SUPPORTING INFORMATION

Mechanistic Characterization of the HDV Genomic Ribozyme: Solvent Isotope Effects and Proton Inventories in the Absence of Divalent Metal Ions Support C75 as the General Acid

Andrea L. Cerrone-Szakal,[1,2] Nathan A. Siegfried,[1] and Philip C. Bevilacqua[1,]*

[1]Department of Chemistry, The Pennsylvania State University, University Park, PA 16802

[2]Present Address: Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute and the National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, MD 20850, USA.

Contents:

S.1: Derivation of the Population-Weighted Gross-Butler Equation

S.2: Limits of the Population-Weighted Gross-Butler Equation

S.1: Derivation of the Population-Weighted Gross-Butler Equation

Equation 9 is a population-weighted form of the Gross-Butler equation, which was used in the main text to describe the proton inventories performed in the high-pL plateau of the rate-pL plots.

$$k_n / k_0 = (1-\mathrm{n})^2 + \phi^{\mathrm{T}}_{\mathrm{C75}}\mathrm{n}(1-\mathrm{n})10^{\Delta \mathrm{p}K_{\mathrm{C75}}} + \phi^{\mathrm{E}}_{\mathrm{OL-}}\mathrm{n}(1-\mathrm{n})10^{-\Delta \mathrm{p}K_{\mathrm{LOL}}} + \phi^{\mathrm{T}}_{\mathrm{C75}}\phi^{\mathrm{E}}_{\mathrm{OL-}}\mathrm{n}^2 10^{\Delta \mathrm{p}K_{\mathrm{C75}}-\Delta \mathrm{p}K_{\mathrm{LOL}}} \qquad (9)$$

Following is a derivation of this equation. We begin by noting that in H_2O-D_2O mixtures at high pL, ribozymes exist with general acid and specific base species in one of four isotopic combinations: HH, DH, HD, and DD, respectively. The reactivity of each isotopic combination can be described as a product of the intrinsic rate constant and the fractional population of the general acid and specific base.[1] The intrinsic rate constant at high pL is the product of $k_{\mathrm{OH-}}$ and the appropriate factor(s), ϕ^{T} and ϕ^{E} for $C75D^+$ and OD^-, respectively.[a] The fractional populations of the general acid and specific base are given by the product of four factors: the fraction of general acid with H or D as appropriate and as determined from its equilibrium constant ($f_{\mathrm{C75L+}}$); the fraction of H_2O or D_2O, as appropriate and determined by factors of 1-n and n,[b] respectively; and two similar terms for the specific base. This leads to the following four-term expression for the observed rate constant, where the four terms on the right-hand side of the equation represent the general acid and specific base in the forms HH, DH, HD, and DD, respectively:

$$\begin{aligned} k_n^{\mathrm{obs}} &= k_{\mathrm{OH-}}(1-\mathrm{n})^2 f_{\mathrm{C75H+}} f_{\mathrm{OH-}} + k_{\mathrm{OH-}}\phi^{\mathrm{T}}_{\mathrm{C75}}\mathrm{n}(1-\mathrm{n}) f_{\mathrm{C75D+}} f_{\mathrm{OH-}} \\ &+ k_{\mathrm{OH-}}\phi^{\mathrm{E}}_{\mathrm{OL-}}\mathrm{n}(1-\mathrm{n}) f_{\mathrm{C75H+}} f_{\mathrm{OD-}} + k_{\mathrm{OH-}}\phi^{\mathrm{T}}_{\mathrm{C75}}\phi^{\mathrm{E}}_{\mathrm{OL-}}\mathrm{n}^2 f_{\mathrm{C75D+}} f_{\mathrm{OD-}} \end{aligned} \qquad (15)^{c}$$

---

[a] As described in the main text, the contribution of $C75D^+$ is most consistent with a transition state fractionation factor (ϕ^{T}), while the contribution of OD^- is most consistent with an equilibrium isotope effect (ϕ^{E}).

[b] n is the mole fraction of D_2O in solution.

[c] Numbering of equations picks up where it left off in the main text.

Note that the contributions of ϕ^{R} (reactant state fractionation factors) are assumed to be unity, except for lyoxide, as is customary.[2] The fraction of the ribozyme having C75L$^+$ at high pL is approximated by $f_{C75L+} = \frac{1}{10^{pL-pK_{R,L}}}$, while the fraction of lyoxide in solution is well approximated by $f_{OL-} = 10^{pL-pK_{LOL}}/55M$[d]

Substituting and factoring gives

$$k_n^{obs} = k_{OH-}\left[\begin{array}{l}(1-n)^2\frac{10^{pH-pK_{HOH}}/55M}{10^{pH-pK_{R,H}}} + \phi_{C75}^{T}n(1-n)\frac{10^{pH-pK_{HOH}}/55M}{10^{pD-pK_{R,D}}} \\ +\phi_{OL-}^{E}n(1-n)\frac{10^{pD-pK_{DOD}}/55M}{10^{pH-pK_{R,H}}} + \phi_{C75}^{T}\phi_{OL-}^{E}n^2\frac{10^{pD-pK_{DOD}}/55M}{10^{pD-pK_{R,D}}}\end{array}\right] \tag{16}$$

The rate constant in pure HOH, which will be used as the reference rate constant, can be obtained by substituting n=0 into eq 16.

$$k_0^{obs} = k_{OH-}10^{pK_{R,H}-pK_{HOH}}/55M \tag{17}$$

Dividing eq 16 by eq 17 gives eq 9 in the main text:[e]

$$k_n^{obs}/k_0^{obs} = (1-n)^2 + \phi_{C75}^{T}n(1-n)10^{\Delta pK_{C75}} + \phi_{OL-}^{E}n(1-n)10^{-\Delta pK_{LOL}} + \phi_{C75}^{T}\phi_{OL-}^{E}n^2 10^{\Delta pK_{C75}-\Delta pK_{LOL}} \tag{9}$$

The ΔpK terms in the exponents of the last three terms on the right-hand side of the equation are for ionization of C75 or autoprotolysis of solvent, and each is equal to the value in DOD minus the value in HOH. Specifically, ΔpK_{C75} is 0.44 (=7.92-7.48) (see Results), while ΔpK_{LOL} is 0.85 (=14.23-13.38).

We note that it is possible to subdivide eq 9 further by taking into consideration the origin of OH$^-$ in HOH or DOH, and the origin of OD$^-$ in HOD or DOD. To a first approximation, the autoprotolysis constant for DOH is intermediate between the constants for DOD and HOH.[3] Simulation of this eight term equation under the four limiting conditions shown in Figures 5B

[d] Here we are taking the fraction of LOL in the OL$^-$ form.

[e] The experiment was conducted by mixing pH 9 and pD 9 buffers, so that pH and pD cancel to a first approximation.

and 5C gave behavior that is similar (not shown) to the simpler four-term eq 9. As such, eq 9 was used in all fitting and simulation.

S.2: Limits of the Population-Weighted Gross-Butler Equation

We consider three limits of eq 9:

Limit 1: If the acid and base species are fully functional, as 'expected' in the plateau region, (*i.e.* $f_{C75L+} = f_{OL-} = 1$) and two protons are transferred, then the base 10 exponential terms in eq 9 go to unity and eq 9 reduces to:

$$k_n^{obs}/k_0^{obs} = (1-n)^2 + \phi_{C75}^{T} n(1-n) + \phi_{OL-}^{E} n(1-n) + \phi_{C75}^{T}\phi_{OL-}^{E} n^2 \qquad (18)$$

This equation can be factored to give

$$k_n^{obs}/k_0^{obs} = (1-n+n\phi_{OL-}^{E})(1-n+n\phi_{C75}^{T}) \qquad (19)$$

which, as expected, is the Gross-Butler equation for two non-equivalent proton transfers (eq 7).

Limit 2: If the acid and base species are fully functional *and* only one proton is transferred (*e.g.* $\phi_{OL-}^{E} = 1$), then eq 18 reduces to:

$k_n^{obs}/k_0^{obs} = (1-n)^2 + \phi_{C75}^{T} n(1-n) + n(1-n) + \phi_{C75}^{T} n^2$, which can be simplified to give

$$k_n^{obs}/k_0^{obs} = 1-n+n\phi_{C75}^{T} \qquad (20)$$

As expected, this is the Gross-Butler equation for one proton transfer (eq 8, m=1).

Limit 3: If there is no intrinsic isotope effect (*i.e.* $\phi_{C75}^{T} = \phi_{OL-}^{E} = 1$), but f_{C75L+} and f_{OL-} are less than unity then eq 9 reduces to:

$$k_n^{obs}/k_0^{obs} = (1-n)^2 + n(1-n)10^{\Delta pK_{C75}} + n(1-n)10^{-\Delta pK_{LOL}} + n^2 10^{\Delta pK_{C75}-\Delta pK_{LOL}} \qquad (21)$$

This equation is used in some of the simulations to see if populations changes-only describe the proton inventories.

References

1. Bevilacqua, P. C. *Biochemistry* **2003**, *42*, 2259-2265.
2. Quinn, D. M.; Sutton, L. D., Theoretical Basis and Mechanistic Utility of Solvent Isotope Effects. In Enzyme Mechanism from Isotope Effects, Cook, P. F., Ed. CRC Press, Inc.: Boca Raton, FL, 1991; pp 73-126.
3. Schowen, K. B.; Schowen, R. L. *Methods Enzymol*. **1982**, *87*, 551-606.