Supporting Information: DNA-mediated signaling by proteins with 4Fe-4S clusters is necessary for genomic integrity

Michael A. Grodick,[1] Helen M. Segal,[1] Theodore J. Zwang,[1] and Jacqueline K. Barton[1,*]

[1]Division of Chemistry and Chemical Engineering,

California Institute of Technology, Pasadena CA 91125

* To whom correspondence should be addressed, E-mail: jkbarton@caltech.edu.

SUPPLEMENTARY ANALYSIS

Atomic force microscopy assay. An alternative way to calculate the binding density ratio is to treat each individual image as a sample, to plot them as a histogram (Figure S3) and to find the mean of the normal distribution. Using this methodology, instead of treating each surface as an individual experiment, a binding density ratio of 1.61 ± 0.08 is found for DinG with mismatched DNA. A binding densitry ratio of 1.40 ± 0.05 is found for DinG mixed with EndoIII and mismatched DNA. Finally a binding density ratio of 0.91 ± 0.04 is found for DinG mixed with EndoIII Y82A.

Because the AFM images are snapshots of the system near equilibrium, the number of bound proteins on strands reflects the apparent non-specific binding affinity of the proteins for DNA. As such, another way to analyze the data would be to qualitatively analyze the apparent binding affinity of each protein in the different experiments by visualizing the occupancy of the DNA i.e. the percent of DNA with a minimum of 1, 2, 3, or 4 proteins on a strand of DNA. Since DNA-binding proteins with 4Fe-4S clusters have a lower DNA binding affinity when the cluster in the reduced vs. oxidized (19), in a collection of proteins, some reduced and some oxidized, CT between proteins should affect the overall effective binding affinity for DNA. Conversely, if only one protein is bound to DNA, its affinity would not be affected by DNA CT. This was observed for mixtures of DNA and DinG (Figure S4), which also indicates that DinG does not preferentially bind a mismatch. The same was observed for mixtures of DinG with WT EndoIII, or with EndoIII Y82A; the percent of long, well-matched strands with one protein bound is the same for DinG/EndoIII and DinG/Y82A mixtures (Figure S5). If more than one protein is bound to a given strand, however, each subsequent protein that binds has a probability to transport charge through the DNA and promote another protein's dissociation, in a way that parallels anti-

cooperative binding. If a mismatch is present, which attenuates CT, however, this anti-cooperative effect would be lessened as it would if EndoIII Y82A is substituted for EndoIII, decreasing the concentration of signaling partners. This is precisely what is observed (Figures S4 and S5). The percentage of long strands with a minimum of 2, 3, or 4 bound DinG proteins is lower for DinG when there is no mismatch, then when there is a mismatch. When DinG is mixed with EndoIII compared to EndoIII Y82A, an anti-cooperative effect is once again observed. There is a decreased percentage of long strands of DNA bound with a minimum of 2, 3, or 4 proteins for DinG and EndoIII compared to DinG and EndoIII Y82A. This observation, especially when considered in conjunction with their similar population of single-bound-protein-DNA complexes, further supports the ability of DinG and EndoIII to cooperate using DNA-mediated CT.

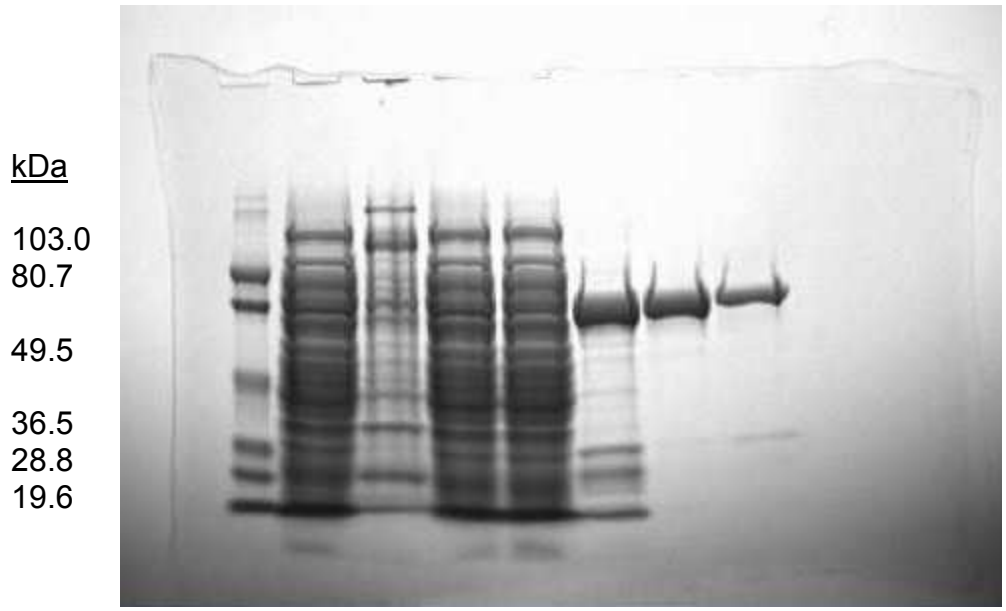

Figure S1. SDS-PAGE gel for purification of DinG. Lanes are referred to as 1 to 8 from left to right. The lanes contain an SDS-PAGE weight standard-low range (Biorad) (Lane 1), the supernatant from the cell lysate (lane 2), the pellet from the cell lysate (lane 3), the filtered cell lysate (lane 4), the the Histrap HP column flow-through (lane 5), the collected fractions from the Histrap HP column (lane 6), the collected fractions from the Superdex 200 column (lane 7), and a 10x dilution of the stored protein after thawing (lane 8). Corresponding molecular weights for each of the six bands in the weight standard lane are designated to the left of the image.

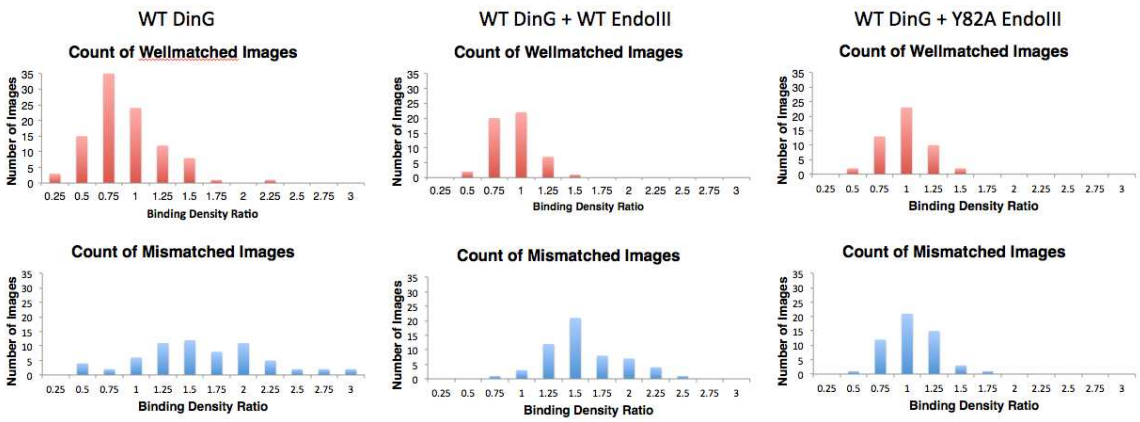

Figure S2. Statistical data for the AFM experiments. Histograms showing the distribution of binding density ratios within the population of sample images. The value on the x-axis is the upper range of a .25 unit wide separation, meaning that the column plotted above "1" is the number of images with a binding density ratio between 0.75 and 1.

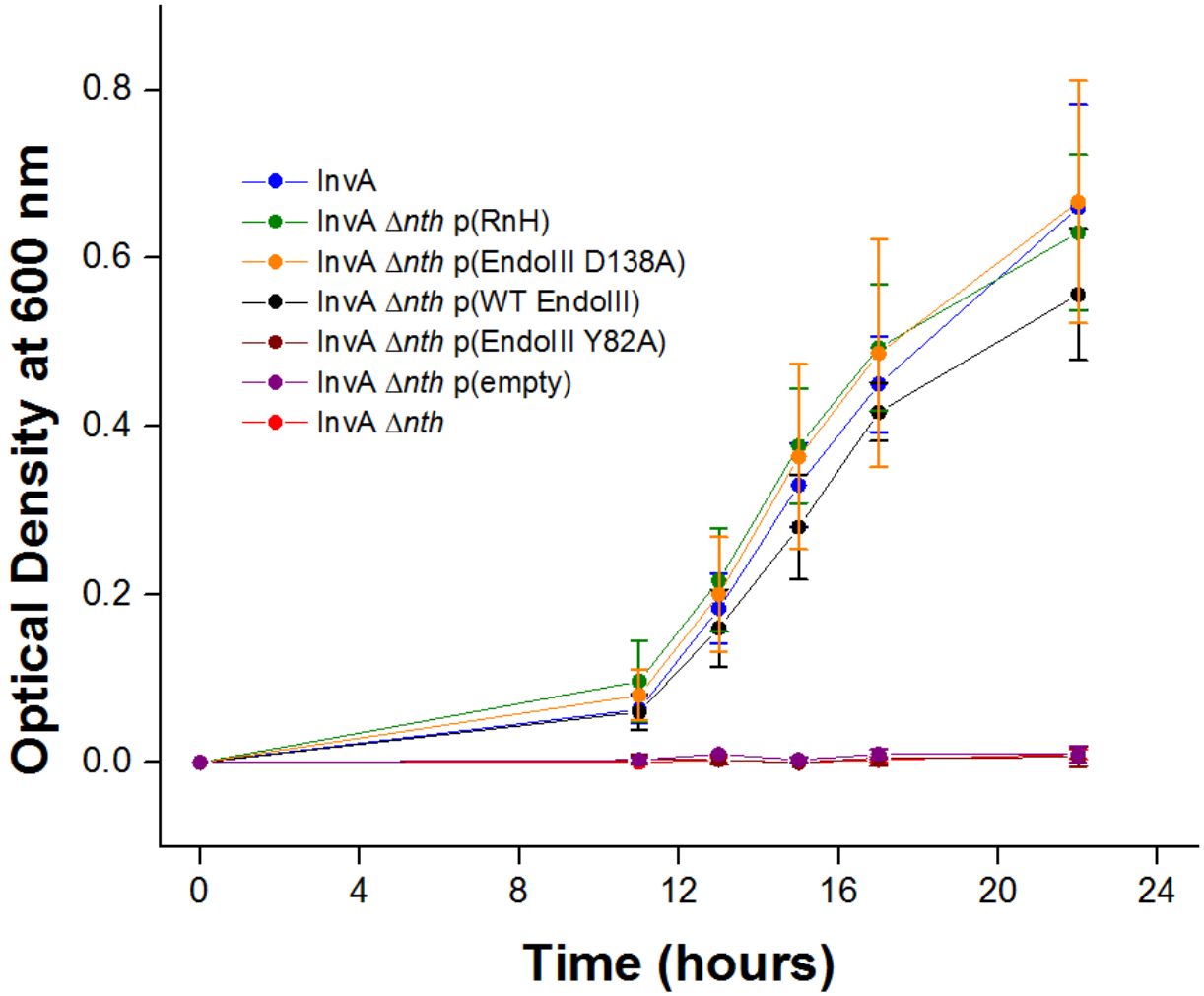

Figure S3. Effect of *nth* deletion in InvA (InvA Δ*nth*). Cultures of MM (M9 + 0.2% glucose) or MM + ampicillin (100 µg/mL) were inoculated with single colonies of the InvA parent strain (blue), InvA Δ*nth* p(RNaseH) (green), InvA Δ*nth* p(EndoIII D138A) (orange), InvA Δ*nth* p(WT EndoIII) (black), InvA Δ*nth* p(EndoIII Y82A) (brown), InvA Δ*nth* p(empty) (purple), or InvA Δ*nth* (red). Growth was monitored over time. ± SEM

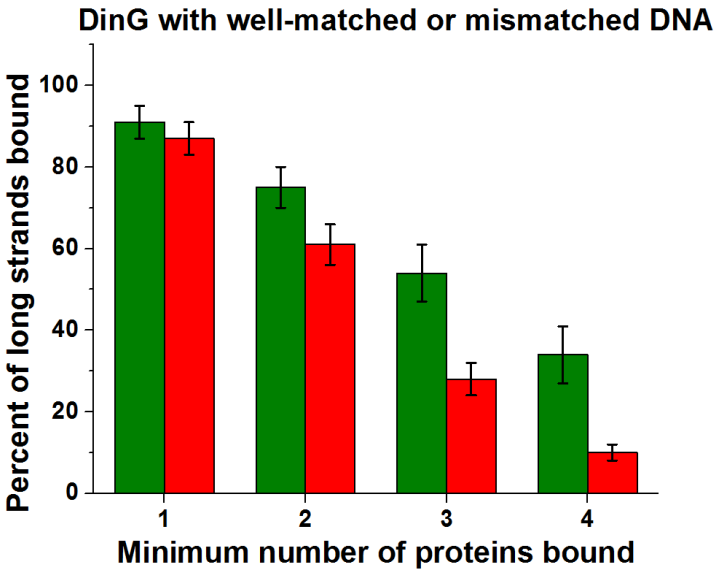

Figure S4. Occupancy of DinG on long strands of DNA. The percent of long strands of either well-matched (green) or mismatched (red) DNA is plotted against the minimum number of proteins bound to a strand.

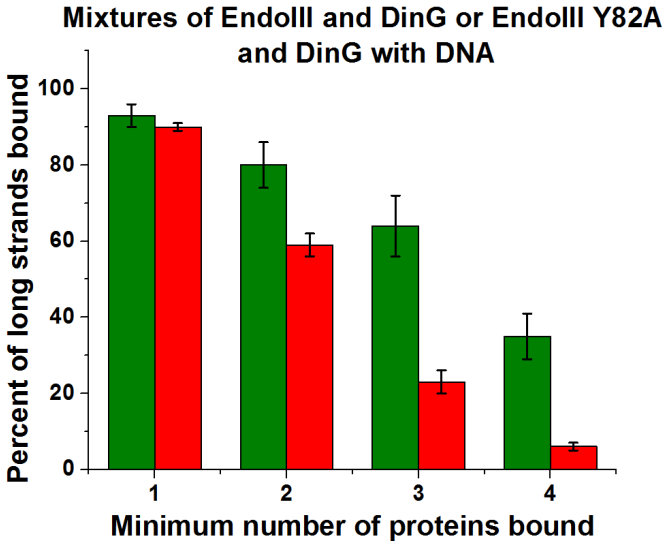

Figure S5. Occupancy of either EndoIII and DinG or EndoIII Y82A and DinG on long strands of well-matched DNA. The percent of well-matched long strands of either EndoIII and DinG (green) or EndoIII Y82A (red) DNA is plotted against the minimum number of proteins bound to a strand.

Supplementary Table 1. Oligonucleotides for electrochemistry substrates, gene replacements, sequencing or colony PCR, and site-directed mutagenesis

Note: All oligonucleotides except for the modified strand for electrochemistry were purchased from IDT

Designation and use	**Sequence (5' to 3')** (bases highlighted in red yield the point mutation after SDM)
Primer for cloning *dinG* forward	GGTTTTCCCATGGCATTAACCGCC
Primer for cloning *dinG* reverse	CATCATTAAAGCTTCCGACGGCGT
pET28b-*dinG* insert sequencing forward 1	T7 promoter primer
pET28b-*dinG* insert sequencing forward 2 (within gene)	ACTGACGCCGAACAATCAGGA
pET28b-*dinG* insert sequencing reverse 1	T7 terminator primer
pET28b-*dinG* insert sequencing reverse 2 (within gene)	TTCGGCAAATGACTGTAAGCCCAC
Substrate for electrochemistry – 20-mer thiolated modified strand	HS-C6-GTGCTGCAACGTGTCTGCGC (annealed with either the well-matched complement or abasic complement to yield the substrate used in experiments)
Substrate for electrochemistry – 35- mer complementary strand for well-matched substrate	AGACTGCAGACGAGAGCGCAGACACGTTGC AGCAC
Substrate for electrochemistry – 35- mer complementary strand for abasic substrate	AGACTGCAGACGAGAGCGCAGACACGTTGC A_CAC ("_" represents an abasic site)
AFM substrates	3.8 kb long strands and 1.6 or 2.2 kb shorts strands were prepared as described previously (22, 35)
Δ*dinG*::cm^R, forward primer	CCGAAAAATGCCACAATATTGGCTGTTTAT ACAGTATTTCAGGTTTTCTCGTGTAGGCTGG AGCTGCTTC
Δ*dinG*::cm^R, reverse primer	CCGAAAAATGCCACAATATTGGCTGTTTAT ACAGTATTTCAGGTTTTCTCGTGTAGGCTGG AGCTGCTTC
Δ*dinG*::cm^R, sequencing forward	GATGGTGTCTTGCATGACGTG
Δ*dinG*::cm^R, sequencing reverse	TCAATACGCCGCCCAACTCA
SDM reverse primer for generation of pBBR1-MCS4 *nth* Y82A	CGATTGGGCTTGCTAACAGCAAAGCAGAAA ATATCATCAAAACCTGC
SDM forward primer for generation of pBBR1-MCS4 *nth* D138A	CTATTCGTGTCGCCACGCACATTTTCCGCGT TTGTAATC
SDM reverse primer for generation of pBBR1-MCS4 *nth* D138A	CGGAAAATGTGCGTGGCGACAGCAATAGTC GGCCAGC

pBBR1MCS-4-*nth* sequencing forward	GGTGCTGATGCCGCTGGCGATTCAG
pBBR1MCS-4-*nth* sequencing reverse	TGTGCTGCAAGGCGATTAAGTTGG
Genomic *nth* region check forward 1	GAGATCCGCATTCCCATTTA
Genomic *nth* region check reverse 1	GGCTTAACGGCGATATGTTC
InvA check 1 from (33)*	CCAGTCATTTGGCGAAAG
InvA check 2 from (33)*	GGCGTAATAGCGAAGAGG

*Used together as primers in PCR to amplify a ~4250 bp product if *rrnA* is inverted

Supplementary Table 2: Plasmids used in this study

Plasmid Designation	Description	Source, reference, or method to construct
pBBR1MCS-4 or p(empty)	pBBR1MCS-4, a vector for the constitutive expression of genes placed in the MCS (multiple cloning site)	(22)
pBBR1MCS-4-*nth* or p(WT EndoIII)	pBBR1MCS-4 carrying the *nth* gene in the MCS, constitutively expresses WT EndoIII	(22)
pBBR1MCS-4-*nth* D138A or p(EndoIII D138A)	pBBR1-MCS4 carrying the *nth* D138A gene in the MCS, constitutively expresses EndoIII D138A	Site-directed mutagenesis of pBBR1MCS-4-*nth* using SDM primers listed in Table S3.
pBBR1MCS-4-*nth* Y82A or p(EndoIII Y82A)	pBBR322-MCS4 carrying the *nth* Y82A gene in the MCS, constitutively expresses EndoIII Y82A	Site-directed mutagenesis of pBBR1MCS-4-*nth* using Y82A forward and reverse primers listed in Table S2
pET-28b-*dinG*	Overexpresses DinG in presence of IPTG	Insertion of *dinG* amplicon into pET-28b(+) (Novagen) as described above (30)
pEM-Ap^R or p(RNaseH)	pACYC184 derived vector carrying the rnh^+ gene and an Ap^R gene, overexpresses RNaseH	Gift from Dr. Bénédicte Michel (Le centre de la Recherche Scientifique, Gif-sur-Yvette, France) (33)

Supplementary Table 3. Strains used in this study

Strain	Genotype designation	Source, reference, or method to construct
***Lac*$^+$ forward reversion assay, a GC:TA transversion assay to probe MutY activity**		
CC104	F128-(CSH104) *lacI373*, *lacZ574*, *ara-600*, *Δ(gpt-lac)5*, *λ*$^-$, *relA1*, *spoT1*, *thiE1*	(22)
CC104 Δ*dinG*	CC104 Δ*dinG*::cm^R	Inactivation of *dinG* by replacement with cm^R using pKD3 and pKD46 (CGSC & 7)
InvA Δ*nth* R-loop assay to probe DinG activity		
JW1625-1	F-, *Δ(araD-araB)567*, *ΔlacZ4787*(::rrnB-3), *λ*$^-$, *Δnth-782::kan*, *rph-1*, *Δ(rhaD-rhaB)568*, *hsdR514*	Obtained from Coli Genetic Stock Center (Yale University, New Haven, CT)
BW16847	F-, *Δ(codB-lacI)3*, *ΔphoA532*, *pdxH15*(Am), *purR106::Tn10*, *Δ(phnP-phnD)3330(phnC?)*	Obtained from Coli Genetic Stock Center (Yale University, New Haven, CT)
InvA	MG1655 ΔlacZ ΔattB::spcR Inv (attL15-cm^R attR75-kan^R)	Gift from Dr. Bénédicte Michel at the Centre de Génétique Moléculaire (CNRS) (33)
MG001	BW16847 Δ*nth*::kan^R	BW16847 * P1 JW1625-1
InvA Δ*nth*	InvA Δ*nth*::kan^R , *pdxH15*(Am), *purR106::Tn10*	MG001 * P1 InvA and colony PCR to find colonies with the Δ*nth*:: kan^R locus
InvA Δ*purR*	InvA *purR106::Tn10*	MG001 * P1 InvA and colony PCR to find colonies without the Δ*nth*:: kan^R locus