# DNA Looping Provides for "Intersegmental Hopping" by Proteins: A Mechanism for Long-Range Site Localization

**Adam J. Pollak, Aaron T. Chin, Frank L.H. Brown and Norbert O. Reich**

***Department of Chemistry and Biochemistry,*** *University of California at Santa Barbara, Building 557, Room 1432, Santa Barbara, CA 93106-9510, USA*

***Correspondence to Norbert O. Reich:*** *reich@chem.ucsb.edu*
http://dx.doi.org/10.1016/j.jmb.2014.08.002
***Edited by K. Severinov***

## Abstract

Studies on how transcription factors and DNA modifying enzymes passively locate specific sites on DNA have yet to be reconciled with a sufficient set of mechanisms that can adequately account for the efficiency and speed of this process. This is especially true when considering that these DNA binding/modifying proteins have diverse levels of both cellular copy numbers and genomic recognition site densities. The monomeric bacterial DNA adenine methyltransferase (Dam) is responsible for the rapid methylation of the entire chromosome (with only ~100 Dam copies per cell) and the regulated methylation of closely spaced sites that controls the expression of virulence genes in several human pathogens. Provocatively, we find that Dam travels between its recognition sites most efficiently when those sites are ~500 bp apart. We propose that this is manifested by Dam moving between distal regions on the same DNA molecule, which is mediated by DNA looping, a phenomenon we designate as intersegmental hopping. Importantly, an intermediate found in other systems including two simultaneously bound, looped DNA strands is not involved here. Our results suggest that intersegmental hopping contributes to enzymatic processivity (multiple modifications), which invoke recent reports demonstrating that DNA looping can assist in site finding. Intersegmental hopping is possibly used by other sequence-specific DNA binding proteins, such as transcription factors and regulatory proteins, given certain biological context. While a general form of this mechanism is proposed by many research groups, our consideration of DNA looping in the context of processive catalysis provides new mechanistic insights and distinctions.

Published by Elsevier Ltd.

## Introduction

Site-specific DNA binding proteins, such as transcription factors, DNA methyltransferases, restriction endonucleases (ENases), and DNA repair enzymes, locate their recognition sites with remarkable speed and efficiency. While nonspecific DNA surrounding these sites aids in the search [1], the mechanism of how proteins move passively (without ATP) and rapidly along nonspecific DNA remains incompletely understood. In particular, current translocation models that largely invoke sliding and hopping processes, which are based on extensive *in vitro* [2] and recent *in vivo* studies [3], are inadequate to solely account for searches over thousands and hundreds of thousands of base pairs routinely achieved by these proteins. With sliding [4,5], the protein maintains continuous contact with DNA, like a roller coaster car on a track (Fig. 1). Each movement is a single-base-pair step in either direction, likely constraining each protein to short stretches of DNA (<100 bp). Hopping involves a rapid series of dissociation and reassociation events, where the protein remains loosely associated with the DNA (Fig. 1b) [6–8]. Details about the average trajectory of each hop, however, remain incompletely characterized. In general, hopping allows for larger movements along DNA than for sliding, but neither mechanism has been shown to contribute to kilobase searches. Sliding/hopping mechanisms are likely coupled with other, presently ambiguously described long-range 3-D (*3-d*imensional) movements [9], which we attempt to further elucidate here.

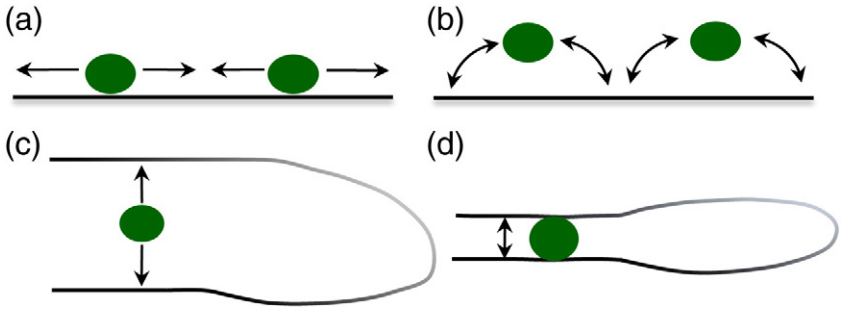


**Fig. 1.** Translocation models. (a) Sliding involves a close association between the protein and the DNA. (b) Hopping includes small jumps. (c) Intersegmental hopping (proposed here) combines aspects of hopping and intersegmental transfer for fluid, long-range protein translocation. (d) A ternary complex is obligatory for proteins to move between distal DNA segments for intersegmental transfer.

The highly processive *Escherichia coli* DNA adenine methyltransferase (Dam) is an ideal model system to study long-range translocation mechanisms: relatively few enzymes (~100 Dam enzymes per cell) methylate ~20,000 recognition sites (5′-GATC-3′) at the $N^6$ position of adenine in approximately 20 min, an essential component of mismatch repair [10]. Dam is interestingly involved in several other processes in the cell, including epigenetic gene regulation [10,11]. Other low-copy-number proteins (e.g., transcription factors) also must transverse long segments of the genome to locate their site of action (see Discussion). Here, we use an activity-based assay and a series of two-site DNA substrates to study processivity (Fig. 2) to deduce translocation mechanisms. Processivity ($F_p$) is quantified as the probability that Dam methylates both sites within a single binding event between the enzyme and the DNA.

Changes in processivity with increases in intersite distances are commonly used and are a powerful tool to deduce translocation mechanisms and the details concerning them [2,12,13]. With sliding, processivity will decrease more sharply with increasing intersite distances than for hopping [12,14] (see below). While the translocation kinetics of some proteins are well described by the sliding and/or hopping processivity [Eq. (2)], several reports contain data that cannot be reconciled with these models (see Discussion). Both mechanisms involve movement following the contour of the helix and predict that processivity should decrease as intersite

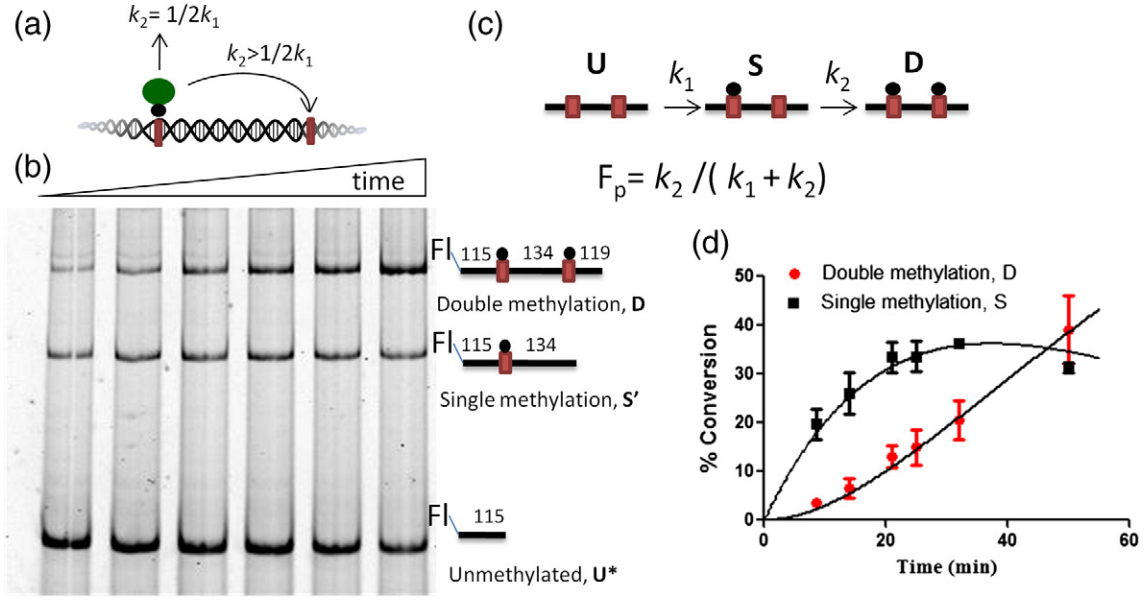


**Fig. 2.** Processivity assay. (a) Dam (green circle) immediately following the first methylation (black circle) at a GATC site (red bold lines). Processivity is based on the probability of a second methylation event following an initial one during a single binding event. (b) The reaction includes 7 nM Dam, 400 nM DNA (symmetric substrate, fluorescein labeled), 30 μM SAM, 30 mM NaCl, and buffer at 37 °C. Heat-quenched samples were digested with DpnII endonuclease (cuts unmethylated GATC sites), and reaction products were separated by nondenaturing PAGE and imaged using a typhoon phosphoimager. (c) The relative amount of each species of the reaction is fit to a sequential reaction mechanism model to derive the rate constants $k_1$ and $k_2$ and, ultimately, the processivity value ($F_p$) [Eqs. (4)–(6)]. (d) Fit of data to model for substrate with 134 bp between sites (5 trials, mean and standard deviation are plotted, Substrate 4B-2, Table S1).