A sequence homology between the pX genes of HTLV-I/II and the murine IL-3 gene

Takashi Gojobori*+, Shin-ichi Aota*, Tadashi Inoue* and Kunitada Shimotohno

*National Institute of Genetics, Mishima, Shizuoka-ken 411, Japan, +Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110, USA and National Cancer Center Research Institute, Tsukiji, Chuo-ku, Tokyo 104, Japan

Received 31 July 1986; revised version received 19 September 1986

Searching the protein sequence database for amino acid sequences homologous to the *x-lor* sequence in the pX region of human T-cell leukemia virus types I and II (HTLV-I/II), we found that there is a region of 38 amino acids where the murine interleukin 3 (IL-3) sequence has a 40% homology with the *x-lor* sequence. A statistical analysis shows that this homology is highly significant with a probability of 1.57×10^{-10}. The biological implication of this homology is discussed.

HTLV-I *HLTV-II* *pX gene* x-lor *gene* *Interleukin 3* *Sequence homology* *Leukemogenesis* *Oncogene*

1. INTRODUCTION

Human T-cell leukemia virus type I (HTLV-I) and type II (HTLV-II) are exogenous human retroviruses associated with certain diseases of T-cell malignancy [1–5]. In particular, HTLV-I is thought to be a causative agent for adult T-cell leukemia (ATL) [6,7]. One of the unique features of HTLVs that distinguish them from other chronic leukemia viruses is the presence of a unique gene called *x-lor* or *tat* in the pX region between the *env* and 3′-long terminal repeat (LTR) [8]. The *x-lor* gene does not seem to be a typical cell-derived oncogene because DNA of the pX region does not hybridize significantly with normal human DNA [9]. Since the *x-lor* gene sequence has been conserved among HTLV-I [8], HTLV-II [10], and even simian T-cell leukemia virus type I (STLV-I) [11], the *x-lor* gene product must mediate the neoplastic transformation of normal human T cells [12]. In fact, proteins of 41 and 38 kDa were found to be encoded from this region in HTLV-I- and HTLV-II-infected cells, respectively [13–16]. However, biological functions of the *x-lor* gene have not been identified.

2. METHODS AND RESULTS

With the aim of elucidating the biological function of the *x-lor* gene, we searched for amino acid sequences homologous to the *x-lor* gene sequence in the NBRF protein sequence database [17]. In our computer search, we found that the murine interleukin 3 (IL-3) sequence [9,18–21] has a region of 38 amino acids where the *x-lor* gene products of HTLV-I/II and STLV-I have a 40% homology with the murine IL-3 (see fig.1). A statistical analysis shows that this homology is highly significant with a probability of 1.57×10^{-10}. If the similarity between different amino acids is taken into account, the homology increases to a level of about 60% (see table 1).

For this homologous region of 38 amino acids, we also compared the hydrophobicity profile and the predicted secondary structures for *x-lor* with those for the murine IL-3. As shown in fig.2, the overall patterns are very similar to each other. In particular, the hydrophobicity profiles of both sequences exhibit four peaks roughly at the corresponding sites. Moreover, the predicted secondary structures show that for both sequences

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

α-helix structures can be located at the N- and C-terminal sides in the homologous regions, though the possible positions of β-structures for *x-lor* are slightly different from those for the murine IL-3. It is thus conceivable that murine IL-3 and *x-lor* possess similar secondary structures for the homologous region of 38 amino acids.

IL-3 is one of a number of colony stimulating factors which are known to regulate haematopoiesis [18]. The colony stimulating factors are hormone-like glycoproteins which are biologically active at low concentrations. IL-3 is produced by mitogen or antigen-activated T cells, and it is involved in regulating the growth and differentiation of pluripotent stem cells leading to the production of all the major blood cell types [20]. It also has a broad range of biological activities and appears to be identical with a number of other factors named on the basis of their biological activities, such as multi-colony stimulating factor CSF, haematopoietic growth factor, mast cell growth factor and so forth [18].

Southern hybridization analysis using a probe derived from a murine IL-3 cDNA clone revealed the presence of a single IL-3 gene in the haploid genome [20]. The murine IL-3 gene is made up of five exons interrupted by four introns [19,20]. The conserved region between the murine IL-3 and

Table 1

Homologies of amino acid sequences between the murine IL-3, and the region of 38 amino acids from the 46th to the 83rd residues in the *x-lor* of HTLV-I [8] and the corresponding regions of other viruses in the HTLV-I family

	IL-3	HTLV-I	STLV-I	HTLV-II	BLV[a]
IL-3	–	55%	58%	55%	31%
	–	(21)	(22)	(21)	(11)
HTLV-I	40%	–	92%	97%	53%
	(15)	–	(35)	(37)	(19)
STLV-I	40%	89%	–	89%	50%
	(15)	(34)	–	(34)	(18)
HTLV-II	37%	87%	76%	–	58%
	(14)	(33)	(29)	–	(21)
BLV	17%	39%	36%	33%	–
	(6)	(14)	(13)	(12)	–

[a] BLV has 36 amino acids, since the two gaps are necessary for the alignment with corresponding regions of other viruses in the HLTV-I family [27]

The percentages in the lower-left domain represent homologies for identical amino acids only, whereas those in the upper-right represent homologies for similar amino acids. The values within parentheses show the number of identical amino acids

```
IL-3     (74) RRVNLSKFVESQGEVDPEDRYVIKSNLQKLNCCLPTSA (111)
              *:: *   :* *  :** *  ** * ** *   **: :
HTLV-I   (46) RHALLATCPEHQITWDPIDGRVIGSALQFLIPRLPSFP ( 83)

              *:: *:  :* *  :** *  ** * ** *   **: :
STLV-I   (46) RHALLTTCPEHQITWDPTDERVIGSALQFLIPRLPSLP ( 83)

              *:: *   :* *  :** *  *: * ** *   **: :
HTLV-II  (48) RHALLATCPEHQLTWDPIDGRVVSSPLQYLIPRLPSFP ( 85)
              #++ #   +# #  +## #  #+ # ## #   ##+ +

              *  :          : : *  :   * : *   *  *
BLV      (42) RIDTTLTCETHRINWTA-DGRPCGLN-GTLFPRLHVSE ( 77)
              #             + + #  +     + #   #
```

Fig.1. Alignment of the amino acid sequences between the murine IL-3 [18] and the *x-lor* regions of HTLV-I [8], STLV-I [11], HTLV-II [10] and BLV [26] for the conserved region of 38 amino acids. * and : respectively represent amino acids of the HTLV-I family identical and similar to those of the murine IL-3. The probabilities that the obtained alignment with IL-3 occurs by chance are 1.57×10^{-10}, 1.57×10^{-10}, 1.88×10^{-9}, and 8.34×10^{-3} for HTLV-I, STLV-I, HTLV-II and BLV, respectively. # and + represent sites where all genes have identical and similar amino acids, respectively. The numbers in parentheses show residue numbers of starting and ending sites for the conserved region in a given *x-lor* sequence.

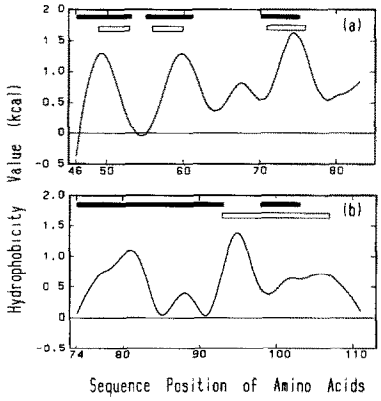

Fig.2. Hydrophobicity profiles and predicted secondary structures of the regions of 38 amino acids from the 46th to the 83rd residues of the *x-lor* of HTLV-I (a) and the corresponding region (the 74th to the 111th residues) of the murine IL-3 (b). The hydrophobicity profiles were obtained by the method of Rose and Roy [37]. The ordinates represent free energy of transfer from aqueous to organic solvent. The secondary structures were predicted from conformational parameters of amino acids computed by Chou and Fasman [38,39]. Solid and open bars represent the possible regions of α-helix and β-structures, respectively.

HTLV-I/II occupies almost entire domains of exons 3 and 4. Analysis of a chemically synthesized protein of the murine IL-3 suggests that an amino terminal fragment of 79 amino acids in the mature murine IL-3 protein is sufficient to stimulate growth of IL-3-dependent cell lines [22]. Interestingly, the conserved region is corresponding to the last 32 amino acids of the amino terminal fragment. Thus, this conserved region may have a function common to IL-3 and *x-lor*.

Recently, the rat IL-3 gene sequence has also been determined [23]. Although the predicted amino acid sequences for the mature rat IL-3 shows a relatively low homology (54%) with its murine counterpart, the rat IL-3 still retains an about 30% homology (and an about 50% homology for similar amino acids) with HTLV-I/II at the corresponding conserved region of the murine IL-3 with HTLV-I. However, Southern hybridization analysis of mammalian DNAs, using a murine IL-3 cDNA probe, fails to detect homologous sequences in most mammalian species including human (except rat), even under conditions for relatively low stringency [9]. This apparent low conservation of mammalian IL-3 genes may be related to the fact that the sequence homologous to the pX region of HTLV-I is represented in the genomes of mouse and rat but not in other species including primates and human [9]. This suggests that a part of the pX sequence may have derived from the rodent gene [9]. Since HTLV-I can infect primates, rabbit and rat, HTLV-I might have been prevalent among a wide variety of mammals and exchanged genetic segments (zoonotic) like influenza virus [9]. If the original host of HTLV-I is a rodent rather than man, the *x-lor* sequence should be homologous to host cellular sequences such as the rodent IL-3 genes and it can be reminiscent of the v-onc gene.

Our finding that the *x-lor* gene has a significant homology with the rodent IL-3 gene is also consistent with the following facts. To date, various T-cell lines established by HTLV-induced transformation have been demonstrated to produce several lymphokines including IL-3 [24,25]. This suggests that HTLV-I-induced transformation of T cells could induce the production of various lymphokines and that the transformed cells retain their stable gene expression. Moreover, Ymer et al. [26] have reported that the constitutive synthesis of IL-3 by the murine myelomonocytic leukemia cell line, WEHI-3B, is due to the insertion of an endogenous retrovirus-like element close to the 5′-end of the IL-3 gene. It is possible that the activation of the IL-3 gene may have been an important step in the production of this leukemia [26]. Thus, biological functions of the pX gene involved in the mechanism of HTLV-induced leukemogenesis may be related to those of IL-3.

Bovine leukemia virus (BLV) also possesses a potential transforming gene, X_{BL}, between the *env* gene and 3′-LTR [27]. The amino acid sequence of the X_{BL} gene of BLV is known to be homologous to that of HTLV-I only for the N-terminal region of about 80 amino acids [27,28]. The homologous region of the murine IL-3 with HTLV-I/II corresponds exactly to the last 36 amino acids of the homologous region between HTLV-I and BLV, though the homology between the murine IL-3 and BLV is not high (see fig.1). This lower homology between the murine IL-3 and BLV may be connected with the fact that enzootic bovine leukosis caused by BLV is mainly for B cells (not T cells) [29,30].

An IL-2 autocrine hypothesis for ATL leukemogenesis was previously proposed on the basis of the observation that an HTLV-carrying cell line, HUT102, produces IL-2 [31]. However, the inadequacy of this hypothesis was demonstrated by the finding that the IL-2 gene was not transcribed in any other cell lines carrying HTLV-I [31,32]. In fact, we could not find any significant homology between the coding region of IL-2 and the *x-lor* of HTLV-I/II. In contrast, it is suggested that IL-2 receptor (IL-2R) expression by ATL leukemia cells could be involved in the mechanism of their leukemogenesis, since Sugamura et al. [33] obtained direct evidence for the IL-2R-inducing ability of HTLV-I. Recently, Birchenall-Sparks et al. [34] found that expression of the IL-2 receptor gene in hematopoietic cells can be regulated by IL-3. Thus, our findings are consistent with those observations. However, the *x-lor* gene products in all of HTLV-I/II are localized mostly in the nucleus of an infected cell [35]. Since the IL-3 protein may be secreted to the outside of the T cells, whether the *x-lor* gene has exactly the same function as the IL-3 gene remains to be solved.

Recent evidence for a direct role of *x-lor* in trans-acting transcriptional regulation was provided by the experiment of transcriptional activation of an LTR-linked CAT gene in cells transfected with only the *x-lor* gene [36]. Thus, it has been speculated that the *x-lor* gene product not only activates transcription of viral genes, but also positively or negatively regulates transcription of some cellular genes [12]. Activation of these genes on virus infection would then lead to immortalization and subsequently transformation of these cells. This mechanism of leukemogenesis may be similar to the mechanism of T-cell activation by growth factors. However, the biological implication of the homology between the murine IL-3 and *x-lor* should be examined experimentally. We hope that our finding can facilitate many experimentalists to pay attention to IL-3 for elucidation of molecular mechanism of leukemogenesis.

ACKNOWLEDGEMENTS

We thank Drs H. Hayashida, T. Ikemura, T. Maruyama, T. Miyata, T. Yasunaga and S. Yokoyama for their helpful comments. This study was supported by a Grant-in-Aid from the Ministry of Education, Science and Culture of Japan to T.G., T.I. and K.S., and NIH grant GM-28672 and NIMH grant MH-31302 to Dr Shozo Yokoyama.

REFERENCES

[1] Poiesz, B.J., Ruscetti, F.W., Gazdar, A.F., Bunn, P.A., Minna, J.D. and Gallo, R.C. (1980) Proc. Natl. Acad. Sci. USA 77, 7415–7419.
[2] Reitz, M.R., Poiesz, B.J., Ruscetti, F.W. and Gallo, R.C. (1981) Proc. Natl. Acad. Sci. USA 78, 1887–1891.
[3] Hinuma, Y., Nagata, K., Hanaoka, M., Nakai, M., Matsumoto, T., Kinoshita, K., Shirakawa, S. and Miyoshi, I. (1981) Proc. Natl. Acad. Sci. USA 78, 6476–6480.
[4] Yoshida, M., Miyoshi, I. and Hinuma, Y. (1982) Proc. Natl. Acad. Sci. USA 79, 2031–2035.
[5] Watanabe, T., Seiki, M. and Yoshida, M. (1984) Virology 133, 238–241.
[6] Yoshida, M., Seiki, M., Yamaguchi, K. and Takatsuki, K. (1984) Proc. Natl. Acad. Sci. USA 81, 2534–2537.
[7] Yoshida, M., Seiki, M., Hattori, S. and Watanabe, Y. (1984) in: Human T-cell Leukemia/Lymphoma Viruses (Gallo, R.C. et al. eds) pp.141–148, Cold Spring Harbor, NY.
[8] Seiki, M., Hattori, S., Hirayama, Y. and Yoshida, M. (1983) Proc. Natl. Acad. Sci. USA 80, 3618–3622.
[9] Fukui, K., Noma, T., Takeuchi, K., Kobayashi, N., Hatanaka, M. and Honjo, T. (1983) Mol. Biol. Med. 1, 447–456.
[10] Shimotohno, K., Takahashi, Y., Shimizu, N., Gojobori, T., Golde, D.W., Chen, I.S.Y., Miwa, M. and Sugimura, T. (1985) Proc. Natl. Acad. Sci. USA 82, 3101–3105.
[11] Watanabe, T., Seiki, M., Tsujimoto, H., Miyoshi, I., Hayami, M. and Yoshida, M. (1985) Virology 144, 59–65.
[12] Wang-Staal, F. and Gallo, R.C. (1985) Nature 317, 395–403.
[13] Miwa, M., Shimotohno, K., Hoshino, H., Fujino, M. and Sugimura, T. (1984) Gann 75, 752–755.
[14] Kiyokawa, T., Seiki, M., Imagawa, K., Shimizu, F. and Yoshida, M. (1984) Gann 75, 747–751.
[15] Lee, T.H., Coligan, J.E., Sodroski, J.G., Haseltine, W.A., Salahuddin, S.Z., Wang-Staal, F., Gallo, R.C. and Essex, M. (1984) Science 226, 57–61.

[16] Slamon, D.J., Shimotohno, K., Cline, M.J., Golde, D.W. and Chen, I.S.Y. (1984) Science 226, 61–64.
[17] NBRF Protein Database (1986) National Biomedical Research Foundation, Georgetown University Medical Center, Washington, DC.
[18] Fung, M.C., Hapel, A.J., Ymer, S., Cohen, D.R., Johnson, R.M., Campbell, H.D. and Young, I.G. (1984) Nature 307, 233–237.
[19] Miyatake, S., Yokota, T., Lee, F. and Arai, K.-I. (1985) Proc. Natl. Acad. Sci. USA 82, 316–320.
[20] Campbell, H.D., Ymer, S., Fung, M.C. and Young, I.G. (1985) Eur. J. Biochem. 150, 297–304.
[21] Yokota, T., Lee, F., Rennick, D., Hall, C., Arai, N., Mosmann, T., Nabel, G., Cantor, H. and Arai, K.-I. (1984) Proc. Natl. Acad. Sci. USA 81, 1070–1074.
[22] Clark-Lewis, L., Aebersold, R., Zilterner, H., Schrader, J.W., Hood, L.E. and Kent, S.B.H. (1986) Science 231, 134–139.
[23] Cohen, D.R., Hapel, A.J. and Young, I.G. (1986) Nucleic Acids Res. 14, 3641–3658.
[24] Salahuddin, S.Z., Markham, P.D., Lindner, S.G., Gootenberg, J., Popovic, M., Hemmi, H., Sarin, P.S. and Gallo, R.C. (1984) Science 223, 703–707.
[25] Hinuma, S., Sugamura, K., Tsukamoto, K. and Hinuma, Y. (1984) Microbiol. Immunol. 28, 935–947.
[26] Ymer, S., Tucker, W.Q.J., Sanderson, C.J., Hapel, A.J., Campbell, H.D. and Young, I.G. (1985) Nature 317, 255–258.
[27] Sagata, N., Yasunaga, T., Ohishi, K., Tsuzuku-Kawamura, J., Onuma, M. and Ikawa, Y. (1984) EMBO J. 3, 3231–3237.
[28] Sagata, N., Yasunaga, T. and Ikawa, Y. (1985) FEBS Lett. 192, 37–42.
[29] Burny, A., Bruck, C., Chantrenne, H., Cleuter, Y., Dekegel, D., Ghysdael, J., Kettmann, R., Leclercq, M., Mammerickx, M. and Portetelle, D. (1980) in: Viral Oncology (Klein, G. ed.), pp.231–289, Raven, New York.
[30] Deschamps, J., Kettmann, R. and Burny, A. (1981) J. Virol. 40, 605–609.
[31] Sugamura, K. and Hinuma, Y. (1985) Immunol. Today 3, 83–88.
[32] Arya, S.K., Wong-Staal, F. and Gallo, R.C. (1984) Science 223, 1086–1087.
[33] Sugamura, K., Fujii, M., Kobayashi, N., Sakitani, M., Hatanaka, M. and Hinuma, Y. (1984) Proc. Natl. Acad. Sci. USA 81, 7441–7445.
[34] Birchenall-Sparks, M.C., Farrar, W.L., Rennick, D., Kilian, P.L. and Ruscetti, F.W. (1986) Science 233, 455–458.
[35] Goh, W.-C., Sodroski, J. and Rosen, C. (1985) Science 227, 1227–1228.
[36] Sodroski, J.G., Rosen, C.A. and Haseltine, W.A. (1984) Science 225, 381–385.
[37] Rose, G.D. and Roy, S. (1980) Proc. Natl. Acad. Sci. USA 77, 4643–4647.
[38] Chou, P.Y. and Fasman, G.D. (1974) Biochemistry 13, 211–222.
[39] Chou, P.Y. and Fasman, G.D. (1974) Biochemistry 13, 222–245.