The pseudo EcoRI site GAATTA in the U3 region of the long terminal repeat of human immunodeficiency virus, which is flanked by a 26-base pair oligopurine tract, is readily nicked by either EcoRI or RsrI. The strand-specific nick occurs predominantly between the G and A residues and is independent of negative supercoiling. Other GAATTA sites surrounded by random (non-oligopurine) sequences are not nicked by these restriction endonucleases. However, other types and lengths of oligopurine tracts are effective in inducing the nicking in neighboring GAATTA sites. Hence, we propose that the flanking oligopurine tracts induce an altered DNA conformation on the GAATTA target site which may be similar to the transition state induced by EcoRI when binding to its canonical recognition site. Gel retardation analyses on restriction fragments containing the oligopurine-GAATTA-oligopyrimidine sequences suggest the presence of helical axis distortions which are consistent with this interpretation.

In recent years, interest in non-B DNA structures adopted by oligopurine-oligopyrimidine (Pur-Pyr) sequences has increased greatly. These sequences occur widely in nature and are often present within regulatory regions of genomes (reviewed in Ref. 1 and references therein) and in recombination hot spots (2, 3). Several studies revealed that these sequences can adopt a number of alternative structures, including intramolecular triplexes (4-10), tetraplexes (11, 12), and other multistranded conformations (13, 14).

The potential biological importance of these sequences has been emphasized recently with studies on protein factors which interact with regions of alternating Cs and Ts found near heat shock genes (15), Pur-Pyr sequences near the c-myc gene (16, 17), and with a Pur-Pyr domain in the c-Ki-ras promoter (18).

The human immunodeficiency virus genome base composition is very highly biased towards stretches of purines on one strand. The long terminal repeat sequences are important for the viral life cycle and contain sequences responsive to both virus-associated trans-acting factors as well as potential negative regulatory elements (19). Moreover nuclear host factors bind the LTRs (20-22).

As part of our efforts to determine if any unusual DNA structural features in this region participate in the regulation of the viral cycle, we conducted chemical, enzymatic, and physical probing (1) of the polypurine sequence in the U3 portion of the LTRs. The GAATTA EcoRI-like site flanking that sequence was nicked by EcoRI under conditions that normally would not allow cleavage or nicking.

EcoRI is a well studied model for DNA-protein interactions. Recent work emphasized that the local DNA helical structure, in addition to the sequence, greatly contributes to the specificity of the DNA-enzyme contacts. Crystallographic studies showed that the DNA in the DNA-EcoRI complex is bent (23-25); even in solution, the EcoRI recognition sequence is significantly distorted, as shown by gel retardation assays (26-28). Also, systematic studies on mutated EcoRI sites (26, 29) revealed that several kinds of changes in the recognition sequence can greatly increase the energetic barrier for the cleavage step and that changes in different positions can differently affect the EcoRI-DNA interaction (29). The GAATTA site is energetically the least unfavorable of a large number of non-canonical sites (29).

Herein, we investigated the behavior of GAATTA sites when embedded in different types and lengths of Pur-Pyr tracts as well as random sequences. The nature of the nicking reaction was studied. Furthermore, gel retardation analyses were performed to evaluate potential structural distortions in the DNA helix.

MATERIALS AND METHODS

Plasmids—pRW790 is a 1990 bp derivative of pBR322 which contains the pUC12 polylinker as an EcoRI-HindIII fragment (30). pRW791 was made by digesting pRW790 with EcoRI, filling in the 5' overhangs with the DNA polymerase Klenow fragment (Boehringer Mannheim or United States Biochemicals Corp.) and dNTPs, and ligating with T4 DNA ligase (United States Biochemicals Corp.) (9). A 33-bp sequence (AGAGAATTAGAAGAAGCCAACAAAGGAGAAGAC) from the HIV LTR (Fig. 1, top line) was synthesized with an Applied Biosystem 380A oligonucleotide synthesizer. The oligonucleotide was purified, annealed, and cloned as described previously (8). The recombinant plasmid was called pRW2201.

pRW1713 was described (31). The insert of pRW1713 was isolated by digesting pRW1713 with BamHI and purified on a 12% polyacryl-amide gel. The fragment then was recloned into the BamHI site of pRW790. Using a large molar excess (approximately 10-fold) of insert over the vector in the ligation reaction, it was possible to favor the insertion of two copies of the insert. This plasmid is called pRW2213.

The inserts in all of the recombinant plasmids were characterised by DNA sequencing of both strands using the Maxam-Gilbert method (32).

Bluescript was purchased from Stratagene and pACYC184 was described (33).

Restriction Endonucleases and Assay Conditions—The EcoRI used for these studies was purchased from three sources (Bethesda Research Laboratories, Boehringer Mannheim, and New England Bio-labs) or was a highly purified preparation (generous gift of Dr. L. Jen-Jacobson, University of Pittsburgh) (34, 35). 3 μg of plasmid were incubated with 10 units of enzyme in the presence of 50 mM

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 266, No. 25, Issue of September 5, pp. 16786-16790, 1991
Printed in U.S.A.

A DNA Conformational Alteration Induced by a Neighboring Oligopurine Tract on GAATTA Enables Nicking by EcoRI*

(Sabrina Venditti and Robert D. Wells
From the Department of Biochemistry, Schools of Medicine and Dentistry, University of Alabama, Birmingham, Alabama 35294)

*This work was supported by Grants GM 30822 and CA 13148 from the National Institutes of Health and National Science Foundation Grant 86-07785. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number(s) K03455.

The abbreviations used are: Pur, Pyr, oligopurine-oligopyrimidine; bp, base pair(s).
Conformationally Altered GAATTA Is Nicked by EcoRI

RESULTS

Unusual Nicking at GAATTA by EcoRI—The HIV LTRs contain several polypurine stretches. The largest tract (33 bp) was synthesized and cloned into the BamHI site of pRW790 to give pRW2201 (Fig. 1). Due to the important biological functions (see the Introduction) of this sequence, we undertook enzymatic, chemical, and physical probe studies. The strong susceptibility of this tract to nicking by EcoRI at one EcoRI-like recognition site (GAATTA) present in the Pur.Pyr stretch was discovered. Fig. 2 shows the results of treatment of pRW2201 with EcoRI at 37 °C for 2 h, subsequent 3' end labeling, and 12% polyacrylamide gel electrophoresis of the isolated fragment. In addition to the full-length fragment, a strong band (arrow) was found unexpectedly that maps between the G and the A of GAATTA. Moreover, the same band was visible in the sequencing lanes (lanes 1 and 2), since EcoRI was used to prepare the fragment for the Maxam and Gilbert reactions. We ruled out the possibility of preferential cleavage of the GA bond by the sequencing reactions, because in this case we should not see the extra bands in lanes 3 and 5 where the sequencing chemistry was not applied. Lanes 3 and 5 also show that the nicking does not depend on piperidine treatment. In addition if the fragment was isolated using EcoO109 instead of EcoRI (lanes 4 and 6), only the full-length fragment was present. Two less intense bands that map

Fig. 2. EcoRI nicking of the insert in pRW2201. The EcoRI reactions were performed as described under "Materials and Methods." The insert was isolated using either EcoRI and PsI or EcoO109 and PsI. All the digestions were performed under standard conditions of salts and temperature. Lanes 1 and 2, A + G and T + C reactions, respectively, conducted following the Maxam and Gilbert procedure. The sequence is shown on the left side of the panel. Lanes 3 and 5, the EcoRI digest was loaded with or without piperidine pretreatment, respectively. Lanes 4 and 6, the EcoO109 digest was loaded with or without piperidine pretreatment, respectively. The band corresponding to the nick introduced by EcoRI is designated by the arrow.

Fig. 1. Sequences employed containing GAATTA sites. pRW2201 and pRW2213 were made by cloning the indicated sequences into the BamHI site of pRW790, pRW791 is a derivative of pRW791. The GAATTA sites in pRW790 (29), pACYC184 (32), and Bluescript are located at positions 1423, 1492, 3840, and 780 of the respective maps.
Conformationally Altered GAATTA Is Nicked by EcoRI

Fig. 3. Analysis of the nicking by highly purified EcoRI and RsrI. 3 μg of pRW2201 were incubated under the conditions described under “Materials and Methods” with either a highly purified preparation of EcoRI (panel A) or with RsrI (panel B). Panels A and B: lane 3, pRW2201 incubated with 44 units of the enzymes. Lanes 4–8, 1:2, 1:5, 1:10, 1:20, 1:40 dilutions of the enzymes, respectively. Lanes 1 and 2, A + G and T + C sequencing, respectively. The site of the nicking is indicated by the arrow in both panels.

between the second A and the first T and between the two Ts were observed occasionally (not shown). Their nature, though, is uncertain since they do not appear reproducibly. In addition, the identity of the faint and more slowly migrating band in lanes 3 and 5 is uncertain.

Parallel studies were conducted to determine if cleavages were made by EcoRI on the complementary strand. pRW2201 was treated with EcoRI from several commercial sources and labeled in order to read the oligopyrimidine strand. We were not able to detect any bands corresponding to nicking at the TAATTC site. Thus, the oligopurine strand containing GAATTA is nicked but the oligopyrimidine strand containing TAATTC is uncleaved.

EcoRI, as well as other restriction endonucleases, has a so-called “star” activity, which is an altered specificity for its substrate observed under particular conditions like the absence of salt or the presence of high concentrations of organic solvents or enzyme (38). In fact, incubation of pRW2201 with EcoRI under star conditions results in the full cleavage of the GAATTA site. However, the star activity is not responsible for our results, since the reactions were performed under standard conditions of salt, pH, and temperature (as described under “Materials and Methods”). Also, the nicking activity was present at low ratios of enzyme to DNA, although to a lesser extent (data not shown).

The nicking reaction was found also with a highly purified noncommercial preparation of EcoRI (34, 35). The results of these experiments are depicted in Fig. 3 (panel A). Lanes 3–8 show the products of the enzymatic reactions performed in the presence of dilutions (from 1 to 1:40) of the enzyme. The band corresponding to the unusual nicking between G and A on the oligopurine strand is indicated by the arrow. Hence, the nicking reaction likely is caused by EcoRI, since it was found with at least four preparations of the enzyme.

Studies with linearized as well as supercoiled substrates revealed similar nicking reactions. Thus, the scission was not due to a supercoil-dependent non-B DNA conformation such as a triplex.

RsrI Also Catalyzes the Nicking Reaction—RsrI, an isoschizomer of EcoRI, was recently characterized (36, 37) and purified. The salt and temperature requirements for this enzyme are different from EcoRI. We tested RsrI for the nicking activity on pRW2201. Fig. 3 (panel B) shows the result of this experiment that was performed using the same dilutions of enzyme as in panel A. The DNA was incubated with the enzyme under the conditions indicated under “Materials and Methods.” The extra band corresponding to nicking between G and A appears at the higher concentrations.

Interestingly, RsrI seems to show the same DNA structural requirements as EcoRI (DNA bending) in order to cut its recognition site. Therefore, the observation of the same nicking reaction even with a different enzyme suggests the existence of a structural feature intrinsic to the DNA site that makes it susceptible to nicking.

Specificity of the Nicking—It has been suggested that EcoRI first recognizes structural features intrinsic to the binding site and that more specific interactions than are obtained through the formation of kinks in the substrate (26). Our data support the hypothesis that the flanking oligopurines confer a conformation on the GAATTA target site which makes it a more suitable substrate for EcoRI. Hence, studies were conducted with other GAATTA sites present in random, or non-oligopurine, tracts. Two GAATTA sites in pACYC184 and one in Bluescript were investigated. Moreover, another GAATTA site is present on pRW790, representing a valuable internal control for our determinations. Fig. 1 lists these sites along with 15 bp of bilateral flanking sequence. It is clear that all the sites are located in random sequences.

The DNAs were treated with several commercial EcoRI under the same conditions as the previous experiments, and the fragments containing the sites were isolated and labeled. For each fragment two sites were alternatively labeled in order to read both strands. The electrophoretic analyses of the products did not show any nicking activity at any of the sites (data not shown). Thus, the presence of a suitable

2 R. I. Gumport, personal communication.
oligopurine tract surrounding the GAATTA seems to be required for the nicking reaction.

Length Requirements of the Flanking Sequence—The potential effect of the length of the flanking oligopurine regions was investigated with pRW1713 and pRW2213 (Fig. 1). pRW1713 was described (31) and was shown to form an intramolecular triplex in vitro. pRW2213 was made by dimerization of the pRW1713 insert; it contains two GAATTA sites and the entire insert is 68 bp long versus the 31-bp insert in pRW1713.

Fig. 4 shows that EcoRI reacts differently with the two plasmids. Lane 3 in panel B indicates that no nicking occurred at the GAATTA site in the shorter insert, whereas the two arrows in panel A point to two extra bands present in the longer insert. Thus, both GAATTA sites on pRW2213 were nicked by EcoRI.

Thus, we propose that both the length and type of oligopurine sequence are important for inducing the GAATTA site to adopt a structure suitable for nicking by EcoRI. The 31-bp sequence of pRW1713 was insufficient in length to stabilize the non-B conformation, whereas the 68-bp tract in pRW2213 was long enough. That the 33-bp region in pRW2201 was nicked, whereas the 31-bp tract in pRW1713 was uncleaved, may be due to the type (nonrepeating for pRW2201) of sequence as well as to the small increment in length. The result in Fig. 4 also suggests that there is no influence of the position of the sites on nicking. In fact the first site on pRW2213 has a longer oligopurine tract at its 3′ side, whereas the second has it at the 5′ side, but despite this difference, both sites are nicked.

_Determination for Bent DNA—Structural distortions of the helical axis, like intrinsic and protein-induced bending or kinks, can be detected by gel retardation experiments on linear fragments (26, 27, 39, 40). Therefore, we performed gel retardation experiments to determine if these fragments showed altered gel mobility. The insert in pRW2201 was isolated using two sets of digestions: EcoO109 and BstNI giving a fragment 178 bp long and HaeIII giving a 167 bp fragment. In both cases, the GAATTA site was located approximately in the middle of the fragment. The products of the digestions then were electrophoresed on 8% polyacrylamide gels at 4 °C and 23 °C.

Fig. 5 shows that at both temperatures a small amount of retardation was observed. Lanes 2 of both panels show that the 178-bp fragment generated by EcoO109 and BstNI actually migrated between the 190- and 201-bp fragments. Similarly, lanes 3 of both panels indicate that the 167-bp fragment migrated like a 180-bp fragment. We think, from this and from the fact that there is no supercoil dependence, that a helical distortion similar to that described previously (24, 26) may be responsible for the anomalous behavior of the fragments. We repeated the experiment with the insert of pRW1713, using the same set of digestions (data not shown). This sequence, which does not show any nicking, also had a similar slight retardation in migration. This was not unexpected, since different polyuridine sequences can undergo different distortions of the helix axis (41). For pRW2201, though, the variations of the helix parameters could be more favorable for the interaction with the enzyme allowing it to nick the site.

**DISCUSSION**

A pseudo EcoRI site GAATTA embedded within a 32-bp oligopurine tract from the U3 region of the LTRs of HIV as well as other GAATTA sites within oligopurine stretches are
nicked between the G and the A residues by EcoRI and RsrI.

Weinreb, A., Collier, D. A., Birshtein, B. K., and Wells, R. D. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 3781-3786

Kobwi, Y., and Kobwi-Shigematsu, T. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 3781-3786

Henderson, E., Harding, C. C., Walk, S. K., Tincoco, L. J., and Blackburn, E. H. (1987) Cell 61, 899-908

Williamson, J. R., Raghuraman, M. K., and Cech, T. R. (1989) Cell 60, 871-880

Shimizu, Min, and Harvey, J. C., and Wells, R. D. (1990) Biochemistry 29, 7044-7047

Wohlrab, F., and Wells, R. D. (1989) J. Biol. Chem. 264, 8207-8213

Gilmour, D. S., Thomas, G. H., and Elgin, S. C. R. (1989) Science 245, 1487-1490

Davis, T. L., Firulii, A. B., and Kinninbigh, A. J. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 9692-9696

Kaminbigh, A. J. (1989) Nucleic Acids Res. 27, 7711-7717

Hoffman, E. K., Truuk, S. P., Murphy, M., and George, D. L. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 2700-2704

Kosen, T. A., Sodrofi, A., and Hasetine, W. A. (1985) Cell 4, 813-823

Gaynor, R., Soltanakisis, E., Kuwabara, M., Garcia, J., and Sisner, D. S. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 4858-4862

Okamoto, T., and Wong-Staal, F. (1986) Cell 47, 29-35

Jones, A. K., Kadonagona, J. T., Lucw, P. A., and Tjian, R. (1986) Science 233, 753-759

Frederick, C. A., Grable, J. M., Samudi, C. J., Jen-Jacobson, L., Wang, B.-C., Greene, P. J., Boyer, H. W., and Rosenberg, J. M. (1984) Nature 309, 527-531

McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P. J., Boyer, H. W., Grable, J. C., and Rosenberg, J. M. (1986) Science 234, 1526-1541

Kim, Y., Grable, J. C., Love, R., Greene, P. J., and Rosenberg, J. M. (1990) Science 245, 1352-1359

Bishop, J. (1990) J. Biol. Chem. 265, 1352-1359

REFERENCES

1. Wells, R. D., Collier, D. A., Harvey, J. C., Shimizu, M., and Wohlrab, F. (1989) FASEB J. 2, 3920-3929

2. Collier, D. A., Griffin, J. A., and Wells, R. D. (1988) J. Biol. Chem. 263, 7907-7915

3. Weinreb, A., Collier, D. A., Birzhan, B. K., and Wells, R. D. (1990) J. Biol. Chem. 265, 1352-1359

4. Mirkin, S. M., Lysnychev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A., and Frank-Kamenetskiy, M. D. (1987) Nature 330, 485-488

5. Voloshin, O. N., Mirkin, S. M., Lysnychev, V. I., Belotserkovskii, B. P., and Frank-Kamenetskiy, M. D. (1988) Nature 333, 475-476

6. Hsu, H., and Dahlberg, J. E. (1989) Science 244, 1791-1796

7. Hsu, H., and Dahlberg, J. E. (1989) Science 245, 1571-1576

8. Harveys, J. C., Kiyunjik, J., and Wells, R. D. (1987) J. Biol. Chem. 263, 7386-7390

9. Harveys, J. C., Shimizu, M., and Wells, R. D. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 6292-6296

10. Kobwi, Y., and Kobwi-Shigematsu, T. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 3781-3786

11. Henderson, E., Harding, C. C., Walk, S. K., Tincoco, L. J., and Blackburn, E. H. (1987) Cell 61, 899-908

12. Williamson, J. R., Raghuraman, M. K., and Cech, T. R. (1989) Cell 60, 871-880

13. Shimizu, Min, and Harvey, J. C., and Wells, R. D. (1990) Biochemistry 29, 7044-7047

14. Wohlrab, F., and Wells, R. D. (1989) J. Biol. Chem. 264, 8207-8213

15. Gilmour, D. S., Thomas, G. H., and Elgin, S. C. R. (1989) Science 245, 1487-1490

16. Davis, T. L., Firulii, A. B., and Kinninbigh, A. J. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 9692-9696

17. Kaminbigh, A. J. (1989) Nucleic Acids Res. 27, 7711-7717

18. Hoffman, E. K., Truuk, S. P., Murphy, M., and George, D. L. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 2700-2704

19. Kosen, T. A., Sodrofi, A., and Hasetine, W. A. (1985) Cell 4, 813-823

20. Gaynor, R., Soltanakisis, E., Kuwabara, M., Garcia, J., and Sisner, D. S. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 4858-4862

21. Okamoto, T., and Wong-Staal, F. (1986) Cell 47, 29-35

22. Jones, A. K., Kadonagona, J. T., Lucw, P. A., and Tjian, R. (1986) Science 233, 753-759

23. Frederick, C. A., Grable, J. M., Samudi, C. J., Jen-Jacobson, L., Wang, B.-C., Greene, P. J., Boyer, H. W., and Rosenberg, J. M. (1984) Nature 309, 527-531

24. McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P. J., Boyer, H. W., Grable, J. C., and Rosenberg, J. M. (1986) Science 234, 1526-1541

25. Kim, Y., Grable, J. C., Love, R., Greene, P. J., and Rosenberg, J. M. (1990) Science 245, 1352-1359

26. Bowers, T., and Lisk, D. T. (1990) J. Biol. Chem. 265, 1352-1359

27. Weinreb, A., Collier, D. A., Birshtein, B. K., and Wells, R. D. (1990) J. Biol. Chem. 265, 1352-1359

28. M. Shimizu and R. D. Wells, unpublished data.