suggested space group $P2_12_12_1$. Integrated intensities of 396 reflections were extracted with use of the program EXTRA [A. Altomare *et al.*, *J. Appl. Crystallogr.* **28**, 842 (1995)]. Direct methods in the program SIRPOW [A. Altomare *et al.*, *ibid.* **27**, 435 (1994)] were used to obtain a set of metal coordinates. These coordinates could be approximately related to those of the cubic phase, and initial oxygen positions were inferred from the low-pressure structure. Combined x-ray and neutron Rietveld refinement, initially with metal-oxygen distances heavily restrained to chemically sensible values to prevent divergence, followed by manual shifts of certain O atoms led eventually to a chemically sensible model, in excellent agreement with both x-ray and neutron data. Refinements were performed in the GSAS suite of programs [A. C. Larson and R. B. Von Dreele; Los Alamos National Laboratory (1994)].

10. For final refinement, higher quality diffraction data were used. The x-ray data was collected from 5° to 100° (d_{min} = 1.005 Å) with a step size of 0.02° and a counting time of 45 s per step (4522 data points). Neutron data from SEPD Bank 1, with a time of flight of 6 to 29.5 ms (0.803 to 3.95 Å) and 4699 data points, were used. Because temperature factors refined from powder data are subject to a number of systematic errors, temperature factors were initially set at 0.01 Å^2 for metal atoms and 0.015 Å^2 for O atoms. An absorption correction for neutron data and a surface roughness correction [P. Suortti, *J. Appl. Crystallogr.* **5**, 325 (1972)] for x-ray data were applied. Absorption parameters were then fixed, and equated isotropic temperature factors on individual elements were allowed to refine. On subsequent free refinement of temperature factors, individual values remained within acceptable limits. Impurities of 1.1% WO_3 and 0.8% ZrO_2 were included in the refinement as additional phases.
11. All bond distances and angles lie within normally observed ranges. Bond valences [(*18*); I. D. Brown and D. Altermatt, *Acta Crystallogr. B* **41**, 244 (1985); N. E. Brese and M. O'Keefe, *ibid.* **47**, 192 (1991)] for metal atoms in the structure are chemically reasonable: Zr1, 4.3; Zr2, 4.3; Zr3, 4.4; W1, 6.1; W2, 6.5; W3, 5.6; W4, 5.9; W5, 5.6; and W6, 5.8. The O valences range from 1.8 to 2.2 with only "terminal" oxygens O104 and O105 having slightly low calculated valences (1.6 and 1.7, respectively).
12. There is some question as to what constitutes a true W-O bond in the strictest sense and what is a weaker W···O interaction. We choose here to define a W-O bond as one less than 2.3 Å. Using the valence method of Brown and Wu (*18*), this definition corresponds to interactions contributing greater than 6% of the total valence sum of W being considered as full bonds. A coordination of 4+1 is intended to imply four short (<2.3 Å) and one longer (2.3 to 2.6 Å) interaction.
13. For both cubic and orthorhombic forms, α_l (*2*) is defined as $\frac{1}{3}\alpha_V$, where $\alpha_V = (V_{T2} - V_{T1})/[(T_2 - T_1)V_{T1}]$.
14. A. P. Giddy, M. T. Dove, G. S. Pawley, V. Heine, *Acta Crystallogr. A* **49**, 697 (1993); I. P. Swainson and M. T. Dove, *Phys. Chem. Miner.* **22**, 61 (1995); K. D. Hammonds, M. T. Dove, A. P. Giddy, V. Heine, B. Winkler, *Am. Mineral.* **81**, 1057 (1996); and references therein.
15. Compressibility, defined as $\beta = -(1/V)(dV/dP)$. Linear regression of six cell parameters between 0 and 6 kbar yielded linear compressibilities $-(l_1/l_2)(dl/dP)$ of 0.53 × 10^{-3}, 0.47 × 10^{-3}, and 0.47 × 10^{-3} kbar^{-1} for *a*, *b*, and *c*, respectively.
16. J.-E. Jorgensen, J. D. Jorgensen, B. Batlogg, J. P. Remeika, J. D. Axe, *Phys. Rev. B* **33**, 4793 (1986).
17. Rietveld refinement of high-pressure data was performed with the use of Bank 2 data of SEPD. Time of flight ranged from 4.5 to 26 ms (d = 0.806 to 4.56 Å), yielding 3071 data points. Data at 0, 1.0, 3.1, 5.2, and 6.2 kbar were refined to χ^2/wRp values of 1.1/6.5, 1.5/4.0, 1.4/3.8, 1.5/3.9, and 1.4/3.8%, respectively. Precise determination of individual bond distances and angles for a structure this complex (33 atoms in an asymmetric unit) is difficult given the lower resolution of the data obtained in the high-pressure cell. Average distances: d_{avg}(Zr-O) = 2.069 − (9 × 10^{-4} × P); d_{avg}(W-O) = 1.809 − (2 × 10^{-4} × P); d_{avg}(Zr-W) = 3.80 − (2 × 10^{-4} × P); d_{avg}(W-W) = 3.904 − (3 × 10^{-3} × P) (pressure in kilobars yields distances in angstroms). Average Zr-O-W angle: 160.3 − (0.14 × P).
18. I. D. Brown and K. K. Wu, *Acta Crystallogr. B* **32**, 1957 (1976).
19. The work at Oregon State was supported by NSF (DMR-9308530) and the Oregon Metals Initiative. The work at Argonne National Laboratory was supported by the U.S. Department of Energy [Office of Basic Energy Sciences, Materials Sciences contract U-31-109-ENG-38 (J.D.J., Z.H., S.S.)] and NSF [Office of Science and Technology Centers grant DMR-91-20000 (D.N.A.)].

28 August 1996; accepted 12 November 1996

Disease Extinction and Community Size: Modeling the Persistence of Measles

M. J. Keeling and B. T. Grenfell

A basic issue in ecology is the relation between extinction and population size. One of the clearest manifestations of a population threshold for extinction is the critical community size below which infections like measles do not persist. The current generation of stochastic models overestimates the observed critical community size for measles, generating much less persistence of infection than is observed. The inclusion of a more biologically realistic model for the duration of infection produced a much closer fit to the actual critical community size and explains previously undescribed high-frequency oscillations in measles incidence.

The relation between disease persistence and community size can be explored through the pattern of fadeouts of infection [three or more weeks without reported cases (*1–5*)]. The observed critical community size (CCS) for measles is about 250,000 to 400,000 (Fig. 1A). These figures are based on a large prevaccination data set for 60 towns in England and Wales for the years 1944 to 1968, but they are also typical of the pattern observed for U.S. cities (*2*) and islands (*3*). Below the CCS, the infection often becomes extinct in the troughs between epidemics and must be reintroduced from an external source.

Fadeout pattern predictions from the best current nonspatial stochastic model (*6–13*) significantly overestimate the CCS (Fig. 1A), generating many more fadeouts than observed for towns with populations over 250,000. This discrepancy is even more marked in a comparison of the observed and expected total weeks of fadeout per year (Fig. 1B). Although this standard model [the realistic age-structured (RAS) model (*10–14*)] captures the deterministic dynamics of measles epidemics very well (*5, 10, 13*), its stochastic dynamics are unstable in populations below about 1 million, generating many more fadeouts than observed (*4–6*). A number of authors have sought an explanation for this failure of current models in terms of spatial heterogeneities in transmission, on both large spatial scales (*4, 15–17*) and at the individual (family and school) level (*16*). Although inclusion of spatial heterogeneities reduces the predicted degree of fadeout, even quite complicated spatial models cannot currently capture the low observed level of the CCS (*4, 15–25*).

The fact that the CCS applies to a wide range of communities, from cities to islands, indicates that we should seek a more generic explanation, one rooted in the biology of transmission. We propose that the problem arises because current models are too sensitive to stochastic fluctuations, which arise from the use of long-tailed exponential distributions for the incubation and infectious periods. This exponential formulation arises from the standard assumption that movement from the exposed to the infectious class and then into the recovered class occurs at constant rates a and g, respectively (*6*).

We can modify the standard model to allow for these effects by assuming normal distributions for the incubation and infectious periods (*19*). More complex distributions could be used (*6*), but evidence from the detailed study of transmission in families (*20*) indicates that the periods show limited variation, and the data are well described by infectious periods normally distributed about their means. The use of more discrete periods has been considered previously (*21*) but seldom in this context of seasonally forced stochastic models. The revised stochastic model tends to produce more concentrated pulses of infection, so we will call it the pulsed realistic age-structured (PRAS) model.

Support for the new model is provided by a comparison of Fourier spectra for simulated epidemic time series from the standard (RAS) and modified (PRAS) models with the observed pattern for England and

Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK.

Wales (Fig. 2). The observed spectrum (Fig. 2A) shows the well-known peaks associated with biennial and annual (seasonal) epidemics (*4*, 6, 8, 9). In addition, the analysis reveals a previously unreported high-frequency peak, reflecting low-amplitude epidemics with periods of 2 to 3 weeks. Both models capture the annual and biennial peaks accurately (Fig. 2B) (6, *10*, *13*); however, the standard RAS model cannot reproduce the high-frequency pulses, whereas the PRAS model does reflect variation at this time scale. Essentially, the more uniform infectious period of the PRAS model produces "generation pulses" of infection (*22*), which are apparent in the real data. These pulses are especially apparent in the data for small communities.

The impact of this uniformity of infection on the predicted CCS is shown in Fig. 1. The match to the observed number and total duration of fadeouts is much closer than with the standard RAS model. This correspondence occurs essentially because the normally distributed infection periods of the PRAS model cause infectiousness to be more evenly distributed among individuals, and the model is therefore much less vulnerable to stochastic fadeout. The mechanism can be illustrated with a simple analytical model (Fig. 3). Figure 3A shows the difference between the assumptions about the duration of infection made for the two models: The PRAS model, with its more discrete infectious period, has the majority of infections lasting between 4 and 6 days, whereas the standard RAS model assumes a lower proportion of people infected at the early stages, with a long exponential tail generating an appreciable probability of infection after 20 days. The associated amount of variation in the infectious periods is reflected in the variance of the basic reproduction ratio of infection R_0 (6). The parameter Var(R_0) represents the stochasticity in the spread of infection and is therefore related to the probability of zero transmission (Fig. 3B), which is much higher with the exponential distribution produced by the standard model (*23*). This difference occurs because the RAS model relies on a few infectious individuals who retain the disease for a long time to spread the infection. Therefore, the RAS model has far more individuals who do not produce any secondary cases, as compared to the PRAS model (Fig. 3B). The greater persistence of the pulsed model is unaffected if we equate the average generation gap of the infection [serial interval between cases (*24*)] rather than the infectious period.

All of these patterns of infection were generated using isolated populations with a low stochastic influx of infectives to reintroduce infection (*5*, *9*, *13*); however, changing this influx level within realistic limits does not qualitatively alter the improvement made by the PRAS model. Of course, explicit spatial heterogeneity will be

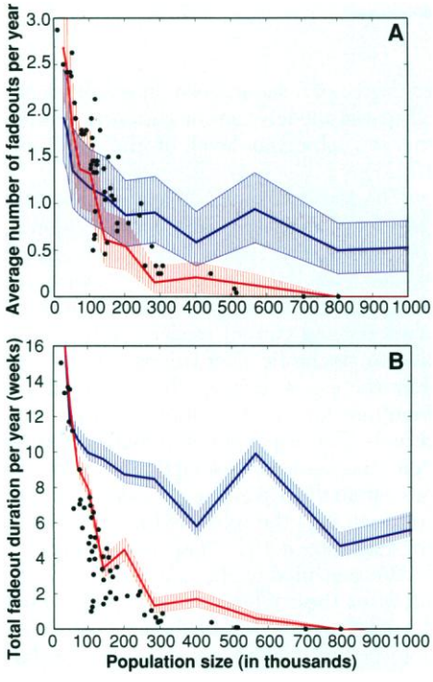

Fig. 1. Comparison of fadeout in real data with predictions from the RAS and PRAS models. Solid black dots represent data from 60 towns in England and Wales for the prevaccination era (1944 to 1968); blue gives results for the standard RAS model; and red, for the new PRAS model. (**A**) Average number of fadeouts (three or more consecutive weeks without case notifications) per year. (**B**) Total number of weeks per year that are part of a fadeout. The solid lines in (B) are the mean fadeout results from a 100-year simulation of the stochastic RAS and PRAS models, with 10 infected imports per year. The shaded regions represent the 95% confidence limits. From the simulations we can calculate the probability of a week being part of a fadeout; these probabilities were then used to find the confidence limits, assuming a 24-year sample size (corresponding to 1944 to 1968) rather than 100 years.

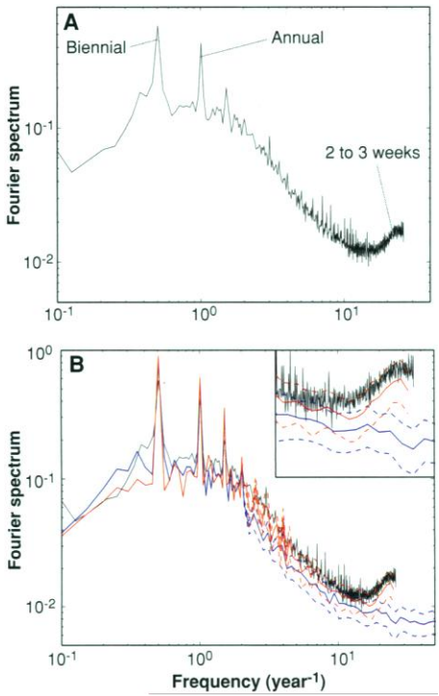

Fig. 2. The average Fourier spectra from the England and Wales data and from simulations. Before the Fourier spectrum was taken, each series was normalized, setting the mean to zero and the variance to unity; therefore, when the average was taken, all simulated and observed data contributed equally, irrespective of the population size. (**A**) Average for the 60 towns. In addition to the strong annual and biennial peaks, there is a marked increase in power at high frequencies. (**B**) Results from 10 stochastic simulations with a population size of 50,000. The solid lines are a smoothed average of the spectra for the two models (the red line is for the PRAS model, blue for the RAS model) and the dashed lines show the expected standard deviation. (Inset) Close-up of the high-frequency end of the spectra, where the improvements from the revised model are clear at around 2 to 3 weeks.

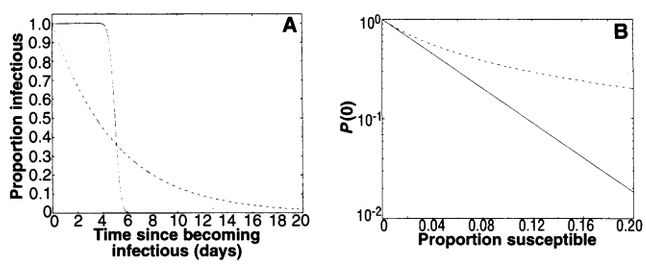

Fig. 3. Comparison of infectiousness in the RAS and PRAS models. The solid lines represent calculations for the normally distributed infectious times of the PRAS model, and the dashed lines are for the exponentially distributed times of the RAS model. The models have the same basic reproduction ratio of infection R_0, which is proportional to the area under the curves. (**A**) The expected proportion of individuals still infectious as a function of time since entering the infectious class. If the time spent in the exposed class is also incorporated into the graph, then for the PRAS model, infectious individuals exist as a distinct pulse, whereas the RAS model exhibits an even slower decay than shown. (**B**) The probability of an individual not causing any secondary infections [$\mathbb{P}(0)$] as a function of the proportion susceptible for the two models. The values were computed taking $\beta = 4$ and $\mathbb{E}(P_i) = 5$ days and demonstrate that the differences between the two models are most pronounced when there is a large number of susceptibles in the population, although a sizable difference is seen for all realistic densities of susceptibles.

required to explain fully the persistence of the disease in a metapopulation where there is no influx from an external source (*5*).

In summary, inclusion of a more realistic infection period in childhood disease models generates the high-frequency pulsing seen in the real data and produces more realistic levels of persistence, as reflected in the lower CCS. This improved fit is likely to be a generic result for infections that occur as self-extinguishing epidemics. More generally, this well-documented example underlines (*25*) the idea that the assumption of constant transition rates (and therefore exponentially distributed times), which is often made in ecology, may need to be reevaluated if we are to fully understand patterns of stochastic fluctuations and extinctions.

REFERENCES AND NOTES

1. M. S. Bartlett, *J. R. Stat. Soc. Ser. A* **120**, 48 (1957); B. M. Bolker and B. T. Grenfell, *Proc. R. Soc. London Ser. B* **251**, 75 (1993).
2. M. S. Bartlett, *J. R. Stat. Soc. Ser. A* **123**, 37 (1960).
3. F. L. Black, *J. Theor. Biol.* **11**, 207 (1966); A. P. Cliff, P. Haggett, J. K. Ord, G. R. Versey, *Spatial Diffusion: An Historical Geography of Epidemics in an Island Community* (Cambridge Univ. Press, Cambridge, 1981).
4. B. T. Grenfell, *J. R. Stat. Soc. Ser. B* **54**, 383 (1992).
5. ———, B. M. Bolker, A. Kleczkowski, *Proc. R. Soc. London Ser. B* **259**, 97 (1995).
6. R. M. Anderson and R. M. May, *Infectious Diseases of Humans* (Oxford Univ. Press, Oxford, 1992).
7. H. E. Soper, *J. R. Stat. Soc. Ser. A* **92**, 34 (1929).
8. P. E. M. Fine and J. A. Clarkson, *Int. J. Epidemiol.* **11**, 5 (1982); J. L. Aron and I. B. Schwartz, *J. Theor. Biol.* **110**, 665 (1984).
9. L. F. Olsen and W. M. Schaffer, *Science* **249**, 499 (1990).
10. D. Schenzle, *IMA J. Math. Appl. Med. Biol.* **1**, 169 (1984).
11. B. T. Grenfell, A. Kleczkowski, S. P. Ellner, B. M. Bolker, *Philos. Trans. R. Soc. London Ser. A* **348**, 515 (1994).
12. B. T. Grenfell, A. Kleczkowski, C. A. Gilligan, B. M. Bolker, *Stat. Methods Med. Res.* **4**, 160 (1995).
13. B. M. Bolker, *IMA J. Math. Appl. Med. Biol.* **10**, 83 (1993).
14. The standard RAS formulation categorizes individuals into four distinct age groups: under 6 (preschool), 6 to 10 (primary school), 10 to 20 (adolescents), and over 20 (adults). Within each class, the population is further subdivided into susceptible (S), exposed (E), infectious (I), and recovered (R) groups. There is a constant birth rate into the youngest age group, but movement between the other age classes occurs annually at the start of the school year. The transmission between and within each age class occurs at different rates and is controlled by the WAIFW (who acquires infection from whom) matrix β (*6, 10, 13*).

$$\beta = \begin{pmatrix} \beta_1 & \beta_1 & \beta_3 & \beta_4 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \beta_3 & \beta_3 & \beta_3 & \beta_4 \\ \beta_4 & \beta_4 & \beta_4 & \beta_4 \end{pmatrix}$$

To account for the school year, which plays a very important role (*8, 9*), the value of β_2 is decreased during school holidays. The WAIFW matrix (for both the RAS and PRAS models) is estimated by obtaining the best fit to the average biennial cycle from the England and Wales data. Full details are given in (*12, 13*).
15. R. M. Anderson and R. M. May, *IMA J. Math. Appl. Med. Biol.* **1**, 233 (1984); R. M. May and R. M. Anderson, *Math. Biosci.* **72**, 83 (1984); B. M. Bolker and B. T. Grenfell, *Philos. Trans. R. Soc. London Ser. B* **348**, 309 (1995).
16. M. J. Keeling and D. A. Rand, preprint (1996).
17. A. M. Lloyd and R. M. May, *J. Theor. Biol.* **179**, 1 (1996).
18. R. Engbert and F. R. Drepper, *Chaos Solitons Fractals* **4**, 1147 (1994); N. M. Ferguson, preprint (1996).
19. If we let P_E and P_I be the probability distribution functions of times in the two classes, then (ignoring mortality in the exposed and infectious classes) the standard SEIR model can be replaced by two time-delayed differential equations

$$dS/dt = m(N - S) - \gamma$$

$$\frac{d\gamma}{dt} = \frac{\gamma}{\beta}\frac{d\beta}{dt} + \frac{\gamma}{S}\frac{dS}{dt} + \beta S \int_0^\infty \gamma(t - \tau)A(\tau)d\tau \quad (1)$$

where

$$A(\tau) = P_E(\tau) - \int_0^\tau P_I(T)P_E(\tau - T)dT$$

$\gamma = \beta SI$, m is the birth and death rate, and N is the population size. Although Eqs. 1 are useful from a mathematical perspective (*23*), for computational simplicity we further subdivide the E and I classes, so that individuals move into a new subclass after short time intervals. Deterministic simulations using these two methods give identical results.
20. R. E. Hope Simpson, *Lancet* **ii**, 549 (1952); N. J. T. Bailey, *Biometrika* **43**, 15 (1956).
21. N. J. T. Bailey, *The Mathematical Theory of Infectious Diseases* (Griffin, London, 1975); F. C. Hoppensteadt, *J. Franklin Inst.* **297**, 325 (1974); Z. Grossman, *Theor. Popul. Biol.* **18**, 204 (1980).
22. M. Burnet and D. O. White, *Natural History of Infectious Diseases* (Cambridge Univ. Press, Cambridge, ed. 4, 1972).
23. For both the standard and revised modes, the average number of secondary infections caused by an infectious individual [$\mathbb{E}(R_0)$] are the same, as is the average infectious period $\mathbb{E}(P_I)$. However when we consider the general form for the variance in R_0

$$\mathrm{Var}(R_0) = \mathbb{E}(R_0) + \mathbb{E}(R_0)^2 \frac{\mathrm{Var}(P_I)}{\mathbb{E}(P_I)^2}$$

the standard RAS model with its exponential form for P_I (Fig. 3A) generates far greater variability in R_0. This in turn leads to greater stochasticity and more extinctions, which can be highlighted by examining $\mathbb{P}(0)$, the probability that an infectious individual will not produce any secondary cases (Fig. 3B)

$$\mathbb{P}_{RAS}(0) = \frac{1}{1 + \beta S\overline{P_I}}, \qquad \mathbb{P}_{PRAS}(0) = e^{-\beta S\mathbb{E}(P_I)}$$

24. D. Mollison, *Math Biosci.* **107**, 255 (1991).
25. ———, *Nature* **310**, 224 (1984).

26. We thank B. Bolker, S. Ellner, P. Harvey, and A. Kleczkowski for stimulating discussions. This research was supported by the Wellcome Trust and the Royal Society.

25 June 1996; accepted 15 October 1996

How Thiamine Diphosphate Is Activated in Enzymes

Dorothee Kern,*† Gunther Kern,‡ Holger Neef, Kai Tittmann, Margrit Killenberg-Jabs, Christer Wikner, Gunter Schneider, Gerhard Hübner

The controversial question of how thiamine diphosphate, the biologically active form of vitamin B_1, is activated in different enzymes has been addressed. Activation of the coenzyme was studied by measuring thermodynamics and kinetics of deprotonation at the carbon in the 2-position (C2) of thiamine diphosphate in the enzymes pyruvate decarboxylase and transketolase by use of nuclear magnetic resonance spectroscopy, proton/deuterium exchange, coenzyme analogs, and site-specific mutant enzymes. Interaction of a glutamate with the nitrogen in the 1′-position in the pyrimidine ring activated the 4′-amino group to act as an efficient proton acceptor for the C2 proton. The protein component accelerated the deprotonation of the C2 atom by several orders of magnitude, beyond the rate of the overall enzyme reaction. Therefore, the earlier proposed concerted mechanism or stabilization of a C2 carbanion can be excluded.

Coenzymes exert their catalytic activity after binding to a specific protein component. Therefore, it is crucial to understand how the reactivity of distinct groups of coenzymes is increased by interaction with the protein.

The coenzyme thiamine diphosphate (ThDP; Fig. 1), the biologically active derivative of vitamin B_1, is used by different enzymes that perform a wide range of catalytic functions. These include decarboxylation of 2-oxo acids and transketolation. Although the free coenzyme can assist some of these reactions, the protein environment potently accelerates the overall enzyme reaction by up to a factor of 10^{12}, as determined for pyruvate decarboxylase (PDC; E.C. 4.1.1.1) (*1*). The reactive C2 atom, located between the nitrogen and sulfur in the thiazolium ring, is the nucleophile that attacks the carbonyl carbon of the different substrates (*2*). For this reaction, the C2-ThDP atom must be activated by the enzyme environment. The deprotonation of C2 (Fig. 1) is the key reaction, because (i) this initial reaction is the only common step for all ThDP enzymes and (ii) the rate constant for this C-H dissociation is far too small in the free coenzyme compared with that of the entire enzyme reaction (*3*). The C2-ThDP activation in enzymes has been discussed for decades. In an early model, stabilization of ThDP C2-