^1-anions, at N1 (5XU-anion ↔ 5XU), can influence rates in the phosphorolysis direction.

Santi and co-workers in 1980 reported on the phosphorolysis of diverse, 5-substituted deoxyuridines catalysed by TP from horse liver [1]. The derived K_m and relative V_{max} values were subjected to quantitative structure-reactivity (QSAR) analysis. No correlation could be detected between a variety of parameters and $1/K_m$, but $\log(V^{rel}_{max}/K_m)$ yielded "by far the most significant" correlation and it required the use of only a single parameter, the field-effect parameter, F, of the 5-substituents: $\log(V^{rel}_{max}/K_m) = 2.20(\pm 0.36)F - 0.3$ (this last value is incorrect, it should be +0.25). The authors expressed no surprise about the equation but it can now be seen to be a very surprising result which is made more apparent by re-writing the equation thus: $\log V^{rel}_{max} = \log K_m + 2.20(\pm 0.36)F + 0.25$. The implication of this equation is that the environment of the 5-substituent changes, between the Michaelis complex and the transition state, in a way that is similar to the change from the Michaelis complex - now known to be predominantly non-polar hydrocarbon - to free substrate in water. A much-expanded version of Scheme 1 incorporating cleft opening and closure, and variable substrate/product ionization states (see Appendix 1; Scheme A1.2) results in a very complex kinetic scheme that could allow of several explanations. The following re-analysis of the Santi data represents a likely explanation.

The nucleofugic ability of 5XUs in this system, and the pK_a values of N^1-H of bound uracils, are expected to be correlated with a combination of σ_I and σ_R-parameters for the 5-substituents [10,11]. The ratio of the coefficients is estimated as 1: 0.5 based on that being, approximately, the ratio relevant to the aqueous pK_as of N^1-H in 5-substituted cytosines; the same ratio is used here in the correlation of uracil N^1-H aqueous pK_a values (see Appendix 2). Unless otherwise indicated, published pK_a values are taken from reference works [12–14]. Cytosines were chosen in part because of their structural similarity to uracils, but also because they exist, in water, essentially exclusively as the N^1-H tautomer and its ionization is the only detectable contributor to observed acidic pK_as. Uracils in contrast ionize in water at both N^1-H and N^3-H with the ratio depending on the nature of the 5-substituent. However, the 5-substituents in TP-bound uracils and uridines are in a mainly non-polar region of the enzyme while the C^2=O and C^4=O carbonyl oxygens are highly solvated by hydrogen-bonding and cationic groups, so the quantitative effects are unlikely to correlate exactly with those determined in water. The only polar components of the 5-substituent binding site are the cationic head group of Arg202 (H-bonded to O^{C4}) and the π-face of Thr118 carbonyl group, the carbonyl oxygen atom of which is almost coplanar with the uracil ring, located between uracil C5 and C6, and the carbonyl group as a whole contacts groups larger than hydrogen at C5 and C6. A plot of $\log V^{rel}_{max}$ for the above compounds, plus that for the 5-nitro compound (estimated from published data [15]), against substituent $[\sigma_I + 0.5\sigma_R-]$ and, equivalently, predicted 5-X-uracil N^1-H aqueous pK_a values, is shown in Figure 2 (the curve is added to assist subsequent qualitative discussion of these data).

With some estimations and assumptions, the potentially triphasic relationship can be explained. The pK_as of TP-bound 5XUs can be estimated (*vide infra*) and those estimates clearly indicate that 5XUs with electron withdrawing 5-substituents are bound very predominantly as N1-anions at the pH of the study. The assumptions are that the chemical steps involving C1′-N breaking/making and 5XU protonation/deprotonation are faster than the closed-to-open protein conformational switching (cleft opening is essential for product release) and that protonation at

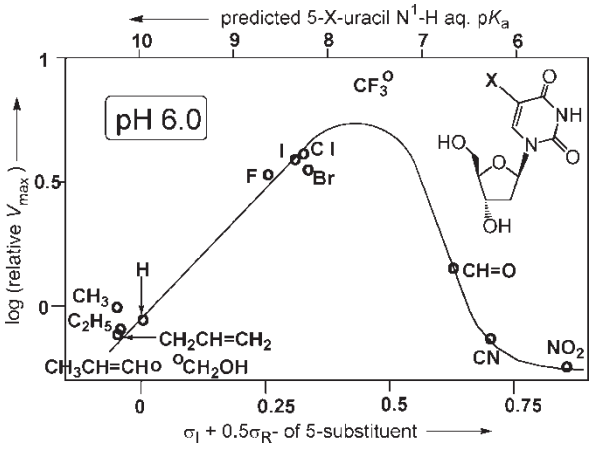


Figure 2. Phosphorolysis of -substituted deoxyuridines by horse TP.

N1 destabilizes the closed form of **[**5XU/dR1P/TP**]** (the bold square brackets are used to indicate multi-component complexes in closed conformations). Under those circumstances, the proportions of TP-bound starting materials and products reach near-equilibrium values during each catalytic event. From the left side of Figure 2, increasing electron withdrawal by the X-substituents in 5-X-deoxyuridines speeds the phosphorolytic step and slows its reversal. The hypothetical enzyme-bound near-equilibrium would increasingly favour products (i.e. the **[**dR1P/5XU-anion/5XU/TP**]** state) and conformational switching increasingly would allow discharge of products rather than starting materials, hence V_{max} increases. However, once the enzyme-bound equilibrium becomes dominated by products, here with X=CF_3, further increases in electron withdrawal would result, mainly, in reducing the proportion of product 5XU-anion that becomes protonated. Since 5XU N^1-anions are more tightly bound than close neutral analogues [16], uncharged (N1 protonated) 5XU release could become rate limiting, in which case V_{max} will decrease as electron withdrawal increases. On the far-right side, it is proposed that the fraction of uncharged 5XU has become so small that release of anionic-5XU has become dominant and V_{max} becomes independent of further increases in 5-substituent electron attracting power. A consequence of this analysis is that the experimental pH will affect the shape of the curve - directly through the extent of N1 protonation, and indirectly through possible changes to the charge status of the protein, of bound phosphate and of bound dR1P, since they must influence the pK_as of TP-bound 5XUs (and vice-versa).

With regard to the rates of substrate protonation/deprotonation, it is shown in the section on modeling that rapid transfer of protonic charge could occur along a chain of hydrogen-bonded waters linking aqueous buffers with bound phosphate/dR1P (see Appendix 6); computational modeling of such processes has recently been much improved [17]. Proton transfer to N1 of product 5XUs is more problematical. It could occur (probably contra-thermodynamically in *all* cases) as suggested by Mendieta et al. [18], from mono-protonated deoxyribose 1-phosphate (dROP(OH)O_2^-), the initial product of phosphorolysis, via His116. But modeling of hTP has failed to reveal a pathway for such a transfer that does not concurrently change the tautomeric form of His116, and such a change would seriously disrupt the local hydrogen-bonding network. Several alternative possibilities can be envisaged but none of them is compelling. Protonation at N1 subsequent to cleft opening is another possibility but this requires that cleft re-closure is faster than 5XU-anion release for all compound in Figure 2 apart from the nitro- and, perhaps, the cyano-compound. Almost certainly, protonation equilibrium is achieved for all substrates in less than the time associated with the major conformational changes these enzymes undergo during each catalytic cycle. For the commonly used substrate, 5-nitro-deoxyuridine (contains the best nucleofuge and is the least reversible), that time will be similar to the turnover time: $k_{cat} = \sim 20\,s^{-1}$ for *E. coli* enzyme and $\sim 3\,s^{-1}$ for hTP [6]. Another reason for questioning proton transfer from dR1P is the need to have that group mono-protonated for the reverse reaction†, if it were not protonated, the leaving group would be the improbable nucleofuge, PO_4^{3-}. This requirement does not deny the possibility that $dR1P^{2-}$ could be an acceptable occupant of that binding site, but the fact that thymidine synthesis from dR1P is only slightly dependent on pH in the range 6.5 – 8.5 (see Appendix 1) (suggesting that the pK_a of TP-bound dR-O-P(OH)O_2^- is >8) implies that the protons come, effectively, from aqueous buffers in experiments conducted at or below pH 8.

The hypothesis, that degree of 5-X-uracil anion protonation can be a contributor to the limiting rate, requires that the shape of the curve in Figure 2 should be pH sensitive: if experiments are conducted at higher pH, compounds on the left side should move down relative to the 5-nitro compound because the proportion of product 5XU existing in N1-protonated form is progressively reduced as the pH increases; the 5-nitro compound, aqueous pK_a at N^1–H = $\sim$5.5, is unaffected since it already leaves as the anion at pH 6.0 (prior hypothesis). Experiments conducted at pH 7.4 with intact human blood platelets produced results [19] which provide evidence for the predicted pH effects and analysis of these and other published data [15,20,21] lends general support to the above proposals (see Appendix 1).

On the evidence so far considered however, it is possible that 5XU-anion protonation is at O^{C2}, or even O^{C4}, generating imino-hydroxy tautomers, rather than at the thermodynamically favoured (in water) N1-position. That neither of these possibilities is supportable in this context is now demonstrated.

Hydroxy tautomers of 5-X-uracils are not the initial phosphorolysis products

A targeted molecular dynamics (MD) and quantum mechanics (QM) study by Mendieta et al., using the open *Escherichia coli.* structure as their starting point (the hTP structure had not then been published), led to a closed model of the enzyme and to the suggestion that phosphorolysis may initially produce the 2-enol form of thymine (C^2—OH) [15]. The activation energy they calculated for the phosphorolysis, 22.2 kcal·mol^{-1} using RHF/3-21G(d), is considerably too large, but the calculated activation energy of 72.5 kcal·mol^1 for the reverse reaction reveals that

very serious deficiencies exist in the QM modeling of the reaction.

Earlier work by Rick et al., also starting with the *E. coli* structure, used desolvation to assist domain movement and closure of the active site around phosphate and thymidine [22]. Their initial QM density functional (DFT) calculations, involving HPO_4^{2-} with O^{C2} or O^{C4} *protonated* deoxyuridine, generated two possible transition states. The O^{C2} protonated form led first to removal a proton from the deoxyribose 3′-OH to generate $H_2PO_4^-$ and a 3′-oxyanion. The O^{C4} protonated isomer led to a lower-energy TS that was used, with added highly-truncated, active-site residues, as the starting point for further calculations (truncation was essential for the large QM calculations).

Their initial structure, on the left of Figure 3, has one NH_4^+ in place of the guanidine group of Arg171 (hTP, R202) and another representing Lys190 (hTP, K195); both those groups, the imidazole ring of His85 (hTP, H116), and the phosphate, are in their expected aqueous, pH 7.4, protonation states. Energy minimization of the ensemble resulted in phosphate deprotonating imidazole and the imidazole deprotonating NH_4^+. Forcing the system to the first TS resulted in O^{C4} taking a proton off the other NH_4^+ to generate a 4-hydroxy group. As will later be demonstrated, even this much more realistic QM calculation predicts an unrealistic outcome.

A much more extreme example of unrealistic predictions from QM/MM calculations involving HPO_4^{2-} was reported by Hillier and coworkers: their calculations predicted that HPO_4^{2-} could remove a proton from C1′ of thymidine to generate a C1′-carbanion that would have no conventional resonance stabilization, and the process was highly favourable (76.6 kcal mol^{-1} with HF/6-31G⋆, or 131.9 kcal mol^{-1} with PM3) [23]. That outcome would be unprecedented in chemistry and the authors expressed their concern about QM calculations involving HPO_4^{2-}. The problems, widely appreciated, centre on inadequate solvation, particularly specific solvation that would allow charge transfer to disperse the high energy electrons over a larger number of atoms. Small, charged species in the gas phase have very high energies: converting trifluoroacetic acid and ammonia into trfluoroacetate and ammonium is unfavourable by 118.3 kcal mol^{-1} (experimental value) [24]. The above three QM examples demonstrate that any worthwhile future QM/MM studies involving TP/HPO_4^{2-} would have to use an extremely large QM region, high-level methodology and a massive computing resource. The authors of a recent QM(AM1)/MD study of PNP resorted to modeling phosphate as the mono-anion, $H_2PO_4^-$, presumably because of problems encountered with the di-anion [25]. It is therefore desirable to investigate the probable nature of the reaction mechanism and the estimation of relevant pK_a values of protein-bound species by other methods. Estimated pK_a values are potentially directly relevant but also they may allow the assessment of tautomer preferences in the bound state.

The aqueous pK_a values of the hydroxy tautomers of uracil can be estimated starting from the work of Poulter and Frederick, aimed at establishing neutral species tautomeric ratios [26]. From their results one can estimate the pK_as of the hydroxy tautomers if the pK_as of their respective NH-forms are known. Estimates of the required pK_as are reported in Appendix 2 so one calculates aqueous pK_as of 6.3 for the 4-hydroxy (p$K_T = 3.58$) and 4.9 for the

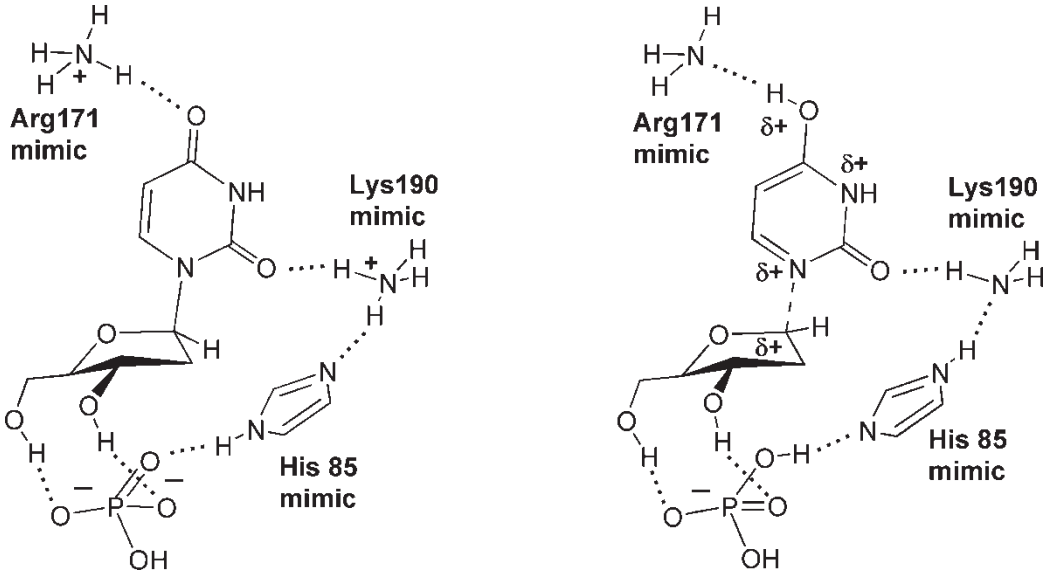


Figure 3. *E. coli* TP active site model by Rick *et al.* [22]: left, is initial state; right, is first TS.

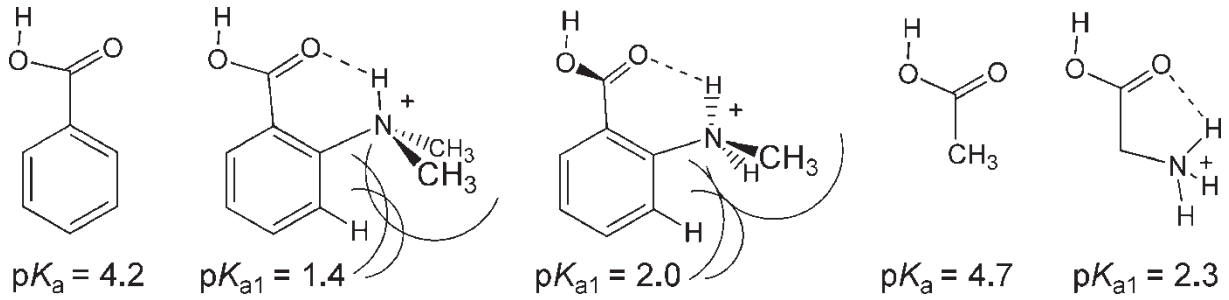


Figure 4. Effects of hydrogen-bonding contact of amine cations on acid pK_as.

2-hydroxy tautomer (p$K_T = 5.04$). In the active site, see Figure 1, they will be substantially more acidic due to the hydrogen-bonded contact with two cations and the NεH of His116. His116 Nδ: is H-bonded to Lys222NH_3^+ and partially shares the latter charge, but the lack of small-molecule experimental data relevant to the $^{\delta+}$NεH--O^{C2} contact precludes its incorporation into the present analysis. The acidity enhancement due to the two cations can be estimated, beginning with the comparisons shown in Figure 4 [12–14,27].

Note these effects, about $2\frac{1}{2}$ pK_a units, largely are due to hydrogen bonding and field effects, not inductive effects through bonds [10]: the interatomic distances and hydrogen bond angles are important and the latter are far from ideal in these examples. In the TP site, one cationic hydrogen-bond contact is directly onto the oxygen of the postulated OH group undergoing ionization, resulting in a much larger change than the above examples would predict [28], while the contact with the other, more-distant carbonyl oxygen that receives extra electron density in the anion, will produce a smaller effect. The geometry of the hydrogen bonds is equal or superior to the average of that in the model compounds in Figure 4. The acidity enhancement due to the two cations is judged to be ≥ 5 pK_a units if the surroundings were like bulk water. In the protein the surroundings are partly organic/non-polar and highly anisotropic, so estimating the influence of the environment change is difficult. Considering the local 'uracil' binding site, the equilibrium under consideration is that where a well delocalized anion involved in four hydrogen bonds, with tight contacts to two cations, and overall charge +1, has its charge increased to +2 by protonation of the anion. That is strongly disfavoured by a low permittivity environment. This is opposite to the conclusion that would be reached if an isolated uracil anion were to be protonated to a neutral uracil in a low permittivity environment (the proton 'escapes', in all cases, to the aqueous phase so no 'correction' arises from this source). Thus the protein environment is likely to enhance the effects relative to water. I leave the 'correction' estimate unchanged at ≥ 5 pK_a units.

The foregoing ΔpK_a estimate is based on aqueous acidities and in water acidic hydrogens are involved in strong, stabilizing hydrogen bonds that have to be broken to form the anion (zwitterion). With the histidine imidazole in the tautomeric form shown in Figure 1, and positioned as shown in the modeling section (and in the hTP crystal structure [29]), no hydrogen bond partner exists for either the 2- or 4-hydroxylic hydrogen so their acidities would be further enhanced: a very conservative estimate of 2 units is applied (hydrogen bond of highly acidic heterocycle-OH--$OH_2^{aq\cdot}$, assigned only 2.8 kcal mol^{-1}).

Therefore the **[**hydroxy-heterocycle/dR1P/TP**]** pK_a values are estimated to be < -1 for the 4-hydroxy and < -2 for the 2-hydroxy group. Protonation of deoxyuridine on O^{C2} or O^{C4}, to generate the starting structures used in the initial calculations of Rick et al., is more disadvantageous than for uracil-anion by more than nine orders of magnitude[‡], giving predicted pK_as of < -10. Protonating the TS on oxygen would form species with pK_as somewhere between < -10 and < -1. These estimates show that protonation on oxygen of 5XU-anions bound to TP, as modeled herein, would, for the equilibrated system, be at a vanishingly small level that would have no relevance to the enzyme mechanism. The foregoing analysis would change if either lysine or arginine residues were present as neutral species - but, in their TP-environments, that situation is so improbable it is not further discussed. Protonation of thymine-anion at pH 7.4 probably is partially rate limiting: this is understandable in light of the above discussion and an estimated hTP-bound thymine(N^1-H) p$K_a = \sim 6$ when bound along with mono-anionic dR1P.

Purine nucleoside phosphorylase facilitates binding of inosine or guanosine as a zwitterionic tautomer en route to the TS

Even very approximate estimation of pK_a values for protein-bound small molecules can lead to new ideas on the binding interactions applicable to the species of interest, and new mechanistic hypotheses when an enzyme is involved. In the case of TP the analysis is

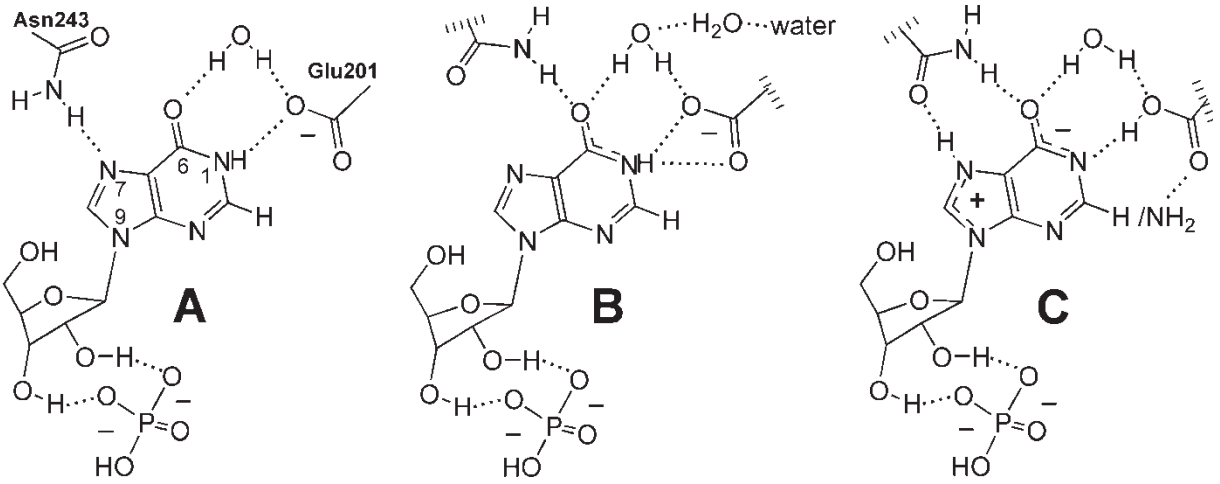


Figure 5. Enzyme-bound state of inosine (/guanosine) just prior to the TS. **A**. As proposed by Mao et al. [31]. **B**. As proposed by Fedorov et al. [32]. **C**. As proposed here.

able to eliminate some proposed mechanisms while for purine nucleoside phosphorylase (PNP) it reveals a mechanism that easily accommodates experimental facts where previous proposals are unconvincing.

PNP is a key enzyme in the purine salvage pathway that provides an alternative to *de novo* biosynthesis of purine nucleobases. At a biochemical level it closely resembles TP in catalysing the reversible phosphorolysis of purine ribonucleosides, inosine and guanosine, to the free bases and ribose 1-phosphate (or dR1P), but at a structural level it is unrelated. Mechanistically there appear to be significant differences and there may be differences between PNPs from different species [30], but those conclusions were mainly based on arsenolysis reactions, not phosphorolysis, and may be misleading. Reactions occur with solvent isotope effects as well as substrate kinetic isotope effects. Mutations Asn243Ala, or Glu201Ala, are extremely deleterious to catalysis so these residues were highlighted by Mao et al. in their mechanistic proposal, shown schematically in Figure 5(A) [31]. The hydrogen bond from Asn243 was suggested to be the main activating influence for breaking the C1′-N9 bond heterolytically, with the reaction proceeding via an oxacarbenium-like TS.

X-ray crystallographic results showing PNP complexed with substrates or a potent inhibitor, led Fedorov et al. to suggest that the water molecule hydrogen bonded to O^{C6}, through its linkage to an ordered water molecule in a narrow channel, that in turn is linked to bulk water, could allow a proton to find its way to N7 [32]. The proposal is presented schematically in Figure 5(B). The written description did not clarify the nature of the proton transfer process, but, since it is not possible for water to be a general acid catalyst to a group with $pK_a < \sim 16$, only small solvent isotope effects would be expected. Lewandowicz et al. have determined various KIEs for human PNP arsenolysis in Tris-HCl (pH 7.5) [30], including solvent D_2O isotope effects that were dependent on the arsenate concentration: 1.8 with 50 mM arsenate and 2.5 with 2.5 mM arsenate; similar results were found with the enzyme from *Plasmodium falciparum*. These results require that two or more molecules of arsenate are involved in the reaction conducted at 50 mM arsenate. Determining whether that is the case at 2.5 mM would require experiments at lower arsenate concentrations. The D_2O isotope effect corresponding to the involvement of a single arsenate molecule is therefore likely to be ≥ 2.5. The solvent isotope effects cannot be explained in the normal manner - through general acid catalysis involving N7 - because there is no enzymic acid close enough to N7 to perform the function.

The scheme shown in Figure 5(C) would enable a substantial solvent isotope effect through donation of a hydron from Glu201 to inosine N1 during the heterolysis of the C1′-N9 bond and because N9 is fully protonated, the scheme provides an excellent rationale for the heterolysis. There remains the vital question of the energetics relating to such a scheme: protonation of inosine in water occurs at N7 and this conjugate acid has $pK_a = \sim 1.2$ [12,13,33] (see structure **a**, Figure 6). Inosine N^1-H (Figure 6.b) has aqueous $pK_a = 8.55$. Gas phase calculations (BHLYP/6-31G(d)) show the zwitterionic tautomer of guanosine to be the least stable of four tautomers considered, (21.1 kcal mol^{-1} less stable than the N^3-H tautomer) [25], so the energetic penalty of forming a zwitterion (see Figure 6.**d**) and maintaining Glu201 non-ionized might seem excessive, but that is not so. In Figure 5(C), only the most relevant polar parts of the protein plus waters are shown. Hydrophobic side chains are predominant around the π-faces of the purine and some are in (near)contact with the carboxylate group of Glu201, thereby decreasing its

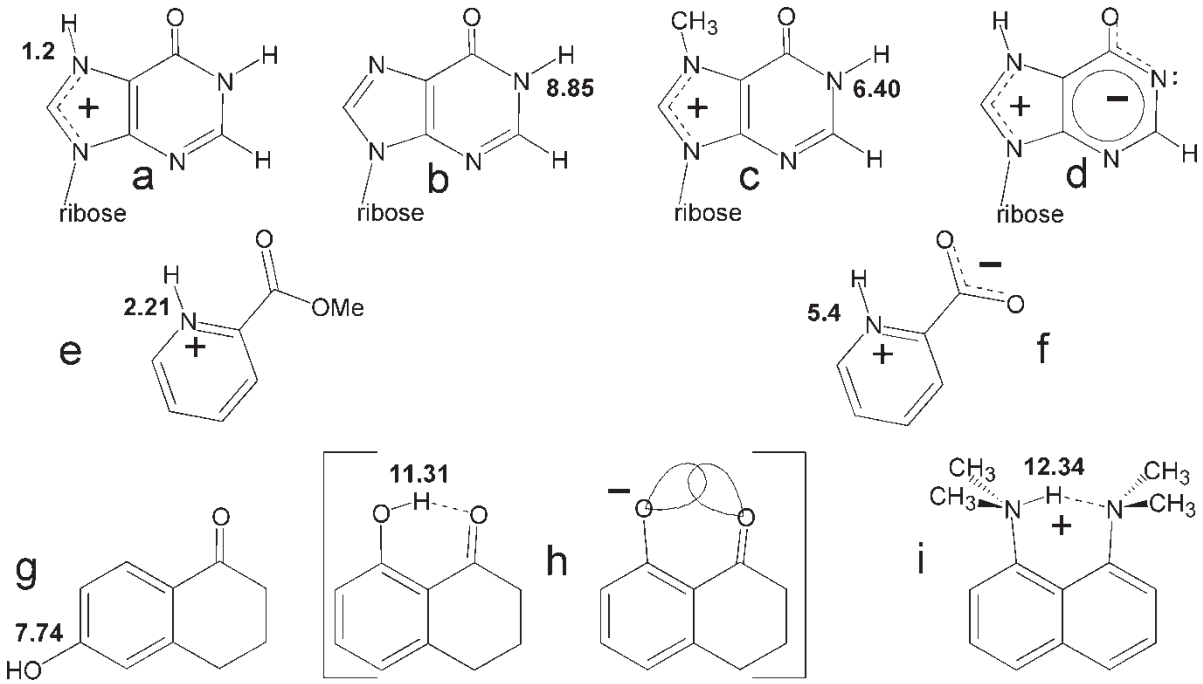


Figure 6. Proposed inosine PNP-binding tautomer, **d**, and pK_as used to assess the proposal.

acidity. Consider the pK_a values of N-7-methylinosine [34] (Figure 6.**c**) and Glu201 when the former compound is hypothetically bound to PNP in exactly the same position as inosine binds. In water, the cationic charge arising from N7-methylation causes N1 to become 2.45 units more acidic; in the PNP bound state the permittivity of the environment is very much lower than that of water so its effect will be greater. The positive charge will also increase the acidity of Glu201, but much less so than N1 because it is farther away and the charge in the N1-anion is delocalized to positions much closer to, and even into, the cationic five-membered ring (one resonance form puts negative charge on C8). So, if Glu201, hypothetically in water, has p$K_a = 4.4$ [35] and 6.**c** is 6.40, in the protein environment the rank acidities are likely to be inverted, or at least they will become very similar. Given the protein structure-enforced juxtaposition of the two groups, one will be a neutral hydrogen bond donor while the other is an anionic acceptor: the alternatives, both protonated (donors) or both anionic (acceptors) are strongly disfavoured.

Under the circumstance of near-equal estimated pK_as, the relative hydrogen bond energies of the two possibilities are relevant. These can very-roughly be estimated from free-energy or enthalpy related structure/hydrogen-bond-energy relationships reviewed by Abraham [36]. The enthalpy correlation is more appropriate here due to the restricted orientational and vibrational freedom of groups inside proteins. Hydrogen bond enthalpies of 1:1 complexes in CCl_4 at 298 K can be calculated from the relation $\Delta H°(\text{kcal}\cdot\text{mol}^{-1}) = 1.2\ E_A{\cdot}E_B$, where E_A is solute hydrogen-bond acidity and E_B the solute hydrogen-bond basicity. Because E_A and E_B are each related to interaction *energies* and the hydrogen bond energy is the *product* (!) of those terms, a mixture of two donor types and two acceptor types achieves its lowest energy when the strongest donor bonds to the strongest acceptor. Hydrogen bonding parameters for charged species are not available but those for neutrals allow one to make qualitative judgments here. The two cases are: (1), (sp^2, pyridine-like) N**:**---HO (acid/phenol/alcohol) and (2), (sp^2, electron-deficient) NH---**:**O=C. Within each donor or acceptor type a strong relationship with pK_a exists. At equal acidities, OH types are better donors than NH types and at equal basicities, carbonyl groups are better acceptors than imines (pyridines, purines, etc.). Regarding the hydrogen bond donors, the following E_A values are relevant: pyrrole, -1.34; methanol, -1.82; benzoic acid, -2.81. Regarding acceptors, the following E_B values are relevant: pyrimidine, 2.14; benzaldehyde, 1.11; ethyl acetate, 1.15. A very strong preference for type (1) bonding can be predicted as long as the geometry of that interaction is not far from ideal. The geometry in PNP is close to ideal [31].

Protonation of inosine N7 in water, structure 6.**a**, occurs only to the extent of one molecule in a million at pH 7.4. That low basicity reflects stabilization of the neutral form by delocalization of N9 π-electron density into the pyrimidine ring, C^6=O in particular, plus the electron withdrawing field effect generated by the pyrimidine ring, plus the slight effect of the sugar ring. All but the last effect is reduced if N^1-H ionizes. A partial estimate of the basicity change is gained from

the pK_as of N-protonated methyl picolinate, 6.**e**, compare 6.**a**, and zwitterionic picolinic acid, 6.**f**, compare 6.**d**. The difference, 3.2 units, underestimates the situation in inosine due to the absence of basicity-reducing resonance in the neutral form of 6.**e**, and the absence of π-resonance electron donation from carboxylate into the pyridine ring: chiefly field effects are important in the ΔpK_a of the pyridine derivatives. The moderate geometrical differences in the compounds: carboxylate oxygen closer to NH^+ in pyridines, requires a small opposing correction. Overall I estimate a ΔpK_a of ≥ 4 units.

In Figure 5, proposals (A) and (B) display strong polarity conflicts between Asn243 amide carbonyl oxygen and inosine O^{C6} in (A) or between amide carbonyl oxygen and N9 in (B). When such conflicts are unavoidable due to steric constraints, large enhancements in basicity are the consequence: protonation of one site can turn a repulsive interaction into an attractive one (see 6.**h**). An extreme case of this is found in the highly congested naphthalene derivative 6.**f**: its pK_a is elevated nearly seven orders of magnitude relative to the bis-monomethylamino compound. In the hydroxy-tetralones example, one uses the pK_a value of 6.**g** to estimate the pK_a of a hypothetical 6.**h** in which there are negligible steric, hydrogen bonding or lone-pair repulsive effects. Both 2- and 4-cyanophenols are largely free from those effects and provide a guide to the ketonic compounds: the ortho-cyano isomer is the more acidic by ~10-fold. Allowing for the different σ_I values of the groups, one calculates a polarity conflict effect of ~4 units for the 8-hydroxy-tetralone. In that compound the hydrogen bond geometry is far from the ideal unstrained, linear O-H--:O arrangement and the hydrogen bond acceptor is only a ketone. In PNP it is an amide carbonyl so the more negative charge on amide carbonyl oxygen produces greater repulsion in the non-protonated state and a stronger hydrogen bond when N7 is protonated. The hydrogen bond geometry in the enzyme is likely to be at least as good as that in the tetralone with either set of substrates bound. The constraints on the local geometry in PNP are far less than in either of the previous examples, but even in PNP a ΔpK_a of $>\sim 3$ units is expected on these grounds. Thus, starting with aqueous, protonated inosine with a pK_a of ~1.2, the corrections so far, $>\sim 7$ units, give an estimated [N^7-H^+, PNP-bound-zwitterion] pK_a of >8. So these crude estimation methods indicate that binding of inosine as a zwitterion may be its predominant binding mode and its exclusive mode en route to the TS. Similar lines of reasoning applied to guanosine, the other important substrate for PNP, result in the same conclusions and predictions.

The pteridine ring system of dihydrofolate is structurally close to the guanine ring system of guanosine and it binds to the enzyme dihydrofolate reductase (DHFR) with a carboxylate group in a position equivalent to the guanine(guanosine)/Glu201 pairing in PNP. In its bound state, N5 of dihydrofolate (equivalent to N7 in purines) needs to be protonated in order that C6 can accept a hydride ion from NADPH, generating tetrahydrofolate. The dihydrofolate N^5-H^+ aqueous pK_a is only 2.5 and, as for PNP, no acid group is close enough to N5 to achieve general acid catalysis [37]. It is possible that here too the reaction may involve (partially) binding of the zwitterionic-pteridine tautomer of dihydrofolate, shown in Figure 7 hydrogen-bonded to a non-ionized DHFR aspartyl residue.

hTP TS structure deduced from arsenate-mediated kinetic isotope effects

The nature of the TS in human TP has been addressed by measuring kinetic isotope effects (KIE) on the arsenolysis of seven isotopomers of thymidine and making the assumption that the arsenolysis TS is very similar to that of the natural phosphorolysis reaction [38]. A second paper by the same authors reported equilibrium binding isotope effects (BIE) for three of those isotopomers to hTP in the presence of 0.5 M ammonium sulphate (sulphate binds in the phosphate binding site) [39]. The KIE results are claimed to demonstrate that reaction of the natural substrates proceeds through a S_N2-like TS (C1′ to N1 and C1′ to O^P, partial bond lengths of 1.68 and 1.76 Å respectively) and analysis of the origins of the isotope effects is presented. The analysis depends on matching, as closely as possible, the kinetic isotope effects deduced from experimental data, with those derived from theoretical calculations of vibrational frequencies of the various isotopomers, in their separate unreacted states and at the TS. For this purpose, the structures of thymidine, phosphate-dianion, the TS and deoxyribose 1-phosphate were optimized in vacuo. The resulting conformations are inconsistent with X-ray defined and all modeled binding modes published heretofore, and taken together they would require extreme conformational changes to occur in the substrates and the protein during movement towards and away from the TS. No parts of the protein, or associated waters, were part of

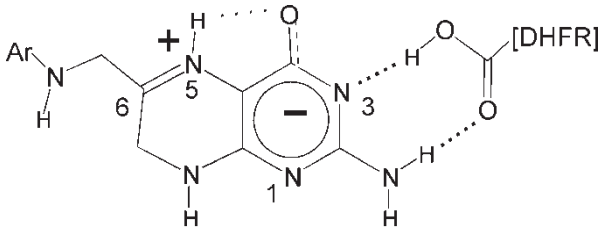


Figure 7. Dihydrofolate reductase may utilize a zwitterionic tautomer of dihydrofolate.

the calculations. Concern over such omissions is increased by recent QM/MD studies which indicate that unreacted substrates bound to enzymes have significantly different vibrational characteristics compared to the unbound state in water [25], and the changed geometry of the substrates and protein at the TS are, in general, most unlikely to leave those interactions unchanged. The S_N2-like transition state structure, claimed for the natural reaction with phosphate rather than arsenate, would correspond to the first case of an S_N2-type reaction for the family of N-ribosyl transferases.

In the above-discussed studies, commitment to catalysis was assessed by methods similar to those reviewed by Northrop [40]. The experiments were performed with arsenate in place of phosphate on the ground that the rate of hydrolysis of 2-deoxyribose 1-arsenate is fast relative to the timescale of relevant experiments, and therefore reverse commitment to catalysis can be ignored. The timescale for the KIE experiments was not reported so that assumption cannot be tested. Graphically-presented results from a forward commitment to catalysis experiment are consistent with the involvement of more than one arsenate molecule when the arsenate concentration is $>\sim 0.1$ mM, but the uncertainties in the positions of the datum points were not indicated so a firm conclusion is not possible (5 mM arsenate was present in the KIE experiments). Analysis of the data, assuming the involvement of a single arsenate, was numerically incorrect, but a revised forward commitment figure would *increase* the reported KIEs, one of which, that for $1'$-^{14}C (KIE = 1.14), is close to the maximum possible theoretical value. That near-maximum figure points unambiguously to a S_N2-type TS involving arsenate while the X-ray derived structure points to the involvement of an oxacarbenium-like TS when the natural substrate, phosphate, is present. As was mentioned above, PNP solvent isotope effects demonstrate arsenate concentration dependence that cannot presently be explained, but the known facts of arsenate chemistry make the observation unsurprising. There needs to be more widespread concern over the use of arsenate as a biochemical substitute for phosphate. The nature of the concerns and the rates of arsenate ester formation and hydrolysis are discussed in Appendix 3. One must therefore consider the possibility that the use of arsenate has fundamentally changed the TP-catalysed reaction away from that which occurs in Nature.

Both CH- --:X hydrogen bonding and sigma resonance can contribute to oxacarbenium ion stabilization

Complexes have been identified in high pressure mass spectrometric studies of water vapor interacting with various carbenium ions. Most of the ions investigated, including O-methyl-acetaldehyde cation (MeO^+=CH—CH_3), cf. Figure 8a (and the *E*-isomer), at ~ 300 K, do not form stable covalent bonds between carbon and water (R—OH_2^+). Rather, they were shown to prefer formation of loose 1:1 complexes, presumed to be stabilized by fluctuating C-$H^{\delta+}$- -:OH_2 hydrogen bonds. These initially unexpected results relate to the much greater entropic freedom of vibrationally loose hydrogen bonded complexes, involving many of the available hydrogens, compared to a relatively rigid, usually-single, covalent adduct. The consequential, experimentally-demonstrated, high entropy at 300 K, together with the enthalpy of the ion-neutral complex, must more than offset the loss of the potential C—O covalent bond energy [41].

Oxacarbenium ions are stabilized by dihedral (θ in Figure 8a) and atom-type dependent hyperconjugative electron release from adjacent bonds into $\pi^\star$, $\sigma^\star_{CO}$, and $\sigma^\star_{CR5}$ orbitals, while a nucleophile may interact with antibonding [$\pi^\star_{CO}$ plus $\psi\pi^\star_{C-CH_3}$] orbitals and/or $\sigma^\star_{CH}$ or $\sigma^\star_{OH}$ orbitals of an oxacarbenium ion, as shown schematically in Figure 8b. The geometry of O-protonated acetaldehyde (*Z*-isomer) has been investigated theoretically and two TSs involving that entity plus water were reported (RB3PW91/6-31+G** or RMP2/6-31+G** methodology). The two transition states, described as 'acute' and 'oblique', have O-protonated 1,1-dihydroxyethane as a local-minimum structure on one side of the TSs, but the nature of the local minima on the other sides of the TSs were not disclosed [42]. Closely related calculations in which water was replaced by imidazole led to a single TS. Some geometric details of the latter TS and the 'acute' water TS are shown in Figure 9 (Z-matrices were provided as Supplementary Material). Both calculated TSs involve very weak bonding between water or imidazole and C1, have C1 with its attached atoms close to their mean plane, and the length of the C1—O bond in each case implies substantial double bond character.

The rationale behind the O-protonated 1,1-dihydroxyethane → oxacarbenium-ion/water TS calculations, and the analogous calculations where protonated imidazole (adenine analogue) replaces the protonated hydroxy group, was to generate TSs relevant to enzyme catalysed reactions, in particular

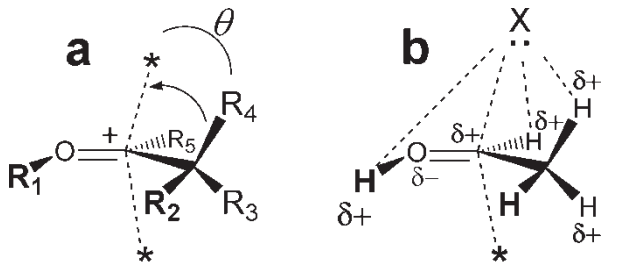


Figure 8. **a**. Sigma resonance interactions. **b**. Diverse nucleophile-oxacarbenium interactions.

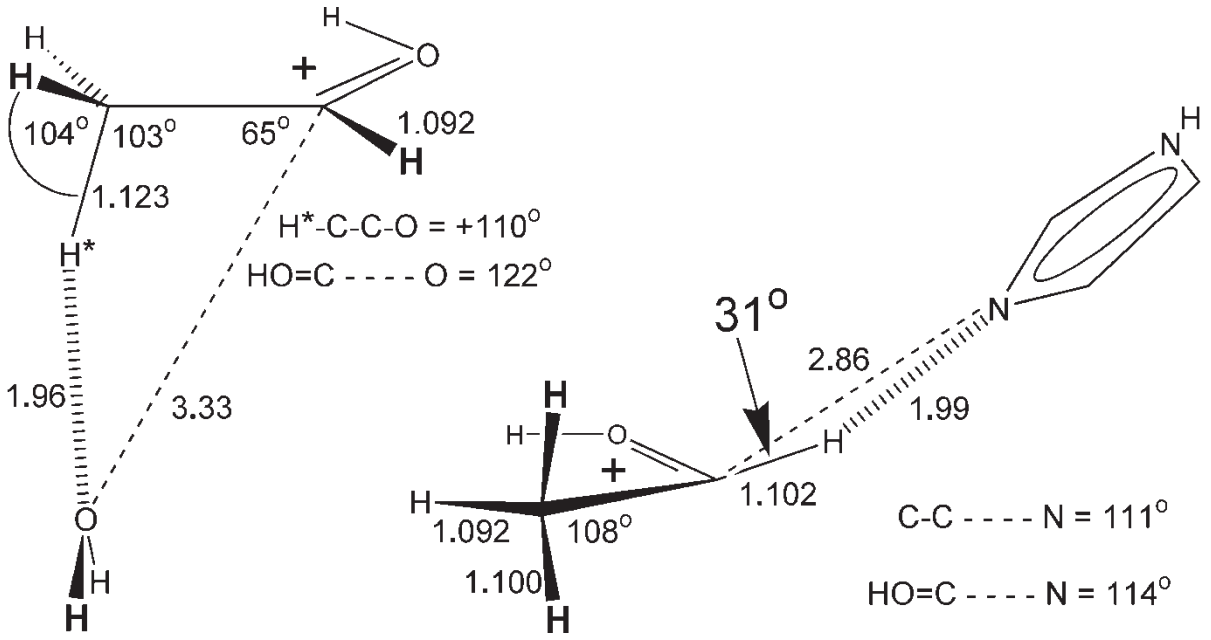


Figure 9. Calculated TS geometries involving O-protonated acetaldehyde with water or imidazole [42].

that for ricin toxin A-chain (RTA) depurination of ribosomal RNA [43]. Detailed analysis of the TS geometries allows one to deduce the nature of the unspecified species, beyond the TSs, as C-$H^{\delta+}$- -:X hydrogen bonded complexes. In the 'imidazole' TS structure, the hydrogen-bonded hydrogen has a lengthened bond to carbon and it is pulled upwards, out of the average plane around C1′, with that group forming a shallow pyramid with its apex pointing *away* from nitrogen! That distortion is expected to be trivial in the projected local-minimum hydrogen-bonded structure beyond the TS (C2, O, C1, H^{C1} and N are expected to be essentially coplanar). In the 'acute' water TS, the water oxygen is slightly beyond van der Waals' contact with the carbenium carbon but is extremely close(1.96 Å) to the hydrogen atom most involved in sigma resonance electron donation: the distance is consistent with a strong CH--O hydrogen bond. The reduced CCH angle and elongated CH bond support that description. The CH length is expected to increase further in the projected C-$H^{\delta+}$- -:OH_2 local-minimum hydrogen-bonded structure beyond the TS. A much more stable structure, the one that can be predicted to develop from the 'oblique' water TS alluded to previously, has water hydrogen bonded to the protonated carbonyl-oxygen of acetaldehyde: H_2O: - -HO^+=$CHCH_3$.

The partially-cationic C-$H^{\delta+}$- -:X hydrogen-bonding interactions seen in the above calculations are almost sure to have important counterparts in the stabilization of carbenium ions and related TSs in the active sites of enzymes, but generally with X being part of the protein or a protein-bound water. Two such possible interactions in the case of TP are described later.

All three TS calculations are consistent with there being no barrier to bond formation, C—O^+H_2 or C-imidazole-H^+, if the geometry of the environment (protein) were to constrain the species in such a way that the C—H hydrogen bonded geometries were inaccessible or energetically unfavourable. Calculations in which the imidazole in Figure 9 was replaced by N^1-H or N^7-H adenine (N-protonated initial covalent structure) and the C--N9 distances progressively changed from that of a full covalent bond, to a maximum of 3.4 Å, showed no TS in either case: no barrier to bond formation was found and no energy minimum corresponding to an ion-molecule complex was detected [42]. If this situation applied to TP, no carbenium ion *intermediate* could exist and the TS would be a double-contact ion-triplet.

Modeling of hTP plus substrates and intermediates/TSs

Until recently, MM [44], QM/MM [22,23,45] and targeted molecular dynamics plus QM [18] modeling of TP has relied for its starting point on the crystal structure of an open form of the (dimeric) enzyme from *E. coli.* [8,46]. It is clear that those studies did not result in structures similar to the hTP structure used as the starting point for this work. That structure, deduced by workers at AstraZeneca, contains the inhibitor TPI (5-chloro-6-[1-(2′-iminopyrrolidin-1′-yl)methyl]uracil hydrochloride, (see Figure 1b), fully enclosed within an enzymically-modified hTP protein (clipping of one external loop had been necessary to obtain crystals diffracting to the required extent, but the clipped product retained its functional activity) [29].

TPI in water exists largely as a zwitterion at near-neutral pH, mainly 'uracil' N^1-anion and iminopyrrolidine cation [7], and the binding mode to TP is consistent with this formulation despite the absence of phosphate from the active site and the consequential build-up of positive charge between the iminopyrrolidine-cation and surrounding groups (presumed-ionized functional groups within 9 Å of the imino-nitrogen belong to residues K115, K157, K221, K222, K235, D123, D233 and, additionally, within 9 Å of the pyrrolidine nitrogen is R202). Phosphate had not been included in the crystallisation solution [29]. Surprisingly, no diffusible anion was discernable in the active site. The deposited protein coordinates (PDB code: 1UOU) correspond to a structure that is likely to be close to the TS-stabilizing conformation, but the number of side chains in high-energy conformations and the unexpected distances and angles between hydrogen-bonding groups and waters, particularly in the active site, gives some cause for concern. Possibly a disordered anion, of variable nature, with partial occupancy exists in the active site.

All modeling, using SYBYL 6.8 (Tripos Inc. St. Louis, MO, US) was performed on one of the subunits of the symmetrical dimer after it had been modified as described in Appendix 5. To locate the binding mode of the phosphate within the closed TP structure, various hypothetical transition-state complexes (see, for example, Figure 1a) were postulated within the active site. Problems arising from the Merck Molecular Force Field (MMFF94s) [47] program's assignment of partial atomic charges, for the substrates particularly, are described in Appendix 5. Energy-refined models having alternative conformations of several side chains, most particularly those of active-site cations Lys222 and Lys115, in combination with variations of the C1′-O^P and C1′-N1 distances, covalent and 'ionic' C1′-N1 bonds, variable deoxyribose conformations, and the two neutral and one protonated forms of His116, were scrutinized. That led to a uniform and remarkably clear preference for very open complexes with geometries expected for a α-oxacarbenium intermediate or TS. In all the minimized structures, HPO_4^{2-} was tightly bound in a manner similar to that of sulphate in *E. coli* TP and phosphate in *B. stearothermophilus* PYNP [9]. A TS-structure with bond lengths C1′ to N1 = 1.68 Å and C1′ to O^P = 1.76 Å, as deduced from the arsenate KIE studies, was modeled in various ways and each was energy minimized. The general characteristics of the KIE-calculated TS structure were adhered to as far as possible but, without massive distortion of the binding site, grossly large changes were unavoidable. No satisfactory binding mode could be found if the above two dimensions were maintained: the thymine ring consistently was pulled well away from the optimum thymine-anion binding position by the strong attraction of the phosphate for a binding site that, in all cases, included at least one tight contact with Lys115, (see Figures 10 and 1a). That dominant attraction would result in much reduced stabilization of the partially anionic thymine subunit. Additionally, no arrangement of the active site waters allowed all their atoms to be involved in hydrogen bonding interactions; this was in stark contrast to all the other models.

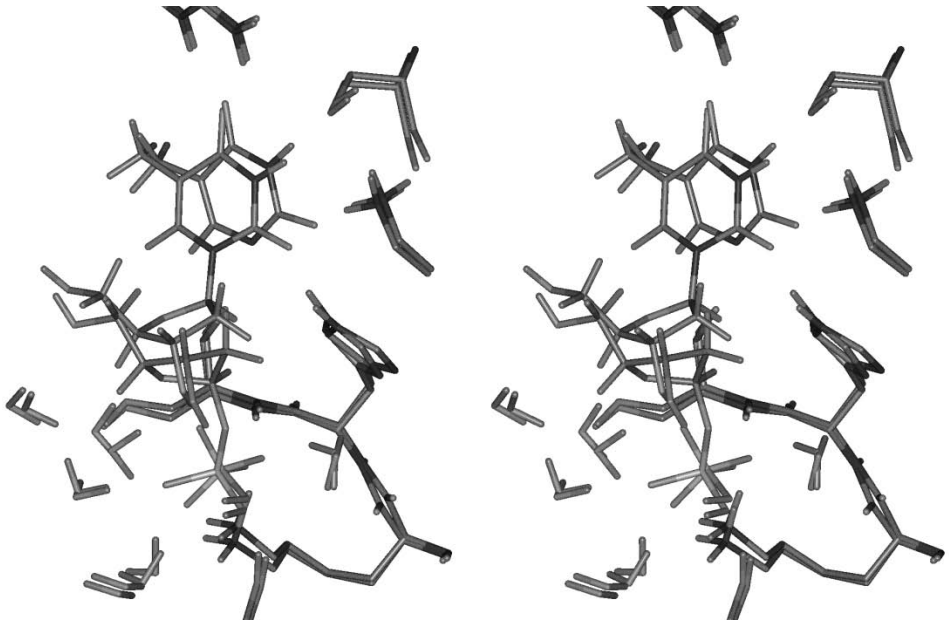


Figure 10. Stereo pair showing, overlaid, the two alternative TP substrates/partial-active-sites. Residues can be identified by reference to Figure 1a.

Since the positions of the thymine-anion and the phosphate group are constrained by strong interactions with parts of the protein that show no signs of substantial flexibility, it is inappropriate to assume that C1′ to N1 and C1′ to O(P) distances are free to assume values corresponding to those predicted by calculations in vacuo. Figure 10 shows an overlay of the [thymidine/P_i] and [thymine-anion/dR1P] Michaelis complexes: hypothetical TSs have geometries between the two extremes.¶

For the open complexes of the oxacarbenium type, all permutations tried resulted in well-defined orientations of the sugar and HPO_4^{2-} as parts of highly structured active sites. With the sugar in a 3′-endo conformation, the phosphate receives hydrogen bonds from Ser126OH, Ser144OH, Ser117NH, Lys115NH_3^+ and four waters: ten hydrogen bonds in total and the POH donates a hydrogen bond, for preference, to Thr154Oγ. If the sugar is in a 3′-exo conformation, (see Figure 1.a), the 3′-OH displaces a water from the phosphate solvation shell. Waters in these tightly-organized active sites are involved in a minimum of three hydrogen bonds, with no atom uninvolved. A chain of hydrogen bonded waters starts in contact with phosphate, contacts and passes Lys157-NH_3^+ and then Glu286-carboxylate, before connecting to bulk water and providing the means to maintain bound phosphate/dR1P ionization states in rapid equilibrium with solvent buffers [17]. Deoxyribose 3′-exo- and endo-forms had, after making allowance for one extra water in the 3′-endo-forms, sufficiently similar energies that a firm preference on that basis alone was not sensible. Currently the 3′-exo forms are preferred on the basis of: (i), the energy calculations; (ii), the more suitable orientation around C1′ in dR1P for uncomplicated motion to a α-oxacarbenium ion positioned similarly to that arrived at when starting from thymidine; (iii), the very stretched hydrogen bond between the 5′-O*H* of 3-exo(but not 3-endo)-deoxyribose 1-phosphate and **:**O(H)Thr151 - which changes to a tight hydrogen bond in all the calculated α-oxacarbenium geometries: in effect, this helps pull the sugar ring away from the phosphate (see Figures 1.a and 10).

The timescale for chemical reactions is $<10^{-13}$s [48], and TSs exist for almost zero time, so immediately surrounding groups have little opportunity to respond to changing local geometry as substrate structures convert to TS structures. Entropy considerations rule out concerted motions of more than a *very* small number of atoms during passage from ground to transition state. It is generally likely that the enzyme/substrates complexes are distorted substantially away from their time-averaged equilibrium geometries immediately prior to reaction occurring. Considerable progress is being made in incorporating quantum mechanical dynamical effects into the study of enzyme kinetics, but the requirement to average reaction coordinates over a myriad of pathways demands high-priority access to a supercomputing facility [49–51]. The nature of some reaction-promoting vibrations in PNP has been addressed by QM/MD/MM methods [25]. Non-dynamic molecular mechanics programs produce energy-lowering repositioning of a very large number of atoms during the protracted minimization processes. Widespread atomic repositioning occurred in some of the studies reported here. Attempts were made to mitigate this deficiency by performing some calculations with no, or very limited, relaxation of the protein as the substrates/intermediates were forced to change their characters and adopt new C1′--N1 distances. In some calculations, only thymine-anion and deoxyribose carbenium ion were allowed to move; another set was generated by allowing those two and phosphate to move. Other sets were generated by allowing those events to be followed by relaxation of the intermediates, together with local protein-residues and waters, for a limited number of iterations. Still other sets were generated by allowing the substrates to change their characters and adopt new C1′--N1 distances within a protein environment that had been partially relaxed in the presence of one of the previous TS mimics, then that structure was further partially relaxed. A more detailed presentation of this work is available as Appendix 5.

A possible α-oxacarbenium-like transition state with His116 in the cationic form and strongly hydrogen bonded to both phosphate dianion and the O^{C2} of thymine-anion, as proposed by Mendieta et al. [18], had the lowest MM-energy, but, since MM energies for differently-bonded structures are not directly comparable, this has minimal significance. More importantly, the positively-charged Cε of His116 (C2 of imidazolium cation) was forced into tight contact with the α-oxacarbenium carbon, and the thymine ring was pulled (N1--C1′ constrained length) away from its optimum position. So the thymine anion and the sugar cation would both be destabilized. This possible TS is rejected: its low energy, in part, is due to the problems of correctly quantifying electrostatic interactions involving HPO_4^{2-} in a very low permittivity environment (permittivity close to gas-phase: $\varepsilon = 1.3$ to 1.5 was required in order to reproduce the X-ray diffraction-derived hTP geometries, especially the hydrogen bond distances).

Because of the uncertainty in the electrostatic components of the energies, particularly those between substrate TS-components (HPO_4^{2-}/α-oxacarbenium / 5XU-anion), a protein perturbation approach to defining an intermediate carbenium-ion geometry, or possible TS, was investigated. The distance between the reacting O^P and thymine N1 was varied incrementally over the range 5.6 to 4.7 Å and the reacting structures (HPO_4^{2-}, oxacarbenium ion, thymine-N^1-anion), plus the waters and the (initially

entire) protein were minimized. Final energies were designated *(PL)* ^{n}. Then the phosphate-deoxyribose-thymine complex was extracted (waters not extracted) with all geometries retained: its energy, designated *(L)* ^{n}, and that of the residual substrates-depleted [protein plus waters], *(P)* ^{n}, were calculated at the fixed geometries. The energies associated with the partially-empty protein lack the van der Waals' and electrostatic energies that existed between the TS-complexes and the protein, so the *(P)* ^{n} energies are of questionable relevance. However they have increasing energies as the O^P to N1 separation goes below 5.4 Å (see Tables in Appendix 5).

More relevant is the energy *(PL)* ^{n} – *(L)* ^{n} which retains the van der Waals' and electrostatic energies previously not present but normalizes all the *(L)* ^{n} energies. This procedure tests the relative goodness of fit of hypothetical intermediate/TS-complexes (L), at particular geometries, for the protein (P) after mutual minimization (no constraints apart from one, sometimes two, 'TS' distance constraints). The results again show that O^P to N1 distances below ~5.4 Å are disfavoured. In a particular TP model, the relative energies at 5.6, 5.4, 5.1 and 4.7 Å O^P to N1 separations were 0, 0, 3 and 6 kcal mol^{-1} respectively (Appendix 5, Table A5.5), with the differences largely being due to stretched hydrogen and ionic bonds. Similar results apply to TP variants having alternative conformations of active-site side chains and deoxyribose ring-flip isomers. In the particular case just presented, C1′ to O^P varied from 2.7 to 2.3 Å, while C1′ to N1 varied from 2.9 to 2.4 Å, C1′ remaining roughly equidistant between O^P and N1. Calculations in which the O^P to N1 distance was held constant at 5.4 Å while the distances of C1′ to O^P and N1 were inverted (to 2.8 and 2.6 Å respectively), resulted in 2 kcal mol^{-1} unfavourable changes in *(PL)* and *(L)*, but *(PL) – (L)* changed by <0.2 kcal mol^{-1}, showing that the normalized system was insensitive to moderate changes in the O^P--C1′--N1 distances at that overall O^P-N1 distance. The long intermediate/TS partial-bond lengths investigated here, together with the fact that the MM-minimized species have partial-bond angles and dihedral angles that generally deviate only modestly from expected values, suggest that the neglected covalent partial-bond orders, being small, would only marginally influence the molecular mechanics defined geometries and *(PL) – (L)* energies for the range of distances examined. Deficiencies in the force field are of greater concern.

If an optimum O^P--N1 distance of 5.4 Å is assumed, it suggests that two TSs may be involved, and that raises the question of what structure the enzyme is best able to stabilize. If it were 'evolved' to stabilize an intermediate ion-triplet, it is reasonable to assume it would substantially stabilize both TSs, e.g. [O^P 2.2 Å C1′ 3.2 Å N1] and [O^P 3.2 Å C1′ 2.2 Å N1]. The model produces larger changes in *(PL) –(L)* energies as C1′ to N1 is reduced below 2.5 Å at constant O^P--N1 = 5.4 Å. The reason for this can be understood when bond and dihedral angles are considered along with distances: these considerations are discussed in the next section. It is worth noting that protein 'evolution' towards greater stabilization of the TS(s) is unlikely if protein conformational switching (closed to open) is rate limiting for the natural substrates at viable intracellular pH values.

Details of some models and discussion of some of their deficiencies

If TP reacts through a deoxyribose carbenium-like TS (similar to Figure 8.b), the making and breaking of the C—N bond will involve $\pi^\star$ orbitals [52] corresponding to the partial double bond between C1 and O4 and the slight double bond character between C2 and C1 due to hyperconjugation, so, discounting steric effects, one expects bond angles around an increasingly planar C1′ to be in the order O4′—C1′---N > C2′—C1′---N > > H1′—C1′---N. Combined electrostatic and steric effects generate the same expectations so it is unsurprising that the MMFF94s calculated 'TS' geometries shown in rows 7a-c, Table 1, fits those expectations in all cases bar the shortest where steric effects are most pronounced.

What was not predictable prior to this modeling is the movement of the thymine ring that brings about the largest part of the angle changes (see Figure 10 and Table 1). In the thymidine complex (dThd column), thymine O^{C2} is 344pm away from the cationic nitrogen of Lys221, so they are not significantly hydrogen bonded, but even for the shortest investigated 'TS' that distance shrinks to allow hydrogen bonding and it becomes tighter with increasing N1--C1′ distances. The decreasing distance very largely results from rotation of the thymine ring about a point close to O^{C4} (not movement of Nζ of K221) and with N1 moving in the initial C4′/C1′/N1 plane, well away from its original location; tightening of the hydrogen bonds to N^3-H and O^{C4} accompany this movement. Deoxyribose C1′ moves substantially away from its initial position (2.3 Å in dR1P), almost along the original N1-C1′ direction, with smaller movements of the other ribose atoms. One result of these motions is to dramatically change the angles between C1′--N1 and the thymine ring (rows 8a,b in Table 1). This stabilization of the incipient thymine anion is accompanied by tightening of the contact between H1′ and the carbonyl oxygen of Ser117 (row 6): this occurs despite the much too small positive charges assigned by the MMFF94s program to C1′ and H1′ and the much too large negative charge assigned to O4′ (see Appendix 5, Figure A5). It is likely that a strong and short stabilizing hydrogen bond exists between H1′ and this carbonyl oxygen at the true TS (or TSs), very much along the lines of that discussed above for the RTA beyond-the-TS complex involving

Table I. Geometric details (pm and degrees) of substrates/TP complexes and carbenium-ion TS-models. The TS-models resulted from full minimization of [N1--C1′]-distance-constrained TS components within a rigid [thymidine + HPO_4^{-2}]-optimized TP binding site followed by iteration-limited joint relaxation of the protein/ 'TS-components. See Figure 1 for atoms involved. Constrained distances have asterisks after the distance.

row	atoms involved	structure: TS-type :	dThd	TS200	**TS220**	**TS240**	TS260	**TS280**	TS300	dR1P
						distances				
1	N1...C1′		149	200⋆	220⋆	240⋆	260⋆	280⋆	300⋆	418
2	N1...O^P		510	512	520	525	542	545	565	542
3	C1′...O^P		364	313	**301**	**293**	284	**278**	271	141
4	O2...Nζ/ K221		**344**	306	306	298	296	292	289	**282**
5	O5′...Oγ/ T151		**276**	280	280	281	281	282	281	**307**
6	H1′...**O**=C/ S117		**294**	278	274	**269**	265	262	258	277
						angles				
7a	N1-C1′-O4′		111	102	102	103	104	107	110	111
7b	N1-C1′-C2′		114	104	101	99	96	92	89	56
7c	N1-C1′-H1′		106	77	74	72	71	70	69	62
8a	C1′-N1- C6		120	114	112	108	106	102	98	100
8b	C1′-N1- C2		119	130	131	133	135	136	136	135

imidazole (Figure 9). The best geometry for an intermediate carbenium ion is judged to be close to that shown for TS280 with, possibly, a thymidine-derived TS near to TS220 – TS240. At these distances the phosphate *di*anion/carbenium-ion attraction makes a very large contribution to lowering the energy of the system - that for TS280 is probably greater than any other single interaction. Hydrogen bonding between the 5′-OH and Oγ of Thr151 changes substantially only with formation of deoxyribose 1-phosphate (row 5): the 307pm distance shown in the Table fails to reflect the true strain here because the threonine hydroxyl has also moved well away from its former position in the joint attempt to maintain a hydrogen bond: the 5′-OH moves ~1.0 Å in total.

The other major deficiency in MMFF94s charge assignments is the zero charges assigned to H(2′*S*) and H(2′*R*) (β and α respectively) in any of the structures, so the program is blind to the possibility of hydrogen bonding to these atoms. In the dR1P complex, H(2′*S*) is in tight contact with $N1^{\delta-}$ of the thymine anion (253pm) and the C2′-H--N1 angle is 157° so, as the C1′--O^P bond stretches this contact may develop into a stabilizing hydrogen bond; progression towards thymidine leaves N1 always tight on H(2′*S*) but the angle C2′-H--N1 becomes progressively smaller, so stabilization of the decreasing partial positive charge on H(2′*S*) would increasingly depend on electrostatics alone rather than in combination with $n_N \rightarrow \sigma\star_{CH}$ charge transfer. The electrostatic potential, arising from the protein, in the region of H(2′*R*) is dominated by its proximity to the amide N-$H^{\delta+}$ of Gly145, ~260pm distance in all TS structures and only slightly longer in the thymidine (278pm) and dR1P (268pm) complexes: some inhibition of hyperconjugative electron donation from the C2′-H(2′*R*) bond to the carbenium carbon and its attached atoms is predicted - but overlap of the most relevant orbitals is non-optimal in the deoxyribose-3′-exo conformation so this is not an important consideration; that would not be true for a 3′-endo TS. A hydroxyl group in place of H(2′*R*) - therefore a ribose derivative - would provoke only minor changes to the binding geometry in any TS or starting state, but hydrogen bonding of the hydroxy-hydrogen would be less than ideal in all structures. For the thymidine case only O^{C2} would be a reasonable partner but the O-H---O angle would be poor and His116εN-*H* would be very close to the hydroxy-hydrogen. For the others the phosphate oxygen destined to bond with C1′ is the only reasonable partner, but even here the O-H--O angles are poor and the acceptor O^P in TS models is already involved in three strong interactions: with 3′OH, H_2O and C1′-cation. The last two limitations would not apply to ribose 1-phosphate (as judged by the deoxyribose 1-phosphate model) which has only one, relatively weak hydrogen bond from the 3′O-*H* to the C1′-O-P oxygen.

Waters around dR1P and the orientation of the HPO_4 group differ substantially from the situation where thymidine or thymidine-derived TS models are bound. The presence of a covalent bond between C1′ and O^P strictly limits conformational space for the new entity and the closeness of C1′ to covalently bound O^P forces a water molecule away from the newly bonded O^P-oxygen to a position where it bonds to O4′ instead. Other waters also change binding partners but with little accompanying translational motion. Re-ionizing the C1′-O^P bond and allowing its elongation did not spontaneously change the HPO_4^{2-}

or water binding modes back to that in the thymidine or thymidine-derived TS models. Molecular dynamics studies would be required to determine if the observed hysteresis has possible relevance to reality.

Conclusions

The extent of protonation of 5-substituted uracil anions is partially rate determining in the phosphorolysis of many 5-substituted deoxyuridines at neutral pH, but experiments avoiding this limitation may be possible for thymidine at pH ≤ 5. Pitfalls due to the use of arsenate as a phosphate mimic in biochemistry need to be more widely appreciated and conclusions arising from past uses of this device need to be reassessed. Extensive molecular modeling studies suggest that TP-catalysed phosphorolysis of thymidine may occur through the involvement of an intermediate carbenium-ion flanked by two TSs, but a single, very-open, carbenium-like TS is also possible. Comprehensive studies with a range of 5-X-deoxyuridines and 5-X-uracils over a wide pH range should greatly improve the mechanistic understanding of reactions catalysed by TP. Examples are presented of the very approximate estimation of pK_a values of enzyme-bound forms of substrates and possible intermediates. Such estimates can be useful in assessing ideas on the likely protonation states and tautomeric forms of those species, on the nature of possible ligand binding interactions, and on possible reaction mechanisms. Application of the ideas to other enzymes may reveal further examples of mechanistically-relevant, large deviations of enzyme-bound substrate pK_a values, and tautomeric ratios, from those observed in dilute aqueous solutions.

Notes

†The second pK_a of α-D-ribose 1-phosphate in water at 25°C is 6.28 [53], ~8-fold more acidic than dihydrogen-phosphate (pK_a = 7.2); α-D-2-deoxyribose 1-phosphate is expected to be about 3-fold more acidic (aqueous pK_a = ~6.7).

‡Poulter and Federick [26] report the *aqueous* pK_a of O^{C2} protonated uracil as −4.46, (not −2.98 as misquoted in reference 38) which then has to be adjusted for attachment of the electron withdrawing sugar ring and then 'site-effects' similar to those discussed in the main text for protonation of the anion.

¶Substrate/'TP' and 'TS'/'TP' structure coordinate files are available, in pdb or mol2 format, by email from the author.

Appendices are available as supplementary online material by accessing the article on the journal webpage.

References

[1] Nakayama C, Wataya Y, Meyer Jr, RB, Santi DV, Saneyoshi M, Ueda T. Thymidine phosphorylase. Substrate specificity for 5-substituted 2′-deoxyuridines. J Med Chem 1980;23:962–964.

[2] el Kouni MH, el Kouni MM, Naguib FN. Differences in activities and substrate specificity of human and murine pyrimidine nucleoside phosphorylases: Implications for chemotherapy with 5-fluoropyrimidines. Cancer Res 1993;53:3687–3693.

[3] Schwartz M. Thymidine phosphorylase from *Escherichia coli*. Properties and kinetics. Eur J Biochem 1971;21:191–198.

[4] Brown NS, Bicknell R. Thymidine phosphorylase, 2-deoxy-D-ribose and angiogenesis. Biochem J 1998;334:1–8.

[5] Cole C, Foster AJ, Freeman S, Jaffar M, Murray PE, Strafford IJ. The role of thymidine phosphorylase/PD-ECGF in cancer chemotherapy: A chemical perspective. Anticancer Drug Des 1999;14:383–392.

[6] Matsushita S, Nitanda T, Furukawa T, Sumizawa T, Tani A, Nishimoto K, Akiba S, Miyadera K, Fukushima M, Yamada Y, Yoshida H, Kanzaki T, Akiyama S. The effect of a thymidine phosphorylase inhibitor on angiogenesis and apoptosis in tumors. Cancer Res 1999;59:1911–1916.

[7] Reigan P, Edwards PN, Gbaj A, Cole C, Barry ST, Page KM, Douglas KT, Stratford IJ, Jaffar M, Bryce RA, Freeman S. Aminoimidazolylmethyluracil analogues as potent inhibitors of thymidine phosphorylase and their bioreductive nitroimidazolyl prodrugs. J Med Chem 2005;48:392–402.

[8] Pugmire MJ, Cook WJ, Jasanoff A, Walter MR, Ealick SE. Structural and theoretical studies suggest domain movement produces an active conformation of thymidine phosphorylase. J Mol Biol 1998;281:285–299.

[9] Pugmire MJ, Ealick SE. The crystal structure of pyrimidine nucleoside phosphorylase in a closed conformation. Structure 1998;6:1467–1479.

[10] Charton M. The Validity of the revised-*F* and revised-*R* electrical effect substituent parameters. J Org Chem 1984;49:1997–2001.

[11] Charton M. Electrical effect substituent constants for correlation analysis. Prog Phys Org Chem 1981;13:119–251.

[12] Perrin DD. Dissociation constants of organic bases in aqueous solution. London: Butterworths; 1965.

[13] Perrin DD. Dissociation constants of organic bases in aqueous solution, supplement. London: Butterworths; 1972.

[14] Serjeant EP, Dempsey B. Ionisation constants of organic acids in aqueous solution. Oxford: Pergamon Press; 1979.

[15] Wataya Y, Santi DV. Continuous spectrophotometric assay of thymidine phosphorylase using 5-nitro-2′-deoxyuridine as substrate. Anal Biochem 1981;112:96–98.

[16] Niedzwicki JG, el Kouni MH, Chu SH, Cha S. Structure-activity relationship of ligands of the pyrimidine nucleoside phosphorylases. Biochem Pharmacol 1983;32:399–415.

[17] Konig PH, Ghosh N, Hoffmann M, Eistner M, Tajkhorshid E, Frauenheim Th, Cui Q. Towards theoretical analysis of long-range proton transfer kinetics in biomolecular pumps. J Phys Chem A 2006;110:548–563, and references cited therein.

[18] Mendieta J, Martin-Santamaria S, Priego EM, Balzarini J, Camarasa MJ, Perez-Perez MJ, Gago F. Role of histidine-85 in the catalytic mechanism of thymidine phosphorylase as assessed by targeted molecular dynamics simulations and quantum mechanical calculations. Biochemistry 2004;43:405–414.

[19] Desgranges C, Razaka G, Rabaud M, Bricaud H, Balzarini J, De Clercq E. Phosphorolysis of (*E*)-5-(2-bromovinyl)-2′-deoxyuridine (BVDU) and other 5-substituted-2′-deoxyuridines by purified human thymidine phosphorylase and intact blood platelets. Biochem Pharmacol 1983;32:3583–3590.

[20] Gbaj, AM. Inhibition studies of thymidine phosphorylase. MPhil thesis. University of Manchester (UK). 2003.

[21] Wataya Y, Matsuda A, Santi DV. Interaction of thymidylate synthetase with 5-nitro-2′-deoxyuridylate. J Biol Chem 1980;255:5538–5544.

[22] Rick SW, Abashkin YG, Hilderbrandt RL, Burt SK. Computational studies of the domain movement and the catalytic mechanism of thymidine phosphorylase. Proteins: Struct, Funct, Genet 1999;37:242–252.

[23] Burton NA, Harrison MJ, Hart JC, Hillier IH, Sheppard DW. Prediction of the mechanisms of enzyme-catalysed reactions using hybrid quantum mechanical molecular mechanical methods. Faraday Discuss 1998;463–475.

[24] Taft RW, Bordwell FG. Structural and solvent effects evaluated from acidities measured in dimethyl sulfoxide and in the gas phase. Acc Chem Res 1988;21:463–469.

[25] Nunez S, Antoniou D, Schramm VL, Schwartz SD. Promoting vibrations in human purine nucleoside phosphorylase. A molecular dynamics and hybrid quantum mechanical/molecular mechanical study. J Am Chem Soc 2004; 126:15720–15729.

[26] Poulter CD, Frederick GD. Uracil and Its 4-Hydroxy-1(H) and 2-Hydroxy-3(H) Protomers-pK_as and Equilibrium-Constants. Tetrahedron Lett 1975;2171–2174.

[27] Tramer A. Tautomeric and protolytic properties of *o*-aminobenzoic acids. I. Ground electronic states. J Mol Struct 1969;4:313–325.

[28] Asaad N, Kirby AJ. Concurrent nucleophilic and general acid catalysis of the hydrolysis of a phosphate triester. J Chem Soc, Perkin Trans 2002;2:1708–1712.

[29] Norman RA, Barry ST, Bate M, Breed J, Colls JG, Ernill RJ, Luke RWA, Minshull CA, McAlister MSB, McCall EJ, McMiken HHJ, Paterson DS, Timms D, Tucker JA, Pauptit RA. Crystal structure of human thymidine phosphorylase in complex with a small molecule inhibitor. Structure 2004;12:75–84.

[30] Lewandowicz A, Schramm VL. Transition state analysis for human and Plasmodium falciparum purine nucleoside phosphorylases. Biochemistry 2004;43:1458–1468, and references cited therein.

[31] Mao C, Cook WJ, Zhou M, Federov AA, Almo SC, Ealick SE. Calf spleen purine nucleoside phosphorylase complexed with substrates and substrate analogues. Biochemistry 1998;37:7135–7146, and references cited therein.

[32] Fedorov A, Shi W, Kicska G, Fedorov E, Tyler PC, Furneaux RH, Hanson JC, Gainsford GJ, Larese JZ, Schramm VL, Almo SC. Transition state structure of purine nucleoside phosphorylase and principles of atomic motion in enzymatic catalysis. Biochemistry 2001;40:853–860.

[33] Dawson RMC, Elliot DC, Elliot WH, Jones KM. Data for biochemical research. 3rd ed. Oxford UK: Clarendon Press; 1986.

[34] Michelson AM, Pochon F. Polynucleotide analogues. Methylation of polynucleotides. Biochim Biophys Acta 1966; 114:469–480.

[35] Nozaki Y, Tanford C. Intrinsic dissociation constants of aspartyl and glutamyl carboxyl groups. J Biol Chem; 1967;242:4731–4735.

[36] Abraham MH. Hydrogen-bond descriptors for solute molecules. In: Wipff G, editor. Computational approaches in supramolecular chemistry. Amsterdam: Kluwer Academic Publishers; 1994. p 63–78.

[37] Deng H, Callender R. Structure of dihydrofolate when bound to dihydrofolate reductase. J Am Chem Soc 1998;120:7730–7737.

[38] Birck MR, Schramm VL. Nucleophilic participation in the transition state for human thymidine phosphorylase. J Am Chem Soc 2004;126:2447–2453.

[39] Birck MR, Schramm VL. Binding causes the remote [5′-H-3] thymidine kinetic isotope effect in human thymidine phosphorylase. J Am Chem Soc 2004;126:6882–6883.

[40] Northrop DB. The expression of isotope effects on enzyme-catalyzed reactions. Annu Rev Biochem 1981;5:103–131.

[41] Meot-Ner (Mautner) M, Ross MM, Campanat JE. Stable hydrogen-bonded isomers of covalent ions. Association of carbonium ions with n-donors. J Am Chem Soc 1985;107:4839–4845.

[42] Chen XY, Berti PJ, Schramm VL. Ricin A-chain: Kinetic isotope effects and transition state structure with stem-loop RNA. J Am Chem Soc 2000;122:1609–1617.

[43] Chen XY, Berti PJ, Schramm VL. Transition-state analysis for depurination of DNA by ricin A-chain. J Am Chem Soc 2000;122:6527–6534.

[44] Cole C, Marks DS, Jaffar M, Stratford IJ, Douglas KT, Freeman S. A similarity model for the human angiogenic factor, thymidine phosphorylase/platelet derived-endothelial cell growth factor. Anticancer Drug Des 1999;14:411–420.

[45] Sheppard DW, Burton NA, Hillier IH. Ab initio hybrid quantum mechanical/molecular mechanical studies of the mechanisms of the enzymes protein kinase and thymidine phosphorylase. THEOCHEM 2000;506:35–44.

[46] Walter MR, Cook WJ, Cole LB, Short SA, Koszalka GW, Krenitsky TA, Ealick SE. Three-dimensional structure of thymidine phosphorylase from *Escherichia coli* at 2.8.ANG. resolution. J Biol Chem 1990;265:14016–14022.

[47] Halgren TA, MMFF VI. MMFF94s option for energy minimization studies. J Comput Chem 1999;20:720–729.

[48] Dewar MJS, Jie C. Mechanisms of pericyclic reactions: The role of quantitative theory in the study of reaction mechanisms. Acc Chem Res 1992;22:537–543.

[49] Alhambra C, Corchado J, Sanchez ML, Garcia-Viloca M, Gao J, Truhlar DG. Canonical Variational Theory for Enzyme Kinetics with the Protein Mean Force and Multidimensional Quantum Mechanical Tunneling Dynamics. Theory and Application to Liver Alcohol Dehydrogenase. J Phys Chem B 2001;105:11326–11340.

[50] Garcia-Viloca M, Truhlar DG, Gao J. Canonical Variational Theory for Enzyme Kinetics with the Protein Mean Force and Multidimensional Quantum Mechanical Tunneling Dynamics. Theory and Application to Liver Alcohol Dehydrogenase. Biochemistry; 2003;42:13558–13575.

[51] Marti S, Moliner V, Tunon I, Williams IH. QM/MM calculations of kinetic isotope effects in the chorismate mutase active site. Org Biomol Chem 2003;1:483–487.

[52] Burgi HB, Dunitz JD, Shefter E. Geometrical Reaction Coordinates .2. Nucleophilic Addition to A Carbonyl Group. J Am Chem Soc 1973;95:5065–5067.

[53] Schneider IC, Rahmy PJ, Fink-Winter RJ, Reilly PJ. High-performance anion-exchange chromatography of sugar and glycerol phosphates on quaternary ammonium resins. Carbohyd Res 1999;322:128–134.

Appendices:

Appendix 1

Protonation states of N1 in product 5-X-uracils can affect phosphorolysis rates of 5-X-deoxyuridines: additional analysis of literature data. The hypothesis that degree of 5-X-uracil anion protonation can be a contributor to the limiting rate requires that the shape of the curve in Figure 2 (the curve is only a visual guide through the points) should be pH sensitive: if experiments are conducted at higher pH, compounds on the left side should move down, relative to the 5-nitro compound, which already at pH 6.0 has its phosphorolysis product, 5-nitrouracil, largely (~76%) in the form of the anion when it is free from the enzyme. That expectation is supported by a report from 1983 of experiments using a similar but slightly larger set of deoxyuridines to those of Santi et al., but with human platelet TP as the catalyst [19]. Despite the fact that these experiments were done at fixed, non-saturating concentrations of substrates (100μM), the relative initial rates of phosphorolysis of the 5-X-deoxyuridines, using purified enzyme at pH 5.7, can be interpreted in the same general way as the Santi data that had been generated using horse enzyme at pH 6.0. A selection of the pH 5.7 results, normalised to the nitro-compound, is shown in column *2* of Table A1.1. A more detailed analysis needs the inclusion of K_m values and, in most cases, the use of assumed values. But that more complex and very uncertain analysis is not essential: the authors reported half-lives for intact platelet catalysed phosphorolysis of substrates at pH 7.4(buffer pH) relative to half lives using purified enzyme. Those ratios, with appropriate assumptions, allow a rough estimation of relative initial rates in platelets at pH ~7.4; those estimates are shown in column *6* of Table A1.1. Columns *5* and *8* of Table A1.1 show relative initial rates corrected for %-ionization of substrates at N^3-H. The corrections assume that the N^3-anions do not bind to TP and that the concentrations of non-ionized 5-X-deoxyuridines are much lower than their respective K_m values; this cannot be quantitatively true but the errors introduced are unlikely to exceed a factor of two and generally the errors are likely to be much less since the phosphorolysis of substrates necessarily keeps their intracellular concentration lower than the extracellular concentration. The N^3-H pK_as of 5-X-deoxyuridines are estimated to be ~0.3 to 0.4 units lower than corresponding N^3-H in 5-XUs (see Table A2.2 for pK_as of 5-XUs).. Interpretation of such data is seldom straightforward as the change in pH may change substrates' ionization states and K_m values, and variable membrane penetration rates could be all-important. The first of these concerns is probably minor in this case because only the 5-nitro-dUrd and 5-cyano-dUrd substantially change their degree of ionization in the pH region studied. Changes brought about by the variable ionization state of the protein are expected to influence all substrates to a similar extent and those changes are in any case likely to be small: for example, thymidine K_m decreases only ~2.9-fold when pH is changed from 6.0 to 7.4 (a 25-fold change in acidity) [15]. Over the same pH range, the K_m for 5-nitro-dUrd changes by threefold in the opposite sense [15], but that is due to the increased proportion of N^3-anion at the higher pH: the N^3-anion should be very weak-binding since its ionization turns an attractive hydrogen bond, N^3H--:$OH(CH_2)$-Ser217, into a repulsive lone pair clash; allowing the serine hydroxyl to become a donor to the N3-lone pair does not solve the problem because the serine-$CH_2O\mathit{H}$ is involved in a strong hydrogen bond to the carboxylate anion of Asp195, and turning the OH around generates at least as bad a repulsive interaction between the Asp195 carboxylate anion and the Ser217 oxygen lone pairs. Allowing for the change in degree of ionization, the nitro compound changes K_m by ~2.3-fold in the same direction as the thymidine change of 2.9-fold.

That differential membrane penetration rates are unlikely to be important and can be inferred from: a) the presence in mammalian cell membranes of nucleoside and nucleobase carrier proteins that possess broad specificities [54;55]; b), the extremely high surface-area to volume ratio of platelets; c), the long timescale of the platelet experiments - reaction half lives ranged from ~1 to 12 hours; d) experimentally, the sum of the extracellular nucleoside and base concentrations was nearly constant and equal to the initial nucleoside concentration, and independent of the extent of reaction.

The Wataya and Santi study that showed variation of 5-NO_2dUrd and dThd K_ms with pH, mentioned above, arose from their discovery that 5-NO_2dUrd is a conveniently monitored substrate for TP. Some of the data pertinent to its use, comparative accompanying data for thymidine (dThd), and new additional analysis of their data (in italics) are shown in Table A1.2. The slopes of the V^{rel}_{max} / $[H^+]$ relationships (see Table footnote for slope definition) are consistent with product 5-nitrouracil dissociating from the enzyme as the N1-anion at all three pH values coupled with pH-independent dissociation rates (there is no compelling reason to expect invariance of dissociation rates with pH and a combination of effects could result in the same findings), while dThd is displaying V^{rel}_{max} rates consistent with the involvement of one or more groups with pK_as in the pH region studied. The two-fold V^{rel}_{max} ratio of the two substrates at the highest pH differs substantially from that which can be deduced from data in Table A1.1 using human platelet enzyme, ~6-fold (K_m values are similar [20]), and the ratio in Table A1.3 using the

Table A1.1. Relative hTP, pH 5.7, phosphorolysis initial rate data from C. Desgranges, et al. [19] and the estimation of relative initial rates applicable to phosphorolysis using intact human platelets at pH 7.4.

	pH ~5.7[a]				**pH ~7.4**[b]		
1	*2*	*3*	*4*	*5*	*6*	*7*	8
5-X-dUrd 5-subst.	V_0^{rel} pH 5.7	N^3-HpK_a [c]	%N3-anion at pH 5.7	V_0^{rel} [logV_0^{rel}] pH 5.7 corrected[d] for ionization at N^3-H	approx. V_0^{rel} pH ~ 7.4	%N3-anion at pH 7.4	V_0^{rel} [logV_0^{rel}] pH 7.4 corrected[d] for ionization at N^3-H
iodo	2.80	8.0	0.5	2.42 [0.38]	0.96	20	0.13 [−0.89]
bromo	2.80	7.8	0.8	2.43 [0.39]	1.50	28	0.23 [−0.64]
chloro	2.80	7.8	0.8	2.43 [0.39]	1.63	28	0.25 [−0.60]
CF_3	2.40	7.7	1.0	2.09 [0.32]	2.36	33	0.39 [−0.41]
formyl	2.16	7.6	1.2	1.89 [0.28]	0.88	39	0.16 [−0.80]
cyano	1.92	7.0	4.8	1.74 [0.24]	2.36	72	0.92 [−0.04]
fluoro	1.60	7.8	0.8	1.39 [0.14]	0.28	28	0.043 [−1.37]
methyl	1.04	9.6	0.01	0.90 [−0.05]	0.17	0.6	0.019 [−1.72]
nitro	**1.00**	6.5	14	**1.00 [0.00]**	**1.00**	89	**1.00 [0.00]**
ethyl	0.80	9.6	0.01	0.69 [−0.16]	0.081	0.6	0.0089 [−2.05]
hydrogen	0.76	9.5	0.02	0.66 [−0.18]	0.063	0.8	0.0070 [−2.15]
$HOCH_2$	0.48	9.1	0.04	0.41 [−0.39]	0.10	2.0	0.011 [−1.96]
amino	0.20	9.9	0.01	0.17 [−0.77]	0.018	0.3	0.0020 [−2.70]

[a] Purified enzyme. 5-X-deoxyuridines (0.1 mM), purified human thymidine phosphorylase(10 U/mg of protein), 10 mM phosphate, 37°C (K_m for thymidine = 0.14 mM, $V_{max} = 55\,s^{-1}$).
[b] Intact platelets. Incubation mixtures, at 37°C, contained 1 mM EDTA, 0.15 M NaCl, 1 mM sodium phosphate, 10 mM Tris-HCl, 0.1 mM 5-X-deoxyuridine and $3{\cdot}10^8$ platelets/mL.
[c] Estimated N^3-H pK_as of 5-X-deoxyuridines are ~0.3 to 0.4 units lower than corresponding N^3-H in 5XUs.
[d] Corrected for ionization at N^3-H by assuming non-ionized 5-X-deoxyuridine concentrations $\ll K_m$.

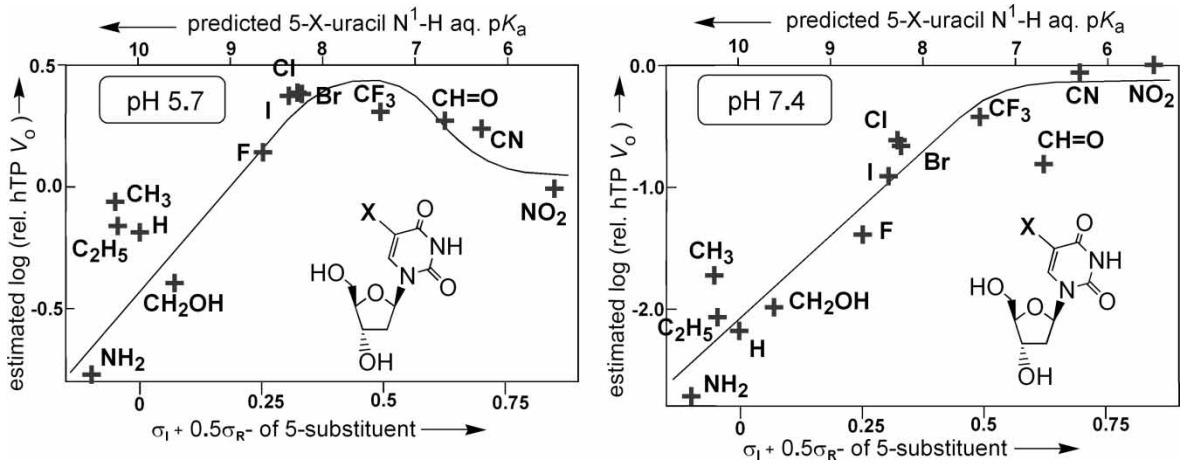


Figure A1.1. Graphical presentation of the ionization-corrected relative logV_0 values in Table A1.1 against $\sigma_I +$ of 5-substituents $0.5\sigma_R-$ and, equivalently, estimated N^1H pK_a values of 5XUs.

enzyme from *E. coli* where the ratio is ~90-fold. This example of marked species differences indicates the need for caution in generalizing particular observations. Differences are also likely in the kinetics of protein conformational change between enzymes from different species, especially bacterial versus mammalian; pH/rate profiles for TPs from a variety of species have been published, but some used very crude enzyme preparations and none of the studies examined variations of substrate K_m with pH [3;57;58]. Determining the variation of K_m with pH is essential for a better understanding of synthesis reaction rates.

The enzyme-bound thymine-N^1H pK_a is estimated to be ~6, when bound along with dR1P mono-anion, while the aqueous value is ~10.2. Enzyme-bound uracil N^1-pK_as in uracil DNA glycosylase have been determined experimentally by J.T. Stivers and coworkers [56]. Depending on the nature of other bound entities, the uracil pK_as were reduced to 7.5 (no abasic DNA analog bound), 6.4 (a tetrahydrofuran analogue of an abasic DNA fragment bound) and ≤ 4.5 (a protonated pyrrolidine analogue of an abasic DNA fragment bound). In the first two cases, the uracil anion is stabilized by hydrogen bonds from backbone amide NHs, bound water and one non-protonated imidazole(His187) which H-bonds to uracil-O^{C2}; significantly, no cationic residues contact the uracil anion. The non-protonated imidazole must make a large contribution to the stabilization of the anion and a similar stabilization of 5-XU-anions bound to TP likely arises from their contact with the functionally important residue His116.

Finally on this topic, results obtained recently in this University by A. M. Gbaj are relevant [20]. Table A1.3 records K_m and V_{max} values for the phosphorolysis of four 5-substituted deoxyuridines catalysed by *E. coli* TP at pH 7.4. Compared to methyl or hydrogen, the

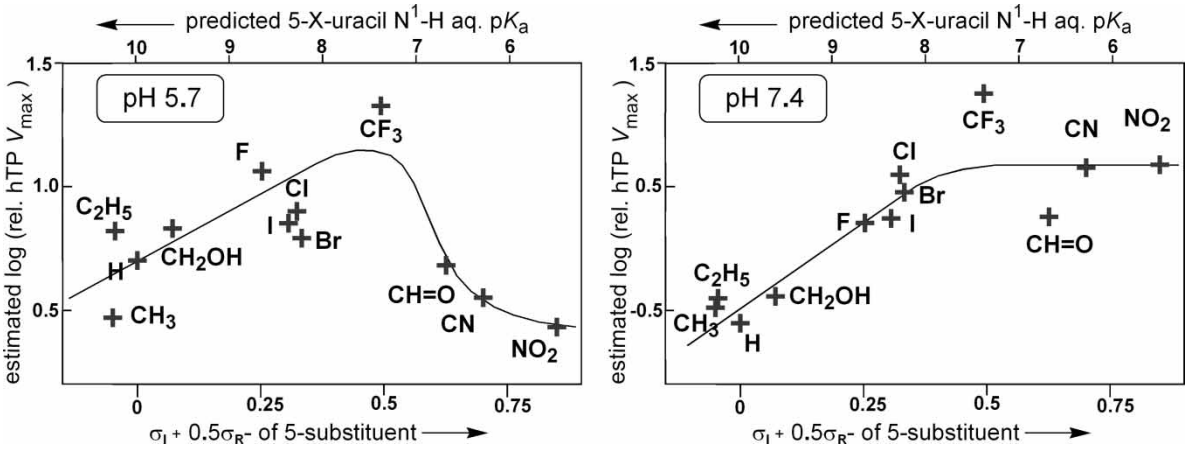


Figure A1.2. Graphical presentation of ionization-corrected relative logV_{max} values derived from the logV_0 values in Table A1.1; the analysis uses the following assumptions: (i) for these compounds, K_m values for human TP at pH 5.7 / 37°C and for horse TP at pH 6.0 / 25°C are identical; (ii) the pH related changes in K_m values of the non-ionized forms of 5-X-deoxyuridines for horse-TP and hTP change in parallel and the K_m value for the nitro-compound at pH 7.4 is 100 μM [15].

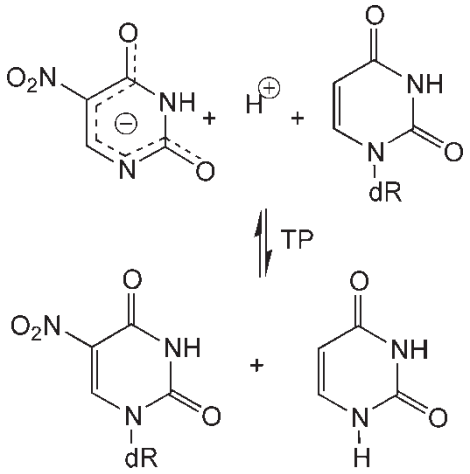
Table A1.2. Kinetic and derived data from horse liver TP phosphorolysis of NO_2dUrd and dThd.[a]

		pH 6.0	pH 7.0	pH 7.4
NO_2dUrd	% neutral	76	24	11
	K_m (mM)	0.10	0.17	0.30
	[ES] / [E_t]	0.67	0.46	0.375
	V^{rel} [b]	65.1	43.1	38.8
	V_{max}^{rel} [b]	98	94	103
	slope [c]		**- 0.02**	**0.04**
dThd	% neutral	100	100	99
	K_m (mM)	0.29	0.11	0.10
	[ES] / [E_t]	0.41	0.65	0.67
	V^{rel} [b]	68.3	43.5	32.6
	V_{max}^{rel} [b]	167	67	49
	slope [c]		**- 0.40**	**- 0.34**

[a] Standard assays, at 25°C, contained saturating potassium phosphate (0.1M), substrate (0.2 mM) and a negligible enzyme concentration.
[b] nmol/min·mg enzyme preparation.
[c] slope = ($\log V^{rel}_{max}$ higher pH) − ($\log V^{rel}_{max}$ lower pH) / ΔpH.

Table A1.3. Kinetic parameters for *E. coli* TP phosphorolysis of 5-X-dUrd in 0.1M phosphate buffer at 25°C and pH 7.4

5-substituent	K_m (mM)	V_{max}(mM/min)
hydrogen	0.55 ± 0.03	1.93 ± 0.04
methyl	0.24 ± 0.02	1.15 ± 0.04
bromo	0.25 ± 0.02	0.20 ± 0.01
nitro	0.23 ± 0.02	0.013 ± 0.002



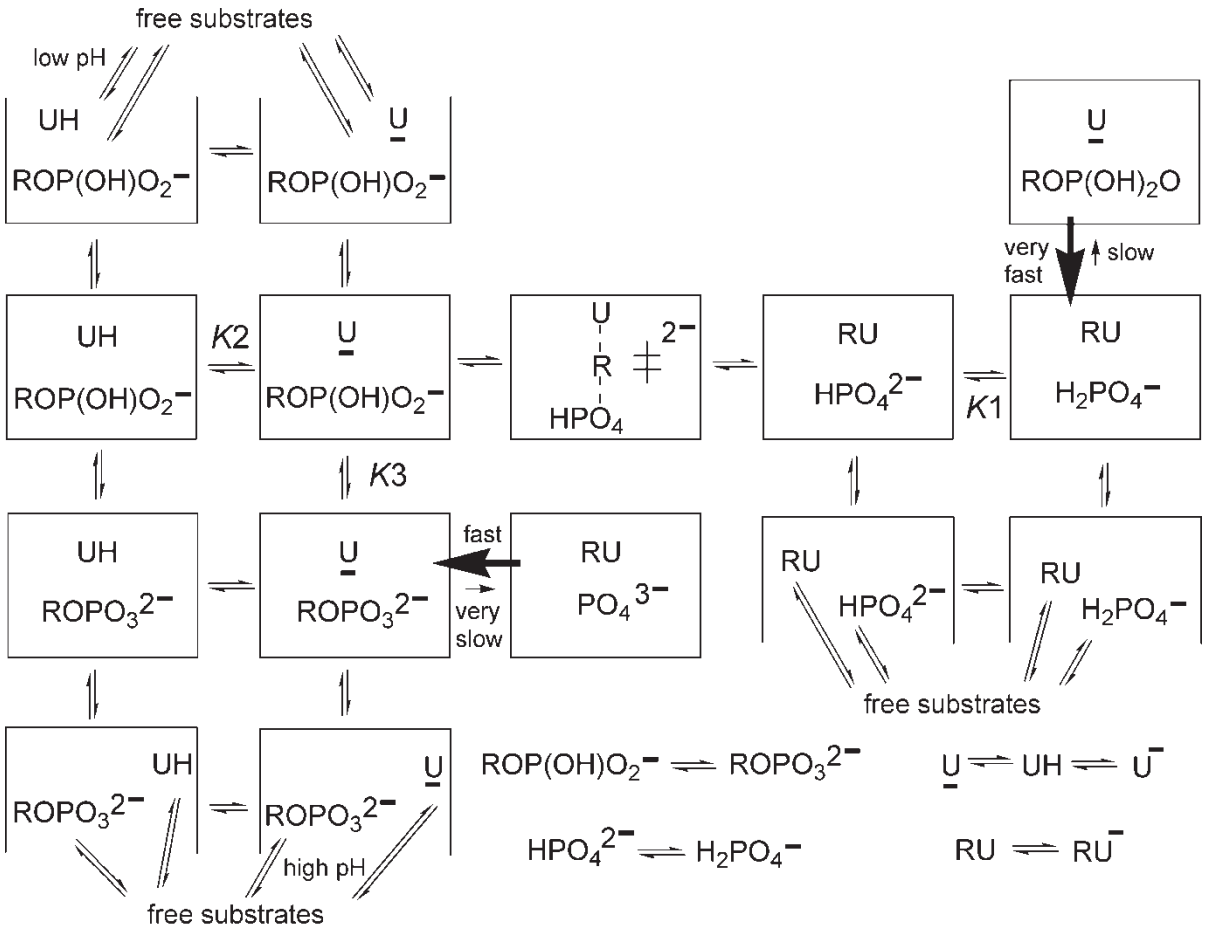
Scheme A1.1. 5-Nitro-deoxyuridine formation is facilitated by reducing the %-ionization of 5-nitrouracil.

result for the nitro compound is more extreme than that of the platelet experiment described above, but it does not suffer from some of the previous uncertainties of interpretation. The small V_{max} for the 5-bromo compound, relative to 5-methyl and 5-hydrogen, suggests that the position of the near-equilibrium involving bound substrates may have shifted from near X = CF_3 for horse and human TP, to a position close to X = CH_3; such a shift would put the bromo

Table A1.4. Numerical values used to construct Figure 2 and K_m (μM) values for the same set (horse-TP; pH 6.0). There were no standard deviations or similar information given for the $\log V^{rel}_{max}$ or K_m values. Also included in the Table are the predicted ionization states of 5-X-uracils at the pH of normal plasma.

5 substituent in 5-X-dUrd	σ_I[a] + 0.5σ_R^-	log V^{rel}_{max}	K_m, μM	fraction of 5XU as N1 anion at pH 7.4	fraction of 5XU as N3 anion at pH 7.4	fraction of 5XU neutral at pH 7.4
H	**0**	0.054	559	0.0028	0.0034	0.994
CH_3	−0.055	**0**	187	0.0015	0.0027	0.996
F	0.25	0.534	631	0.045	0.144	0.810
Cl	0.32	0.623	186	0.099	0.130	0.771
Br	0.33	0.545	121	0.113	0.124	0.763
I	0.31	0.586	157	0.094	0.084	0.822
CF_3	0.49	0.931	798	0.393	0.214	0.393
CH=O	0.625[b]	0.163	125	0.799	0.041	0.160
CN	0.70	−0.134	49	0.855	0.076	0.069
NO_2[c]	0.855	−0.240	100	0.948	0.040	0.012
CH_2OH	0.07	−0.220	577	0.0064	0.0084	0.985
C_2H_5	−0.045	−0.100	719	0.0017	0.0028	0.995
$CH_2CH=CH_2$	−0.05	−0.117	943	0.0015	0.0033	0.995
$CH=CHCH_3$	0.025	−0.241	314	0.0038	0.0057	0.990

[a] Sigma values are those of M. Charton [11].
[b] Sigma values are those of M. Charton [60].
[c] $\log V^{rel}_{max}$ estimated from data of Wataya and Santi [15]; K_m value is from that paper. Correction of K_m for %ionized at pH 6.0 gives a better comparison with the other data: 'non-ionized K_m' = 77 μM.



Boxed species correspond to closed forms of the enzyme/substrates complexes
Partially enclosed species correspond to open forms of the enzyme/substrates complexes

Open forms with a single, or no substrate bound are not shown but are needed for a more complete analysis. Variable ionisation states of protein side chains are not depicted despite His116 and Glu225 probably having roles in the pH profiles. Rate reductions below pH 6 may primarily be due to protonation of His116 in the open form, while the likely ionisation of Glu225 above pH ~7-8 will influence the rate profiles on the alkaline side.

R represents 1-linked deoxyribose; UH represents 5-substituted uracils;

RU represents 5-substituted deoxyuridines; U represents 5-substituted uracils ionised at N1;

U⁻ represents 5-substituted uracils (or uridines) ionised at N3.

Scheme A1.2. Ionization states of substrates and substrate/TP-complexes most relevant to effects of substrate-structure and pH on enzyme-catalysed phosphorolysis and deoxyribosyl transferase rates. No binding constants are available for any of the species into the complexes depicted and the only pK_as presently known are for unbound substrates. As a basis for the analysis of present data, including pH/rate profiles for (crude)TP from various species [3;57;58], I suggest the following very approximate pK_a values: p$K1 \approx 6$, p$K2 \approx 6$(thymine), p$K3 \geq \sim 8$(enzyme-source dependent). It presently seems that the binding site prefers the presence of one P-OH hydrogen bond donor, probably H-bonded to a protein residue that acts, very preferentially, as an acceptor.

substrate into the category of rate being limited by slow or low degree of product protonation at pH 7.4, but not limited by the position of the enzyme-bound equilibrium, which, on this hypothesis, must lie well on the side of products (cf. Figure 2: pH 6.0 using horse enzyme and Figures A1.1 and A1.2 using hTP). However, consideration of Scheme A1.2 shows that there are several plausible explanations of these results. No conclusion is possible on the nature of the product dissociation processes of the 5-methyl or 5-hydrogen uracils. It is clear that this prokaryotic enzyme differs substantially from the two mammalian enzymes.

The foregoing analysis strongly suggests that a protein conformational change, associated with N1 protonation, is (partially) rate limiting for many substrates. If that is the case, an enzyme-bound near-equilibrium is established between the two states

immediately on either side of the phosphorolysis transition state (or states) - see Scheme 1 for the thymidine/thymine case. The position of the enzyme-bound equilibrium (and off-enzyme equilibrium) will depend on pH, on the nature of the uracil 5-substituent and on whether phosphate is replaced by arsenate or sulphate. The fact that deoxyribose 1-sulphate is not detected when sulphate and thymidine are exposed to the enzyme can be explained by the enzyme-bound equilibrium, and the two rates that define the equilibrium, being closely related to the pK_as of the species involved: hydrogen-sulphate in water (p$K_a = 1.9$) is 200,000-times stronger an acid than dihydrogen-phosphate and that acidity ratio is unlikely to change much in the enzyme-bound state. Since the off-enzyme equilibrium (pH 7.5; 38°C) favours [thymidine + phosphate] over [deoxyribose 1-phosphate + thymine] by three to one at equal starting concentrations of thymidine and phosphate [59], the huge acidity difference should result in enzyme-bound deoxyribose 1-sulphate reverting back to thymidine extremely rapidly and vastly faster than release from the enzyme. The result could be almost undetectably slow formation of the sulphate ester. In the presence of a high sulphate concentration, its off-enzyme equilibrium position is expected to allow, eventually, small amounts of sulphate ester to be produced, but its spontaneous hydrolysis in the aqueous phase is likely to be overwhelming and in that case the enzyme would appear to act as a hydrolase. In the absence of sulphate (and phosphate and arsenate), hTP was reported to slowly hydrolyse thymidine (Asp123 may provide general base catalysis, or active-site-bound (adventitious) (bi)carbonate may be the nucleophile, or a buffer molecule may act as a nucleophile or as a general base for water attack in the open form of the enzyme) and that reaction could be inhibited by 0.5 M ammonium sulphate [39]. That result is consistent with the above analysis.

This type of analysis applies also to 5-X-uracil acidities. One can assume that the enzyme is designed to be efficient in both directions with its natural substrates since both reactions contribute to its utility to the organism. Selective pressures, particularly in micro-organisms, could lead to structure evolution to satisfy that efficiency criterion (optimizing for equal use in both directions would require equalization of the off-rates from the two sides of a $K_s(1) \leftrightarrow K_s(2)$ equilibrium, not equalization of their energies). The N^1-H of 5-nitrouracil is 50 000 times stronger acid than thymine so, here too, the enzyme-bound equilibrium should be hugely unfavourable to formation of 5-nitro-dUrd from deoxyribose 1-phosphate. Consistent with this, the enzyme-mediated synthesis of 5-nitro-dUrd from 5-nitrouracil and deoxyuridine(an insitu source of deoxyribose 1-phosphate) used an 80-fold excess of deoxyuridine at pH 6.0 and 37°C for *48 hours*! [21] The fact that the modest yield of 30% is not lower is in part due to the use of the somewhat below-optimum pH of 6.0 (deoxyribosyl transferase activity is optimal at pH ~7). The low pH makes the off-enzyme equilibrium position less unfavourable to 5-nitro-dUrd formation (pK_a ~6.5, so largely neutral) by increasing the proportion of 5-nitrouracil (pK_a ~5.5) existing in the neutral form: see Scheme A1.1.

Of course the poor result could be ascribed to the use of insufficient enzyme, but that is an extremely unlikely explanation.

The data for thymine, together with its hTP affinity [61], suggest that thymine N^1-anion binds to the enzyme with high nano-molar affinity (pK_s ~ 6.3 ± 0.5); therefore it is possible that protonation at N1 follows cleft opening, but in that case cleft reclosure must be faster than loss of thymine N^1-anion from the open form. K_m values for 5-X-dUrd synthesis, for most of the 5XUs in Table A1.4 are not available but the values for some 5XUs acting as inhibitors of 5-F-dUrd arsenolysis by *E. coli* TP at pH 5.9 [62], show that electron withdrawing groups at C-5 improve affinity when measured by *total* 5XU concentrations. However, consideration of the fraction-ionisation figures in Table A1.4 shows that electron withdrawing groups at C-5 strongly decrease affinity when it is measured by N^1-anion concentrations: withdrawing some of the negative charge from O^{C2} and O^{C4} in the N^1-anions leads to reduced binding of those atoms to the Arg202 and Lys221 cations.

Appendix 2

Discussion of σ_I and σ_R^- coefficients applicable to pK_a correlations. The differing σ_I and σ_R^- coefficients of 5-X-cytosines, 5-X-uracils at N^1-H and 5-X-uracils at N^3-H deserve comment. For the discussion, X is assumed to be electron withdrawing compared with X = H. In the resonance structures depicted, orbitals notionally containing two electrons are shown as circles or pear-shapes, see Figure A2. Antibonding orbitals are shown as dashed bonds to asterisks, as in structure (**s**).

Cytosine ($5.6\sigma_I + 2.9\sigma_R^-$) exists essentially exclusively in the form shown in structure (**a**) (three orders of magnitude separate this from the next-most stable tautomer, the N^3-H) so pK_a data cleanly reflects ionization at N1. The high dipole moment associated with the NH_2-C=N-C=O system enhances the ability of X to create a field complementary to it, and causes that field to deviate slightly away from the C-X direction towards the NH_2 function as shown in the figure. Resonance forms (**e**) (mainly σ_I) and (**f**) (mainly σ_R^-) are facilitated by the C-X field and by the

Table A2.1. Correlation of literature 5-substituted cytosines pK_a data with σ_I plus σ_R^-.

5-substituent	σ_I	σ_R^-	predicted pK_a of N^1—H=12.2 − 5.6σ_I+2.9σ_R^-	lit. pK_a N^1H
H	**0**	**0**	**12.20**	~12.2
CH_3	−0.01	−0.09	12.52	12.4
F	0.54	−0.58	10.86	10.87
Cl	0.47	−0.30	10.44	10.36
Br	0.47	−0.28	10.38	10.33
NHCH=O[a]	0.33	−0.33	11.31	10.92
CO_2Et[b]	0.30	0.31	9.62	9.85
NO_2[b]	0.67	0.37	7.37	7.40

[a] σ_I and σ_R^- values are conformation dependent (s-cis / s-trans, and twisting out of plane of ring)
[b] σ_R^- values are conformation dependent (but the hydrogen bond to the NH_2 helps keep them near-planar, especially in their anionic states).

Table A2.2. Correlation of literature 5-X-uracil pK_a data with σ_I plus σ_R^- and predictions of approximate pK_a values for other derivatives relevant to this work. All at 20–25°C.

5-X	σ_I	σ_R^-	predicted pK_a of N^1H= 9.95 − 5.2σ_I − 2.6σ_R^-	predicted pK_a of N^3H= 9.85 − 4.0σ_I −0.8σ_R^-	predicted pK_a N^1H and N^3H together[a]	lit. pK_a N^1H and N^3H together
H	**0**	**0**	**9.95**	**9.85**	9.59	9.43;9.45;9.5
			[~9.9]	[~9.7]		
CH_3	−0.01	−0.09	10.24	9.96	9.78	~9.8
F	0.54	−0.58	8.65	8.15	8.03	8.04;8.00;8.15
Cl	0.47	−0.30	8.29	8.21	7.95	7.95
Br	0.47	−0.28	8.23	8.19	7.91	7.83;7.95;8.05
			[8.30]	[8.30;7.84]		
I	0.40	−0.18	8.34	8.39	8.06	8.25
CF_3	0.40	0.18	7.40	8.11	7.32	7.35
CH=O[b]	0.36	0.53	6.70	7.99	6.68	
CN	0.57	0.26	6.31	7.36	6.27	
NO_2[b]	0.67	0.37	5.50	6.87	5.49	~5.5
			[5.70]	[7.35;7.20]		
NH_2[c]	0.17	−0.55	10.50	9.61	9.56	9.35;9.52
C_2H_5	−0.01	−0.07	10.18	9.95	9.75	9.9
CH_2OH	0.11	−0.08	9.59	9.47	9.22	9.4
CH_3S[b]	0.30	−0.24	9.01	8.84	8.62	8.5

Values in square brackets are literature values for mono-N1/N3-methylated uracils. The approximation sign (~) indicates that several determinations have been reported: the tabulated value is a judgment-weighted mean. The predicted pK_a values in the table should, *at best*, be considered reliable to 0.1 unit. Both σ_I coefficients have uncertainties of about 5% of the given value, the σ_R^- for N^1H about 10% and that for N^3H about 50%.

[a] The sum of the predicted K_a values for N^1H and N^3H results in the prediction for the K_a of the parent compound in which both NH groups contribute to the observed K_a (and hence pK_a).

[b] σ_R^- values are conformation dependent. The nitro group is expected to be twisted out of the plane of the ring in the neutral form but may be planar in the N1-anion due to increased charge delocalization. This contrasts with the nitro N3-anion where the twist is expected to increase due to minimal resonance of charge onto the nitro group and increased negative charge on O^{C4}, longer C4-O bond and bending of that bond away from $N3^{\delta-}$ towards the NO_2 group, so the predicted pK_a, which does not take account of such effects, is too low.

[c] The σ_I and σ_R^- values for NH_2 are specially context dependent since they depend on the hybridization and conformational states in the neutral species and the anion(or cation), and on the hydrogen bonding status of the group - particularly that of the lone pair to water. In anionic species, e.g. para-aminophenolate, the amino group becomes more pyramidal and the lone pair can rotate out of conjugation and become well solvated by water. Both effects likely have some relevance to the N1-anion in water. The σ_R^- values used here are derived from anilinium acidities not phenol acidities; NH_2 lone-pair resonance in 4-NH_2-anilinium cation is much enhanced over that in para-aminophenol. Phenol-ionization based σ_R^- for NH_2 is reduced to −0.25 [63], due to the previously described conformational and hydrogen bonding effects. But in the TP-bound state no group is available to hydrogen bond to the NH_2 lone pair, while hydrogen bonding occurs between one amino N-H to Thr118 carbonyl and the other NH to the partially anionic oxygen on C4, the latter oxygen coplanar and the former nearly coplanar with the ring. Also, the dipole generated by resonance donation of the NH_2 lone pair into the ring is complementary to the C4=O dipole. This contrasts with the dipoles associated with electron withdrawing groups at C5 which are repulsive. These factors, combined, suggest that the ΔpK_a between 5-CH_3 and 5-NH_2 uracils bound to TP may be larger than the numbers in this table would predict.

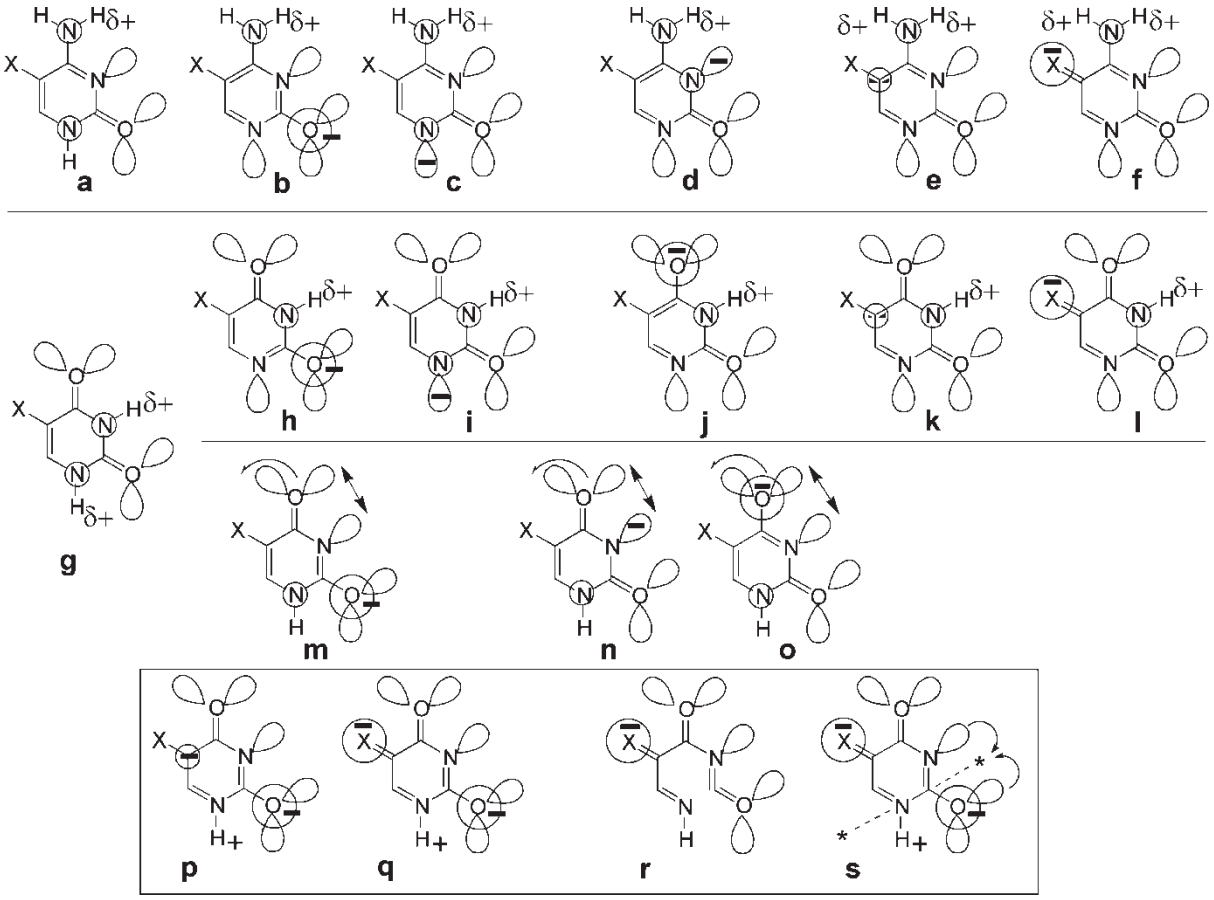


Figure A2. Resonance forms of cytosine and uracil anions.

reduction in lone-pair and $N^{\delta-}/O^{\delta-}$ repulsions present in (**b**), (**c**) and (**d**).

5-Substituted uracils exist essentially exclusively in the tautomeric form shown in (**g**), but ionization of both N^1-H and N^3-H contribute to the observed pK_a. The strong dipole associated with C4═O reduces the ability of X to create a field, because of the resulting dipole repulsion, and the field generated will deviate slightly away from the carbonyl (the reverse applies to a planar 5-NH_2 group).

Ionization at uracil N1 ($5.2\sigma_I + 2.6\sigma_R^-$) increases electron density on both C═O groups and, considering structure (**g**), the increased density and C—O bond length increases the dipole opposition with C-X; resonance forms (**k**) and (**l**) similarly are somewhat inhibited. Lone-pair and $N^{\delta-}/O^{\delta-}$ repulsions are reduced in comparison to the cytosines due to the presence of the hydrogen on N3, so there is less pressure pushing electrons onto C5 and X.

Ionization at uracil N3 ($4.0\sigma_I + 0.8\sigma_R^-$) has by far the smallest coefficients. This was expected for σ_R^- but unexpected in the case of σ_I. Because the dipole associated with C-X will deviate slightly away from the C-X bond, as shown in (**g**), it probably destabilizes the proton on N3 very slightly more than that on N1. Thus the small σ_I coefficient arises almost exclusively from C-X effects in the anion. The main resonance forms, (**m**), (**n**) and (**o**), result in severe lone-pair and $N^{\delta-}$ / $O^{\delta-}$ repulsions. The increased negative charge on O^{C4}, the longer C4—O bond, and bending of that bond away from $N3^{\delta-}$ towards the X group, increases their mutual repulsion. The latter fact, together with the attraction of $N1H^{\delta+}$ for the adjacent lone pair and the charge on O^{C2} in (**m**), and the very minor contribution of (**p**) and (**q**), results in the average position of the negative charge being substantially farther away from C-X than for the other cases considered. The use of σ_R^- rather than σ_R^0 for resonance involving this meta-like relationship is based on two arguments: firstly, the much greater electron density on O^{C2} and O^{C4} in the anion allows electrons on N1 to interact more strongly with C5-X, as in structures (**p**) and (**q**); secondly, the high electron density on O^{C2} and N3 in the anion, and their mutual repulsion, encourages sigma resonance donation into the N1-C2 antibonding orbital with concomitant elevation of the density and energy of N^1-associated electrons. That further enhances the interaction with C5X. The sigma resonance can be depicted as in structure (**r**), the bond/ nobond

representation, but structure (**s**) better represents the very small amount of electron density that is actually involved in the interaction.

Appendix 3

Arsenate ester formation and hydrolysis: kinetics and implications. The assumption that arsenate monoesters in water hydrolyse spontaneously at a high rate is still being made despite long-standing publications to the contrary. In 1973, Long and Ray demonstrated that aqueous solutions of glucose and inorganic arsenate at pH 7.5 and 30°C produce, along with other products, glucose 6-arsenate with $k_{\text{form}} = 1.4 \times 10^{-6}\ \text{M}^{-1}\text{sec}^{-1}$. Hydrolysis of the glucose 6-arsenate under the same conditions had $k_{\text{hydrol}} = 4.2 \times 10^{-4}\,\text{sec}^{-1}$: that corresponds to a half-life ($t_{\frac{1}{2}}$) of 28 minutes and the equilibrium constant for the aqueous reaction: glucose + $As(OH)O_3^{2-}$ ↔ (glucose 6-arsenate) + H_2O, has $K_{\text{eq}} = 0.19$. They also showed that ester formation and hydrolysis reactions are both faster at lower pH, but with little effect on K_{eq} [64]. At pH 7.0 and 25°C, glucose 6-arsenate and 6-arsenogluconate are reported to have $t_{\frac{1}{2}}$ of 6 and 30 minutes respectively: long enough to measure their K_{m} values and inhibitory characteristics towards a number of enzymes [65]. Under the same conditions of temperature and pH, nucleoside 5′-arsenates and 2′-deoxynucleoside 5′- or 3′-arsenates have half-lives of 30 to 45 minutes [66]. The earliest relevant work, from 1959, showed that at pH 8.0 and 32°C the exchange rate of oxygen isotopes between water and $As(OH)O_3^{2-}$ proceeds with a $t_{\frac{1}{2}}$ of 44 minutes [67].

The kinetics of hydrolysis of deoxyribose-1-arsenate has not intentionally been the subject of experiment. However, experimental results reported graphically by Friedkin and Roberts in 1954 can be used to give some indication of its rate of hydrolysis [58]. The utility of their graph for this purpose only becomes apparent when one converts the ordinate from their absorbance change (ΔE_{300}), to % reaction using results from a subsequent paper by the same authors [59]. The resulting graph is shown as Figure A3.

Thymidine phosphorolysis by *pure* TP leads to an equilibrium mixture of [thymidine plus phosphate] and [thymine plus dR1P] in a ratio of ~75:25 when starting from equal concentrations of reactants [59], but the crude enzyme used for this experiment also had weak hydrolase activity. The curve for arsenolysis is consistent with approach to an equilibrium mixture of [thymidine plus arsenate] and [thymine plus deoxyribose 1-arsenate] in a ratio of ~1:1 with slow non-enzymic hydrolysis of the arsenate ester (added phosphate reduced the hydrolase activity of the crude enzyme preparation but the extent to which added arsenate modified that activity can not be estimated). These results are consistent with deoxyribose 1-arsenate hydrolysis occurring on a timescale similar to or even longer than the 6-hour timescale of the experiment. But until results from specifically-designed experiments are forthcoming, the estimate must be regarded as very uncertain.

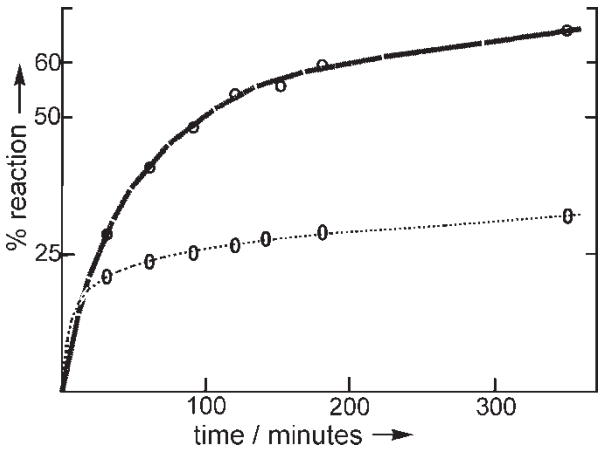


Figure A3. Phosphorolysis (lower curve), or arsenolysis (upper curve), of thymidine at 38°C and pH 7.5 catalysed by crude TP from horse liver. Initial concentrations were all 3.2mM.

Slow deoxyribose 1-arsenate hydrolysis is consistent with many reports of efficient, arsenate-catalysed, pyrimidine deoxyribosyl-transferase activity with TP and related enzymes, dating back at least as far as 1967 [68]. Equally long established is the fact that arsenate, unlike phosphate, very inefficiently stimulates ribosyl transfer with purine nucleoside phosphorylase (PNP) [69]. The latter observations have been claimed to demonstrate the hydrolytic lability of arsenate esters, but the observed absence of transferase activity could be accounted for by the migration of the arsenate to the more favoured 2-, then 3-, and even 5-hydroxyl groups via 5-membered and 6-membered cyclic pentavalent diester intermediates. Longer As-O bonds than P-O bonds and their greater flexibility allow for less strain in the 5- and 6-membered rings and even faster internal transfer than their relative, acyclic rates of hydrolysis would indicate. The intramolecular transfer is expected to be fast for those migrations involving 5-membered rings since they incur minimal strain when bonded in axial plus equatorial positions of trigonal-bipyrimidal intermediates. The spontaneous phospho-diester hydrolysis and rearrangement that occurs in RNA (contra DNA wherein no five-membered ring intermediate is possible) at neutral pH provides an excellent, well-documented precedent [70].

The equilibrium position of the arsenate should disfavour arsenic attachment to the 1-oxygen on the

grounds that it is a secondary alcohol, is the most acidic, and the large ester is expected to exist, on stereoelectronic grounds, mainly in a sterically-hindered pseudo-axial conformation - as seen for ribose 1-phosphate bound to bovine PNP [32], and calculated for 2-deoxyribose 1-phosphate in vacuo [38]. Attachment to the primary 5-oxygen is expected to be most favoured but much the slowest to form. Molecular modeling suggests that ester hydrolysis could be speeded by adjacent hydroxyl groups hydrogen bonding to equatorial anionic oxygens of the trigonal-bipyrimidal intermediates involved in hydrolysis. That type of intermediate linked to a 2-ribofuranose oxygen could be stabilized by hydrogen bonds, in a O^{C2}-pseudo-axial conformer, from both cis-related, 1- and 3-hydroxy groups to two of the equatorial oxygens (not possible with tetrahedral arsenate). By analogy with phosphoranes [70,71], a mono-anionic penta-coordinate intermediate, with an equatorial oxygen carrying a formal charge, is expected to be basic, pK_a ~7–11 (most likely ~9), and therefore subject to general acid catalysis by buffer species and hydrogen bonding catalysis from the, entropically-favoured, adjacent hydroxy groups that are more acidic than water. Stabilization of a possible di-anionic intermediate (pK_2 *very* approximately 13) would be substantially greater. The ribose hydroxyls also may facilitate proton transfers during formation, rearrangement and breakdown of the intermediates. In marked contrast to the ribose 1-ester, transfer of arsenate intramolecularly in the deoxy series would have to involve a strained bicyclo[3,2,1]-heterocyclic, trigonal-bipyrimidal intermediate, so this transfer is unlikely to compete with normal hydrolysis, and there are no adjacent hydroxy groups with the potential to facilitate hydrolysis.

The easy formation of arsenate esters from hydroxy groups and inorganic arsenate at neutral pH has been utilized to generate a very long-lived, active holoenzyme by exposing the apoenzyme of aspartate aminotransferase to pyridoxal and inorganic arsenate [72]. This and related observations noted above raise concerns that similar ester formation could occur in other proteins which bind, by design or otherwise, arsenate(phosphate) close to a hydroxyl-bearing cofactor, or a substrate, or indeed the hydroxy groups of the protein itself. Human TP, as modeled herein, has two serine and one threonine hydroxyl groups in tight contact with phosphate/arsenate and phosphate binding sites generally are likely to have serine/threonine residues contacting the phosphate.

The interpretation of results arising from use of arsenate as a substitute for phosphate is fraught with uncertainties. Appropriate qualifications should surround any conclusions based on its use. Many results and conclusions arising from past uses of this device in biochemistry need to be reassessed.

Appendix 4

Concerns over KIE interpretations arising from related QM calculations in vacuo. High level QM calculations indicate that substantial kinetic, binding and equilibrium isotope effects may result from conformational changes in ^{n}H-C-O-H angles and dihedral angles, differences in hydrogen bonding to the oxygen or its attached hydrogen or both, and other types of interactions between a ligand and a protein that differ from those in the starting aqueous solution. Mainly very simple model compounds were used in the calculations, but the results demonstrate that the limited range of functionality, and even more limited investigation of geometrical parameters, leaves much room for speculation concerning the situations in which the investigated interactions will dominate effects associated with covalent bond changes [73–75].

One particularly striking example has the oxygen atom of formaldehyde aligned with the C^2-H bond of propan-2-ol as depicted in Figure A4. Tritium equilibrium isotope effects (EIE) were calculated at several CH–O distances up to the maximum value of 350pm (3.5 Å), which is well beyond van der Waals contact. C-H bond lengths also were calculated [75]. Changes to these bond lengths, relative to unperturbed length, are plotted below the structures in Figure A4. The length changes were correlated with the EIE effects over the separation distances 230-350pm. The difference in EIEs over that range was ~8.2%. As can be judged from the graph, the EIEs relative to an isolated propan-2-ol molecule are unexpectedly large. Most remarkably, the effects are substantial at distances well beyond the van der Waals contact distance and these effects arise with formaldehyde maintained orthogonal to the C-H bond. One can only speculate on the magnitude of the effects if a very-much more polar amide oxygen atom, or even a carboxylate anion, is aligned so its most negative aspect is in line with the C-H bond. There seems to be only one possible explanation for the effects at medium and longer distances: the negative field due to the carbonyl oxygen inhibits $n_O \rightarrow \sigma^{\star}_{CH}$ due to that interactions enhancement of electron density beyond the H^{C2} nucleus in the direction of the carbonyl oxygen atom.

A related example comes from very recent KIE/QM calculations of the TS structure of 5′-methylthioadenosine nucleosidase: these required that the attacking water nucleophile be separated from the oxacarbenium carbon by substantially more than 350pm since including a water molecule at that distance altered the calculated KIEs beyond the errors of the experimental KIEs. So, as with the isopropyl alcohol/formaldehyde example, there are substantial calculated effects due to an oxygen atom that is well

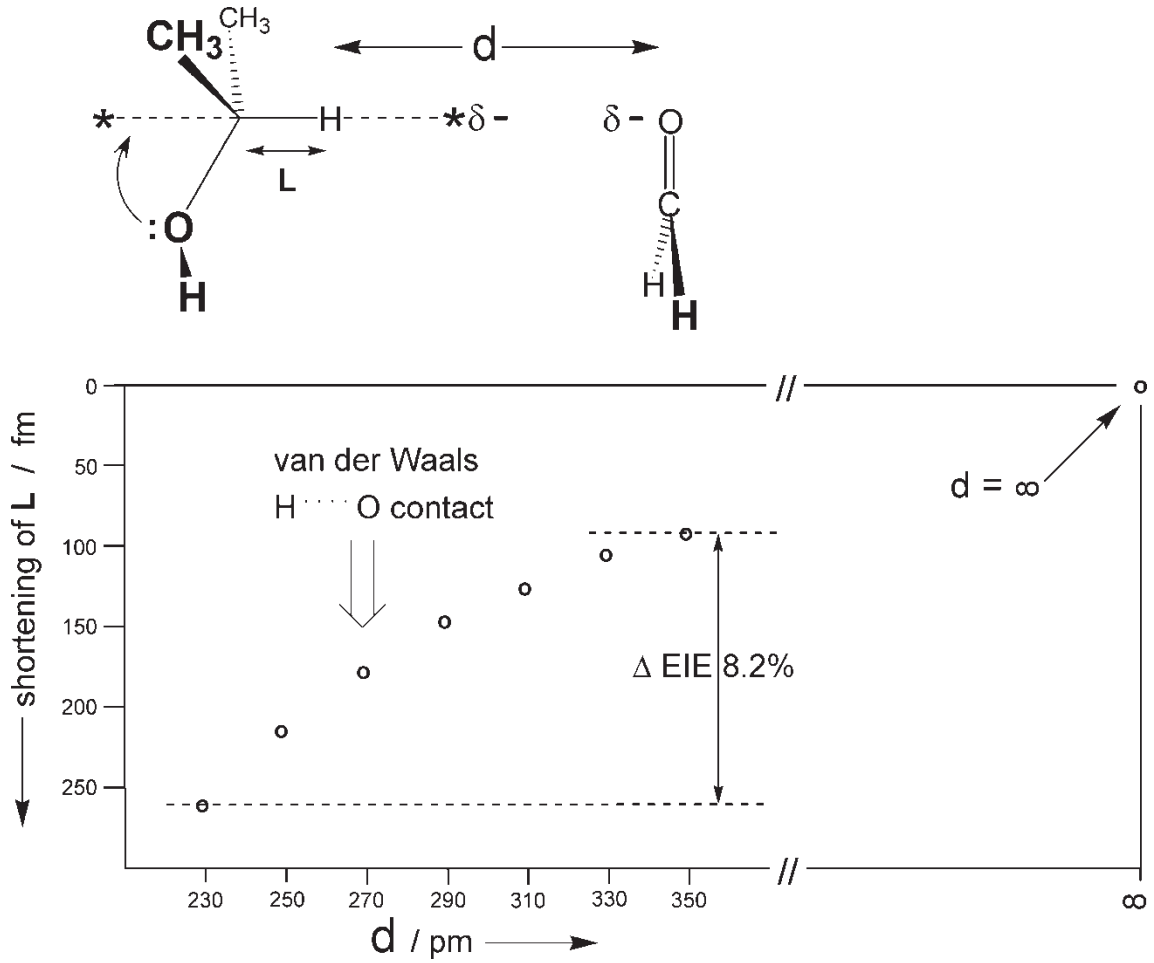


Figure A4. Propan-2-ol C^2-H bond length changes in response to variable H^{C2} distances from formaldehyde oxygen.

beyond the van der Waals contact distance with a relevant atom [76]. On the basis of these results, the effects due to the oxygens of water in the unbound states of ligands containing hydroxy-groups, ether oxygens, amines, and some other functionalities, are likely to be substantial and structure dependent, while the effects in bound states would be extremely ligand-position and protein specific.

There are many other concerns about the validity of combined KIE/QM calculations of transition state structure as currently performed but one more example will suffice: there often are wide variations in KIE predictions according to the methodology used. Sometimes results from a number of methods are reported and one or more deviate significantly from the remainder [74;43].

The KIE/QM-calculated TS geometry for TP is particularly uncertain since the KIE-matching S_N2-like structure, determined at the B1LYP/6-31G* level of theory, required very significant constraints on the C1′ to N1 and C1′ to O^P distances in order to maintain the structure some distance from a stationary point. An unconstrained calculation at the same level of theory predicted an oxacarbenium-like TS structure, with partial bond lengths of 2.14 and 2.51 Å. In both constrained and unconstrained calculations, the HPO_4^{2-} is stabilized by a single hydrogen bond from the deoxyribose-3′-OH while the partially-anionic thymine ring is involved in a single hydrogen bond from the deoxyribose-5′-OH to O^{C2} (the authors show computer-generated, space-filled pictures, and capped stick pictures in their Supporting Information, but no coordinate files or other quantitative stereochemical information were published) [38]. In the waters plus protein modeled 'TS'-environments discussed herein, the phosphate anion and the substantially-anionic thymine leaving group are both hugely stabilized by multiple extremely strong interactions, ionic and hydrogen bonding, the result of which will be greatly to stabilize a D_NA_{Nint} or $D_N{\star}A_{Nint}$ [77], (S_N1-like) TS relative to a tight A_ND_N (short partial bonds, S_N2-like) TS.

Kinetic isotope effects using *E. coli* thymidine phosphorylase have been determined for the arsenolysis of [$1'^{13}$-C], [$2'^{13}$-C], [$1'^{2}$-H], [$2'^{2}$-H], [$5'^{2}$-H], and [1^{15}-N] uridines. The KIEs obtained were matched to a calculated gas-phase transition state using bond energy bond order vibrational analysis.

The transition state was deduced to have substantial oxacarbenium ion character with the C-N glycosidic linkage being largely cleaved and with the phosphate oxygen nucleophile 'just within bonding distance' of C1′ [78]. Oxacarbenium ions derived from uridines are very much less easily formed than those from their deoxy-equivalents, so these results make the KIE findings for hTP/dThd even more surprising.

Appendix 5

Modification of hTP (PDB: 1UOU) structure and TS modeling. One half of the hTP dimeric structure (PDB 1UOU) was deleted then hydrogens were added and external waters were deleted if they seemed to have little relevance to maintaining the X-ray-defined protein conformation. New waters were added where they could help to maintain the protein conformation during the energy refinement calculations. On the same logic, external ionized side chains were either truncated to remove the ionized functional group, or were conformationally pre-adjusted to minimize electrostatics-induced distortions at the chosen low, uniform permittivity ($\varepsilon = 1.3$ or 1.5). The very low permittivity was required in order to reproduce the X-ray diffraction derived hTP geometries, especially the hydrogen bond distances. Some of the high energy side chains were remodeled to lower energy forms when that resulted in minimal changes to their contacts with surrounding residues. Glu225 was modeled in non-ionized form due to its very tight contact with the carboxylate group of Asp114. Except for His116, all histidines, tyrosines and cysteines were modeled in non-ionized forms with imidazole tautomeric forms chosen, along with the conformations of all other hydrogen bonding groups and waters, so as to generate dense hydrogen bonding networks wherever possible. His116 was modeled in cationic and both tautomeric neutral forms. All non-truncated arginine and lysine residues were modeled in their protonated forms. Two unresolved residues, Glu238 and Ala239, were added but not so the unobserved residues at chain termini. It was considered desirable to add amino-acids 238–239 because His150εNH can form a hydrogen bond to the carbonyl oxygen of Ala239 and that interaction may be important for holding in place the long, flexible, glycine-rich loop that, in part, closes the active site; His150 is at the far end of that loop. The two new residues were modeled in various ways and the energies minimized. No wholly satisfactory arrangement was found but changing the local structure had little effect elsewhere. Two arrangements were used in further work; latterly the choice was to have the Gly237-Gly238 bond modeled with an s-cis amide as this caused least distortion of the amide bonds in the completed loop. Disordered Met273 was modeled in different ways according to the side chain conformation of Lys235; the preferred arrangements are Met273(C1 to Cε: a,a,a) and Lys235(C1 to Nζ: a,a,a,g^+) with hydrogen bonds to Asp123, Asp233 and a tight contact/bent-H-bond with Ser117Oγ. The active site lysines, Lys115 and Lys222 in the 1UOU structure would be in high energy conformations if they existed in a small polypeptide exposed to solvent water, therefore alternative conformations were investigated: Lys115(C1 to Nζ: g^+,a,g^+,g^+) and Lys222(C1 to Nζ: $\sim$140,g^+,a,a) were those most studied but the 1UOU conformations are preferred.

Energy refinements were performed using the Merck Molecular Force Field, MMFF94s [47] as supplied by Tripos Inc. Partial bonds are not accommodated by the program so the input TS structures were drawn with either five covalent bonds to the sugar C1′ or with C1′ having a double bond to cationic O4′ and no bonds drawn to N1 or O^P. The charges assigned to thymidine, phosphate, 1,2-di-deoxyribose 1-carbenium ion, O-protonated acetaldehyde, and the four program-acceptable representations of thymine ionized at N1 are shown in Figure A5. The charges assigned to C4′ and O4′ of the oxacarbenium ion are absurdly large but

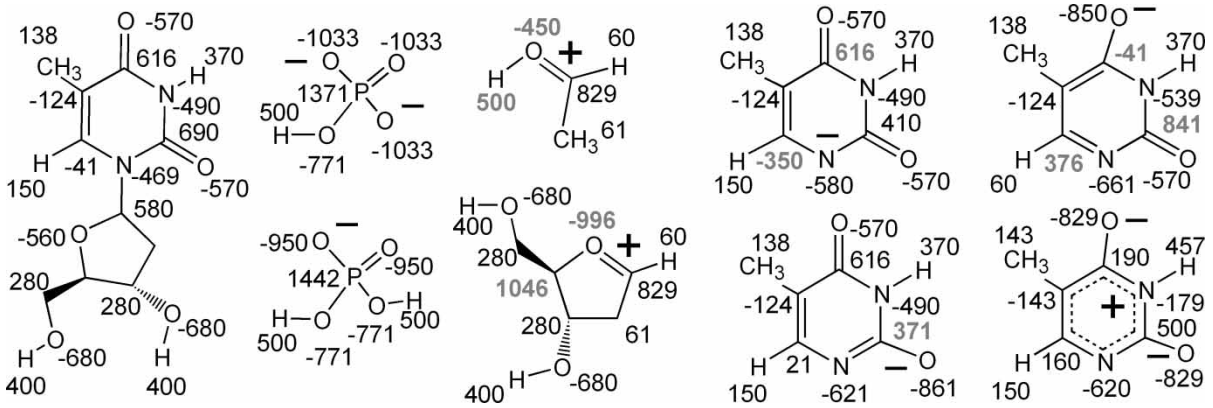


Figure A5. Charges(millielectron) assigned to TPrelated structures.

Table A5.1. Structural and energetic properties (kcal; pm) of MMFF94s force-field refined models of carbenium-type hTP transition states and substrates bound to a model of hTP based on a recent X-ray structure of human TP (PDB: 1UOU).

	column:	**a**	**a** ionic	**b**	**c**	**d**	**e**	**f**	**g**	**h**	**i** ionic	**i**
	structure:	dThd	**a** ionic	TS200	TS220	TS240	TS260	TS280	TS free	TS300	**i** ionic	dR1P
row												
1.0	*constraint*	none	none	200	220	240	260	280	none	300	none	none
	N1−− C1′	149	149						291		418	418
2.0	C1′−−O^P	364	364	313	301	293	284	271	274	271 ionic	141	141
2.1	N1−−O^P	510	510	512	520	525	542	545	560	565	542	542
3.0	**P.L E_t**	(−746.2)	−524.0	−773.9	−795.8	−807.8	−813.4	−815.7	−815.9	−815.8	−568.7	(−718.0)
3.1	**normalized**		291.9	**42.08**	**20.12**	**8.13**	**2.56**	**0.2**	**0**	**0.16**	247.2	
4.0	L E_t	(−125.8)	106.9	−141.3	−162.7	−173.9	−178.8	−180.7	−180.6	−180.5	49.7	(−118.5)
4.1	**normalised**		287.6	39.39	17.95	6.79	1.84	0	0.05	0.14	230.3	
5.0	**(P.L−L) E_t**		−630.9	−632.6	−633.1	−633.9	−634.6	−635.1	−635.3	−635.3	−618.4	
5.1	**normalized**		**4.4**	**2.7**	**2.2**	**1.3**	**0.7**	**0.2**	**0**	**0.0**	**16.9**	
5.2	(P.L−L) E_e		−7054.0	−7062.8	−7064.2	−7065.4	−7066.6	−7068.1	−7069.0	−7069.5	−7052.2	
5.3	normalized		15.5	6.73	5.33	4.12	2.86	1.42	0.52	0	17.25	
6.0	**P E_t**			−497.65	−497.52	−498.51	−498.49	−498.58	−498.43	−498.33		
6.1	**normalized**			**0.93**	**1.06**	**0.07**	**0.09**	**0**	**0.15**	**0.25**		
6.2	P E_e			−6937.7	−6937.7	−6938.3	−6938.4	−6938.6	−6938.6	−6938.7		
6.3	normalized			1.00	1.04	0.38	0.20	0.11	0.15	0		

Abbreviations: P.L, protein plus ligand/substrates complex; L, ligand/substrates; E_t, molecular mechanics ***t***otal energy; E_e, molecular mechanics ***e***lectrostatic energy

Table A5.2. Geometric details of structures **(a)** to **(f)** and **(h)** and **(i)** in Table A5.1

	Structure							
Distance/angle	**a**	**b**	**c**	**d**	**e**	**f**	**h**	**i**
	dThd	TS 200	TS 220	TS 240	TS 260	TS 280	TS 300	dR1P
N1––C1′	149	200	220	240	260	280	300	418
N1––O^P	510	512	520	525	542	545	565	542
N1––dR––H(2′*S*)	***255***	256	258	261	263	265	267	***253***
N1––H(2′*S*)––C2′	76	90	96	101	107	112	118	**157**
N1––H116Nε	***418***	394	392	378	371	358	348	***311***
O2––H116Nε	***304***	311	313	312	313	314	317	***324***
O2––K221Nζ	***344***	306	306	298	296	292	289	282
K221Nζ––D1195Oδ	262	**269**	269	269	269	268	267	266
O2––C2′	297	336	348	364	377	390	400	440
O2––H2′	264	268	276	285	294	303	311	359
N3––S217Oγ	319	313	311	307	303	299	295	297
O4––R202Nη1	296	289	288	287	285	283	280	286
η2	294	291	288	287	285	284	282	281
C6-N1-C1′	120.0	114.1	111.8	108.5	105.6	101.9	98.4	100.2
C2-N1-C1′	119.3	129.9	131.1	133.3	134.9	136.1	135.7	135.2
N3–C2–N1–C1′	181.7	180.5	183.2	188.9	194.5	204.0	222.3	210.7
C1′––O^P	364	313	301	293	284	278	271	141
H1′––S117,**O**═C	***294***	278	274	**269**	265	262	258	277
C1′–H1′––S117,**O**═C	119	122	124	124	126	127	129	146
O4′–C1′–N1	111.4	102.4	102.3	102.6	104.4	106.6	109.8	100.8
C2′–C1′–N1	113.8	104.1	101.4	99.0	95.7	92.3	88.6	55.9
H1′–C1′–N1	106.5	77.3	74.2	72.0	70.6	69.8	69.1	62.5
C1′––O–P	142.2	135.0	135.2	132.2	131.2	129.0	128.0	121.4
O^P–C1′– N1	163.3	173.6	172.6	172.1	170.0	167.3	163.2	146.6
O3′––O^P	259	256	255	255	255	255	255	279
O3′-H–––O^P	173.1	166.0	165.3	164.1	163.6	163.3	163.4	***137.6***
O3′––T154N.am	***338***	314	310	308	304	303	303	304
O3′––R146,**O**═C	324	311	311	310	309	307	305	***285***
C4′––O^P	342	324	324	322	322	322	324	310
O5′––T151Oγ	***276***	280	280	281	281	282	281	***307***

Table A5.3. Structural and energetic properties (kcal; pm) of MMFF94s force-field refined models of carbenium-type TP transition states bound to a model of hTP: method varied from that in Table A5.1 as indicated in black/bold/italic.

Structure:	**A5.3.1**	**A5.3.2**	**A5.3.3**	**A5.3.4**	**A5.3.5**	**A5.3.6**	**A5.3.7**
constraint: N1––C1′	200	220	240	260	/**xxx**	free	300
6. annealed: thy⁻ & deox⁺; immobile: TS 220 site plus Pi, as in column (c) Table A5.1							
6.0 P.L.E_t	−773.7	−795.9	−807.4	812.8	−814.5	−814.5	−813.9
6.1 L.E_t	−141.5	−162.9	−173.9	−178.7	−180.1	−179.6	−179.2
6.2 (P.L.-L) E_t	−632.2	−633.0	−633.5	−634.0	−634.4	−634.9	−634.6
6.3 (P.L -L) E_t normalized	**2.69**	**1.9**	**1.38**	**0.84**	**0.51**	**free 0**	**0.28**
6.4 C1′––O^P / N1––C1′ **(xxx)**					/**280**	265/**282**	
6.5 O^P ––N1	514	521	528	533	542	541	550
7. annealed: thy⁻ & deox⁺; immobile: dThd site plus Pi, as in column (a) of Table A5.1							
7.0 P.L E_t			−801.6	−806.5	−807.7	−808.1	−807.2
7.1 L.E_t			−172.4	−177.1	−177.3	−178.0	−177.3
7.2 (P.L. -L) E_t			−629.2	−629.4	−630.3	−630.1	−629.9
7.3 (P.L -L) E_t normalized			**1.19**	**0.86**	**0**	**free 0.28**	**0.40**
7.4 C1′––O^P / N1––C1′ **(xxx)**			284		/**270**	263/**278**	
8. annealed: thy⁻ & deox⁺& Pi; immobile: dThd site as in column (a) of Table A5.1							
8.0 P.L E_t		−791.2	−802.2	−807.5	−808.6	−809.2	−808.8
8.1 L E_t		−162.5	−173.4	−178.1	−179.3	−180.7	−180.6
8.2 (P.L -L) E_t		−628.6	−628.8	−629.4	−629.3	−628.5	−628.2
8.3 (P.L -L) E_t normalized		**0.75**	**0.56**	**0**	**0.09**	**free 0.92**	**1.22**
8.4 C1′––O^P / N1––C1′ **(xxx)**		298	288	278/**260**	**/270**	268/**281**	264

Table A5.4. Structure **A5.3.4; rows 8.n**, from Table A5.3 was the starting point for the new structure **A5.4.1**; constraints were applied as in the **A5.4.1** column and relaxation was as follows: *substructures/waters having an atom within 9 Å of substructures/waters within 3.7 Å of substrates (32 substructures/waters @ ≤ 3.7 Å) annealed with immobile substructures/waters outside the 3.7 Å /9 Å set and having an atom within 17 Å of the 3.7 Å set (→ 5720 atoms out of 6699 total). Nonbonded cutoff = 15 Å, ε = 1.5, Powell minimization for 43 iterations with nonbonded reset every 4 iterations.* All energies in the table are for complete end structures, all 6699 atoms, constraints removed, nonbonded cutoff > max. dimension of protein, ε = 1.5. Structure **A5.4.2** was obtained from **A5.3.4; rows 8.n** by relaxation in the same (*italicized*) way. Structure **A5.4.3** used structure **A5.4.2** as its start point and it was relaxed as previously. Structures **A5.4.4** and **A5.4.5** each used the previous structure as start points and each was relaxed as in the other cases. In this way the structures suffering the largest changes were allowed longer and more gradual relaxation.

		structure				
		A5.4.1	A5.4.2	A5.4.3	A5.4.4	A5.4.5
row						
1.0	constraint N1--C1′	260	250	240	235	230
1.1	Constraint C1′--O^P	275	265	255	245	235
2.0	**N1--O^P**	**532**	**511**	**491**	**476**	**461**
3.0	P.L E_t	−810.4	−806.9	−802.2	−797.5	−791.4
4.0	L E_t	−178.2	−175.4	−171.2	−167.5	−162.7
5.0	**(P.L -L) E_t**	−632.2	−631.5	−631.1	−630.0	−628.7
5.1	**normalized**	**0**	**0.76**	**1.16**	**2.28**	**3.56**
5.2	(P.L -L) E_e normalized	0	2.00	3.49	5.71	7.70
6.0	P E_t	−496.4	−496.1	−496.8	−497.2	−497.6
6.1	normalized	1.21	1.51	0.75	0.39	0

those assigned to O-protonated acetaldehyde are, for this program, roughly in line with expectations; also note the very small charge assigned to H1′ and C2′ and zero charges on all sp^3C-*H* atoms, including both H2′ in the carbenium ion. The gross variation in atom-centered charges with the depiction of the thymine anion also presented problems: the 2-oxyanion-3,4-dihydro-4-oxo form was used throughout this work. Angles attaching an oxyanion to the ring were a concern, especially for the 4-position: the minimized 2-oxo-4-oxyanion had external angles at C4 of 109.1° and 133.3°. Changing the substituent at C5 modifies the charge at C5 but at no other atom in or attached to the ring. Atom-centered charges are automatically assigned by the program and changes to specific atoms′ charges are not possible in the Sybyl MMFF94s program. In consequence, the energy-refined structures will change according to the depiction of the compounds. The electrostatic contributions to total energies, and to minimized intermolecular contact distances, need careful consideration. When HPO_4^{2-} is present this caution needs special emphasis since polarizability and charge transfer effects also are not accommodated by the program.

To locate the binding mode of the phosphate within the closed TP structure, transition-state complexes were postulated within the active site, based on the orientation of the inhibitor and initially using 2′-deoxyribose in a 3′-endo conformation with C1′···O(P) and C1′···N1 distances of 2.1 and 2.2 Å, respectively.

Table A5.5. Structural and energetic properties (kcal;Å; degrees) of MMFF94s force-field refined models of carbenium-type TP transition states bound to an entire-structure minimized model of hTP based on human TP, PDB code: 1UOU. The most relevant alterations to side chain conformations were: Lys115 (C1 to Nζ: g^+,a,g^+,g^+) and Lys222 (C1 to Nζ: a,g^-,a,a). Deoxyribose was 3′-endo with 3′OH - O = C Arg146 and 5′OH - OγT151. TS geometries were changed by constraints on both C1′/N1 and C1′/O^P distances. Minimizations (all atoms) were performed with nonbonded cutoff > max. dimension of protein, ε = 1.3, to gradients of 0.02 kcal·mol^1·$Å^1$ (< 2 cal/iteration). All energies in the table are for complete end structures, constraints removed, with nonbonded cutoff > maximum dimension of protein, ε = 1.3.

	Distance; *angle*; *species/energy-type*	Structure			
		A5.3a	A5.3b	A5.3c	A5.3d
row					
1	**N1--O^P**	**5.56**	**5.44**	**5.13**	**4.72**
2	C1′--N1	2.89	2.83	2.63	2.41
3	C1′--O^P	2.68	2.62	2.50	2.31
4	C1′--O--P angle	126.6	129.0	131.6	135.6
5	**P.L E_t**	**−5984.6**	**−5983.7**	**−5978.6**	**−5965.1**
6	**L E_t**	**−203.3**	**−202.6**	**−199.9**	**−189.5**
7	**P E_t**	**−5530.9**	**−5530.1**	**−5529.2**	**−5525.0**
7.1	**P E_t normalized**	**0**	**0.8**	**1.7**	**5.8**
8	(P.L--L) E_t	−5781.3	−5781.0	−5778.7	−5775.5
8.1	**(P.L--L) E_t normalized**	**0**	**0.3**	**2.6**	**5.8**
8.2	(P.L--L) E_e	−12759.4	−12759.8	−12755.7	−12753.4

The 3′-endo conformation was chosen because that would produce the largest hyperconjugative difference in TS stabilization energy between 2′-deoxyribose and ribose(extremely poor substrates) derivatives. This would specially be so if the TS had significant carbenium-ion character: if that character predominates, related reactions in water indicate that, with the ribose or deoxyribose rings free to adopt optimum conformations for reaction, the presence of a 2α-hydroxy group would be expected to result in a rate retardation of two to three orders of magnitude [79;80]. That ratio would, if the enzyme-mediated forward chemical step were rate determining, be enough to account for the observed substrate-type differences.

The waters and side-chains of active site residues were remodeled to accommodate the new structure. For one set of models the half-dimer structure was minimized at $\varepsilon = 1.3$. For another set, residues with any atom within ~ 15 Å of any substrate atom were minimized at $\varepsilon = 1.5$; residues beyond the inner group were periodically allowed to relax slightly but their relative α-carbon positions remained extremely close to those in the PDB file, and the changed geometry of the main chain in the region involved in dimer formation was barely perceptible in overlaid structures. The orientations, placements and conformations of active site waters and side-chains were reviewed every 20 to 50 iterations during the initial stages of refinement. No cutoffs were applied to nonbonded interactions in the case of full-structure minimizations. For the partially minimized structure a cutoff of 35 Å was applied. Termination was at $<$ 0.02 kcal·mol^{-1}·Å^{-1} ($<$ 2 cal/iteration).

Goodness of fit analysis of hTP TS models. The common starting structure, **(a)**, Table A5.1, was the result of annealing a partially minimized model of hTP (His116 neutral with εNH; Lys115 and Lys222 conformations as in the 1UOU structure; deoxyribose in 3′-exo conformation), described in the latter part of Modification of hTP (PDB: 1UOU) structure and TS modeling. Substructures/waters having an atom within 9 Å of the 32 substructures/waters that were within 3.7 Å of substrates, were annealed with immobile substructures/waters outside the 9 Å set and having an atom within 17 Å of the 3.7 Å set ($\rightarrow$ 5722 atoms out of 6699 total). Nonbonded cutoff was at 15 Å, $\varepsilon = 1.5$; Powell minimization was performed to a gradient of 0.03 kcal·mol^{-1}·Å^{-1}. All energies in the table are for complete end structures, all 6699 atoms, constraints removed, nonbonded cutoff $>$ maximum dimension of protein, $\varepsilon = 1.5$. Each structure, **(b)** to **(f)**, was derived in turn from structure **(a)** by breaking the C1′–N1 bond, redefining the structures/atom-types/bond-types as discussed above, applying the C1′--N1distance constraint shown in Row 1.0, at 90,000 kcal·mol^{-1}·Å^{-2}, annealing the thymine-anion plus deoxyribose carbenium ion into the immobile existing site to a gradient of 0.05 kcal·mol^{-1}·Å^{-1}, then annealing that structure as described for **(a)** but for only 97 iterations. The iteration limits were used because transition states do not have time to equilibrate with their surroundings, but some immediately pre-TS movement of the protein plus substrates, favourable to TS crossing, is very likely and is to some extent allowed for by the final annealing with an arbitrary limit on the number of iterations. The results are presented in rows 1.0 – 5.3 of Table A5.1. Also in Table A5.1 are values for thymine-anion/deoxyribose-1-phosphate (column **i**) that was minimized as for (**a**), and values for both (**a**) and (**i**) wherein the covalent bond C1′-N1 or C1′-O^{P} was broken and the structures made identical to the TS structures in terms of atom types and charges, but all atoms remain where they were: data in rows 5.0–5.3 of columns (**a** ionic) and (**i** ionic) are thereby made comparable to those containing TS models.

The pattern of electrostatic to non-electrostatic energy changes is that expected for changing distances between hydrogen bonded and/or ion-paired atoms: as such bonds/interactions get shorter, the electrostatics improve while the van der Waals compression gets worse - a new equilibrium is achieved. Therefore electrostatic energy changes typically are more pronounced than the total-energy changes.

To examine the influence of the final annealing (97 iterations) a second method was applied: the protein/phosphate/waters environment generated from the partial minimization with N1–C1′ = 220pm, i.e. TS220, column **(c)**, Table A5.1, was kept immobile while thymine-anion and the carbenium-ion were minimized (to gradient 0.02 kcal·mol^{-1}·Å^{-1}, or $<$2 cal./iteration) in the site with the same or similar N1–C1′ distance constraints as previously (Table A5.2). The results are in Table A5.3 rows 6.0–6.5.

Restricting completely the movement of the protein and waters away from the TP/PO_4/thymidine starting structure **(a)**, yielded two further TS data sets (Tables A5.4 and A5.5):

1. phosphate was immobile while thymine-anion and the carbenium ion were allowed to minimize to gradient 0.02 kcal·mol^{-1}·Å^{-1}, or $<$2 cal./iteration.
2. phosphate, thymine-anion and the carbenium ion were allowed to minimize as before. The results are in Table A5.3 rows 7.0–7.4 and 8.0–8.4 respectively.

Appendix 6

Examples of transfer of protonic charge. Transfer of protonic charge occurs via hydrogen bonded-partner

switching within unbroken hydrogen bonds and transfer of protonic charge along [H-O(H) ····H-O(H) ····]$_n$ water bridges can, in effect, amount to proton transfer if waters within the bridge can rotate back to their starting orientations after the charge has been passed. The rotation process also allows groups such as R-NH_2, R-OH, Ar-OH, R-SH, and R-CO_2H to transfer protons along hydrogen bonded chains without necessarily having substantially to readjust the pathway/matrix interactions. Often however, reversion to something similar to the starting state, without reversing the charge transfer, can be energetically unfavourable: this might be the case if, in Example 1, group Z was not present. Interruption of the chain by X:····H-O-H····:Y entities, as in Example 2, or related entities, or ionized groups such as R-NH_3^+ or R-CO_2, may introduce higher energy barriers that could slow charge transfers. Example 2, with X-H being lys-NH_3^+, corresponds to the situation that occurs in the TP models; here transfer rates may be enhanced through stabilization of HO^- ions. Transfer through a histidine residue always changes the tautomeric form and that change cannot be reversed by rotating the ring system. In solution, transfer of protonic charge is expected to be non-concerted except, perhaps, when a single water is involved; therefore long-range charge transfer involves transient H_3O^+, or HO^- ions; inside a protein however, the diminished entropy of tightly-bound water molecules may allow much longer concerted pathways to compete with processes involving transient H_3O^+, or HO^- ions.

Long-range charge transfer processes along water 'wires' in proteins and membranes have been the subject of debate for decades but only very recently have theoretical methods looked likely to provide believable quantitative insights into these widespread phenomena [17].

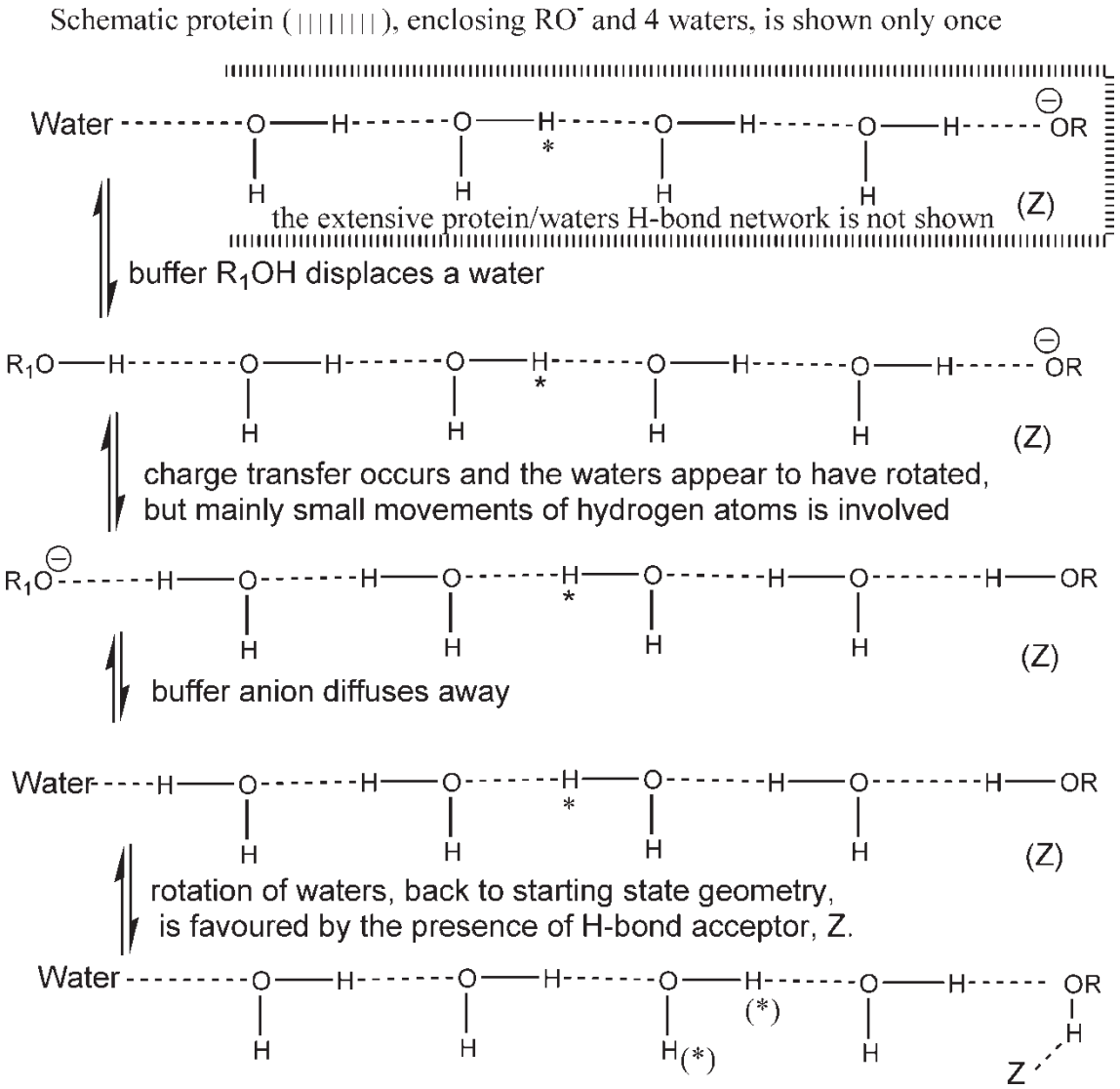


Figure A6.1. Transfer of protonic charge. Example: the simplest case.

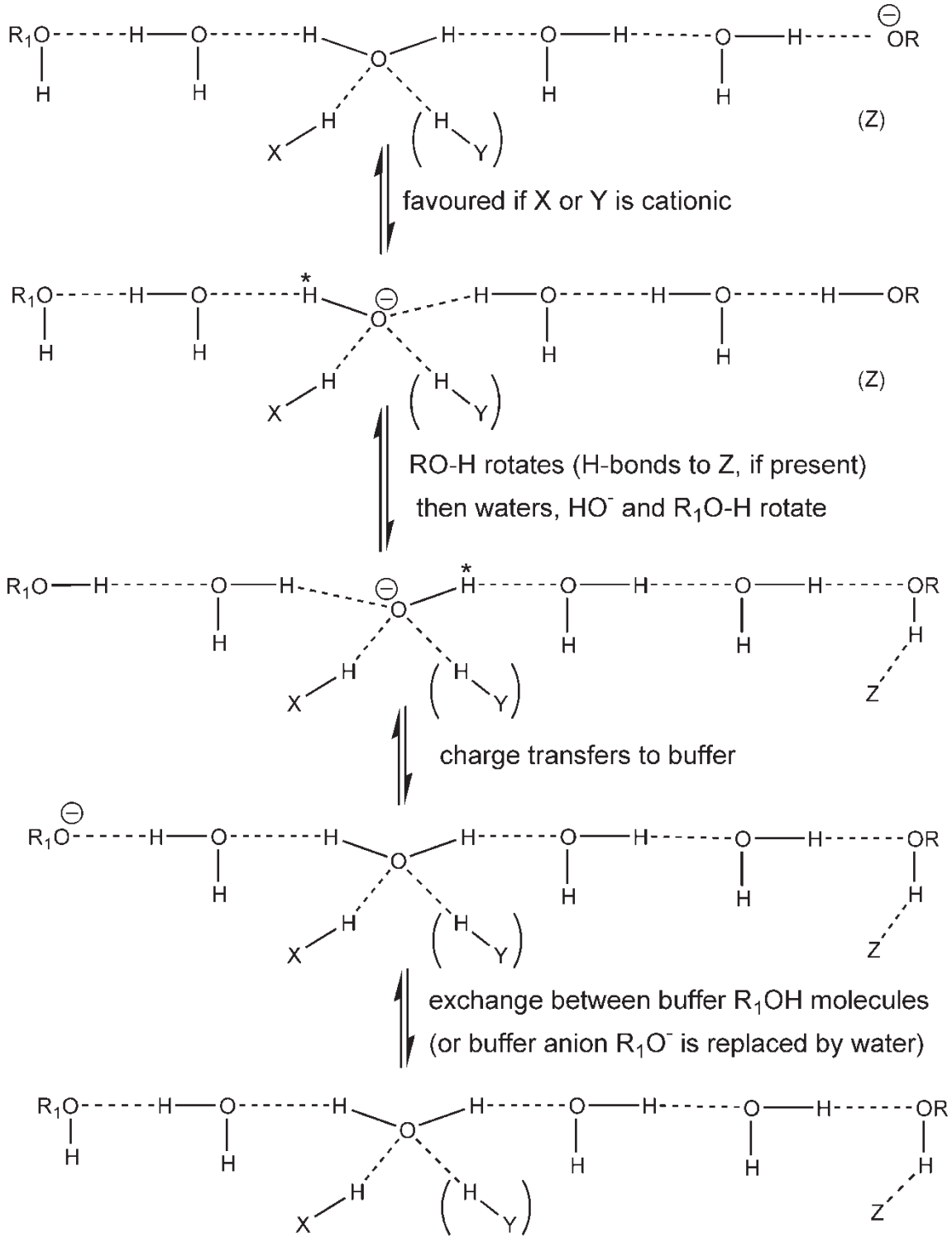


Figure A6.2. Example 2: Transfer of protonic charge: an example similar to that suggested to occur in TP.

Appendices references

[54] Plagemann PGW, Wohlheuter RM, Woffendin C. Nucleoside and nucleobase transport in animal cells. Biochim Biophys Acta 1988;947:405–443.

[55] Zimmerman TP, Domin BA, Mahony WB, Prus KL. Membrane transport of nucleoside analogues in mammalian cells. Nucleosides Nucleotides 1989;8:765–774.

[56] Jiang YL, Drohat AC, Ichikawa Y, Stivers JT. Probing the limits of electrostatic catalysis by uracil DNA glycosylase using transition state mimicry and mutagenesis. J Biol Chem 2002;277(18):15385–15392.

[57] Zimmerman M, Seidenberg J. Deoxyribosyl Transfer.I. Thymidine Phosphorylase + Nucleoside Deoxyribosyltransferase in Normal + Malignant Tissues. J Biol Chem 1964;239(8):2618–2621.

[58] Friedkin M, Roberts D. The enzymatic synthesis of nucleosides .1. Thymidine phosphorylase in mammalian tissue. J Biol Chem 1954;207(1):245–256.
[59] Friedkin M, Roberts D. The enzymatic synthesis of nucleosides .2. Thymidine and related pyrimidine nucleosides. J Biol Chem 1954;207(1):257–266.
[60] Charton M. Substituent effects in nonaromatic unsaturated systems. Prog Phys Org Chem 1973;10:81–204.
[61] Gallo RC, Breitman TR. Enzymatic mechanisms for deoxythymidine synthesis in human leukocytes .3. Inhibition of deoxythymidine phosphorylase by purines. J Biol Chem 1968;243(19):4943–4951.
[62] Baker BR, Kawazu M. Irreversible Enzyme Inhibitors .79. Inhibitors of thymidine phosphorylase.V. mode of pyrimidine binding. J Med Chem 1967;10(3):313–316.
[63] Ehrenson S, Brownlee RTC, Taft RW. Generalized treatment of substituent effects in the benzene series. Statistical analysis by the dual substituent parameter equation. I. Prog Phys Org Chem 1973;10:1–80.
[64] Long JW, Ray WJ. Kinetics and thermodynamics of formation of glucose arsenate – reaction of glucose arsenate with phosphoglucomutase. Biochemistry 1973;12(20):3932–3937.
[65] Lagunas R. Sugar-arsenate esters – thermodynamics and biochemical behavior. Arch Biochem Biophys 1980;205(1):67–75.
[66] Lagunas R, Pestana D, Diezmasa JC. Arsenic mononucleotides – separation by high-performance liquid-chromatography and identification with myokinase and adenylate deaminase. Biochemistry 1984;23(5):955–960.
[67] Kouba RF, Varner JE. Properties of the Arsenate-Water O-18 exchange reaction. Biochem Biophys Res Commun 1959;1(3):129–132.
[68] Gallo RC, Perry S, Breitman TR. Enzymatic mechanisms for deoxythymidine synthesis in human leukocytes. I. Substrate inhibition by thymine and activation by phosphate or arsenate. J Biol Chem 1967;242(21):5059–5068.
[69] Krenitsky TA. Purine nucleoside phosphorylase: Kinetics, mechanism, and specificity. Mol Pharmacol 1967; 3(6):526–536.
[70] Beckmann C, Kirby AJ, Kuusela S, Tickle DC. Mechanisms of catalysis by imidazole buffers of the hydrolysis and isomerization of RNA models. J Chem Soc, Perkin Trans 1998;2(3):573–582.
[71] Davies JE, Doltsinis NL, Kirby AJ, Roussev CD, Sprik M. Estimating pK_a values for pentaoxyphosphoranes. J Am Chem Soc 2002;124(23):6594–6599.
[72] Ali BRS, Dixon HBF. Pyridoxal arsenate as a prosthetic group for aspartate-aminotransferase. Biochem J 1992; 284:349–352.
[73] Gawlita E, Lantz M, Paneth P, Bell AF, Tonge PJ, Anderson VE. H-bonding in alcohols is reflected in the C alpha-H bond strength: Variation of C-D vibrational frequency and fractionation factor. J Am Chem Soc 2000; 122(47):11660–11669.
[74] Lewis BE, Schramm VL. Conformational equilibrium isotope effects in glucose by C-13 NMR spectroscopy and computational studies. J Amer Chem Soc 2001;123(7):1327–1336.
[75] Lewis BE, Schramm VL. Binding equilibrium isotope effects for glucose at the catalytic domain of human brain hexokinase. J Amer Chem Soc 2003;125(16):4785–4798.
[76] Singh V, Lee JE, Nunez S, Howell PL, Schramm VL. Transition state structure of 5′-methylthioadenosine / S-adenosylhomocysteine nucleosidase from *Escherichia coli* and its similarity to transition state analogues. Biochemistry 2005;44(35):11647–11659.
[77] Guthrie RD, Jencks WP. IUPAC Recommendations for the representation of reaction mechanisms. Acc Chem Res 1989;22(10):343–349.
[78] Rezaei M, Kearny M, Kline PC. Kinetic isotope effects of thymidine phosphorylase, American Chemical Society Book of Abstracts, 219th ACS National Meeting, San Francisco, CA, March 26–30, 2000, Biol-041. 2000.
[79] Venner H. Studies on nucleic acids. IX. Stability of the N-glycosidic linkage in nucleosides. Hoppe-Seyler's Z Physiol Chem 1964;339(1):14–27.
[80] Huang XC, Surry C, Hiebert T, Bennet AJ. Hydrolysis of (2-Deoxy-β-D-Glucopyranosyl)Pyridinium Salts. J Am Chem Soc 1995;117(43):10614–10621.