A human immunoglobulin kappa orphon without sequence defects may be the product of a pericentric inversion

Christian Huber, Rainer Thiebe, Horst Hameister¹, Hans Smola, Erika Lötscher+ and Hans G. Zachau*
Institut für Physiologische Chemie, Physikalische Biochemie und Zellbiologie der Universität München, and ¹Abt. Klinische Genetik, Universität Ulm, FRG

Received April 19, 1990; Accepted May 18, 1990
EMBL accession no. X51887

ABSTRACT

The Vᵦ gene segments that have been transposed from the kappa locus on the short arm of chromosome 2 at 2p11–12 to other chromosomal sites are called orphans. The 18 Vᵦ orphans sequenced up to now carry defects and are to be considered pseudogenes. We now describe the VᵦI gene segment V108 whose sequence is without any defects and which was localized to the long arm of chromosome 2 at 2q12–14 by in situ hybridization. The V108 region may have been transposed from the short to the long arm of chromosome 2 by a pericentric inversion. Possible reasons for the conservation of its sequence are discussed. In spite of its bona fide sequence V108 is considered to be an unlikely candidate for a Vᵦ-Jᵦ rearrangement and subsequent functional expression.

INTRODUCTION

The human immunoglobulin locus coding for kappa light chains is located on the short arm of chromosome 2 at 2p11–12 (1,2). The locus comprises a single constant (Cᵦ), five joining (Jᵦ) and less than 100 variable gene and pseudogene (Vᵦ) segments (reviews 3,4). A number of cloned Vᵦ gene segments were localized outside the kappa locus: five of them on chromosome 22, one on chromosome 1 (5,6), five on still unspecified chromosomes other than chromosome 2 (6,7) and 11 on the long arm of chromosome 2 (8). These gene segments are called orphons in analogy to the histone and ribosomal RNA genes found outside of the respective gene clusters (9).

All Vᵦ orphans were found to contain an intron in their leader regions indicating that they are not products of transcription, splicing and retrotranscription. Various structural features of the orphans and their surroundings were studied in order to find out how the orphans might have been transposed in evolution from the kappa locus to their present sites (8,10, 11) but no specific mechanism could be formulated that is valid for all orphons.

A common feature of the 18 Vᵦ orphans sequenced up to now is that they carry one or several defects in their sequences as stop codons, insertions or deletions (5–7,12). The orphons were therefore considered to be pseudogenes on double account: because they are separated from the Jᵦ-Cᵦ region and because their reading frames are damaged. We now report on a Vᵦ orphon whose coding sequence is without apparent defect and discuss the mechanistic and possible functional implications of the finding.

MATERIALS AND METHODS

The cosmid clone cos108 was isolated by M. Pech from the cosmid library IIIa by screening with a VᵦI probe (13). Restriction mapping and subcloning followed standard procedures (e.g.13). DNA fragments, subcloned in M13 phages (14), were sequenced by the dideoxy chain termination method (15) using the 'Sequenase DNA Sequencing Kit' (United States Biochemical Corporation). The pulsed field gel blot, prepared similar to the ones of ref. 8, was kindly donated by G. Weichhold.

The rodent-human cell hybrid DNAs which were used for the chromosomal localization analyses were the same as in refs. 5,6 and were donated by K.-H. Grzeschik.

The procedure to combine chromosomal banding techniques and in situ hybridization has been described (16). In brief: lymphocyte cultures were labeled with bromo-deoxyuridine (20 µg/ml) and the chromosomes prepared according to standard techniques. Denaturation was done for 1.5–2 min in 70% formamide/30% 2 X SSC (v/v) at 70°C. 5–20 ng of 3H-labeled probe DNA (spec. activity 2–5 X 10⁷ dpm/µg) were hybridized in 20 µl 50% formamide, 10% dextran sulfate in 2 X SSC and addition of 0.05 µg/µl salmon sperm DNA for 16 h at 40°C. After extensive washing steps the slides were exposed for autoradiography for 2 weeks. After development staining was done according to the fluorochrome photolysis technique (17). For evaluation only the signal on chromosome 2 was recorded.

RESULTS

The cosmid clone cos108 was isolated early in our work on the kappa locus by screening a cosmid library that had been prepared

* To whom correspondence should be addressed

+ Present address: Cancer Research Laboratory and Department of Microbiology and Immunology, University of California, Berkeley, CA 94720, USA
The cosmid clone cos108 and the V_{K}I gene V108
The restriction map of cos108 is presented in Fig. 1. The first subclones were the V_{K}I containing clones and the clone m108-1 that was designed as a search clone for chromosomal walking experiments. The clone was not fully unique and the cosmid clones isolated with its help from the cosmid libraries HI, IVa, and IVb (13) did not contain regions of restriction map homology to cos108. But the clone m108-1 was important in the first chromosomal location experiments (see below). The subclone m108-13 is not well suited for chromosomal walking experiments because it is located near the center of cos108. But it turned out to be practically unique and therefore well suited for the chromosomal location experiments.

Early indications that cos108 may contain in addition to the V_{K}I gene some sequences hybridizing with V_{K}II and V_{K}III probes could not be substantiated. The subclones in question (p108-2 and p108-3) as well as the whole cos108 hybridized even from a Sau3A partial digest of placental DNA (13). The cosmid clone went then through several hands for restriction mapping (18,19), isolation of unique subclones for chromosomal location and walking (19) and for sequencing. The cosmid clone remained interesting because the sequencing had shown it to contain a bona fide V_{K}I gene and the location was determined to be on chromosome 2 in hybridization experiments with human-rodent cell hybrid DNAs (19). But all attempts failed to link cos108 to the emerging contigs that constitute the kappa locus on the short arm of chromosome 2. The data are now being published since it became clear that the clone is located on the long arm of chromosome 2 and, with that, some interesting points of discussion arose as to the evolution and possible function of the sequence.
Figure 2. DNA sequence and formal translation product of V108. The decanucleotide (22), TATA (23) and V-J joining signal sequences are underlined as well as the invariant amino acids (24). L, L', FR and CDR designate leader, framework and complementarity determining regions, respectively. The sequence including 225 bp downstream of the one shown in the figure has been transmitted to the EMBL Data Library, Heidelberg (accession number X51887).

under very relaxed conditions (200 mM phosphate) so weakly with the V_{KII} probe m607-3 and the V_{KIII} probe m41-7 that we did not pursue the point. In our experience with UHOs (unidentified hybridizing objects; e.g. refs. 20,21) we do not expect to find V_{K}-like sequences in these regions.

The V_{KI} gene (called V108) was sequenced following the strategy shown in the lowest panel of Fig. 1. According to its sequence (Fig. 2) VI08 is a potentially functional V_{KI} gene segment. The regulatory sequences, splice sites and recombination signals conform to the rules and so-called 'invariant amino acids' (24) appear in the proper positions. The sequence of the V_{KI} gene region of VI08 (including the four 5' adjacent leader codons) is about 90% identical to the corresponding parts of other V_{KI} genes including genes from the so-called L region (25). Cosl08 was also found to hybridize with the low repetitive probe m20-16 that is derived from and hybridizes with the 3' flank of V_{KI} genes of this region (26).

The possible significance of this observation will be discussed below.

Chromosomal localization of cosl08

The first experiments were carried out with the subclone m108-1 and DNAs from human-rodent cell hybrids kindly provided by K.-H. Grzeschik. In BglII digest of the cell line DNAs the probe hybridized with a 2 kb and a 3 kb band, the first one segregating with chromosome 2, the other one with chromosome 9. Since only the 2 kb band corresponds to the situation in cosl08 this can be taken as a good indication that the insert of cos 108 is derived from chromosome 2.

The localization of the cosl08 region was achieved by in situ hybridization experiments with the \(^3\)H-labeled probe ml08-13. 15 metaphases were scored and a specific signal was noticed on chromosome 2 in the region 2q12-14 with the peak at 2q13 (Fig. 3). Therefore the localization on the long arm of chromosome 2 seems to be unequivocal.

Since the 11 V_{K} pseudogenes of the so-called W regions were also localized to the long arm of chromosome 2 (2cen-q11) (8) it was of interest to check whether the large scale restriction maps of W and cosl08 overlap. A pulsed field gel electrophoresis experiment showed, however, that this is not the case.
Hybridization of a blot with the probe m108-13 gave the following fragments: NotI, 1220 kb; NruI and NruI/NotI, 1050 and 450 kb (not shown). This hybridization pattern is clearly different from the patterns seen on the same blot with a W region specific probe (8) and with kappa locus specific probes.

DISCUSSION

The finding of a bona fide V<sub>K</sub> gene segment separate from the kappa locus is surprising. We therefore searched for experimental artifacts and excluded the obvious ones: the cosmid clone did apparently not arise from ligation of independent genomic DNA fragments and the cos108 region is distinct from the kappa locus both according to the chromosomal localization studies and the pulsed field gel experiments. We therefore have to consider the findings at face value.

We have discussed the location of the W regions with their 11 V<sub>K</sub> orphans at 2 cen-q11 as the result of a pericentric inversion that has taken place in the evolutionary past (8). An analogous argumentation can be applied to the cos108 region. Because the sequence of V108 is without apparent defects one would have to assume, however, that the inversion happened not long ago or one would have to invoke gene conversion like processes that prevented the sequence from diverging but did not do the same for the orphans of the W regions or other orphans. The characteristics of the cos108 and the W regions are so different that it is easier to assume two distinct inversion events than one inversion that transposed both regions together. Also the in situ hybridization experiments tend to indicate neighboring but not identical sites on the long arm of chromosome 2.

It is interesting to note that pericentric inversions of chromosome 2 involving the chromosomal segment 2p11-2q13 are observed by cytogenetic methods in about 0.1% of today's population (27 and earlier literature). Also de novo inversions have been observed at chromosome 2 (28), indicating that the inversion is a spontaneous event. The V108 region is perhaps the product of such an inversion that had occurred fairly recently.

The latter point touches on the question whether V108 on the long arm and the J<sub>K</sub>-C<sub>K</sub> gene segment on the short arm of chromosome 2 may combine to form a potentially functional V<sub>K</sub>-J<sub>K</sub>-C<sub>K</sub> gene. This would have to happen by a somatic pericentric inversion. The 5'-3' orientation of V108 relative to the centromere is not known but it should be noted that an inversion leading to V<sub>K</sub>-J<sub>K</sub>-C<sub>K</sub> is possible only if V108 is organized such that its 5' end points towards the centromere. Large scale inversions that involve the immunoglobulin heavy chain and the T cell receptor loci on opposite ends of the long arm of the human chromosome 14 have been observed in T cell lymphomas, the products being V<sub>H</sub>-Ja C<sub>C</sub> hybrid genes (29). Yet pericentric inversions involving V108 and J<sub>K</sub>-C<sub>K</sub> should be so rare events that they, if they occur at all, would not contribute significantly to the light chain repertoire.

The finding of an immunoglobulin orphan with a bona fide sequence is not unique. Two human germline V<sub>H</sub> genes with bona fide sequences were recently found on chromosome 16, i.e. outside of the heavy chain locus which is located on chromosome 14 (T. Honjo, pers. communication).

ACKNOWLEDGEMENTS

We thank M. Pech for help in the early mapping experiments of cos108, R. Lamm for assistance in the sequencing, K.-H.

Grzeschik for cell hybrid DNAs and G. Weichhold for a pulsed field gel blot. The work was supported by Bundesministerium für Forschung und Technologie (Center Grant 0316200A2) and Fonds der Chemischen Industrie.

REFERENCES