MOLECULAR AND CELLULAR BIOLOGY, Mar. 1989, p. 1324–1326
0270-7306/89/031324-03$02.00/0
Copyright © 1989, American Society for Microbiology
Vol. 9, No. 3

NOTES

Nucleotide Sequence of an Unequal Sister Chromatid Exchange Site in a Mouse Myeloma Cell Line

DANIEL R. KATZENBERG, SHERMAINE A. TILLEY,† AND BARBARA K. BIRSHTEIN*

Department of Cell Biology, Albert Einstein College of Medicine, Bronx, New York 10461

Received 15 July 1988/Accepted 7 December 1988

The mouse myeloma cell line MPC 11 carries two Cγ2a immunoglobulin heavy-chain genes on the expressed chromosome, a duplication shown to have occurred through unequal sister chromatid exchange (USCE). In the present report, we present the nucleotide sequence of the USCE joint and show that both breaks occurred within tracts of repeated TC dinucleotides. Additional TC dinucleotide tracts and two oligonucleotide segments (N sequences) were inserted at the USCE site.

A single copy of each of the eight immunoglobulin heavy-chain constant region genes is present in the germ line of BALB/c mice. The myeloma cell line MPC 11 carries a tandem duplication of the Cγ2a gene on the expressed chromosome (Fig. 1). The expressed chromosome contains a deletion of γ2a-associated tandem repeats, while its homolog maintains a germ line configuration in this region; since both copies of γ2a involved in the duplication carry the deletion, unequal sister chromatid exchange (USCE), rather than mitotic recombination between homologs, is the molecular mechanism responsible for the duplication (13). Furthermore, USCE appears to be an ongoing process in MPC 11; variants in heavy-chain production derived from MPC 11 carry the reciprocal products of further USCE (either loss or further amplification of the duplicated γ2a gene).

The DNA sequences of the germ line regions involved in the crossover, one located approximately 3 kilobases 3′ to Cγ2b and the other the same distance 3′ of Cγ2a, are virtually identical (15), strongly suggesting homologous pairing of the two regions as an initial step. The size of the USCE joint, however, implies that if crossing over followed precise pairing, the nicks in the two strands did not occur at equivalent positions. We measured an offset of about 50 nucleotides between the two cuts. The crossover site was further localized by genomic Southern blot analysis within a region of juxtaposed tracts of TC and TG dinucleotide repeats. Alternating purine-pyrimidine tracts, such as $(TG)_n$, may assume a Z-DNA conformation (11), whereas regions of TC simple sequence may assume a novel non-B, non-Z, triplex structure (10) and have been suggested to be structural elements of S1 nuclease-hypersensitive sites (3).

To precisely localize the site of recombination, we sought to clone and sequence the USCE joint from a cell line carrying the Cγ2a duplication. The USCE joint was cloned from the MPC 11 variant ICR 16, the γ2a genes (and hence the USCE joint) of which are amplified approximately 5- to 10-fold (13). Since previous cloning attempts with large fragments in cosmid and phage vectors were unsuccessful, perhaps due to recombinogenic properties of the USCE joint, we sought to clone this region as a small fragment in a plasmid vector. In order to enrich the library for the USCE joint, a double-fractionation strategy was employed. The USCE joint was present on a 6.4-kilobase *Bgl*II fragment and on an internal *Hin*dIII fragment of 2.5 kilobases (Fig. 1a). Genomic DNA was digested with *Bgl*II and size fractionated; DNA of the appropriate size was recovered, digested with *Hin*dIII, and size fractionated a second time. DNA of the desired size range was recovered and ligated with *Hin*dIII-digested pBR322. Transformants were screened with γ2IVS 3.6 (Fig. 1). Two independent positive colonies, NJ25 and NJ41, were isolated; restriction maps of the plasmid they carried were consistent with the map of the USCE joint determined by genomic Southern blot analysis.

The sequence of the USCE joint, present on the 700-base-pair (bp) *Hin*dIII-*Eco*RI fragment (Fig. 1a), was determined by the chemical modification method (9). Both strands were read from the *Hin*dIII site of γ2a-3′ to the *Nsi*I site present in γ2b-3′ sequences; the sequences of NJ25 and NJ41 were identical (Fig. 1b). As expected, the sequence was identical to γ2a-3′ for a maximum of 49 bases at its 5′ end and identical to γ2b-3′ for a maximum of 106 bp 5′ to the *Nsi*I site. Between the two germ line sequences lay at least 64 bp with the general structure $(TC)_{11}$ N_6 $(TC)_5$ N_8 $(TC)_9$. The N segments were not present within either γ2a-3′ or γ2b-3′. In γ2a-3′, the USCE site lay either at the 3′ edge of or somewhere within the TC tract. Similarly, the break in γ2b-3′ was either at the 5′ edge of or anywhere within the TC dinucleotide repeat.

The potentially Z-DNA-forming TG-rich region of γ2a-3′ was eliminated from the USCE joint, whereas that of γ2b-3′ lay at least 16 bases distal to the USCE site. Studies on the Ustilago RecI protein (5–7) have shown that Z-DNA segments are important in the pairing step of the recombination catalyzed by this enzyme, while other studies (14) identified a role for TG sequences in yeast meiotic recombination. In the MPC 11 USCE, the potential Z-DNA may serve as a recognition site for recombinases involved in the pairing of the sister chromatids with the actual nicks introduced some distance away.

The mechanism of the USCE appears to be complex. In both clones which were sequenced (NJ25 and NJ41), there

* Corresponding author.
† Present address: Department of Developmental and Structural Biology, The Public Health Research Institute, New York, NY 10016.

```
b          10        20        30        40        50        60
γ2a-3'  -------------------------------------------------[
NJ-25   CTAGAGACCATGAATGTCTTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
γ2b-3'  ------gt------------aaagtgt--------c-            [

           70        80        90        100       110       120
γ2a-3'                                                       ]
NJ-25   CTCTCTCTCTCTTTCTGTCTCTCTCTCTTGCTCACTCTCTCTCTCTCTCTCTCTCTCTCT
γ2b-3'                                                       ]------

           130       140       150       160       170       180
γ2a-3'                     c-------g-------g-------g------tgc---cat-c
NJ-25   CTCTCTCTCAGTGTCTCTTTGTGTGTATGTGTGTATGTGTGTTTGTGTGCATGTGTGCGT
γ2b-3'  ------------------------------------------------------------

           190       200       210
γ2a-3'  a-g-tatt-ttttggc-ttt-aaaca-taa
NJ-25   GCACATGCGGGCGCGCACAGGTGCACACTC
γ2b-3'  ------------------------------
```

FIG. 1. (a) Map of USCE in MPC 11. Top line shows the region 3′ to Cγ2a involved in the USCE, while the second line depicts the region 3′ to Cγ2b involved in the exchange. Previous studies (15) localized the crossover site in γ2b-3′ between the *Xba*I and *Nsi*I sites. A restriction map for one product of the USCE, carrying two Cγ2a genes, is shown on the bottom line. Sequences derived from γ2a-3 are shaded, and those originating in γ2b-3′ are not. Restriction sites are indicated as follows: Bg, *Bgl*II; H, *Hin*dIII; N, *Nsi*I; S, *Sph*I; Nc, *Nco*I; E, *Eco*RI. *, TC repeats; ○, TG repeats. (b) DNA sequence of the USCE joint (NJ-25) is shown in capital letters. The sequences of the germ line regions are depicted by dashes (if identical) or in lowercase letters. The site has been located to maintain the entire TC tract from both germ line sequences; it could actually be located further 5′ in γ2a-3′ or further 3′ in γ2b-3′, resulting in the insertion of additional TC dinucleotides within the intervening DNA. Brackets demarcate the joining sequence itself, depicted in bold letters and having the general structure $(TC)_{11}N_6(TC)_5N_8(TC)_9$. The N sequences are overlined. Attempts to determine the sequence of the USCE joint by the dideoxy-chain termination method (12) were unsuccessful since termination in all sequencing reactions occurred at the start of the TC tract.

was an identical insertion of 64 bp at the USCE site with the entire tract, including N sequences, either derived from a donor sequence elsewhere or generated de novo. The USCE is, therefore, not a simple, single crossover event. If the 64-bp joint fragment is present on a donor sequence, the entire insertion could have occurred during a single exchange. An insertion of similarly sized (51-bp) but dissimilar sequence was found to cause a truncated μ chain in the BW leukemic cell line (1). Alternatively, the USCE joint could have been generated by a multistep mechanism. For example, TC tracts could have been expanded through several rounds of crossing over or from slippage of DNA polymerase during replication. Furthermore, enzymes capable of adding additional residues to the ends of DNA have been found. Among these are telomerase (4), which adds tracts of simple sequence telomeric repeats to telomere sequences, and terminal deoxynucleotidyl transferase, whose expression in the pre-B-cell stage of differentiation is correlated with the addition of novel nucleotides to V-D and D-J junctions in the immunoglobulin heavy-chain variable region gene (2, 8). Novel nucleotides in the USCE joint are comparable in size and sequence properties with those added during VDJ joining.

In MPC 11 variants, such as ICR 16, from which the USCE joint was isolated, USCE accounted for both the initial duplication of the Cγ2a genes and subsequent amplification-deletion (13). With 5 to 10 tandemly arrayed γ2a genes in ICR 16, 4 to 9 USCE joints were present. Should NJ25 and NJ41 represent different USCE joints, it seems plausible that the identical 64-bp insertion that each contains occurred during the initial γ2a gene duplication. As genomic Southern blot mapping indicated (13), further USCE would then take place either within the 13 kilobases of sequence identity shared by regions flanking the duplicated γ2a genes, or if within the original USCE joint, through precisely equal exchanges.

This report presents the first molecular characterization of a USCE. Regions of TC dinucleotide repeats, which are capable of assuming a unique triple helix structure with one AG strand unpaired (10), were the site of exchange in this recombination event. Either the single-strand or the "triplex" structure could serve to initiate complex recombination events such as those described here. Furthermore, two segments of 6 and 8 nucleotides were inserted de novo at the USCE site. Additional USCE sites must be identified and sequenced to determine whether our findings, both TC dinucleotide repeat expansion and N sequence insertion, are a general feature of USCE.

This work was supported by grants AI13509, AI10702, and PC30CA13330 from the National Institutes of Health and by the Irvington House Institute for Medical Research. B.K.B. is an Established Investigator of the American Heart Association. D.R.K. is supported by a training grant from the National Institutes of Health, T32GM7288.

We thank Peter Ruvolo for his excellent tutelage in the techniques of Maxam and Gilbert sequencing, Nancy Martinez for expert technical assistance, Julius Marmur for stimulating discussion of our results, and Ari Weinreb and Sandra Giannini for critical reading of the manuscript.

LITERATURE CITED

1. **Bakhshi, A., P. Guglielmi, U. Siebenlist, J. V. Ravetch, J. P. Jenson, and S. J. Korsmeyer.** 1986. A DNA insertion/deletion necessitates an aberrant RNA splice accounting for a μ heavy chain disease protein. Proc. Natl. Acad. Sci. USA **83:**2689–2693.
2. **Desiderio, S. V., G. D. Yancopoulos, M. Paskind, E. Thomas, M. A. Boss, N. Landau, F. W. Alt, and D. Baltimore.** 1984. Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. Nature (London) **311:**752–755.
3. **Evans, T., and A. Efstradiatis.** 1986. Sequence-dependent S1 nuclease hypersensitivity of a heteronomous DNA duplex. J. Biol. Chem. **261:**14771–14780.
4. **Greider, C. W., and E. H. Blackburn.** 1985. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. Cell **43:**405–413.
5. **Kmiec, E. B., K. J. Angelides, and W. K. Holloman.** 1985. Left-handed DNA and the synaptic pairing reaction promoted by Ustilago RecI protein. Cell **40:**139–145.
6. **Kmiec, E. B., and W. K. Holloman.** 1984. Synapsis promoted by Ustilago RecI protein. Cell **36:**593–598.

7. **Kmiec, E. B., and W. K. Holloman.** 1986. Homologous pairing of DNA molecules by Ustilago RecI protein is promoted by sequences of Z-DNA. Cell **44**:545–554.
8. **Landau, N. R., D. G. Schatz, M. Rosa, and D. Baltimore.** 1987. Increased frequency of N-region insertion in a murine pre-B-cell line infected with a terminal deoxynucleotidyl transferase retroviral expression vector. Mol. Cell. Biol. **7**:3237–3243.
9. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
10. **Mirkin, S. M., V. I. Lyamichev, K. N. Drushlyak, V. N. Dobrynin, S. A. Filippov, and M. D. Frank-Kamenetskii.** 1987. DNA H form requires a homopurine-homopyrimidine mirror repeat. Nature (London) **330**:495–497.
11. **Nordheim, A., E. M. Lafer, L. J. Peck, J. C. Wang, B. D. Stollar, and A. Rich.** 1982. Negatively supercoiled plasmids contain left-handed Z-DNA segments as detected by specific antibody binding. Cell **31**:309–318.
12. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
13. **Tilley, S. A., and B. K. Birshtein.** 1985. Unequal sister chromatid exchange: a mechanism affecting Ig gene arrangement and expression. J. Exp. Med. **162**:675–694.
14. **Treco, D., and N. Arnheim.** 1986. The evolutionarily conserved repetitive sequence $d(TG \cdot AC)_n$ promotes reciprocal exchange and generates unusual recombinant tetrads during yeast meiosis. Mol. Cell. Biol. **6**:3934–3947.
15. **Weinreb, A., D. R. Katzenberg, G. L. Gilmore, and B. K. Birshtein.** 1988. A site of unequal sister chromatid exchange contains a potential Z-DNA-forming tract. Proc. Natl. Acad. Sci. USA **85**:529–533.