Published in final edited form as:

Angew Chem Int Ed Engl. 2008 ; 47(52): 10192–10194. doi:10.1002/anie.200803588.

Electron Spin Resonance shows common structural features for different classes of EcoRI-DNA complexes

Katherine M. Stone⁺,

Department of Chemistry, University of Pittsburgh, 219 Parkman Ave, Pittsburgh, PA 15260

Jacqueline E. Townsend⁺,

Department of Biological Sciences, University of Pittsburgh, 320 Clapp Hall, Pittsburgh, PA 15260

Jessica Sarver,

Department of Chemistry, University of Pittsburgh, 219 Parkman Ave, Pittsburgh, PA 15260

Paul J. Sapienza,

Department of Biological Sciences, University of Pittsburgh, 320 Clapp Hall, Pittsburgh, PA 15260

Sunil Saxena^{*} [Prof], and

Department of Chemistry, University of Pittsburgh, 219 Parkman Ave, Pittsburgh, PA 15260

Linda Jen-Jacobson^{*} [Prof]

Department of Biological Sciences, University of Pittsburgh, 320 Clapp Hall, Pittsburgh, PA 15260

Keywords

EPR spectroscopy; DEER; DNA binding protein; structure elucidation; protein flexibility; non-cognate complex

In this communication, we show that the EcoRI restriction endonuclease binds different classes of DNA sites in the same binding cleft. EcoRI generates widespread interest because it exhibits an extraordinary sequence selectivity to carry out its function of cleaving incoming foreign DNA without causing potentially lethal cleavage of cellular DNA. For example, EcoRI binds to its correct recognition site GAATTC up to 90,000-fold better than to miscognate sites that have one incorrect base pair.[1,2] The ~650 specific sites in the *E. coli* genome are protected from cleavage by double-strand methylation. The ~21,000 miscognate sites are not methylated, but are still cleaved by EcoRI with a second-order rate constant that is ~10⁹-fold lower.[1,2] EcoRI forms only non-specific complexes with no cleavage at sites that differ from GAATTC by two or more base pairs.[1,2]

In order to understand the source of such high specificity, it is necessary to determine how the structures of EcoRI complexes differ at specific, miscognate (5/6 bp match), and non-specific ($\leq 4/6$ bp match) DNA sites. This effort is timely given the extensive genetic, biochemical and biophysical data on EcoRI.[1-9] Footprinting results[1] suggest that the three classes of complexes are “structurally” distinct, and thermodynamic profiles (ΔG° , ΔH° , ΔS° , ΔC°_p)[3,4] suggest that the specific complex has more restricted conformational-vibrational mobility of the protein and DNA. There are crystal structures of the free protein,

^{*}Department of Chemistry, University of Pittsburgh, 219 Parkman Ave, Pittsburgh, PA 15260, Fax: (+1) 412-624-8611, sksaxena@pitt.edu, www.chem.pitt.edu/saxena, Department of Biological Sciences, University of Pittsburgh, 320 Clapp Hall, Pittsburgh, PA 15260, Fax: (+1) 412-624-4759, ljen@pitt.edu.

⁺These authors contributed equally to this work.

Supporting information for this article is available on the WWW under <http://www.angewandte.org> or from the author.

[6] and the metal-free specific protein-DNA complex.[6,7] Miscognate and non-specific complexes however, have not been readily accessible to crystallographic analysis. Indeed, for the ~3600 known restriction endonucleases, there are currently 73 crystal structures of 38 distinct enzymes in complex with specific DNA. However, there are only 4 structures of miscognate or non-specific complexes in the protein data bank.[10] In this communication, we demonstrate the utility of pulsed ESR distance measurements to shed light on miscognate and non-specific complexes.

Figure 1 shows the structure of the EcoRI specific complex.[6,7] The protein contains a large, relatively rigid and structured globular “main” domain and a smaller “arm” region. The protein arms are invisible in the free protein[6] but become ordered and enfold the DNA in the specific complex, where they play a role in modulating specificity.[2,4] Mutations R131C, S180C, and K249C-S180C were chosen based on the crystal structure. [6,7] These sites are solvent accessible and therefore likely to spin label with minimal perturbation to protein structure. Residues R131 and S180 lie in the inner and outer arms, respectively. Residue K249 is in the main domain, which has very restricted movement[6] and acts as a reference point. Since EcoRI is a 62 kDa homodimer, single cysteine mutations provide two sites for spin labeling, and double mutations provide four sites.

The proteins were spin labeled at the cysteines with the methanethiosulfonate spin label (MTSSL). There is an intrinsic cysteine at position 218, but it is buried, leading to <10% labeling even with a 100-fold molar excess of the spin label. The mutant proteins and their spin labeled derivatives catalyze DNA cleavage and have DNA binding affinities similar to that of wild type EcoRI, indicating that they are functionally active (Supporting Information).

DEER experiments[11] were performed on spin labeled S180C specific and non-specific complexes, and on R131C and K249C-S180C specific, miscognate, and non-specific complexes. The DEER experiment[11] is now well established for measuring distance constraints in membrane proteins,[12,13] soluble proteins,[14,15] peptides,[16] oligonucleotides,[17,18] and synthetic oligomers.[19-21] Recently, the DEER experiment has been used to probe structural rearrangements upon metal binding in the anthracis repressor, a DNA binding protein.[22]

The four-pulse DEER data for each of the mutant complexes are shown in Figure 2a. The time domain signals were inverted to obtain the distance distribution functions, using a Tikhonov regularization method in the DEERAnalysis2006 program.[23] The resulting distance distribution functions are shown in Figure 2b.

The most probable distances between the spin labels for the R131C EcoRI specific, miscognate, and non-specific complexes are 35 Å, 36 Å and 35 Å, respectively. The R131 C β -C β distance in the crystal structure of the specific complex is 32 Å.[6,7] The interspin distance measured by ESR is expected to differ because of the added length of the spin label. The most probable distance for the S180C mutant in the specific and non-specific complexes is 64 Å. To enable measurement of such a large distance, a large volume of S180C in 30% deuterated glycerol, 65% deuterated water, and 5% protonated water was used, and the temperature was lowered to 40K. With the enhanced signal and increased phase memory time (3 μ s), a sufficiently long dipolar evolution time could be collected (Figure 2a, middle panel).

For specific, miscognate and non-specific complexes of the K249C-S180C mutant protein, the most probable experimental distance was 33 Å in all cases (Figure 2b, lower). In principle, multiple distances corresponding to S180C-S180C, K249C-K249C, and S180C-K249C distances are anticipated for the K249-S180C double mutant. The corresponding C β -

C β distances in the specific complex crystal structure are 27 Å (S180C-K249C intra-monomer), 59 Å (S180C-S180C), 60 Å (K249-K249), and 57 Å (S180-K249 inter-monomer).[6, 7] It is likely that the larger distances were not detected in this series of experiments given that only ~1.5 μ s of the data could be collected due to short phase memory times. The 33 Å peak for the double mutant can thus be assigned to the S180C-K249C intra-monomer distance.

Strikingly, the experimental point-to-point distances are very similar for specific, and non-cognate (i.e. miscognate and non-specific) EcoRI-DNA complexes. The data show preservation of the distances between the inner arms (R131C data), outer arms (S180C data), and from the outer arm (S180C) to a fixed reference point (K249C) in the main domain. For both the R131C and K249C-S180C mutant proteins, the distance distribution is narrower for the specific complex than for the corresponding non-cognate complexes. This might indicate a greater flexibility of the arms in the EcoRI complex with non-cognate DNA. Further ESR experiments that probe dynamics are underway to confirm this hypothesis. In addition, the distributions for both the R131C inter-arm distance and the K249C-S180C distance show asymmetries in the non-cognate complexes. However, it is unclear if this represents an asymmetric set of accessible conformations of the arms or different orientations accessible to the spin labels.

Taken together the data suggest that on average the EcoRI arms envelop the DNA and are similarly oriented in non-cognate and specific DNA complexes. This implies that the DNA in the specific and non-specific complexes occupies roughly the same binding cleft of the EcoRI dimer. In addition, slopes of the salt dependence for formation of specific and non-specific complexes are the same ($d \log KA/d \log [NaCl] \sim -11$)[24] and are consistent with the number of Coulombic interactions observed in the specific complex. This provides additional strong evidence that the arms enfold the DNA in the non-specific complex. This enfolding may contribute to processivity as the protein slides along non-specific DNA[8,9,25,26] to locate its specific recognition site. Our results on a DNA-protein complex by pulsed ESR establish a methodology that can measure the solution structure and range of conformational states for complexes with different classes of DNA sites for which there is little or no prior structural information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by an NSF CAREER grant (MCB 0346898) to S.S. and an NIH MERIT 5R37GM029207 grant to L.J.-J.

References

1. Lesser D, Kurpiewski M, Jen-Jacobson L. *Science*. 1990; 250:776–786. [PubMed: 2237428]
2. Sapienza P, dela Torre C, McCoy W, Jana S, Jen-Jacobson L. *J Mol Bio*. 2005; 348:307–324. [PubMed: 15811370]
3. Jen-Jacobson L, Engler L, Ames J, Kurplewski M, Grigorescu A. *Supramol Chemistry*. 2000; 12:143–160.
4. Sapienza P, Rosenberg J, Jen-Jacobson L. *Structure*. 2007; 15:1368–1382. [PubMed: 17997963]
5. Kurpiewski M, Engler L, Wozniak L, Kobylanska A, Koziolkiewicz M, Stec W, Jen-Jacobson L. *Structure*. 2004; 12:1775–1788. [PubMed: 15458627]
6. Grigorescu, A.; Horvath, M.; Wilkosz, P.; Chandrasekhar, K.; Rosenberg, J. *Restriction Endonucleases*. Springer-Verlag; 2004.

7. Kim Y, Grable J, Love R, Greene P, Rosenberg J. *Science*. 1990; 249:1307–1309. [PubMed: 2399465]
8. Pingoud A, Fuxreiter M, Pingoud V, Wende W. *Cell Mol Life Sci*. 2005; 62:685–707. [PubMed: 15770420]
9. Wright D, Jack W, Modrich P. *J Biol Chem*. 1999; 274:31896–31902. [PubMed: 10542216]
10. Berman H, Westbrook J, Feng Z, Gilliland G, Bhat T, Weissig H, Shindyalov I, Bourne P. *Nucleic Acids Res*. 2000; 28:235. [PubMed: 10592235]
11. Martin RE, Pannier M, Diederich F, Gramlich V, Hubrich M, Spiess H. *Angew Chem Intl Ed*. 1998; 37:2834–2837.
12. Hilger D, Polyhach Y, Padan E, Jung H, Jeschke G. *Biophys J*. 2007; 93:3675–3683. [PubMed: 17704177]
13. Smirnova I, Kasho V, Choe JY, Altenbach C, Hubbell WL, Kaback HR. *Proc Natl Acad Sci U S A*. 2007; 104:16504–16509. [PubMed: 17925435]
14. Bennati M, Robblee J, Mugnaini V, Stubbe J, Freed J, Borbat P. *J Am Chem Soc*. 2005; 127:15014–15015. [PubMed: 16248626]
15. Galiano L, Bonora M, Fanucci G. *J Am Chem Soc*. 2007; 129:11004–11005. [PubMed: 17705389]
16. Milov A, Samoilova RI, Tsvetkov YD, Formaggio F, Toniolo C, Raap J. *J Am Chem Soc*. 2007; 129:9260–9261. [PubMed: 17616195]
17. Schiemann O, Piton N, Mu Y, Stock G, Engels JW, Prisner TF. *J Am Chem Soc*. 2004; 126:5722–5729. [PubMed: 15125665]
18. Piton N, Mu Y, Stock G, Prisner T, Schiemann O, Engels J. *Nucleic Acids Res*. 2007; 35:3128–3143. [PubMed: 17452362]
19. Pornsuwan S, Bird G, Schafmeister C, Saxena S. *J Am Chem Soc*. 2006; 128:3876–3877. [PubMed: 16551072]
20. Godt A, Schulte M, Zimmermann H, Jeschke G. *Angew Chem Int Ed*. 2006; 45:7560–7564.
21. Bode B, Margraf D, Plackmeyer J, Durner G, Prisner T, Schiemann O. *J Am Chem Soc*. 2007; 129:6736–6745. [PubMed: 17487970]
22. Sen KI, Logan T, Fajer PG. *Biochem*. 2007; 46:11639–11649. [PubMed: 17880108]
23. Jeschke G, Koch A, Jonas U, Godt A. *J Magn Reson*. 2002; 155:72–82. [PubMed: 11945035]
24. Jen-Jacobson L. *Biopolymers*. 1997; 44:153–180. [PubMed: 9354759]
25. Terry B, Jack W, Modrich P. *J Biol Chem*. 1985; 260:13130–13137. [PubMed: 2997157]
26. Halford S, Marko J. *Nucleic Acids Res*. 2004; 32:3040–3052. [PubMed: 15178741]

Figure 1.

X-ray structure of the EcoRI specific complex. a) The bottom view shows monomers in red and blue. b) The side view illustrates the arm domain by circles. The DNA sequence is shown in yellow, and the residues mutated to cysteine are highlighted in green. Coordinates are from a highly refined version[6,7] of PDB entry 1CKQ.

Figure 2.

a) The DEER data for DNA complexes with R131C, S180C, and K249C-S180C EcoRI mutants. Simulated traces based on the distance distributions shown on the right are overlaid on the experimental data. b) Normalized distance distribution functions. Red lines in the crystal structure indicate the distance measured. DNA sequences are: TCGCGAATTCGC (specific); TCGCAAATTCGC (miscognate) and GTGCCTAAGCGCG (non-specific).