Electronic Supplementary Material (ESI) for Nanoscale.
This journal is © The Royal Society of Chemistry 2022

Supplementary materials: Single-molecule evidence for chemical ratchet in binding between *cam* repressor and its operator

Hiroyuki Kabata,[1, 2] Hironori Aramaki,[3, 4] and Nobuo Shimamoto*[, 1]

[1] National Institute of Genetics, and Department of Genetics, School of Life Science, The Graduate University for Advanced Studies, SOKENDAI, 1111 Yata, Mishima, Shizuoka, Japan 411-8540.
[2] Present address: Headquarters for Innovative Society-Academia Cooperation, University of Fukui, 3-9-1 Bunkyo, Fukui, Japan 910-8507.
[3] Department of Molecular and Life Science, Faculty of Pharmacy, Daiichi University of Pharmacy, 22-1 Tamagawa-cho, Minami-ku, Fukuoka, Japan 815-8511.
[4] Present address: Sankyu Drug Co., Ltd. 3-1-13 Kurogawa-Nishi, Moji-ku, Kitakyushu, Japan 801-0825.

*Corresponding Author. E-mail: nshima@nig.ac.jp

1. Definition of chemical ratchet

Chemical ratchet is a general mechanism of a reaction and one of the two alternative ways of describing chemical reactions with rate equations. Therefore, it is not limited to the specific complex formation mediated by one-dimensional diffusion. The following table is the summary of the commentary revealing a historical misuse of transition-state theory in the case of chemical ratchet.[1] The circulating flow occurs among the sub-states involved in a single thermodynamic state, and thus it does not violate thermodynamics.

Way of description with rate equations	Transition-state theory (conventional)	Chemical ratchet
Definitions		
timescale of relevant reaction (timescale of reaction coordinate)	distinctly separated from all the other degrees of freedom	close to another degree(s) of freedom$^{\alpha}$, namely the coupled degree(s) of freedom
Properties		
thermal equilibrium between transition-state and ground state	established	not yet established
level of the potential of mean force at transition state	time-independent	changes according to the coupled degree(s) of freedom
circulating flow among reaction states in the stationary state to which the relevant reaction system converges	does not exist = detailed balance holds. (The antenna effect by one-dimensional diffusion should not exist.)	generally exists. (The antenna effect by one-dimensional diffusion can exist.)
number of required sets of rate equations	single set	two or more sets alternatively switched according to the coupled degree(s) of freedom
converging stationary state	both macroscopically and microscopically equilibrated	macroscopically equilibrated but microscopically non-equilibrated

Table S1 Comparison between transition-state theory and chemical ratchet. Note that detailed balance of reaction rates and microscopic equilibrium are physically tautological.
α: If the timescale of the measurement, such as a time cost in measuring an affinity, is taken into account, the timescale of the coupled degree of freedom is widened to faster than the measurement.

2. Additional specific sites

Among 71 observed trapping events, three-fourths occurred at a site where *camO* had been inserted into the DNA, so far as could be judged within the resolution of microscopy (Fig. S1A). The remainders were observed at two (or more) other sites, near an end and at 4 μm from an end in the DNA belts. There are two DNA sites homologous to *camO*, one at coordinate 1,588 (0.54 μm from the "left" end) with 87% identity to the 15 bp central palindrome; the other at 12,477 (4.2 μm from the "right" end) with 80% identity.

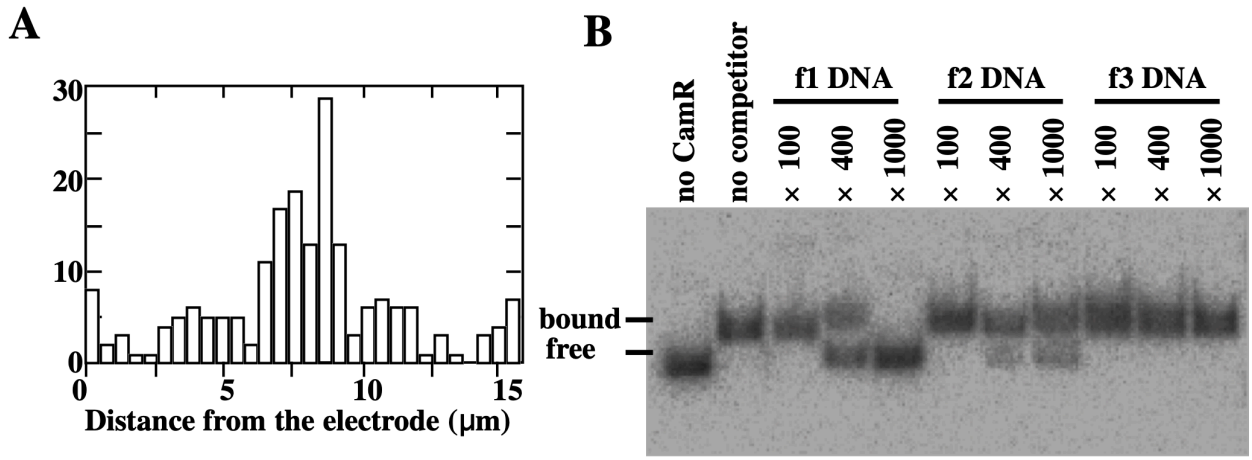

Figure S1. Trapping of CamR molecule on λgt11::*camO* DNA. The trapping at these sites was decreased in the presence of 5 mM D-camphor, the inducer of the cam operon, as was the complex at *camO* itself (the first and the third bars in Fig. 2 in the Main Text). These results indicate that trapping is attributable solely to formation of specific complexes (Fig. S1B). In other words, all nonspecific complexes of CamR can fall into the complex diffusing one-dimensionally. Since the presence of D-camphor does not affect one-dimensional diffusion or jumping (the third bar in Fig. 2), we conclude that it does not affect binding to nonspecific sites. Note that possibly similar inducer-resistant complexes of the *lac* repressor at nonspecific sites were reported using a filter binding assay.[2]

(A) Distribution of positions at which CamR molecules were "trapped". The authentic *camO* site is located at the 6.7 μm and 8.3 μm positions because of the random polarity of the fixed 16 μm DNA. Specificities of two sites on λDNA homologous to *camO*.

(B) The migration of the ^{32}P-labeled *camO* DNA in the presence of competitors. The labeled 126 bp DNA fragment carrying *camO*, the *Hin*cII-*Hin*dIII fragment of pJP3111, was mixed with the indicated molar excess of a competitor DNA (f1, f2, or f3), and the mixture was incubated with 1.3-M CamR for 1 h at 25°C before gel electrophoresis. The competitor f1 DNA: 115 bp *Ava*I-*Bbi*II, f2 DNA: 144 bp BarI, f3 DNA: 290 bp *Ava*I-*Bvi*II fragments of λCI^{857}Sam7 DNA. The f1 and f2 DNA contain the sequence homologous to *camO*. The f3 DNA is a negative control containing no homologous sites. The positions of the free and bound DNA are indicated in the margin. The slightly increased mobilities of the bands in the left three lanes are electrophoretic artifacts.

3. Energetic and geometric terms composing a rate parameter.

According to transition-state theory, the rate constant is expressed as the multiple of the energetic exponential factor with a permeable factor dependent on the structure of reactant under the assumption that all the degrees of freedom except the reaction coordinate are thermally equilibrated. This microscopic assumption is tautology of detailed balance of reaction, and does not hold when there is a degree of freedom with a timescale close to that of the reaction. This assumption tends to be forgotten and detailed balance of reaction tends to be mistakenly considered as a universal truth denying chemical ratchet.[1] We here show that the form as the multiple is valid also in the case of chemical ratchet.

Since movements of small solvent molecules are thermally equilibrated in the timescale of the relevant processes, the probability of receiving energy E to drive a chemical process is the canonical distribution proportional to $\exp(-E/k_BT)$, where k_B is Boltzmann constant and T is absolute temperature. A single collision with large enough energy can result in perturbation of the structure of the DNA-protein complex molecule leading to its dissociation or conformational change. In addition, since multiple collisions happen during the lifetime of the perturbation, a set of collisions with smaller energies can provide a coherent energy driving the processes. We here define "a hit" as a single collision as well as the set of coherent collisions inducing a chemical process. Its efficiency depends on its energy, its orientation, and its position on the surface of the complex. This efficiency is stochastic because the transferred energy will be dissipated into heat with different degrees depending on the configurations of the reactant macromolecule. The efficiency depends on the orientation and position of a hit. These are geometric factors and here supposed to be represented by x. For convenience of calculations, we denote $-E/k_BT$ by a negative variant n.

Let us consider the example that a hit can drive the single-step dissociation (k^U in Fig. 4A in the Main Text). The efficiency per unit time for a single hit is supposed to be a function, $f(k^U, n, x)$. Note that the variant k^U in the parenthesis merely indicates the process but not its value. Thus, the value of the rate constant k^U should equal the integral of the function with respect the energy (n) and all the geometric factors (x). At first, we obtain the integral with respect to the energy. The function f is null for the energy smaller than the potential difference between transition and ground states. It is also null for very large energy (large negative n), because such energy damages the macromolecule rather than drives a process of k^U. If we suppose m as (level of transition-state energy)/k_BT,

$$\int_{-\infty}^{0} f(k^U,n,x)\exp(n)dn = \int_{-\infty}^{-2m} f(k^U,n,x)\exp(n)dn = \int_{-\infty}^{-2m} f(k^U,n,x)dn\exp(-2m)$$

$$+\left[f(k^U,n,x)\right]_{-\infty}^{-2m}\int_{-\infty}^{-2m}\exp(n)dn = \int_{-\infty}^{-2m} f(k^U,n,x)dn\exp(-2m) \qquad \text{(S1)}$$

If the function F is defined as $F(process,x) \equiv \int_{-\infty}^{0} f(process, -n, x)dn$ (S2),

$$k^U = \int\int_{-\infty}^{0} f(k^U,n,x)\exp(n)dndx = \exp(-2m)\int F(k^U,x)dx \qquad \text{(S3)}.$$

This is the separation of rate constant into energy-dependent and energy-independent terms, and the latter integral can be called geometric factor. Similarly, for other processes in Fig. 4A in the Main Text,

$$k^{\mathrm{L}} = \exp(\mathrm{d}-\mathrm{m})\int \mathrm{F}(k^{\mathrm{L}}, x)dx \ , \ p^{\mathrm{UL}} = \exp(-\mathrm{d}-\mathrm{m})\int \mathrm{F}(p^{\mathrm{UL}}, x)dx,$$

$$\text{and } p^{\mathrm{LU}} = \int \mathrm{F}(p^{\mathrm{LU}}, x)dx \qquad \text{(S4).}$$

Since the major dissociation pathway is either k^{U} or $p^{\mathrm{UL}}k^{\mathrm{L}}/(p^{\mathrm{LU}}+k^{\mathrm{L}})$, the condition for the pathway via complex$_{\mathrm{B}}$ to be the major is

$$\frac{k^{\mathrm{L}}p^{\mathrm{UL}}}{k^{\mathrm{U}}\left(p^{\mathrm{LU}}+k^{\mathrm{L}}\right)} > 1, \text{ and } \quad \frac{p^{\mathrm{UL}}k^{\mathrm{L}}}{k^{\mathrm{U}}p^{\mathrm{LU}}} >> 1 \text{ is sufficient} \qquad \text{(S5).}$$

Among many possible models, we here pick up a model where complex$_{\mathrm{A}}$ and complex$_{\mathrm{B}}$ are respectively the most stable CamR complex with straight *camO* DNA and an unstable one with bent DNA, as has been already introduced for the study on binding between TrpR and *trpO*.[3,4] In Fig. S2A, we show potentials of mean force of complex$_{\mathrm{B}}$, and the transition state relative to its level of complex$_{\mathrm{A}}$. These potentials should not be confused with free energy containing entropy on the assumption of thermal equilibria. These potentials are determined by the structure of the macromolecule. Notably all the energy-dependent factors shown in Fig. S2A are compensated with each other except minor potential barrier from nonspecific site to bulk. Then the lower set of rate equation in Fig. 4A in the Main Text becomes the major dissociation pathway, if the parameter shown in Eq. (S6) is larger than unity.

$$\frac{p^{\mathrm{UL}}k^{\mathrm{L}}}{k^{\mathrm{U}}p^{\mathrm{LU}}} = \frac{\int \mathrm{F}(p^{\mathrm{UL}}, x)d\mathrm{x}}{\int \mathrm{F}(k^{\mathrm{U}}, x)dx} \ \frac{\int \mathrm{F}(k^{\mathrm{L}}, x)dx}{\int \mathrm{F}(p^{\mathrm{LU}}, x)dx} \qquad \text{(S6).}$$

In consideration of the relation between a chemical process and the collision by solvent molecules, we must take into account that proteins are generally less rigid than double-stranded DNA, as evidenced by their 100-fold difference in their persistent lengths.[5,6] The collision perturbs their structure only locally, but the perturbation spreads in a larger part. The collision energy sometimes dissipates into heat with no effect on their structure. In another case, the interacting domains in their complex are distorted to lose all their interactions between macromolecules by consuming a part of the energy of the collisions, namely dissociation. In yet another case, the perturbed structure is maintained for a while accompanied with a new solvation structure, namely conformational change. To break an interaction, a moment around the axis penetrating the interacting domain is efficient because of the large rotational inertia of macromolecule.

Considering the interaction of one or two interacting pairs of the bars and the ditches shown in Fig. S2B, the areas of highly effective hit points are speculated and shown in red gradation in Fig. S2C. These areas are either in the vicinity of the interacting domain or the area where a hit produces a rotational moment breaking the interaction. The effective areas of p^{UL} and k^{L} are likely to be larger than k^{U} and p^{LU}, respectively. Namely, for both the fractions in the right-hand side of Eq. S6, the nominators are speculated to be larger than the denominators, and thus the left-hand side is very likely to be larger than unity. This is a possible speculation explaining why TrpR and CamR binding

to their operators are chemical ratchet mechanisms. Therefore, there is at least one possible molecular model consistent with chemical ratchet mechanism, but there may be more at present.

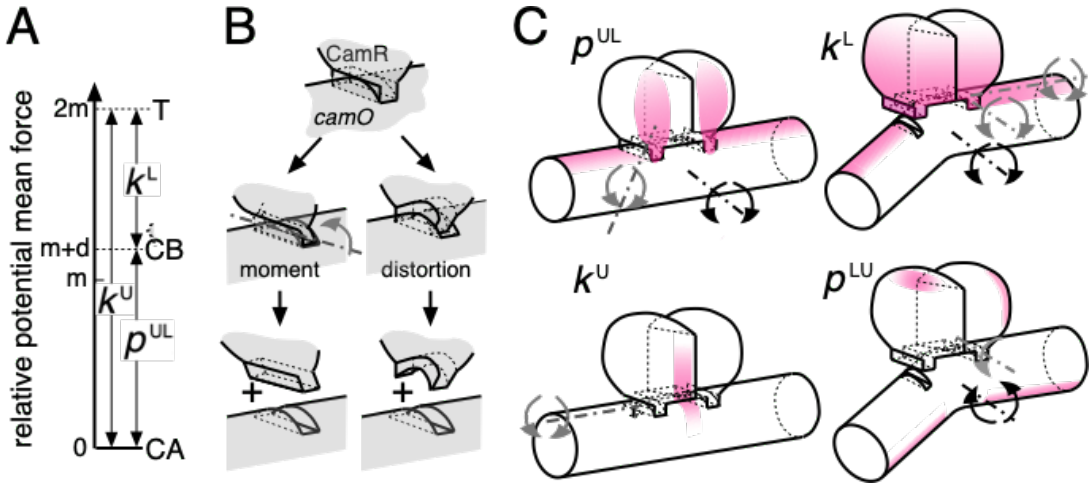

Figure S2. Chemical ratchet of CamR-*camO* binding.

(A) The potentials of mean force of complex$_A$ (CA), complex$_B$ (CB), the transition state of dissociation of a complex (T) which is close to the dissociated state. The bipartite interacting domain is drawn as two pairs of a square bar and a ditch, each pair of which is supposed to generate a stabilization energy of m. The energy required for bending the operator DNA is denoted by d. The definitions of k^U, k^L, p^{UL}, and p^{LU} are given in the kinetic scheme in Fig. 4A in Main Text.

(B) Mechanisms for breaking the interaction between CamR and *camO* by perturbing the structure. A distortion of CamR or DNA molecule by "a hit" is transmitted to the interacting domain and distorts it to destroy the interaction (right pathway). A hit generates a rotational moment (gray arrow in left pathway) around the axis penetrating the interacting domain (chained line) and strips the interacting domain of CamR from that of *camO*. The rotational inertia of a whole CamR molecule contributes to the strip.

(C) Symbolic illustration of a bacterial repressor and its operator. The areas of highly efficient hit points are indicated in red on the surface of the complex. The rotational moment effectively driving a process is indicated as a chain line and arc arrows for protein in gray and DNA in black. p^{UL}: complex$_B$ is produced by breaking either of the interacting pairs of complex$_A$. Because of the rigidity of DNA, the effective area is the large part of the upper side of DNA to bend it downward. In addition, a rotational moment with the axis parallel to DNA may peel off the protein rod from a DNA ditch. k^U: the two interacting domains of complex$_A$ must be broken at once, otherwise complex$_B$ is produced instead of the dissociation. Thus, the effective area is located in a thin plate perpendicular to DNA. Furthermore, most of the effective area on the upper side of DNA is protected by the protein from the hits, making the effective area smaller than that of p^{UL}. k^L: This process is induced both from bending DNA downward and generating rotational moments with the axis parallel and perpendicular to the protein rod. Thus, the area is large. p^{LU}: This process is induced by the rotational momenta of DNA and protein parallel to the protein rod. However, the formation of interaction requires more accurate orientations of the axes of rotation than its break. Thus the area of p^{LU} is smaller than that of k^L.

1. N. Shimamoto, Commentary: The limitation of the combination of transition state theory and thermodynamics for the reactions of proteins and nucleic acids, *Biomolecules*, 2022, **12**, 28.

2. A. Revzin and P. H. von Hippel, Direct measurement of association constants for the binding of *Escherichia coli lac* repressor to non-operator DNA, *Biochemistry*, 1977, **17**, 4769-4776.

3. T. Kinebuchi and N. Shimamoto, One-dimensional diffusion of TrpR along DNA enhances its affinity for the operator by chemical ratchet mechanism, *Sci. Re*p., 2021, **11**, 4255.

4. N. Shimamoto and M. Imashimizu, RNA polymerase and transcription mechanisms: The forefront of physicochemical studies of chemical reactions, *Biomolecules*, 2021, **11**, 32.

5. H. Li, A. F. Oberhauser, S. D. Redick, M. Carrion-Vazquez, H. P. Erickson and J. M. Fernandez, Multiple conformations of PEVK proteins detected by single-molecule techniques, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 10682-10686.

6. K. B Chhetri, A. Sharma, S. Naskar and P. K. Maiti, Nanoscale structures and mechanics of peptide nucleic acids, *Nanoscale*, 2022, **14**, 6620-6635.