Supporting Information

Schopf et al. 10.1073/pnas.1503828112

SI Text

S1. Background

As explained in detail elsewhere (e.g., ref. 1), the most meaningful reference reaction (as much as the effect of the enzyme environment is concerned) is a solution reaction that involves the same mechanism as the one used by the enzyme (namely the reaction described in Fig. 2). What is known from experimental studies is the enthalpy of the Co–C bond cleavage; the bond dissociation energy (BDE) is about 30 kcal/mol (2) and the equilibrium constant for the bond breaking of the AdoCbl is about 8×10^{-18} M (3), which corresponds to a reaction free energy of about 23 kcal/mol (unlike the BDE, this contribution includes a significant entropic component). The rate constant for homolysis of AdoCbl in solution has been estimated to be around 8.9×10^{-9} s^{-1} (see references in refs. 3 and 4), which corresponds to about 31 kcal/mol activation free energy. On the other hand, the activation barrier of the chemical step in B_{12} enzymes is estimated from the corresponding k_{cat} to be about 2–300 s^{-1} ($\Delta g^{\#} = 17$–16 kcal/mol) (see references in ref. 3), and the reaction free energy is estimated by some to be close to zero (3). Because the BDE and the activation barrier for the solution reaction are around 30 kcal/mol, we will assume that this represents a reasonable estimate of the activation barrier for the bond-breaking process. However, it is not clear at all if the bond-breaking reaction is the proper reference reaction. That is, if the reaction in the enzyme is a concerted reaction, then we should compare it to a similar concerted reaction in solution, and, at present, there is no clear information about such a reaction. Fortunately, the recent ab initio study of Kozlowski et al. (5) has shown that the concerted reaction has about 4 kcal/mol lower activation barriers than the stepwise reaction and that the reaction enthalpy of the solution reaction is around 22 kcal/mol. Similar conclusions have also emerged from the study of Sandala et al. (6).

If we estimate the reaction entropy to be about 30 eu (a very reasonable estimate of typical translational and rotational effects), we get a reaction free energy of about 22–10 = 12 kcal/mol for the solution reaction. Comparing these values to the corresponding values in the enzyme, we get $\Delta\Delta g^{\#} = 31 - 4 - 17 = 10$ kcal/mol and $\Delta\Delta G = 10$ kcal/mol. If we take the bond-breaking process as the reference reaction, we obtain, $\Delta\Delta g^{\#} = 31 - 17 = 14$ kcal/mol. The key problem, however, is the identification of the origin of this catalytic effect, and this issue will be addressed below.

S2. The Solution Surface

The evaluation of the potential surface for the reference solution reaction is of great importance because it provides major information on the relative importance of the stepwise and concerted mechanisms. This is crucial in view of the controversy about this subject and the repeated implication of problematic calculations that claim to be more reliable than the ab initio calibrated EVB and insist to prove that the reaction is not concerted (note that low-level quantum calculations might give incorrect results unless they are verified or calibrated on accurate experimental and theoretical information). To help the readers follow our points, we summarize them below.

Our previous study (7) found, in agreement with gas phase work (5), that the reference reaction involves a concerted path. The corresponding results (which are summarized in Fig. S1) were obtained with the density functional theory (DFT) (RB3LYP/LANL2DZ) level and COSMO solvation corrections and a very careful 2D mapping. In this respect, we like to clarify that the possibility that the lowest energy path is concerted cannot be eliminated by current experimental studies, despite inconsistent attempt to argue so (e.g., ref. 8). This point is elaborated in section S4.

S3. The EVB Potential Surface in the Protein

The EVB model considered as quantum system the atoms that are labeled in figure 1 of ref. 7 as EVB atoms. The bonding and charges of theses atoms were assigned according to the following resonance forms:

$$\begin{aligned} \phi_1 &= [\mathrm{Co} - \mathrm{C}_{(6)}\ \mathrm{H} - \mathrm{C}_{(8)}] \\ \phi_2 &= [\mathrm{Co}\ \mathrm{C}_{(6)}\ \mathrm{H} - \mathrm{C}_{(8)}] \\ \phi_3 &= [\mathrm{Co}\ \mathrm{C}_{(6)} - \mathrm{H}\ \mathrm{C}_{(8)}], \end{aligned} \qquad \textbf{[S1]}$$

where the energy of the corresponding diabatic energies was described by a "force field"-type set of functions, and the mixture of these diabatic energies by off-diagonal terms gave the actual potential surface for the reaction. The reaction free energy was then evaluated by the FEP/umbrella-sampling (FEP/US) procedure described elsewhere (e.g., ref. 9).

The charges of the reacting system were taken as the ab initio electrostatic potential (ESP) (10) charges evaluated at the B3LYP/6-31+G* level using the Gaussian 03 (11) program package for the given system in a COSMO solvent model (12). The relevant charges are given in table S1 of the supporting information of ref. 7 for MCM.

The EVB free-energy profiles were evaluated with the free energy perturbation/umbrella sampling (FEP/US) method using the MOLARIS program (13) with an explicit simulation sphere of 18 Å completed to 21 Å by Langevin dipoles, and then extended to infinity by a macroscopic sphere. This system was subjected to the surface constraint all-atom solvent model (13) with its spherical polarization and distance constraints. The long-range electrostatic effects were treated by the local reaction field treatment (13) and described by the ENZYMIX (13) force field. The ESP charges were taken as the EVB charges. The EVB parameters are given in ref. 7.

An important issue of the present study is the analysis of the nature of the solution surface. This is done in Fig. S1, where we start by constructing the relevant generic 2D surface. This surface reflects a combination of the results from our ab initio calculations, and from those of ref. 5 and experimental constraints. More specifically, obtaining an accurate surface for the B_{12} system is very challenging, and even the calculations of ref. 5 have not reproduced quantitatively the gas phase bond energy. However, the trend in such calculations is expected to be reasonable. Thus, the study of ref. 7 first performed hybrid DFT (RB3LYP/LANL2DZ) calculations, and then forced the energy along the $r_1(R_{Co-C_6})$ axis to reproduce the observed bond energy. This study also forced the hydrogen transfer axis for r_2 $(R_{C_6-H_9}) = 4.0$ Å to reproduce the observed difference in radical energy (14), viz., −10 kcal/mol for transfer from A to B. Basically, we introduced a linear interpolation to correct the gas phase calculations, and then add the same corrections to the solution surfaces that were obtained by using the COSMO solvent model. The solution surface also includes the entropic corrections associated with the solute entropy, which were estimated using a restraint release (RR) approach (15).

This generic solution surface (Fig. S1) was then used to generate the corresponding EVB surface, and the same EVB parameters were used in the protein calculations.

S4. The Concerted Option in Water and in the Protein

As stated in section S2, the nature of the 2D reaction surface is still controversial, and thus it is crucial to analyze the current point of views. As a start, we note that arguments in support of the stepwise path by large observed isotope effects (e.g., ref. 16) are problematic, because the bond-breaking step is (*i*) not expected to produce significant isotope effects and (*ii*) a large isotope effect would only be obtained for the concerted path or for the case when the hydrogen transfer is rate limiting. Furthermore, we would like to clarify that it is very difficult to exclude experimentally the concerted mechanism in the solvent cage, because the requirement of bringing the Co–C bond and the hydrogen donor to the same cage makes it extremely challenging to distinguish between the barriers of the stepwise and concerted path. Of course, distinguishing experimentally between these paths in the protein is close to impossible.

As to theoretical attempts to exclude the concerted path, we may start by noting that the recent MTD study of ref. 8 has not been as reliable as one might tend to think, considering the elegance of this approach. In fact, the conclusions (8) that the surface cannot be concerted are problematic, as the actual surface is quite flat in the diagonal range and the presented results involve very short sampling times for the potential of mean force (PMF) calculations (1.5 ps for each simulation window) and very short (5-fs) time intervals between the iterative depositions of Gaussians in the MTD sampling, which might not be sufficient for an accurate evaluation process. It seems to us that the strategy (e.g., ref. 7) of using very careful QM calculations in solution (with a higher level QM model than that used in ref. 8) have produced a concerted path, and that moving the solution surface to the protein environment by a calibrated EVB model is expected to be more reliable (both in terms of sampling and in terms of extrapolation of reliable reference systems) than the direct MTD in the protein site. Here again, it would be interesting to see a comparison between the paradynamics (PD) and the MTD studies, and since the EVB (with very careful calibration) has already reproduced quantitatively the catalysis in B_{12} with a concerted path, we tend to believe that the same results (or at least a flat surface) would be obtained by the PD QM (ai)/MM approach with the EVB reference potential.

Next, it is useful to consider the problematic analysis of ref. 16. First, this work failed again to generate a 2D surface, which is absolutely essential for addressing the difference in energy of the concerted and stepwise paths. Second, it is useful to note that our initial careful studies, summarized in Fig. S1, have not added arbitrarily extra driving forces to the hydrogen transfer step (as implied by of ref. 16) but incorporated well-known experimental findings about the radical transfer, which are extremely difficult to capture by most low-level ab initio models (see discussion in ref. 7). The same concerted features have now also been captured by our higher-level models without the above correction (Figs. S2 and S3).

We would also like to note that the work of ref. 17 used a semiempirical model that has not been calibrated by ab initio surfaces. Furthermore, although ref. 17 seems to support the stepwise mechanism and was interpreted as such by some workers (e.g., ref. 16), no 2D concerted surface has been reported or examined.

In summary, the concerted surfaces were found to be preferable in different levels in solution and also in the gas phase (Fig. S2) and then transferred the solution surface, with the very careful EVB calibration, to the protein and then used with the needed extensive sampling (Fig. S4). On the other hand, the opposing theoretical assertions have not been based on high-level calibration and in most cases have not been supported by 2D surfaces.

S5. Fundamental Problems with Assertions of Large Steric Effects

It is useful to elaborate here on our finding that B_{12} enzymes catalyze their reactions by electrostatic effects and on the problems associated with the assertion of large steric effects [e.g., see works discussed in a recent review (18)]. We start by noting that the pioneering work of Jensen and Ryde (4) identified the electrostatic interactions but concluded that they are converted to van der Waals steric effects. However, as we demonstrated in ref. 7 and discuss below in the context of other less critical works, this finding reflects insufficient sampling. We turn next to the work of Morokoma and coworkers (19), who basically produced an unstable and noncontinuous surface, while not generating any 2D description that would allow to judge the significance of the calculations (see discussion in ref. 7). Briefly, the calculations started by moving toward the TS but could not do so in a continuous way, thus jumped to the intermediate structure (see discontinuity in figure 6 of ref. 19), arbitrarily presuming that the energy of the intermediate is equal to the energy of the end of the first profile.

The implications of large steric effects persisted with the more recent ONIOM study (16) that also has not generated a 2D surface, which is essential for addressing the difference in energy of the concerted and stepwise paths. Furthermore, this work found that most of the catalysis is due to an enormous steric and conformational effect (although the paper states that it has not found RSD, its actual calculated surface shows enormous RSD, which must be due to strain). This inconsistent finding reflects the inability of energy minimization studies to estimate steric effects.

Additionally, this work included in its catalytic effect an inconsistently defined "cage effect." This reflects a common misunderstanding that considers the case when the Co–C bond is not completely dissociated before the hydrogen transfer as a catalytic cage effect, instead of the result of a reaction with a late concerted path. Obviously, the same 2D surface (Fig. 2) should be considered both in the reference state in solution and in the enzyme. Note that, as clarified in the main text, the only cage effect is the effect of bringing the hydrogen donor (substrate) and the cofactor to the same solvent cage to generate a proper reference reaction (nothing to do with enzyme cage). This concentration effect of 2.5 kcal/mol is always considered in our reference state. A reference state with a completely broken Co–C bond without the substrate is not a proper reference state for analysis of the catalytic effect.

We also note that the determination of the contributions of different residues to the catalytic effect is far from simple with the ONIOM method, because it is not clear how to use this method in performing reliable linear response approximation calculations, such as those performed in ref. 7. Moreover, the ONIOM method is unable to separate consistently the relevant electrostatic and van der Waals contributions, a crucial task that has been only met by the EVB.

Because the implication of large steric effect may be very appealing, we also consider here another recent problematic study (20) that tried to support the strain idea by simulating the conformational change upon substrate binding, using targeted molecular dynamics (TMD) for moving between the open and closed structures of the enzyme plus substrate complexes. To clarify the problems with this approach, we start by noting that TMD is clearly not the way to explore catalytically relevant free-energy changes (it simply is forcing the system from one structure to another), whereas the energy of moving between the unbound and bound structures is not related to the catalytic barrier, but simply to the free energy of the binding process. However, the problem is much more serious, because the TMD (that has been performed between the open structure where the substrate is supposed to be unstrained and the closed structures

where the substrate is supposed to be strained) was probably done to imply that the resulting path is relevant to the strain issue. Unfortunately, the strain contribution, which is allegedly relevant for B_{12} catalysis, has little to do with the path between the two assumed initial and final structures but only with the magnitude of the strain in the correct substrate structure in the closed enzyme–substrate (ES) complex form. Ref. 20 basically "assumed" the crucial substrate structure, instead of trying to obtain it by forcing the substrate and the surrounding protein in the ES to relax using a mild constraint on the observed ES structure. Furthermore, the relaxation must be done with the QM/MM surface and not with the MM used in the TMD. Now, determining the strain requires one to evaluate the QM/MM free energy for moving from the unstrained structure in water (Rw) to the structure in the ES complex in the protein [$R(p_{ES})$], where the unstrained structure in the protein [$Rw(p_{ES})$] should be forced to have the least Cartesian displacement relative to the bound structure (Rp_{ES}), which should have the smallest steric interaction with the protein.

Obtaining the free energy of the structural change should be done for both, the RS and TS, and thus provide the effect of the strain on the activation barrier. This type of study can be performed by setting the electrostatic interactions to zero (which converts the substrate to a nonpolar form) and to separate the electrostatic and steric effects. The PMF should then be calculated for moving between the two substrate structures in the protein and in water (as was done in the study of lysozyme in figure 6.4 of ref. 21). However, performing such calculations in a reliable way is very challenging, especially when we have significant structural changes. Fortunately, one can evaluate the strain of the nonpolar substrate by a rigorous thermodynamic cycle, mutating all of the atoms to nothing both in protein and in water, and thus evaluating the steric contributions in both structures [this is exactly the approach taken in our previous work (7), which is summarized in Table S6]. Interestingly, no such study was performed in ref. 20, which estimated the strain contribution simply by considering "unrelaxed" single-point energies of the substrate in the open and closed structures. We believe that this is highly problematic, because the single-point energies are most likely very unstable without proper sampling. In other words, while consistent relaxation will generate configurations with minimal steric repulsions and will reduce most of the steric strain [as was found in our studies (e.g., figure 8 in ref. 22) and of course in MCM (Table S6)], such relaxed configurations were never generated by the TMD (21). In this respect, we also note that the conformation used for the bound substrate (where most of the calculated strain comes from a presumed steric interaction between the ribose, adenine, and corrin) were at best relaxed by the TMD of an MM surface, without any relaxation on the QM/MM surface used for the strain calculations.

Interestingly, the catalytic effect has also been obtained by cluster and QM/MM minimization approaches. However, the problem is that it is very hard to obtain the decomposition to strain and electrostatics without proper sampling, and using just the X-ray structure is not sufficient, because it is crucial to allow the system to relax with the given potential surface where the corresponding (sometimes small) structural relaxation is not always reflected by the X-ray structure. Thus, the issue addressed above is the actual magnitude of the strain contribution. In this respect, we would like to clarify that we have here one type of studies that has reached its conclusions by inconsistent single-point unrelaxed calculations (and other problems), where our study (7) used fully convergent free-energy calculations with careful sampling that allowed the protein to relax around the substrate. Of course, once the alternative view will be based on proper sampling, generating continuous free-energy surfaces and doing so with the QM/MM surface, it will be interesting to see how much strain is still remaining in the corresponding model.

S6. The RR Approach and the Calculations of the Activation Entropies

The RR approach for evaluating configurational entropy is described schematically in Fig. S5. This approach imposes strong harmonic Cartesian restraints on the position of the reacting atoms in the TS and in the RS, and then evaluates the free energy associated with the release of these restraints by means of a FEP approach. The results of the FEP calculations depend on the position of the restraint coordinates, since the RR free energies contain a residual contribution from the enthalpy of the system. However, this contribution approaches zero for restraint coordinates that give the lowest RR energy (for details, see refs. 23 and 24). Accordingly, when we use the restraint position that gives the minimal absolute value of the restraint release free energy, we satisfy $-T\Delta S = \Delta G_{RR}$. Accordingly, we can write the following:

$$-T\Delta S^{\neq} = \min\left|\left(\Delta G_{RR}^{TS}\right)\right| - \min\left|\left(\Delta G_{RR}^{RS}\right)\right|, \qquad \textbf{[S2]}$$

where "min" indicates the minimum value of the indicated ΔG_{RR}.

Generally, in the case of chemical reactions, one is interested in the entropic contribution for a 1 M standard state. This can be obtained, in principle, by choosing a simulation sphere of a volume, which is equal to the molar volume (v_0 = 1,660 Å^3) while allowing the force constant in the final constraint to approach zero. However, such an approach is expected to encounter major convergence problems because the ligand is unlikely to sample the large simulation sphere in a reasonable simulation time. A faster convergence would be obtained by allowing the ligand to move in a smaller effective volume, v_{cage}, by imposing an additional constraint. This is done by the FEP approach, using a mapping potential of the form:

$$U_m^N = (1-\lambda_m)U_{rest,1}^N + \lambda_m U_{rest,2}^N + \left(K_{cage}/2\right)\left(\boldsymbol{R}_{s,i} - \overline{\boldsymbol{R}}_{s,i}\right)^2 + E, \qquad \textbf{[S3]}$$

where $\boldsymbol{R}_{s,i}$ is the position of a specified central atom of the reacting system. Here N designates a specific state (RS or TS), while $U_{rest,1}^N$ and $U_{rest,2}^N$ are Cartesian restraint potentials with an initial and final force constant (K_l and K_n respectively). Using U_m^N leaves v_{cage} unaffected by the change of λ_m. Now, we can let K_n approach zero without a divergence in $\Delta S'$ because the volume of the system is restricted by the K_{cage} term.

In the case of reactions in solutions, we evaluate the entropy associated with the release of K_{cage} analytically by the following:

$$-T\Delta S_{cage} = -\beta^{-1}\ln\left(v_0/v_{cage}\right), \qquad \textbf{[S4]}$$

where

$$v_{cage} = \left(\frac{2\pi}{\beta K_{cage}}\right)^{3/2}. \qquad \textbf{[S5]}$$

Following the above considerations, we can write:

$$-T\Delta S_{conf,w}^{\neq} = \min\left|\Delta G_{RR}^{TS}\right|^w - \min\left|\Delta G_{RR}^{RS}\right|^w - T\Delta S_{cage}^w, \qquad \textbf{[S6]}$$

where "conf" designates the substrate conformational entropy. In the case of reactions in enzymes, we do not need the ΔS_{cage} and the K_{cage} terms, because the enzyme active site restrains the reacting fragments and the probability of finding them out of the enzyme is small. Thus, we use the following:

$$-T\Delta S_{conf,p}^{\neq} = \min\left|\Delta G_{RR}^{TS}\right|^p - \min\left|\Delta G_{RR}^{RS}\right|^p. \qquad \textbf{[S7]}$$

The above approach has been simplified significantly since the work of refs. 23 and 25, where, instead of starting with a large

value of K_1, we save a major amount of computer time by modifying Eq. **S6** and using:

$$-T\Delta S^{\neq}_{conf\,p} = -T\Delta S^{TS}(K=K'_1)_{QH} + \min|\Delta G^{TS}_{RR}(K=K'_1 \to K=0)| + T\Delta S^{RS}(K=K'_1)_{QH} - \min|\Delta G^{RS}_{RR}(K=K'_1 \to K=0)|, \quad \textbf{[S8]}$$

where the $-T\Delta S(K{=}K'_1)_{QH}$ designates the entropy computed by the quasiharmonic (QH) approximation where K'_1 is the initial value of the restraint. In general, the QH approximation tends to be valid when restraints are significant; however, it starts to be very problematic when the restraints become small, resulting in a range of very shallow and anharmonic potential energy surfaces.

The practical RR calculations include in the present study not only to the substrate but also the surrounding residues. The actual simulations involved the following steps: We started with an initial relaxation of the RS and TS using MD runs of 500 ps at 300 K for MCM, and at 245 K for EAL, each with a time step of 0.5 fs. After the first 500 K steps, the restraint coordinates $\overline{\boldsymbol{R}}$ s were collected every 10 K steps leading to 25 different sets of $\overline{\boldsymbol{R}}$ s, of which the 16 $\overline{\boldsymbol{R}}$ s with the lowest total energy were selected for the subsequent RR calculations. The RR contributions for these 16 $\overline{\boldsymbol{R}}$ s were evaluated at both the RS (state I) and the TSs (states II and III) by calculating the QH contribution with $K'_1 = 6.0$ kcal·mol^{-1}·Å^{-2}, followed by the RR-FEP contribution for changing K from 6.0 to 0.03 kcal·mol^{-1}·Å^{-2} in steps of 0.1 kcal·mol^{-1}·Å^{-2} leading to 60 simulation windows, plus an additional two windows at 0.05 and 0.03 kcal·mol^{-1}·Å^{-2} (62 simulation windows in total). For each of these 62 simulation windows, sampling was performed for 10 ps with a time step of 0.5 fs leading to a simulation length of 620 ps (and around 10 ns for all 16 $\overline{\boldsymbol{R}}$ s). Out of these 16 simulations, the minimum value (minΔG) between state I and both state II and state III was determined.

The constraints used for the RR in MCM involved (*i*) the constraints for region1 (the substrate and the leaving group) and (*ii*) residue constraints for residues 87, 88, 89, 90, 92, 98, 101, 102, 116, 117, 118, 119, 137, 138, 139, 141, 166, 197, 207, 243, 244, 245, 246, 248, 249, 287, 328, 329, 330, 331, 332, 333, 334, 335, 336, 348, 351, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 384, and 387. Weak distance restraints of 3.0 kcal·mol^{-1}·Å^{-2} to maintain the system stable throughout the simulations were applied between the substrate carboxylate moiety and ARG 207, the cobalt atom, and HIS 610 (the base), and between the ribose hydroxyl moieties and GLU 370. The FEP averages obtained for the forward and backward RR free energies (without the QH harmonic contribution) were marked by a large statistical variance, which seems to be correlated with the observation of disrupted hydrogen bonding when these distance restraints were not used.

The constraints used for the RR in EAL involved (*i*) the constraints for region 1 (the substrate and the leaving group) and (*ii*) region 2, which includes all residues within an 18-Å sphere from the center of region 1. As in MCM, weak distance restraints of 3.0 kcal·mol^{-1}·Å^{-2} to maintain the system stable throughout the simulations were applied between the substrate's amino group and residues ASP 362 and GLU 287, between the hydroxyl group of the substrate and residues ARG 160, ASN 193, and GLU 287, and the hydroxyl groups of ribose moiety of ado and residues ASN 193 and GLU 287. The weak distance constraints mentioned above for MCM and EAL were applied to all three states, because they rendered the systems more stable, and were subsequently also not released in our RR calculations.

Finally, for all simulations in MCM, residue GLU 370 was ionized, and residues ARG 160, GLU 287, and ASP 362 were protonated/ionized for all simulations in EAL, which were carried out with a modified version of MOLARIS (26).

At this point, we would like to comment on our evaluation of the activation free energies. That is, in principle, we should have evaluated the activation entropy by determining the exact position of the TS and then performing RR at that position. However, in the present case, the TS position is not certain (concerted or stepwise). Thus, we found it convenient to evaluate the RR near the TS of the stepwise path (state II) and between the product state and the TS of the concerted path (state III). If the path is close to the stepwise path, then the entropy of the I-to-II transition gives the needed result. If we have a concerted path, then we can assume a linear correlation (linear free-energy relationship) between the activation entropy of the concerted path with the entropy change of the I-to-III transition (where the entropy for the I-to-III transition provided an upper limit of the absolute value of the activation entropy). Our approximation for the upper limit in the case of the concerted path seems reasonable in particular if we are with the entropy of the environment. Thus, we assume that the value of the correct activation entropy is the range bound by the estimates for the I-to-II and I-to-III transitions.

1. Warshel A, et al. (2006) Electrostatic basis for enzyme catalysis. *Chem Rev* 106(8): 3210–3235.
2. Hay BP, Finke RG (1986) Thermolysis of the Co–C bond of adenosylcobalamin. 2. Products, kinetics, and Co–C bond-dissociation energy in aqueous solution. *J Am Chem Soc* 108(16):4820–4829.
3. Brown KL (2005) Chemistry and enzymology of vitamin B_{12}. *Chem Rev* 105(6):2075–2149.
4. Jensen KP, Ryde U (2005) How the Co–C bond is cleaved in coenzyme B_{12} enzymes: A theoretical study. *J Am Chem Soc* 127(25):9117–9128.
5. Kozlowski PM, Kamachi T, Toraya T, Yoshizawa K (2007) Does Cob(II)alamin act as a conductor in coenzyme B_{12} dependent mutases? *Angew Chem Int Ed Engl* 46(6): 980–983.
6. Sandala GM, Smith DM, Marsh ENG, Radom L (2007) Toward an improved understanding of the glutamate mutase system. *J Am Chem Soc* 129(6):1623–1633.
7. Sharma PK, Chu ZT, Olsson MHM, Warshel A (2007) A new paradigm for electrostatic catalysis of radical reactions in vitamin B_{12} enzymes. *Proc Natl Acad Sci USA* 104(23): 9661–9666.
8. Bucher D, Sandala GM, Durbeej B, Radom L, Smith DM (2012) The elusive 5′-deoxyadenosyl radical in coenzyme-B_{12}-mediated reactions. *J Am Chem Soc* 134(3):1591–1599.
9. Aqvist J, Warshel A (1993) Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches. *Chem Rev* 93(7):2523–2544.
10. Singh UC, Kollman PA (1984) An approach to computing electrostatic charges for molecules. *J Comput Chem* 5(2):129–145.
11. Frisch MJ, et al. (2003) Gaussian 03 (Gaussian, Inc., Pittsburgh).
12. Klamt A, Schuurmann G (1993) COSMO—a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J Chem Soc Perkins Trans* 2(5):799–805.
13. Lee FS, Chu ZT, Warshel A (1993) Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs. *J Comput Chem* 14(2):161–185.
14. Bordwell FG, Zhang X-M (1993) From equilibrium acidities to radical stabilization energies. *Acc Chem Res* 26(9):510–517.
15. Villa J, et al. (2000) How important are entropic contributions to enzyme catalysis? *Proc Natl Acad Sci USA* 97(22):11899–11904.
16. Li X, Chung LW, Paneth P, Morokuma K (2009) DFT and ONIOM(DFT:MM) studies on Co-C bond cleavage and hydrogen transfer in B_{12}-dependent methylmalonyl-CoA mutase. Stepwise or concerted mechanism? *J Am Chem Soc* 131(14):5115–5125.
17. Dybala-Defratyka A, Paneth P, Banerjee R, Truhlar DG (2007) Coupling of hydrogenic tunneling to active-site motion in the hydrogen radical transfer catalyzed by a coenzyme B_{12}-dependent mutase. *Proc Natl Acad Sci USA* 104(26): 10774–10779.
18. Blomberg MR, Borowski T, Himo F, Liao RZ, Siegbahn PE (2014) Quantum chemical studies of mechanisms for metalloenzymes. *Chem Rev* 114(7):3601–3658.
19. Kwiecien RA, et al. (2006) Computational insights into the mechanism of radical generation in B12-dependent methylmalonyl-CoA mutase. *J Am Chem Soc* 128(4): 1287–1292.
20. Pang J, Li X, Morokuma K, Scrutton NS, Sutcliffe MJ (2012) Large-scale domain conformational change is coupled to the activation of the Co–C bond in the B_{12}-dependent enzyme ornithine 4,5-aminomutase: A computational study. *J Am Chem Soc* 134(4):2367–2377.
21. Warshel A (1991) *Computer Modeling of Chemical Reactions in Enzymes and Solutions* (Wiley, New York).
22. Ram Prasad B, Warshel A (2011) Prechemistry versus preorganization in DNA replication fidelity. *Proteins* 79(10):2900–2919.

23. Strajbl M, Sham YY, Villà J, Chu ZT, Warshel A (2000) Calculations of activation entropies of chemical reactions in solution. *J Phys Chem B* 104(18):4578–4584.
24. Singh N, Warshel A (2010) A comprehensive examination of the contributions to the binding entropy of protein-ligand complexes. *Proteins* 78(7):1724–1735.
25. Sharma PK, Xiang Y, Kato M, Warshel A (2005) What are the roles of substrate-assisted catalysis and proximity effects in peptide bond formation by the ribosome? *Biochemistry* 44(34):11307–11314.
26. Warshel A, et al. (2012) MOLARIS-XG, 9.11 (University of Southern California, Los Angeles).

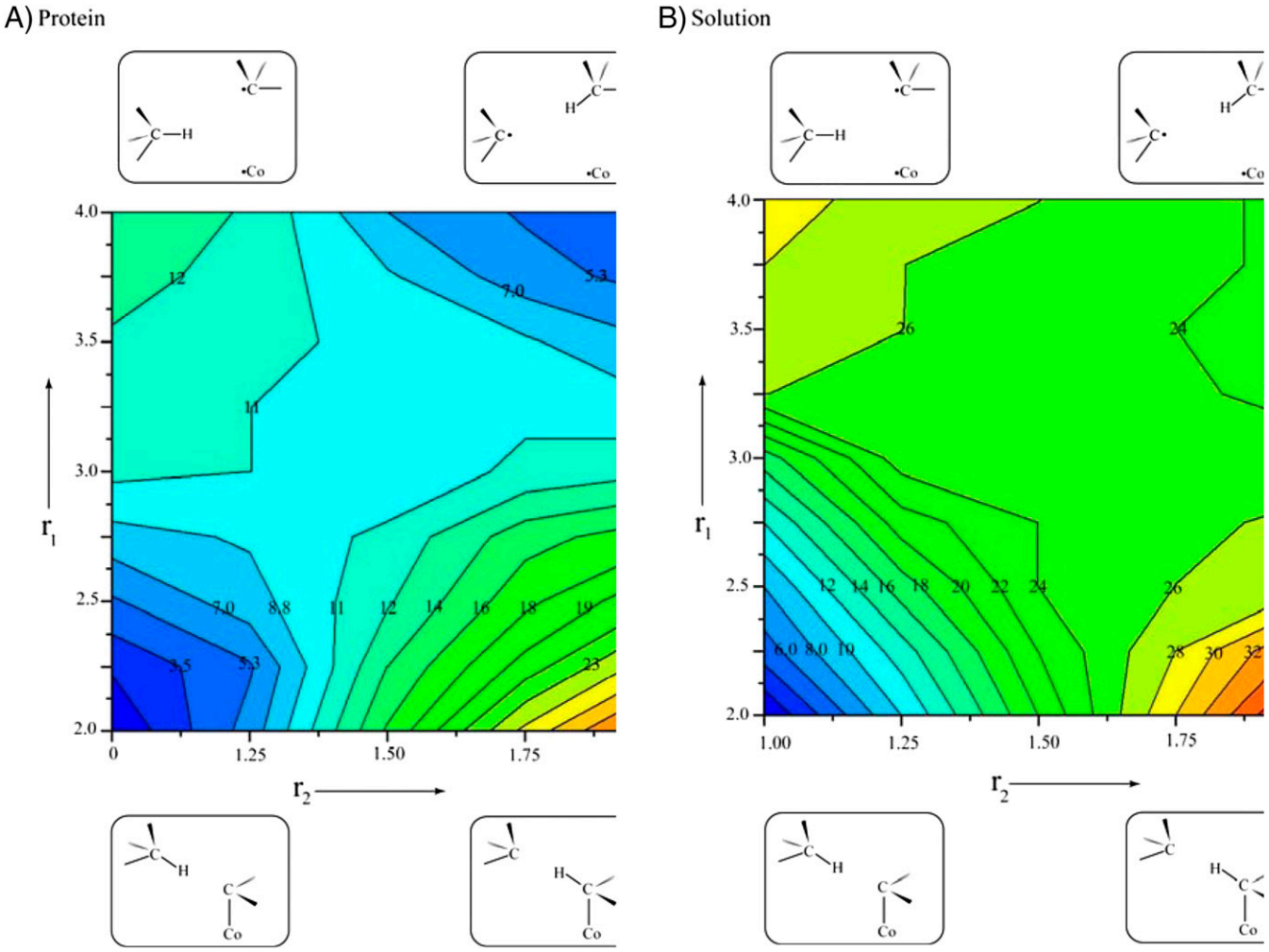

Fig. S1. The potential surfaces of ref. 1, for the protein (*A*) and solution (*B*), evaluated with the B3LYP/6-31-G* level of theory and calculated with the Gaussian 03 program package (2) and the COSMO solvent model (3). Reproduced with permissions from ref. 1.

1. Sharma PK, Chu ZT, Olsson MHM, Warshel A (2007) A new paradigm for electrostatic catalysis of radical reactions in vitamin B_{12} enzymes. *Proc Natl Acad Sci USA* 104(23):9661–9666.
2. Frisch MJ, et al. (2003) Gaussian 03, Revision C.02 (Gaussian, Inc., Pittsburgh).
3. Klamt A, Schuurmann G (1993) COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J Chem Soc Perkin Trans 2* (5):799–805.

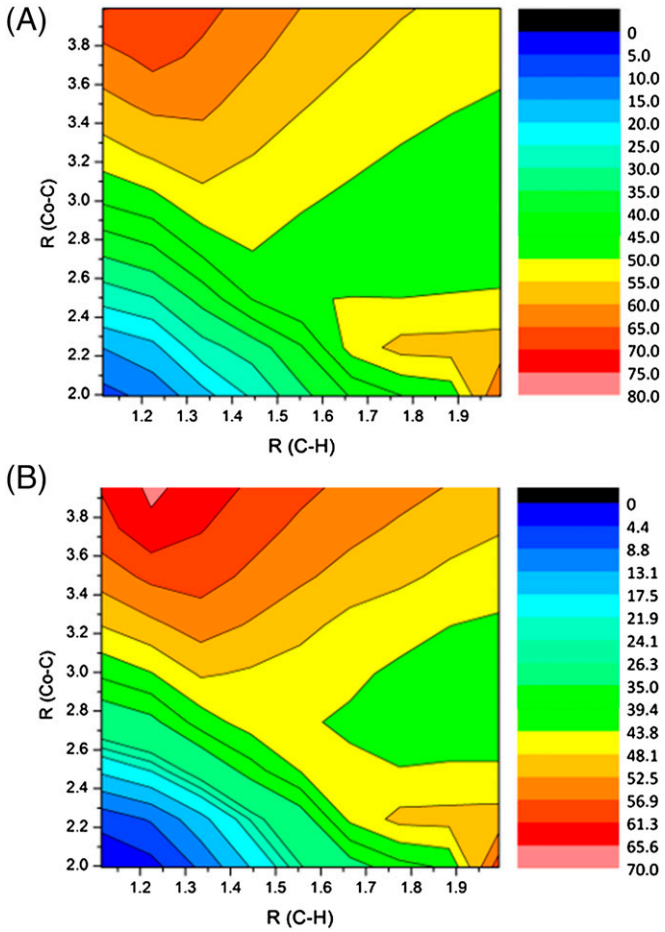

Fig. S2. (*A*) Gas phase 2D contour color plot of surface data obtained at the BP86/6-31G(d) level of theory. The surface was constructed by covering the space with a 10 × 10 grid, and for each grid point the intrinsic reaction coordinate output was searched for the nearest geometry to use as the start point, which was then reoptimized with the two distances fixed at their grid values. (*B*) Single-point BP86/6-31G(d) + COSMO surface for the optimized structures obtained in *A*.

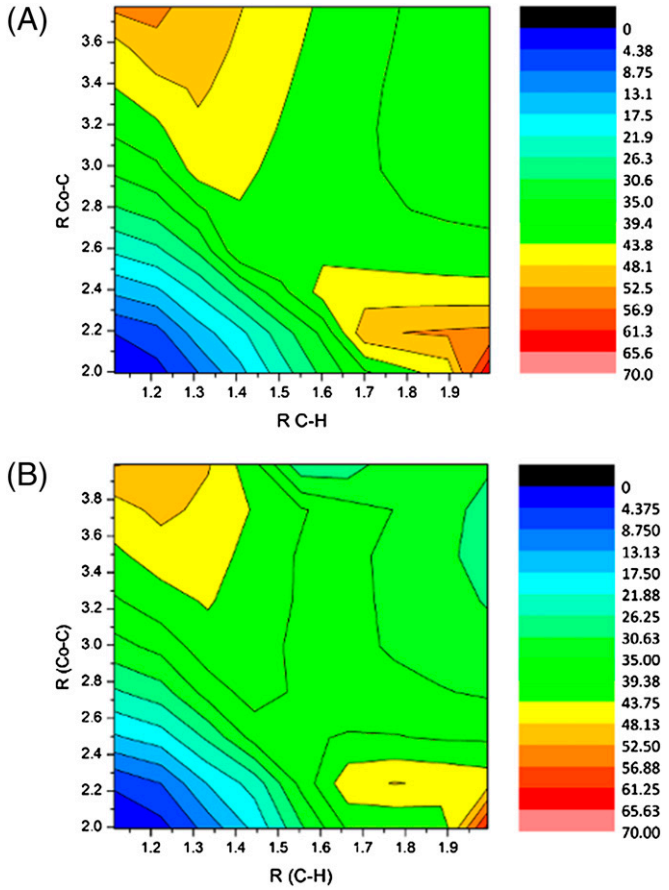

Fig. S3. (*A*) Two-dimensional contour plot of the BP86/6-31G(d)+DFT-D3 reaction surface. Energies are computed at the optimized BP86/6-31G(d) level and corrections are added without optimization of the geometry. (*B*) Two-dimensional contour plot of the BP86/6-31G(d)+DFT-D2 reaction surface. The geometries used in *A* were reoptimized at the BP86/6-31G(d)+DFT-D2 with QChem (1).

1. Shao Y, et al. (2015) Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Molecular Physics: An International Journal at the Interface Between Chemistry and Physics* 113(2):184–215.

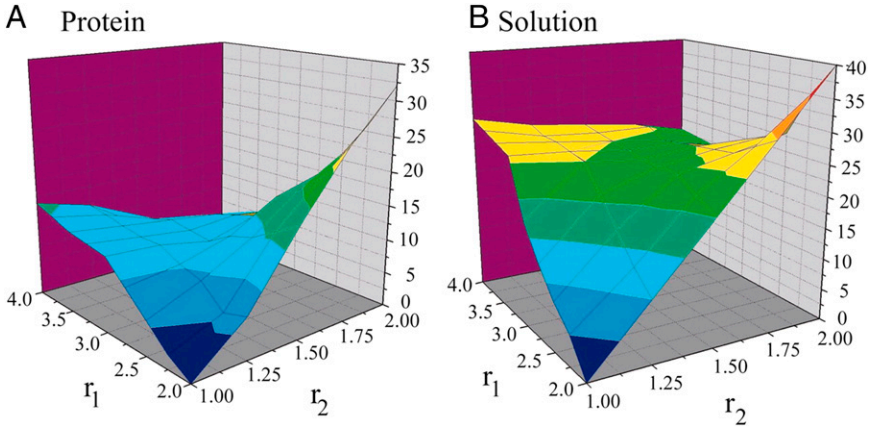

Fig. S4. Describing the free-energy surface on the complete concerted space for the reaction in mutase (*A*) and in solution (*B*). Reproduced with permissions from ref. 1.

1. Sharma PK, Chu ZT, Olsson MHM, Warshel A (2007) A new paradigm for electrostatic catalysis of radical reactions in vitamin B12 enzymes. *Proc Natl Acad Sci USA* 104(23):9661–9666.

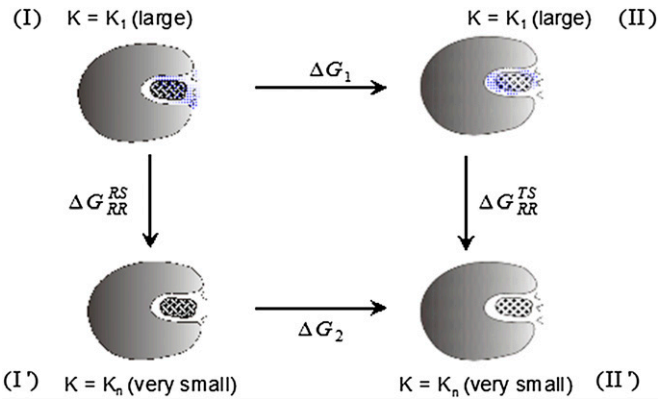

Fig. S5. Thermodynamic cycle for the evaluation of the RR configurational entropy contribution to the activation free energy for the RS and the TS (configuration I and II, respectively).

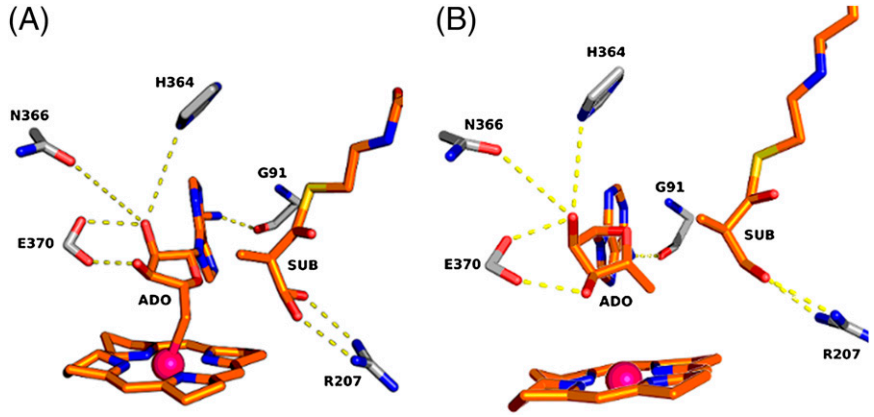

Fig. S6. The change in the active-site structure upon moving from the RS (*A*) to the TS (*B*) in MCM.

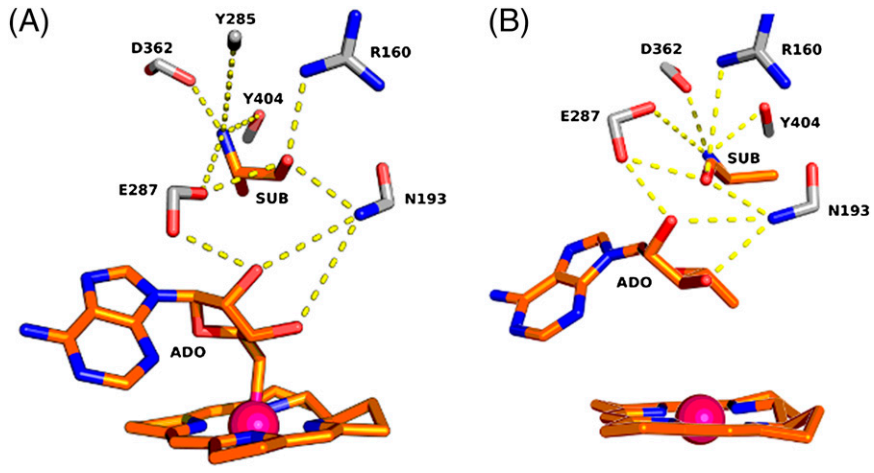

Fig. S7. The change in the active-site structure upon moving from the RS (*A*) to the TS (*B*) in EAL.

Table S1. RR results at 300 K for MCM where all residues are charged*

State I			State II			State III		
∑RR (6.0–0.03)	QH_6	∑RR+QH	∑RR (6.0–0.03)	QH_6	∑RR+QH	∑RR (6.0–0.03)	QH_6	∑RR+QH
−159.3	−206.3	−365.6	−165.0	−201.0	−366.0	−161.5	−205.9	−367.4
−160.4	−200.0	−360.4	−168.8	−200.5	−369.3	−166.0	−201.4	−367.4
−158.2	−203.6	−361.8	−168.7	−203.7	−372.4	−164.0	−194.9	−358.9
−158.4	−193.7	−352.1	−171.4	−201.3	−372.7	−161.0	−200.2	−361.2
−158.6	−195.9	−354.5	−173.4	−198.6	−372.0	−163.0	−196.8	−359.8
−161.0	−197.9	−358.9	−172.7	−200.9	−373.6	−165.3	−196.2	−361.5
−158.1	−195.8	−353.9	−165.2	−200.3	−365.5	−170.7	−203.6	−374.3
−159.4	−198.3	−357.7	−170.4	−199.5	−369.9	−164.7	−197.4	−362.1
−158.6	−193.2	−351.8	−168.5	−197.5	−366.0	−164.7	−201.1	−365.8
−161.5	−195.6	−357.1	−169.2	−199.3	−368.5	−163.4	−201.4	−364.8
−158.3	−192.7	**−351.0**	−164.5	−196.7	**−361.2**	−161.1	−207.3	−368.4
−160.3	−197.0	−357.3	−163.1	−199.0	−362.1	−162.4	−200.3	−362.7
−158.7	−193.6	−352.3	−165.7	−200.6	−366.3	−162.1	−202.6	−364.7
−159.8	−196.4	−356.2	−172.2	−203.3	−375.5	−161.2	−202.0	−363.2
−156.5	−196.4	−352.9	−167.1	−207.8	−374.9	−161.9	−195.7	**−357.6**
−159.2	−195.3	−354.5	−170.8	−197.4	−368.2	−161.7	−205.3	−367.0
$-T\Delta S$ (SII-SI)		**−10.2**						
$-T\Delta S$ (SIII-SI)		**−6.6**						

*Energies in kcal/mol. RR and QH designate restraint release and quasiharmonic, respectively. Values 6.0–0.03 designate the initial and final values of the restraint force constant (in kcal/mol), whereas QH_6 indicates that the QH calculation was done with a constraint force constant of 6 kcal/mol. The regions that were restrained are listed in section S6. The RR energies with the smallest absolute values are given in boldface.

Table S2. RR results at 300 K for MCM where all atoms of region 1 as well as residues in the first solvation shell have no charge*

State I			State II			State III		
∑RR (6.0–0.03)	QH_6	∑RR+QH	∑RR (6.0–0.03)	QH_6	∑RR+QH	∑RR (6.0–0.03)	QH_6	∑RR+QH
−161.8	−200.3	−362.1	−166.3	−199.3	−365.6	−167.3	−202.3	−369.6
−159.5	−206.2	−365.7	−166.2	−202.3	−368.5	−165.7	−198.3	−364.0
−160.1	−200.4	−360.5	−165.2	−203.7	−368.9	−167.6	−203.1	−370.7
−161.4	−198.4	−359.8	−162.6	−207.4	−370.0	−166.1	−199.4	−365.5
−161.2	−209.0	−370.2	−166.7	−206.5	−373.2	−161.5	−200.1	**−361.6**
−161.7	−200.5	−362.2	−167.1	−202.7	−369.8	−164.3	−207.5	−371.8
−158.4	−199.3	**−357.7**	−167.0	−203.0	−370.0	−166.3	−201.0	−367.3
−160.5	−201.4	−361.9	−159.9	−203.0	**−362.9**	−165.6	−198.2	−363.8
−162.6	−204.3	−366.9	−167.3	−207.7	−375.0	−164.1	−202.1	−366.2
−159.9	−200.2	−360.1	−167.5	−208.4	−375.9	−165.3	−203.2	−368.5
−159.0	−203.3	−362.3	−161.7	−205.5	−367.2	−165.1	−200.4	−365.5
−159.3	−207.1	−366.4	−165.3	−205.4	−370.7	−163.8	−202.3	−366.1
−162.4	−200.2	−362.6	−167.4	−205.7	−373.1	−162.3	−203.1	−365.4
−160.0	−202.4	−362.4	−166.1	−202.3	−368.4	−164.5	−205.3	−369.8
−160.4	−204.4	−364.8	−164.8	−205.3	−370.1	−169	−198.2	−367.2
−159.7	−204.0	−363.7	−164.5	−200.0	−364.5	−168.5	−201.1	−369.6
$-T\Delta S$ (SII-SI)		**−5.2**						
$-T\Delta S$ (SIII-SI)		**−3.9**						

*Energies in kcal/mol. RR and QH designate restraint release and quasiharmonic, respectively. Values 6.0–0.03 designate the initial and final values of the restraint force constant (in kcal/mol), whereas QH_6 indicates that the QH calculation was done with a constraint force constant of 6 kcal/mol. The regions that were restrained are listed in section S6. The RR energies with the smallest absolute values are given in boldface.

Table S3. RR results at 245 K for EAL where all residues are charged*

State I			State II			State III		
∑RR (6.0–0.03)	QH_6	∑RR+QH	∑RR (6.0–0.03)	QH_6	∑RR+QH	∑RR (6.0–0.03)	QH_6	∑RR+QH
−266.4	−180.8	−447.2	−276.5	−185.5	−462.0	−285.5	−190.2	−475.7
−255.5	−180.3	**−435.8**	−277.5	−186.7	−464.2	−286.3	−189.5	−475.8
−261.4	−178.5	−439.9	−276.3	−183.3	−459.6	−280.4	−187.4	−467.8
−258.6	−181.3	−439.9	−278.6	−185.4	−464.0	−284.5	−188.4	−472.9
−264.0	−182.5	−446.5	−278.4	−187.3	−465.7	−286.3	−190.3	−476.6
−264.0	−179.6	−443.6	−280.6	−183.3	−463.9	−288.5	−186.4	−474.9
−263.7	−177.5	−441.2	−277.5	−185.6	−463.1	−287.4	−191.3	−478.7
−266.4	−178.9	−445.3	−279.3	−187.8	−467.1	−284.0	−192.0	−476.0
−257.3	−180.4	−437.7	−278.5	−186.5	−465.0	−285.0	−193.7	−478.7
−261.0	−181.9	−442.9	−274.6	−183.4	**−458.0**	−287.2	−188.4	−475.6
−258.2	−178.4	−436.6	−279.7	−185.5	−465.2	−289.6	−187.3	−476.9
−265.3	−178.5	−443.8	−278.6	−184.3	−462.9	−284.3	−191.4	−475.7
−267.2	−179.7	−446.9	−282.3	−184.0	−466.3	−281.5	−190.3	**−471.8**
−265.2	−177.6	−442.8	−280.4	−186.4	−466.8	−285.4	−197.5	−482.9
−255.8	−184.3	−440.1	−276.6	−187.3	−463.9	−286.3	−189.9	−476.2
−260.7	−183.2	−443.9	−275.9	−185.6	−461.5	−282.9	−190.0	−472.9
$-T\Delta S$ (SII-SI)		**−22.2**						
$-T\Delta S$ (SIII-SI)		**−36.0**						

*Energies in kcal/mol. RR and QH designate restraint release and quasiharmonic, respectively. Values 6.0–0.03 designate the initial and final values of the restraint force constant (in kcal/mol), whereas QH_6 indicates that the QH calculation was done with a constraint force constant of 6 kcal/mol. The regions that were restrained are listed in section S6. The RR energies with the smallest absolute values are given in boldface.

Table S4. RR results at 245 K for EAL where all atoms of region 1 as well as residues in the first solvation shell have no charge*

State I			State II			State III		
∑RR (6.0–0.03)	QH_6	∑RR+QH	∑RR (6.0–0.03)	QH_6	∑RR+QH	∑RR (6.0–0.03)	QH_6	∑RR+QH
−270.3	−183.5	−453.8	−275.4	−183.5	−458.9	−280.3	−186.5	−466.8
−271.5	−182.5	−454.0	−274.6	−185.3	−459.9	−279.6	−185.4	−465.0
−265.5	−183.2	−448.7	−277.6	−180.5	−458.1	−280.3	−183.7	−464.0
−271.3	−180.3	−451.6	−273.8	−187.4	−461.2	−282.5	−180.4	**−462.9**
−270.4	−182.4	−452.8	−278.9	−186.3	−465.2	−281.3	−187.9	−469.2
−269.3	−181.9	−451.2	−277.4	−183.4	−460.8	−280.5	−188.4	−468.9
−270.0	−179.2	−449.2	−275.6	−187.4	−463.0	−283.6	−186.4	−470.0
−269.9	−182.5	−452.4	−278.4	−179.2	**−457.6**	−286.4	−183.7	−470.1
−269.5	−179.9	−449.4	−275.3	−190.4	−465.7	−287.5	−184.0	−471.5
−272.4	−184.3	−456.7	−280.3	−184.3	−464.6	−285.4	−185.6	−471.0
−266.1	−178.2	**−444.3**	−275.3	−182.4	−457.7	−289.6	−186.0	−475.6
−272.4	−185.4	−457.8	−280.1	−181.7	−461.8	−284.5	−188.4	−472.9
−271.3	−180.3	−451.6	−275.3	−184.3	−459.6	−283.4	−188.7	−472.1
−269.8	−182.5	−452.3	−272.7	−185.2	−457.9	−282.9	−183.2	−466.1
−271.4	−183.2	−454.6	−274.5	−186.3	−460.8	−279.6	−185.5	−465.1
−274.3	−185.4	−459.7	−276.1	−184.4	−460.5	−280.9	−187.9	−468.8
$-T\Delta S$ (SII-SI)		**−13.3**						
$-T\Delta S$ (SIII-SI)		**−18.6**						

*Energies in kcal/mol. RR and QH designate restraint release and quasiharmonic, respectively. Values 6.0–0.03 designate the initial and final values of the restraint force constant (in kcal/mol), whereas QH_6 indicates that the QH calculation was done with a constraint force constant of 6 kcal/mol. The regions that were restrained are listed in section S6. The RR energies with the smallest absolute values are given in boldface.

Table S5. Key hydrogen-bonding distances in states I and III in the first solvation shell between region 1 and region 2 residues of EAL

Region 2 residue	Region 1 residue	State I	State III
D362/OD2	SUB/NH	2.3	1.9
Q162/OE1	SUB/NH	2.7	4.9
Q162/OE1	SUB/O	4.2	2.2
R160/HH1	SUB/O	1.9	3.1
N193/HD2	SUB/O	2.0	2.0
N193/HD1	ADO/OR1	3.0	2.0
E287/OE1	SUB/OH	1.9	3.5
E287/OE1	SUB/NH	4.0	2.4
E287/OE1	ADO/OH	2.2	1.9

Distances (in angstroms) for individual hydrogen bonds between region 1 residues adenosyldeoxyribose (ADO) and 2*S*-aminoproanole (SUB) plus region 2 residues for configurations I and III. Note that the interaction with R160 is complex because it is repelled by the positive charge of the substrate.

Table S6. The FEP analysis of strain effects

System	$\Delta G^{p}_{np\rightarrow d}$	$\Delta G^{w}_{np\rightarrow d}$	$\Delta G^{w\rightarrow p}_{np\rightarrow d}$	$\Delta G^{w\rightarrow p}_{np\rightarrow d(2A-4A)}$
$Ring_2$ (r1 = 2.0 Å)	15.0	15.3	0.3	1.7
$Ring_2$ (r1 = 4.0 Å)	16.0	14.0	2.0	
$Ring_3$ (r1 = 2.0 Å)	17.4	18.5	1.1	−1.1
$Ring_3$ (r1 = 4.0 Å)	18.5	18.5	0.0	

Energies in kilocalories per mole. The table gives the free energies of mutating the reacting system from its fully nonpolar form to dummy atoms (*d*). The calculations were done for $ring_2$ and $ring_3$ at different Co–C distances. Calculations are taken from ref. 1.

1. Sharma PK, Chu ZT, Olsson MHM, Warshel A (2007) A new paradigm for electrostatic catalysis of radical reactions in vitamin B12 enzymes. *Proc Natl Acad Sci USA* 104(23):9661–9666.