One-dimensional Diffusion of Proteins along DNA

ITS BIOLOGICAL AND CHEMICAL SIGNIFICANCE REVEALED BY SINGLE-MOLECULE MEASUREMENTS

Nobuo Shimamoto‡
From the Structural Biology Center, National Institute of Genetics, Mishima, Japan 411-8540

Evidence for sliding of proteins along DNA has been provided by many kinetic studies, but single-molecule-based measurements have uncovered distinct problems, the solutions of which may lead us to an understanding of new mechanisms for gene regulation. Furthermore, they reveal a deep problem lying between chemistry and physics regarding the seemingly simple binding between DNA and protein. Single-molecule dynamics provides a tool to solve this problem without prejudgments or unsound assumptions.

Importance of Single-molecule-based Analysis

Single-molecule dynamics is the most powerful, and probably the only direct, method for analyzing a biological phenomenon in which the history of a single molecule appears at macroscopic level. The long-standing question as to whether or not proteins slide along nonspecific parts of DNA provides an excellent example of the power of the new method and its complementarity to classic kinetics. The sliding of a protein (a microscopic history) can macroscopically enhance the formation of specific complexes, as described below (see also Ref. 1 for review). The histories of individual molecules have traditionally been considered unimportant in chemistry, because the differences in histories of different molecules are normally obliterated due to numerous random collisions with solvent molecules. The resulting group of molecules with a unified average history is considered as a single chemical species. In practice, however, a chemical species is usually arbitrarily defined, and such a theoretical uncertainty causes a serious problem in kinetics, because kinetic analysis is based on hypotheses of mechanisms. Direct observation of individual molecules, on the other hand, requires no theoretical assumptions. This is the most important reason for the application of single-molecule dynamics to a problem like sliding.

The limitation of space here has urged exclusion of the sliding of DNA clumps and of the energy-driven translocation of enzymes such as helicases and RNA polymerase being engaged in RNA synthesis; the successful single-molecule studies of the latter have been recently reviewed (2).

Practical Definitions of Translocation Mechanisms

To avoid serious confusion due to the variable terminology used in past discussions of protein sliding, specific terms should first be defined clearly. One-dimensional diffusion of a protein here includes all mechanisms of translocation along a single DNA molecule that do not involve a free state of the protein. These mechanisms are classified into sliding and intersegment transfer (Fig. 1). The latter is here supposed to require at least two DNA binding sites on the protein molecule as in the general theory (3, 4), excluding a transient interaction as a secondary binding. Sliding and hopping are taken to imply, respectively, a helical movement due to tracking a groove of DNA or a non-helical movement parallel to the DNA axis (5, 6), although they have been differently defined theoretically (3). These distinctions are sacrificed here for simplicity, because a simple microscopic observation cannot distinguish them.

Single-molecule Dynamics of RNA Polymerase, CamR, and Photolyase

To visualize movement of a single molecule of protein using a commercially available fluorescent microscope, the molecule must be made strongly fluorescent. This can be done by attaching several tens of fluorophores per molecule (7, 8). The clearest visual assay of sliding movement is to let DNA take up a special geometry as in Fig. 2A and to detect the traces of protein molecules moving with the same geometry. A DNA concentration of 10–100 µg/ml is required to observe binding events at a significant frequency. These requirements were satisfied by applying dielectrophoresis to align extended DNA molecules in parallel (9).

Fluorescently labeled Escherichia coli RNA polymerase holoenzyme was injected so as to flow at an angle across the array of DNA molecules (8). Linear motions parallel to DNA were observed in half of the traces passing through the DNA region (Fig. 2A). The linear motions disappeared when holoenzyme was preincubated with heparin or with a short DNA fragment harboring a strong promoter. Furthermore, neither the IgG used for fluorescent labeling nor microrystals of rhodamine showed linear motions. These negative controls, together with the quantitative agreement between the lifetimes of the observed sliding complexes and those previously measured for nonspecific holoenzyme-DNA complexes (10), prove that the observed movement is a true sliding of holoenzyme along DNA and not a hydrodynamic artifact such as rectification of flow in the DNA region. The sliding complex is the only complex involved in nonspecific DNA regions, and stably bound complexes form only at the promoter. Moreover, a typical bacterial repressor, Pseudomonas putida CamR, which is a small homodimeric protein with a helix-turn-helix DNA binding motif, shows very similar movements.7 These results suggest that sliding is a general property of DNA-binding proteins, although not all might be able to slide.

The flow introduced in the above assay, or asymmetric collisions with solvent molecules, converts otherwise bidirectional sliding movements into a large unidirectional travel of several micrometers. The length of travel should not be confused with the sliding distance, which is defined as the mean size of DNA segment scanned per binding event (in the absence of flow), and has been kinetically estimated as 350–1000 base pairs for RNA polymerase (11, 12).

REFERENCES

1. H. Kabata and N. Shimamoto, unpublished results.

‡ To whom correspondence should be addressed. Tel.: 81-559-81-6843; Fax: 81-559-81-6844; E-mail: nshimamoto@LABSTRG-1.LAB.NIG.AC.JP.

© 1999 by The American Society for Biochemistry and Molecular Biology, Inc.

This paper is available on line at http://www.jbc.org

Printed in U.S.A.
Single-molecule Dynamics and Kinetics Are Complementary

The relationship between the new single-molecule-based techniques and classic kinetics resembles that between the new techniques developed in molecular biology and classical genetics, which have become fused and essential to modern biology. Kinetics can compensate for the defects of single-molecule dynamics, namely the heavy modification used for visualizing and detecting target molecules, and the unusual experimental conditions.

One-dimensional diffusion along DNA was first deduced from the finding that *E. coli* LacI binds to its operators on a 48-kilobase pair DNA 100-fold faster than by the fastest threedimensional diffusion theoretically predicted from the sizes of the reactants (14). A similar controversy was found in the binding of *E. coli* RNA polymerase (15). Since then kinetic evidence for sliding has been accumulated by using the assays shown in Fig. 2, B–F. The black box nature of kinetics would blur the meanings of evaluated kinetic parameters, but the most important qualitative conclusion, the existence of sliding, is very likely to be correct for the following proteins: restriction endonuclease EcoRI (5, 16–21), BssHI (6), HindIII (16), BamHI (16, 22), EcoRV (21), *E. coli* RNA polymerase (7, 8, 10–12, 23), LacI (24–30), GalR (30), EcoRI methylase (31, 32), BamHI methylase (22), UvrABC (33), λ Cro repressor (34), and T4 endonuclease V (35, 36). The kinetic length effects (Fig. 2B) observed for LacI were once attributed to intersegment transfer (37), but more critically designed experiments (Fig. 2D) established the coexistence of sliding (24).

**Merits and Demerits of Kinetic Techniques Used in the Test for Sliding**

Several quantitative techniques to detect DNA-protein complexes have been used in the kinetic test for sliding. The filter-binding technique has been widely used (14, 15, 18, 24, 26, 29, 31, 34, 37) but is open to misinterpretation if there are two or more types of complexes that are trapped with different efficiencies (37). If practicable, the gel shift assay (38, 39) is an excellent choice (24, 37, 40), because it can be made highly quantitative by introducing a competition between two DNA molecules trapped on a flat surface at a resolution smaller than 1 nm. By fixing DNA on a mica surface, the random displacements due to sliding of *Aspergillus nidulans* photolysase have been directly detected (13). The drawback of this method is slowing down of the movements of molecules because of interaction with a surface.

**Biological Significance of Sliding: Kinetic Contribution, Processivity, and More?**

Sliding along DNA obviously can kinetically affect a biological process through the acceleration of association, if bimolecular association is the rate-limiting step of the process. This situation, however, may only apply to limited cases in vivo. For example, transcription initiation includes several time-consuming elementary reaction steps. As a result the association of RNA polymerase with a promoter is rarely rate-limiting,
A. two-step dissociation mediated by sliding

B. direct dissociation from a specific site

C. thermodynamic length effect (antenna effect)
   weak affinity
   strong affinity

D. Model of regulation by limiting sliding
   weak affinity
   strong affinity

Fig. 3. The antenna effect and putative role of sliding in gene regulation. There are two mechanisms of dissociation from specific sites (yellow). A protein molecule (green) first slides into nonspecific sites and then dissociates into bulk (A). The dissociation rate of this mechanism depends on the length of nonspecific DNA, whereas it becomes independent by the direct dissociation mechanism (B). If this direct dissociation is combined with an association mediated by the sliding mechanism, a specific site on a longer DNA should have stronger affinity than the identical site on a shorter DNA, which is the antenna effect (C). A similar decrease in the affinity is expected when sliding movements are limited by the bindings of other proteins (pink) near the specific site (D), causing an action at a distance. The seeming violation of the microreversibility in the antenna effect is discussed in the text.

making the acceleration of association by sliding less significant. However, sliding may be more general in the production of processivity, as evidenced in restriction (21, 31, 32) and repair (33, 36) of DNA.

The Antenna Effect and a New Role of Sliding

If a protein slides from its specific site into nonspecific sites (Fig. 3A), longer DNA will accelerate the overall dissociation from the specific site. This length effect compensates the kinetic length effect on association (Fig. 2B), and binding affinity will be unchanged. Indeed such absence of the enhancement of affinity by sliding has been reported for EcoRI (18) and Cro repressor (34).

However, there are several cases showing enhancement of affinity by the extension of DNA (Fig. 3C); the observed enhancements are 20-fold for EcoRI methylase (31, 32), 10-fold for human immunodeficiency virus type 1 integrase (43), more than 100-fold for CamR, and 10–30-fold for LacI (26, 29). Interestingly the existence of sliding by RNA polymerase was detected by preferential occupancies among several identical promoters on the same DNA, using a design similar to that shown in Fig. 2C (12). Nonspecific DNA up to 350 base pairs in length enhanced affinity for its flanking promoter in an equilibrium condition, although this applied only when it was placed downstream of the promoter. These examples indicate that an extended region of nonspecific DNA can work as an antenna to harvest protein for binding to a flanking specific site.

To investigate this antenna effect without making any theoretical assumptions, the movement traces of molecules of CamR and RNA polymerase were compared, following dissociation from their specific sites in the presence of bulk flow. Most of the traces of CamR showed direct dissociation into bulk, or sliding movements too small to be resolved by the microscope, prior to complete dissociation (Fig. 3B). In contrast the traces of RNA polymerase showed sliding during the two-step dissociation (Fig. 3A) almost as extensive as during association. Accordingly the dissociation of CamR will be less enhanced than association by increases in DNA length, consistent with the observed large enhancement of its affinity. The asymmetry of the small antenna effect found for RNA polymerase (12) can also be explained along these lines. If RNA polymerase slides off the promoter more frequently in an upstream than a downstream direction because of the asymmetry of the promoter-holoenzyme interaction, the extension of downstream DNA would enhance the promoter affinity more than extension upstream, making the antenna effect asymmetric.

An alternative explanation for the antenna effect is that the additional length of DNA loops back onto the specific complex, permitting a transient secondary contact, which further stabilizes the specific complex. If longer DNA stabilizes the complex 100-fold by this mechanism, dissociation occurs only from a 1% fraction of complexes that are not looped, reducing the overall dissociation rate to 1% of that in the absence of looping (or less, if unlooping is rate-limiting in the overall dissociation process). This prediction, however, contradicts the observation that the dissociation rate is independent of DNA length in the case of EcoRI methylase (31). Furthermore, this model cannot explain the asymmetric antenna effect seen with RNA polymerase or the apparent absence of sliding upon dissociation of CamR. Therefore, sliding is the primary candidate as the basis of the antenna effect. If so, the role of sliding would be much more important than previously speculated; affinities for specific sites could be reduced by limitation of sliding movements (Fig. 3D). In fact EcoRI endonuclease is inhibited by various protein complexes locating near the cleavage site, as well as unusual DNA structures such as pseudo-recognition sites and triple helices (5). This is “action at a distance” in the true sense, which could be implicated in many gene regulation mechanisms such as nucleosome remodeling in transcription.

A Deep Insight into the Chemistry of Sliding Is Also Gained from Single-molecule Dynamics

In most kinetic studies nonspecific complexes are supposed to form a single chemical species, giving a two-step model: free components ↔ nonspecific complex ↔ specific complex. This assumption has been made arbitrarily, ignoring the definition of a chemical species, which states that every member of the species has an equal probability of conversion. Because nonspecific complexes at distal and proximal sites have different probabilities of sliding into a specific site, the above assumption in effect excludes the possibility of sliding from the outset. Alternatively a nonspecific complex at any position on DNA could be considered a distinct chemical species, as has been supposed in a general theory (see the “Appendix” of Ref. 3) (4).

If so, the energy barrier between contiguous nonspecific positions should be high enough to allow sufficient collisions with solvent molecules to randomize the histories of individual protein molecules while they remain at a nonspecific site. On the other hand the single-molecule assay revealed that this barrier is so small that the speed of sliding is about half of the parallel component of bulk flow (8). In other words, polymerase molecules move “half-freely” along DNA consisting of nonspecific sites. This may or may not contradict the distinct species at each position model but raises the third possibility that a group of nonspecific complexes within the sliding distance would more truly represent a chemical species than one at each position. Therefore, the true chemical status of nonspecific complexes is not yet clear.

The largest problem of the antenna effect, as has already been briefly noted (31), is a seeming contradiction to a basic assumption of chemistry. The principle of microreversibility, which is derived from the second law of thermodynamics and the definition of chemical species, demands an equal influx and efflux between a specific complex and the bulk free pool in
equilibrium if these two states constitute two chemical species that are directly connected by an elementary chemical step. The larger accumulation of a specific complex by the antenna effect should concomitantly increase the efflux from the specific site, $k_a \times [\text{specific complex}]$. In contrast the influx from the free state, $k_a \times [\text{free protein}]$, should be unchanged because both $k_a$ and $[\text{free protein}]$ are independent of the antenna effect, contradicting the principle. Therefore, something is missing in the chemical interpretation of the antenna effect.

This contradiction could be resolved in two ways, both proposing the existence of an additional chemical species. The looped complex model is easier to be understood but cannot explain some experimental results, as mentioned above. In an alternative model, the chemical species of the specific complex is defined similarly to the third possibility mentioned above so that it includes all complexes within the sliding distance from the specific site. In addition, an elongated space surrounding a DNA chain, which can be termed the “DNA domain,” is proposed to exist, and the protein molecules dissociating from the extended specific site initially enter this space and diffuse in it. As a distinct chemical species, the protein in this DNA domain can equilibrate both with free protein and with the extended specific complex. Increasing the length of DNA up to the sliding distance would also elongate the DNA domain and thus increase the influx from free state into the domain. The increased influx satisfies the principle because in equilibrium it becomes equal to the efflux into and the influx from the specific site, which are increased by the antenna effect.

Experimentally a space of this sort was found in rapid UV photo-cross-linking studies of sliding (10, 11, 23). RNA polymerase in a rapid flow was confined to this space yet in a state such that cross-linking was impossible, and the postulation of such a space was needed for numerical simulation.

More experiments would prove or disprove this model in the future, and they are very likely to be single-molecule measurements, because they should be designed so as to be free from traditional chemical assumptions. The above examples claim that the world of molecules may not always harmonize with our ordinary macroscopic experience not only in quantum mechanics but also in the chemistry and biology of seemingly simple DNA-protein interaction.

Acknowledgments—I thank Dr. R. S. Hayward (University of Edinburgh) for critical reading of the manuscript and Dr. J. C. Wang (Harvard University), Dr. P. Modrich (Duke University) and M. Tokunaga (National Institute of Genetics) for insightful discussions.

REFERENCES

1. von Hippel, P. H., and Berg, O. G. (1989) J. Biol. Chem. 264, 675–678
2. Gelles, J., and Landick, R. (1998) Cell 93, 13–16
3. Berg, O. G., Winter, R. B., and von Hippel, P. H. (1981) Biochemistry 20, 6929–6948
4. Berg, O. G., and Ehrenberg, M. (1982) Biophys. Chem. 15, 41–51
5. Jeltsch, A., Alves, J., Wolfs, H., Maass, G., and pingoud, A. (1994) Biochemistry 33, 10215–10219
6. Berkshout, B., and van Wamel, J. (1996) J. Biol. Chem. 271, 1837–1840
7. Shimamoto, N., Kabata, H., Kurosawa, O., and Washizu, M. (1992) in Structural Tools for the Analysis of Protein-Nucleic Acid Complexes (Lilley, D., Heumann, H., and Suck, D., eds) pp. 241–253, Birkhäuser Verlag AG, Basel
8. Kabata, H., Kurosawa, O., Arai, I., Washizu, M., Margarson, S. A., Glass, R. E., and Shimamoto, N. (1993) Science 262, 1561–1563
9. Washizu, M., Kurosawa, O., Arai, I., Suzuki, S., and Shimamoto, N. (1995) IEEE Trans. Ind. Appl. 31, 447–456
10. Park, C. S., Hillel, Z., and Wu, C.-W. (1982) J. Biol. Chem. 257, 6944–6949
11. Singer, P., and Wu, C.-W. (1987) J. Biol. Chem. 262, 14178–14189
12. Ricchetti, M., Metzger, W., and Heumann, H. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 4610–4614
13. van Noort, S. J., van der Werf, K. O., Eker, A. P., Wyman, C., de Grooth, B. G., van Hulst, N. F., and Greve, J. (1998) Biophys J. 74, 2840–2849
14. Riggs, A. D., Bougeois, S., and Cohn, M. (1970) J. Mol. Biol. 33, 401–417
15. Belinster, B. N., Zavviev, S. K., and Shemyakin, M. F. (1980) Nucleic Acids Res. 8, 1391–1404
16. Ehbrecht, H.-J., Pingoud, A., Urebehnke, G., Maass, G., and Gualerzi, C. (1985) J. Biol. Chem. 260, 6160–6166
17. Terry, B. J., Jack, W. E., and Modrich, P. (1985) J. Biol. Chem. 260, 13129–13137
18. Jack, W. E., Terry, B. J., and Modrich, P. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 4010–4014
19. Langowski, J., Alves, J., Pingoud, A., and Maass, G. (1983) Nucleic Acids Res. 11, 501–515
20. Jeltsch, A., Fritz, A., Alves, J., Wolfs, H., and Pingoud, A. (1993) Annu. Rev. Biochem. 62, 233–240
21. Jeltsch, A., Wenz, C., Stahl, F., and Pingoud, A. (1996) EMBO J. 15, 5104–5111
22. Nardone, G., George, J., and Chirikjian, G. J. (1986) J. Biol. Chem. 261, 12126–12133
23. Singer, P., and Wu, C.-W. (1988) J. Biol. Chem. 263, 4208–4214
24. Rusaala, T., and Crothers, D. M. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 4903–4907
25. Barkley, M. D. (1981) Biochemistry 20, 3833–3842
26. Winter, R. B., Berg, O. G., and von Hippel, P. H. (1981) Biochemistry 20, 6961–6967
27. Berg, O. G., and Blomberg, C. (1978) Biophys. Chem. 8, 271–280
28. Lohman, T. M., deHaseth, P. L., and Record, M. T. (1978) Biophys. Chem. 8, 281–294
29. Khoury, A. M., Lee, H. J., Lillis, M., and Lu, P. (1990) Bioclin. Biophys. Acta 1067, 55–68
30. Hsieh, M., and Brenowitz, M. (1997) J. Biol. Chem. 272, 22092–22096
31. Surby, M. A., and Reich, N. O. (1996) Biochemistry 35, 2209–2217
32. Surby, M. A., and Reich, N. O. (1996) Biochemistry 35, 2201–2208
33. Gruskin, E. A., and Lloyd, R. S. (1988) J. Biol. Chem. 263, 12738–12743
34. Kim, J. G., Takeda, Y., Matthews, B. W., and Anderson, W. F. (1987) J. Mol. Biol. 196, 149–158
35. Lloyd, R. S., Hanawalt, P. C., and Dodson, M. L. (1980) Nucleic Acids Res. 8, 5113–5127
36. Gruskin, E. A., and Lloyd, R. S. (1988) J. Biol. Chem. 263, 12728–12737
37. Fickert, R., and Mueller-Hill, B. (1992) J. Mol. Biol. 226, 59–68
38. Fried, M. G., and Crothers, D. M. (1981) Nucleic Acids Res. 9, 6505–6525
39. Garner, M. M., and Revzin, A. (1981) Nucleic Acids Res. 9, 2947–2960
40. Winter, R. B., and von Hippel, P. H. (1981) Biochemistry 20, 6948–6960
41. Yang, S., and Nash, H. A. (1995) EMBO J. 14, 6292–6300
42. Schuett, P. (1997) Annu. Rev. Biophys. Biomol. Struct. 26, 589–584
43. Lee, S. P., Censullo, M. L., Kim, H. G., and Han, M. K. (1995) Biochemistry 34, 10215–10223
44. Lieberman, B. A., and Nordeen, S. K. (1997) J. Biol. Chem. 272, 1061–1068