Year: 1997

A human endogenous retroviral superantigen as candidate autoimmune gene in type I diabetes

Conrad, B; Weissmahr, R N; Böni, J; Arcari, R; Schüpbach, J; Mach, B

A human endogenous retroviral superantigen as candidate autoimmune gene in type I diabetes

Abstract

Microbial superantigens (SAGs) have been implicated in the pathogenesis of human autoimmune diseases. Preferential expansion of the Vveta7 T cell receptor positive T cell subset in patients suffering from acute-onset type I diabetes has indicated the presence of a surface membrane-bound SAG. Here, we have isolated a novel mouse mammary tumor virus-related human endogenous retrovirus. We further show that the N-terminal moiety of the envelope gene encodes an MHC class II-dependent SAG. We propose that expression of this SAG, induced in extrapancreatic and professional antigen-presenting cells, leads to beta-cell destruction via the systemic activation of autoreactive T cells. The SAG encoded by this novel retrovirus thus constitutes a candidate autoimmune gene in type I diabetes.

Cell, Vol. 90, 303–313, July 25, 1997, Copyright ©1997 by Cell Press

A Human Endogenous Retroviral Superantigen as Candidate Autoimmune Gene in Type I Diabetes

Bernard Conrad,* Richard Nicolas Weissmahr,†
Jürg Böni,† Rosanna Arcari,‡
Jörg Schüpbach,† and Bernard Mach*
*Department of Genetics and Microbiology
University of Geneva Medical School
CH-1211 Geneva 4
Switzerland
†Swiss National Centre for Retroviruses
University of Zurich
CH-8028 Zurich
Switzerland
‡Department of Internal Medicine
University of Torino Medical School
10126 Torino
Italy

Summary

Microbial superantigens (SAGs) have been implicated in the pathogenesis of human autoimmune diseases. Preferential expansion of the Vβ7 T cell receptor positive T cell subset in patients suffering from acute-onset type I diabetes has indicated the presence of a surface membrane–bound SAG. Here, we have isolated a novel mouse mammary tumor virus–related human endogenous retrovirus. We further show that the N-terminal moiety of the envelope gene encodes an MHC class II–dependent SAG. We propose that expression of this SAG, induced in extrapancreatic and professional antigen-presenting cells, leads to β-cell destruction via the systemic activation of autoreactive T cells. The SAG encoded by this novel retrovirus thus constitutes a candidate autoimmune gene in type I diabetes.

Introduction

Insulin-dependent diabetes mellitus (IDDM) is a β cell–specific, T cell–mediated autoimmune disease (reviewed by Steinman, 1995; Tisch and McDevitt, 1996). Multiple genes contribute to the familial clustering of the disease, the major histocompatibility complex (MHC) being the most important one (reviewed by Vyse and Todd, 1996). The concordance rate in monozygotic twins is, however, only 50% (reviewed by Tisch and McDevitt, 1996). The incidence of type I diabetes has increased over the last three decades. Moreover, it shows temporal, epidemic-like variations, and the clinical disease exhibits preferential seasonal onset (reviewed by Karvonen et al., 1993). Hence, combined genetic and epidemiological evidence has been interpreted to mean that unknown environmental factors may play a role in either precipitating the disease and/or dominantly shaping its course (Steinman, 1995).

A genome-wide analysis of polymorphism associated with microsatellite markers in multiplex IDDM families has implicated the MHC class II genotype as one of the strongest genetic factors determining disease susceptibility (reviewed by Vyse and Todd, 1996). However, the respective roles of the different MHC class II$^+$ cell types in promoting disease has not yet been clarified. It is now clear that MHC class II expression on pancreatic β cells and the vascular endothelium is irrelevant for the development of insulitis in NOD mice and for the recurrence of the human disease (Lo et al., 1993; Stegall et al., 1996; Tydén et al., 1996). On the other hand, selective binding of self peptides to disease-associated MHC class II molecules presented by professional antigen-presenting cells (APCs) may explain the observed T- and B-cell reactivity toward autoantigens (reviewed by Wucherpfennig and Strominger, 1995a). The natural occurrence of autoreactivity, however, does not explain how autoreactive T cells are activated and gain access to the target tissues (Pette et al., 1990; Mackay, 1993). Two independent, but not necessarily mutually exclusive, mechanisms have been proposed to account for the activation of peripheral autoreactive T cells: molecular mimicry (Oldstone, 1990; Wucherpfennig and Strominger, 1995b) and bacterial or viral superantigens (SAGs) (Brocke et al., 1993; Cole and Griffiths, 1993; Conrad et al., 1994). Both invoke mechanisms that could reconcile the importance of environmental factors with the requirement for a major and systemic T cell stimulatory event. Superantigens (Choi et al., 1989; White et al., 1989) are microbial proteins able to mediate interactions between MHC class II$^+$ and polyclonal T cells resulting in reciprocal activation (Fleischer and Schrezenmeier, 1988; Acha-Orbea et al., 1991; Choi et al., 1991). Their function is restricted by only two absolute requirements: the presence of MHC class II on the surface of the presenting cells and the expression of one or more defined variable (V)β T cell receptor (TCR) chain(s) on T cells.

The potential role of SAGs in human diseases is much less defined. Bacterial SAGs have been proposed to be implicated in the pathogenesis of autoimmune diseases (White et al., 1989). However, although pathogen disease associations have been described, none of these have as yet implicated a pathogen-encoded SAG (Howell et al., 1991; Paliard et al., 1991). A SAG-like activity resembling the one encoded by MMTV has been reported to be associated with herpesvirus infections (Dobrescu et al., 1995; Sutkowski et al., 1996). However, in none of these two systems has it been demonstrated that the SAG activity is actually encoded by the infectious agent.

In two patients with type I diabetes, a dominant pancreatic enrichment of one Vβ family, Vβ7, has been observed (Conrad et al., 1994). The same dominant enrichment of Vβ7 could be mimicked by stimulating T cells of diverse haplotypes with surface membrane preparations derived from the pancreatic inflammatory lesions but not with membranes from MHC-matched healthy control islets. This was taken as evidence for the presence of a surface membrane–associated SAG (Conrad et al., 1994). Given the lack of clear epidemiological evidence linking specific bacterial or viral diseases to type I diabetes, we tested the hypothesis that this

SAG is of endogenous retroviral origin. Here, we show that the SAG identified in these two patients is encoded by a human endogenous retrovirus related to MMTV. Expression of this endogenous SAG in IDDM suggests a general model according to which self SAG–driven and systemic activation of autoreactive T cells leads to organ-specific autoimmune disease.

Results

Cultured Leukocytes from Inflammatory β-Cell Lesions of IDDM Patients Release Reverse Transcriptase Activity

Expression of cellular retroelements may be associated with measurable reverse transcriptase (RT) activity (Heidmann and Heidmann, 1991). An RT assay detected up to a 100-fold increase in RT activity in supernatants from short-term cultures of freshly isolated pancreatic islets derived from two patients (Figure 1A) (Conrad et al., 1994; Pyra et al., 1994). No RT activity above background levels was detected in medium controls, indicating that the RT activity could not be accounted for by a contamination of the synthetic media and sera with animal retroviruses. We can also exclude the possibility that the RT activity represents cellular polymerases released into the supernatant by dying cells. Indeed, no RT activity can be detected in cultures from nondiabetic controls under conditions in which cell death is strongly enhanced, namely mitogen-treated peripheral blood lymphocytes (PBL), splenocytes, and cocultures of islets with allogeneic T cells. Moreover, the IDDM-derived islets were cultured for 5 days, whereas control cultures were sequentially analyzed for up to 4 weeks. Finally the absence of RT activity in the supernatants of the mitogen-treated control PBL also excluded the possibility that the RT activity detected with the IDDM islets was simply due to nonspecific cell activation. Both the islets and the inflammatory infiltration represented potential sources for the enzymatic activity. As shown in Figure 1B, supernatants from cultured spleen cells from the patients contained more RT activity than the inflammatory β-cell lesions. Moreover, the RT activity disappeared together with the local inflammatory lesion in two patients with chronic and long-standing disease, but it persisted in cultured spleen cells from the same patient (Figure 1B). This was interpreted as being compatible with the leukocytes as the most likely source of this RT activity.

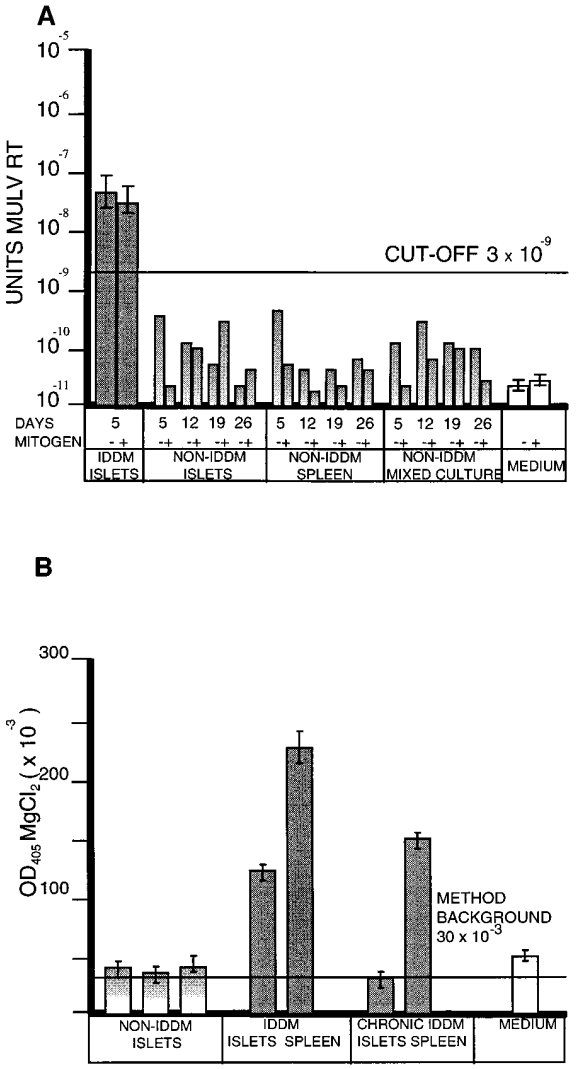

Figure 1. Leukocytes from IDDM Patients Release Reverse Transcriptase Activity

(A) Supernatants derived from cultured islets isolated from two patients (Conrad et al., 1994) were assayed for RT activity, using a half-logarithmic dilution series of purified murine leukemia virus (MULV) RT as a standard (Pyra et al., 1994). Results are expressed as mean ± 1 SD. Islets and spleen cells from nondiabetic organ donors were cultured either alone, in the absence or presence of mitogen (−/+), or together in mixed allogeneic cultures (time as days in culture prior to collection of the supernatant is indicated below the bars).

(B) Islets and spleen cells from three nondiabetic organ donors, from the two patients with acute-onset IDDM, and two patients with chronic IDDM (Conrad et al., 1994) were cultured for 1 week, and supernatants were analyzed for the presence of RT activity. Results are expressed as mean ± 1 SD for at least three individual measurements.

Isolation of a Full-Length Retroviral Genome, IDDMK$_{1,2}$22, from Supernatants of IDDM Islets

A strategy to isolate putative retroviral genomes from polyadenylated RNA extracted from the supernatants of IDDM islets was developed (Figure 2A). This strategy relies on the following three characteristic features of functional retroviruses. First, each retroviral genome contains a primer binding site (PBS) near its 5′ end. Cellular tRNAs anneal to the PBS and serve as primers for reverse transcriptase (reviewed by Whitcomb and Hughes, 1992). Second, the R (repeat) sequence is repeated at the 5′ and 3′ ends of the viral RNA (Temin, 1981). Third, the RT-RNAse H region of the *pol* gene is the most conserved sequence among different retroelements (McClure et al., 1988; Xiong and Eickbusch, 1990). These three features were exploited in a six-step procedure as follows.

(1) To isolate the 5′ ends (5′R-U5) of putative retroviral RNA genomes, a 5′ RACE procedure was performed

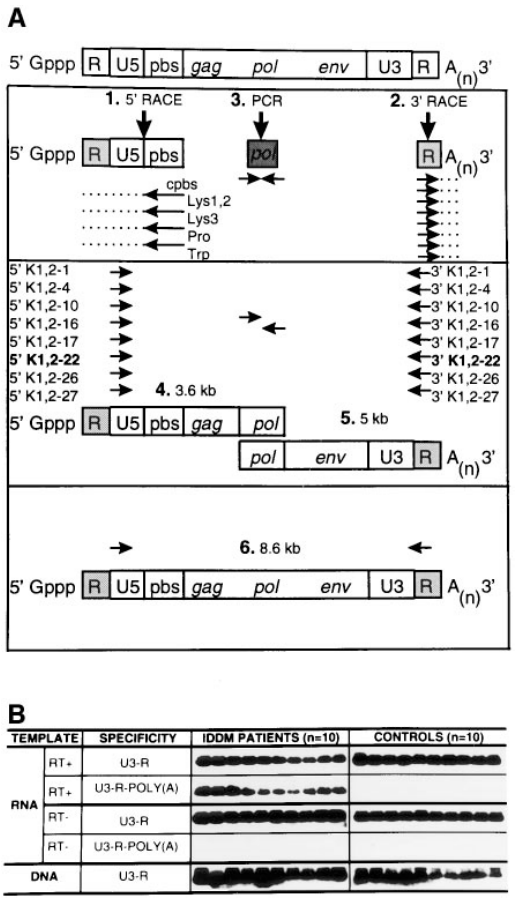

Figure 2. A Single Endogenous Retroviral Genome, IDDMK$_{1,2}$22, Is Associated with Type I Diabetes at Disease Onset

(A) Isolation of a single full-length retroviral genome, IDDMK$_{1,2}$22, with a six-step procedure. (1) cPBS primers (Lys1,2, Lys3, Pro, Trp) were used to perform a 5′ RACE. (2) The eight 5′ R-U5 sequences obtained in (1) were used to perform a 3′ RACE with primers annealing in the R. (3) The conserved RT-RNAse H region was amplified with degenerate primers. (4) The 5′ moiety (the predicted size for full-length HERV-K-retroviruses is 3.6 kb) was amplified by PCR using primers specific for the eight 5′ R-U5 sequences in conjunction with a primer specific for the 3′ of the central *pol* region obtained in step 3. The primer specific for the K$_{1,2}$22 5′ consistently yielded a fragment of this size. (5) The 3′ (the predicted size for HERV-K-retroviruses is 5 kb) was amplified by PCR using a primer specific for the 5′ of the central *pol* region isolated in step 3 and primers specific for the poly(A) signals present in the 3′ R-poly(A) sequences obtained in step 2. The PCR reaction using a primer specific for the 3′ clone K$_{1,2}$22 (amplified in step 4) consistently yielded a fragment potentially representing an intact 3′ HERV-K moiety of 5 kb. (6) The presence of an intact 8.6 kb retroviral genome containing the overlapping 5′ and 3′ moieties isolated in steps 4 and 5 was confirmed by PCR using primers specific for its predicted U5 and U3 regions.

(B) IDDMK$_{1,2}$22 is an endogenous retrovirus found in the plasma of IDDM patients at disease onset but not in the plasma of healthy controls. Signals in the first and third rows correspond to amplification of contaminating DNA present in the plasma of IDDM patients (left-hand columns, 1–10) and controls (right-hand columns, 1–10) and were, as expected, RT independent. In contrast, signals in the second row resulted from the amplification of viral RNA present only in IDDM patients (left-hand columns, 1–10) but not in the nondiabetic controls (right-hand columns, 1–10). This was supported by the absence of amplification products in reactions lacking RT (fourth row, right- and left-hand columns, 1–10). In addition, the signal could be diminished below background by RNAse treatment (data not shown). In the fifth row, the genomic DNA from IDDM patients and controls was amplified with the U3-R-specific primers.

with primers complementary to known PBS sequences (cPBS primers) (Weissmahr et al., 1997). Eight different candidate 5′R-U5 sequences (5′K$_{1,2}$-1, -4, -10, -16, -17, -22, -26, and -27) were obtained with the cPBS-lysine$_{1,2}$ primer. All eight sequences contained features typical of the 5′ ends of retroviral genomes (Temin, 1981). These include at the expected positions the presence of: (i) a PBS region; (ii) conserved and correctly spaced upstream regulatory sequences, such as a poly(A) addition signal and site, and the downstream GT- or T-rich elements (Wahle and Keller, 1996); (iii) a putative 5′ end–specific U5 region; and (iv) a putative R region. Of the eight 5′R-U5 sequences isolated, three (5′K$_{1,2}$-1, -4, and -22) were identified on the basis of sequence homology as belonging to previously identified families of human endogenous retroviruses (HERVs) that are closely related to mouse mammary tumor viruses (MMTV), namely HERV-K(C4) (Tassabehji et al., 1994), HERV-K10, and HERV-K18 (Ono, 1986; Ono et al., 1986). The remaining five sequences exhibited only a distant relationship with HERV-K retroviruses.

(2) The eight 3′R-poly(A) ends (3′K$_{1,2}$-1, -4, -10, -16, -17, -22, -26, and -27) corresponding to the eight different 5′R-U5 regions identified in step 1 were isolated by means of a 3′ RACE procedure using primers specific for the R regions. In each case, the isolated sequences contained the expected R region followed by a poly(A) tail.

(3) The conserved RT-RNase H region within the *pol* gene was next amplified by PCR using degenerate primers (Medstrand and Blomberg, 1993). Fifteen individual subclones were sequenced, and all exhibited approximately 95% similarity at the protein level to the RT-RNase H region of the HERV-K family.

(4) The 5′ moiety (from the U5 region at the 5′ end to the *pol* gene) of the putative retroviral genome was amplified by PCR using primers specific for the eight different U5 regions present in the 5′R-U5 sequences (isolate in step 1) in conjunction with a primer specific for the 3′ end of the central *pol* region (isolated in step 3). The expected size of the PCR product corresponding to the 5′ moiety of full-length HERV-K retroviruses is 3.6 kb (Ono et al., 1986). Only the PCR reaction using the primer specific for the K$_{1,2}$22 5′ end clone consistently yielded a fragment of this size. Sequence analysis of several independent clones confirmed that this 3.6 kb fragment contains the R-U5-PBS region followed by coding regions corresponding to the *gag* and *pol* genes, and thus indeed represents the 5′ moiety of an intact retroviral genome.

(5) The 3′ moiety (from the *pol* gene to the 3′ end) of the putative retroviral genome was amplified by PCR using a primer specific for the 5′ end of the central *pol* region (isolated in step 3) and primers specific for the poly(A) signals present in the 3′R-poly(A) sequences (isolated in step 2). The expected size of the PCR product corresponding to the 3′ moiety of full-length HERV-K-retroviruses is 5 kb (Ono et al., 1986). The PCR

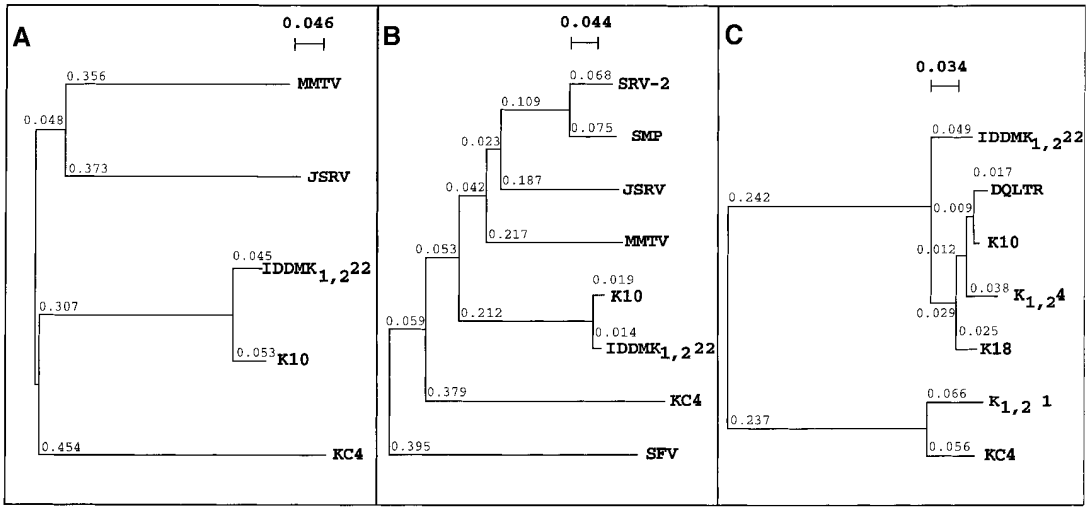

Figure 3. Phylogenetic Trees of Coding and Noncoding Regions Place IDDMK$_{1,2}$22 in the HERV-K10 Family of HERVs

(A) IDDMK$_{1,2}$22 SU-ENV is most closely related to HERV-K10 and is also related to the B-type retroviruses MMTV and JSRV.

(B) The phylogenetic analysis of the RT region shows that IDDMK$_{1,2}$22 belongs to the HERV-K10 family and is more closely related to B-type retroviruses such as MMTV than to D-type retroviruses such as Simian Mason Pfizer (SMP) or Spumaviridae (SFV). Abbreviations used: SRV-2, Simian retrovirus; JSRV, Jaagsiekte Sheep retrovirus; SFV, Simian foamy virus.

(C) The noncoding LTR region was used to construct a phylogenetic tree of the HERV-K family. K$_{1,2}$1 and K$_{1,2}$ 4 (see above) were isolated only as subgenomic or truncated transcripts. K$_{1,2}$1 is related to KC4, while K$_{1,2}$4 and IDDMK$_{1,2}$22 are related to the K10/K18 subfamily. Within this family, K$_{1,2}$4 is closely related to K10, whereas IDDMK$_{1,2}$22 appears to be more distant.

reaction using a primer specific for the 3′ end clone K$_{1,2}$22, which is the one that should correspond to the 3′ end of the retrovirus from which the 3.6 kb 5′ moiety was amplified in step 4, consistently yielded a fragment potentially representing an intact 3′ moiety of 5 kb. Sequence analysis of several independent clones confirmed that this 5 kb fragment indeed contains coding regions corresponding to the *pol* and *env* genes followed by the expected U3-R-poly(A) region.

(6) Finally, the presence of an intact 8.6 kb retroviral genome containing the overlapping 5′ and 3′ moieties isolated in steps 4 and 5 was confirmed by PCR using primers specific for its predicted U5 and U3 regions.

The full-length retroviral genome that was isolated was called IDDMK$_{1,2}$22, where IDDM refers to the tissue source, K$_{1,2}$ refers to lysine$_{1,2}$ cPBS primer, and 22 represents the serial number of the clone. IDDMK$_{1,2}$22 was determined to be a novel retrovirus on the basis of two criteria. First, it has a unique pattern of restriction enzyme cleavage sites that is distinct from that of other known viruses. Second, its nucleotide and amino acid sequences in noncoding and coding regions diverge from other known retroviruses by at least 5%–10%.

IDDMK$_{1,2}$22 was the only full-length virus identified in these experiments, suggesting that it is the only functional retrovirus specifically associated with the supernatants of the cultured IDDM islets. PCR reactions using primers specific for the other 5′R-U5-PBS and 3′U3-R-poly(A) clones isolated in steps 1 and 2 did not yield fragments of the size expected for intact retroviral genomes in steps 4 and 5. In particular, primers specific for the 5′ and 3′ ends corresponding to the ubiquitous HERV-K10 virus did not amplify fragments corresponding to complete genomes, although this virus is known to be released as a full-length genome associated with viral particles from several cell lines and tissues (Tönjes et al., 1996). Our inability to detect full-length HERV-K10 genomes in the IDDM islet supernatant is unlikely to be due to a technical problem because it could be amplified very efficiently from both genomic DNA and a size-selected cDNA library prepared from a B-lymphoblastoid cell line (data not shown). It is more likely that HERV-K10 is not released in significant amounts by the cultured IDDM islets.

PCR primer pairs were designed that are either specific for the U3-R or for the U3-R-poly(A) region of IDDMK$_{1,2}$22 (see Experimental Procedures). The U3-R primer pair amplified both viral RNA and DNA, whereas the U3-R poly(A) primer pair amplified selectively viral RNA. In addition, the signal generated with the U3-R-poly(A) primer pair could be diminshed below background by RNAse treatment and was not detected when genomic DNA was used as a template (data not shown). The amplified material was hybridized with probes generated with the molecularly cloned U3-R region of IDDMK$_{1,2}$22. Using this technique, we confirmed by RNA-specific PCR that sequences identical, or highly similar, to the 3′ U3-R-poly(A) of IDDMK$_{1,2}$22 were present in RT-positive but not in RT-negative samples analyzed; in a preliminary epidemiological study, we detected sequences identical, or highly similar, to the 3′ U3-R-poly(A) of IDDMK$_{1,2}$ only in the plasma of 10 recent onset IDDM patients but not in the plasma of 10 age-matched nondiabetic controls (Figure 2B); and we confirmed the presence of sequences identical, or highly similar, to the U3-R region of IDDMK$_{1,2}$ in genomic DNA of IDDM patients (n = 10) and nondiabetic controls (n = 10) (Figure 2B). In summary, these data indicate that IDDMK$_{1,2}$ is an endogenous retrovirus that is released from leukocytes in IDDM patients but not in nondiabetic controls.

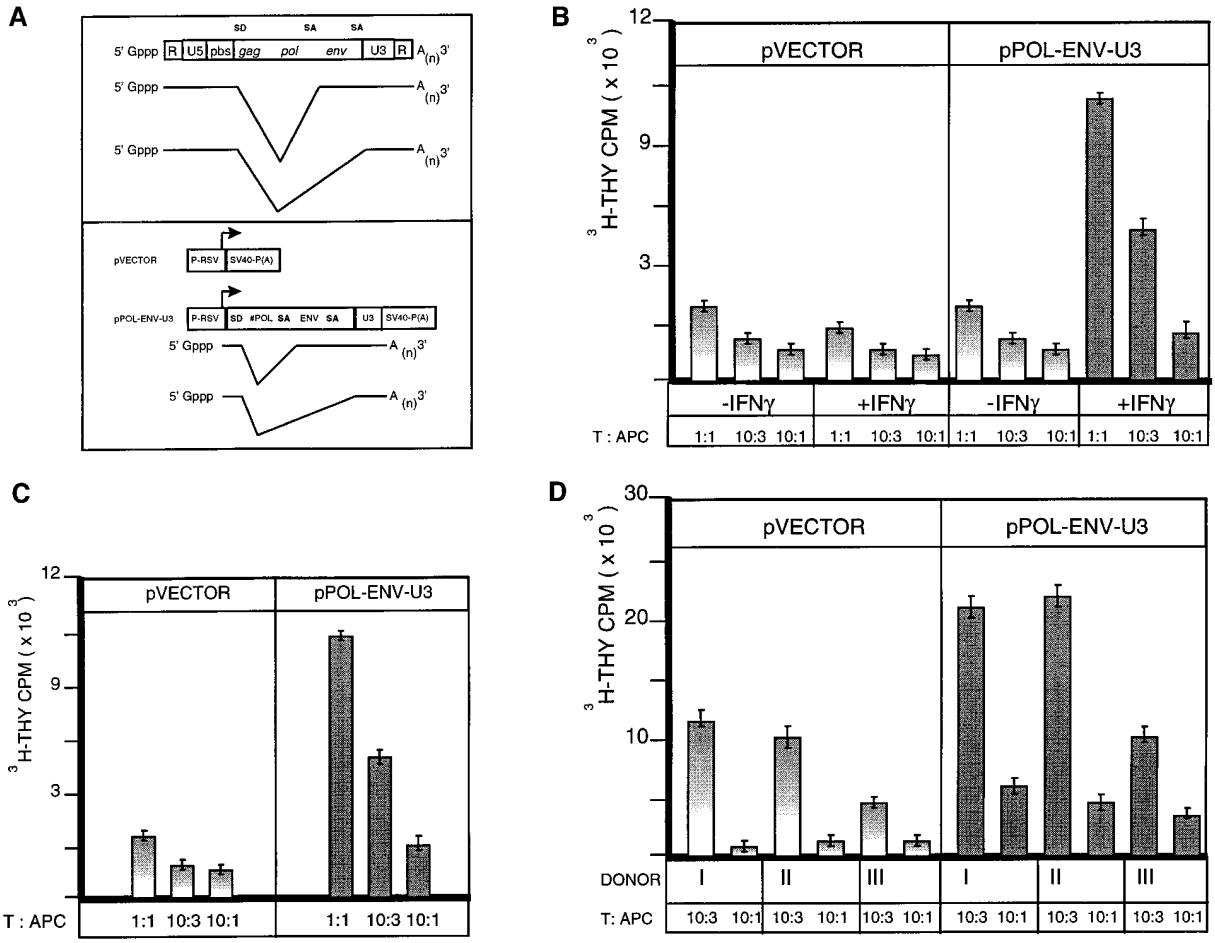

Figure 4. The *pol-env*-U3-R Region of IDDMK$_{1,2}$22 Exerts an MHC class II–Dependent but Not MHC-Restricted Mitogenic Effect upon Transfection in Monocytes

(A) IDDMK1,222 is expected to generate two singly spliced subgenomic RNAs, one encoding ENV and one comprising the U3-R region. The episomal expression vector was engineered to carry a proximal SD downstream of the promoter (pPOL-ENV-U3). Thus, the two naturally expected subgenomic RNAs can also be generated.

(B) Monocytic cell lines do not express MHC class II surface proteins in the absence of induction by IFN-γ (reviewed by Mach et al., 1996). The monocyte cell line THP1 was transiently transfected with pPOL-ENV-U3 or with the expression vector alone (pVECTOR). Mitomycin C–treated transfectants, either induced with IFN-γ for 48 hr or noninduced (+/- IFN-γ, indicated below the X axis) were cultured with MHC-compatible T cells at different responder:stimulator ratios as indicated below the graphs (T:APC). ^{3}H-thymidine incorporation was measured during the last 18 hr of a 72 hr culture and is given on the Y axis as n × 10^3 cpm. Results are presented as mean ± 1 SD.

(C) The MHC class II transactivator CIITA mediates IFN-γ-inducible MHC class II expression (reviewed by Mach et al., 1996). An integrative and stable THP1-CIITA transfectant (THP1-CIITA) was transfected with pVECTOR or pPOL-ENV-U3 and was used in functional assays identical to those described in Figure 4B.

(D) Peripheral blood lymphocytes (PBL) from healthy, MHC-unrelated donors (donors I, II, and III indicated below the X axis) were cultured with retroviral (pPOL-ENV-U3) and control transfectants (pVECTOR) at T:non-T ratios as indicated below the graphs (T:APC).

IDDMK$_{1,2}$22 Is a Novel Member of the MMTV-Related Family of HERV-K and Is Related to HERV-K10

To evaluate the relationship between IDDMK$_{1,2}$22 and other known retroviruses, we derived phylogenetic trees for subregions exhibiting different degrees of conservation (Saitou and Nei, 1987; Thompson et al., 1994; Galtier et al., 1996). The three regions chosen for this analysis were the RT region of the *pol* gene (Figure 3B), the outer region (SU, surface) of the *env* gene (Figure 3A), and the U3 region of the LTR (Figure 3C). The RT and SU regions were selected to construct interspecies phylogenetic trees because they represent, respectively, the most highly conserved and the most variable of the protein coding regions (McClure et al., 1988). The U3 region of the LTR was chosen to construct an intraspecies tree of the family to which IDDMK$_{1,2}$22 belongs because LTR sequences are conserved in size and sequence only within a given species, and the U3 region accounts for most of the intraspecies differences (Temin, 1981). As shown in Figure 3A, the ENV polyprotein of IDDMK$_{1,2}$22 is most closely related to that of HERV-K10. Both proteins are related to those of MMTV and Jaagsiekte sheep retrovirus (JSRV). The same is essentially true for the RT subregion of the POL polyprotein, where IDDMK$_{1,2}$22 and HERV-K10 are most closely related to the B-type retrovirus MMTV (Figure 3B). Figure 3C illustrates that

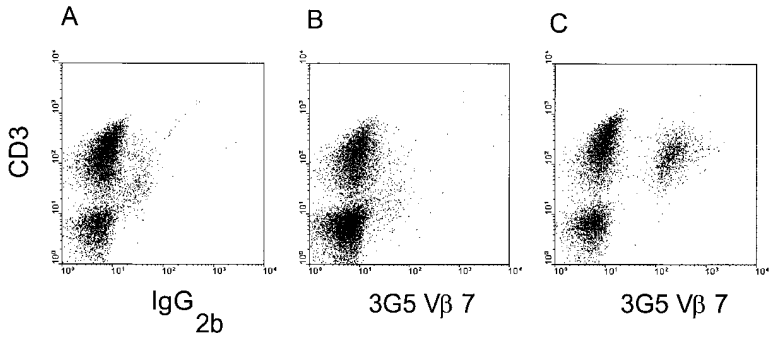

Figure 5. IDDMK$_{1,2}$22 Mediates a Vβ7-Specific SAG Effect

T cells (10^6 cells/ml) were cultured for 3 days with mitomycin-treated pPOL-ENV-U3 and pVECTOR transfectants at T:non-T ratios as indicated. Twenty U/ml of recombinant IL-2 were then added to the cultures, and FACS analysis was performed after 3–4 days of expansion (Conrad et al., 1994).

(A) THP1 cells were transfected with pPOL-ENV-U3, and the stimulated and expanded T cells were stained with anti-CD3 monoclonal antibodies and an isotype control after 7 days of coculture.

(B) T cells stimulated by THP1 transfected with the vector (pVECTOR) alone were stained with anti-CD3 monoclonal antibodies and the anti-Vβ7-specific antibody 3G5.

(C) THP1 cells were transfected with pPOL-ENV-U3, and the stimulated T cells were stained with anti-CD3 monoclonal antibodies and the anti-Vβ7 antibody 3G5.

K$_{1,2}$1 is related to HERV-K(C4), while K$_{1,2}$4 and IDDMK$_{1,2}$22 are related to the K10/K18 subfamily. Within this family, K$_{1,2}$4 is closely related to K10, whereas IDDMK$_{1,2}$22 appears to be more distant.

IDDMK$_{1,2}$22 Encodes a Vβ7-Specific SAG

The strategy used to identify a putative SAG function encoded by IDDMK$_{1,2}$22 was dictated by predictions based on the biology of the MMTV-SAG, general requirements for a protein–protein interaction between a SAG and MHC class II molecules, and intracellular trafficking mechanisms used by proteins encoded by retroviruses. The prototypical retroviral SAG of MMTV is a type II transmembrane protein that is encoded within the U3 of the 3′ LTR (reviewed by Acha-Orbea and MacDonald, 1995). It is targeted into the MHC class II peptide loading compartment and exported to the cell surface. On the basis of potential splice donor (SD) and acceptor sites (SA) present in its sequence, IDDMK$_{1,2}$22 is expected to generate two subgenomic mRNAs, one encoding ENV and a second transcript comprising the U3-R region (Figure 4A). Based on these criteria, we produced an episomal expression construct (pPOL-ENV-U3) with a 5′ SD positioned upstream of the truncated *pol*, *env*, and U3 regions (Figure 4A). It is expected that both of the putative subgenomic mRNAs can be generated from this construct (Figure 4A).

Retroviral and control transfectants of monocyte and B lymphocyte cell lines were generated and tested for their ability to stimulate MHC-compatible and allogeneic T cell lines in a Vβ7-specific manner. Monocytes do not express measurable MHC class II surface proteins in the absence of induction by interferon-γ (IFN-γ); the MHC class II transactivator CIITA mediates IFN-γ-inducible MHC class II expression (reviewed by Mach et al., 1996). As shown in Figure 4A, transient monocyte (THP1, U937) transfectants induced with IFN-γ and expressing the truncated IDDMK$_{1,2}$22 genome (pPOL-ENV-U3) stimulated in a dose-dependent fashion T cell lines from MHC-compatible donors essentially to the same extent. The mitogenic effect was dependent on the presence of MHC class II, since IFN-γ-mediated MHC class II expression specifically induced the stimulatory capacity of retroviral as compared to control transfectants (Figure 4B). The use of THP1 cells rendered constitutively MHC class II positive by transfection with CIITA resulted in a stimulation comparable to IFN-γ-induction, suggesting that the IFN-γ-induced and CIITA-dependent MHC class II expression was indeed responsible for this functional difference (Figure 4C). The mitogenic effect is not MHC restricted, since a response exceeding allostimulation was observed when PBL from several different MHC-disparate donors were tested for proliferative responses to monocytes transfected with pPOL-ENV-U3 (Figure 4D). In essence, these functional data suggest that the truncated IDDMK$_{1,2}$22 (pPOL-ENV-U3) genome is responsible for a mitogenic effect that is MHC class II dependent but not MHC restricted.

Experiments were performed in bulk cultures using TCR-Vβ-specific stimulation and expansion as a readout. Retroviral THP1 transfectants induce a more than 15-fold increase in the number of the Vβ7 family but not of the two control families tested (Vβ8, Vβ12) after specific stimulation and subsequent amplification (Figure 5, Table 1). This was verified by using two different Vβ7-specific monoclonal antibodies, 3G5 and 20E. A comparable effect was also observed when PBL from MHC-disparate donors were tested. This was interpreted as evidence for the presence a Vβ7-specific SAG.

The monocytic cell lines were at least 3 times more efficient in terms of specific TCR Vβ7 amplification as compared to the most efficient B-lymphoblastoid cell line (Table 1). This difference could not be explained by

Table 1. IDDMK$_{1,2}$22 Mediates a Vβ7-Specific SAG Effect

	Vβ Family		
Transfectant	Vβ7	Vβ8	Vβ12
Raji-pPOL-ENV-U3	7%	5%	2.5%
Raji-pVECTOR	1.5%	5.5%	2%
THP1-pPOL-ENV-U3	16%	5.3%	2.8%
THP1-pVECTOR	1%	5.8%	3%

The B-lymphoblastoid cell line Raji was stably transfected with either pPOL-ENV-U3 or pVECTOR and used in functional assays (equivalent to Figure 5) 2 weeks after selection. The monocytic cell line THP1 was cultured for 48 hours after transfection with the same constructs. The percentages of double positive (CD3 and Vβ7, Vβ8, and Vβ12) T cells are indicated that were obtained after 1 week of coculture with the respective transfectants (pPOL-ENV-U3 or pVECTOR).

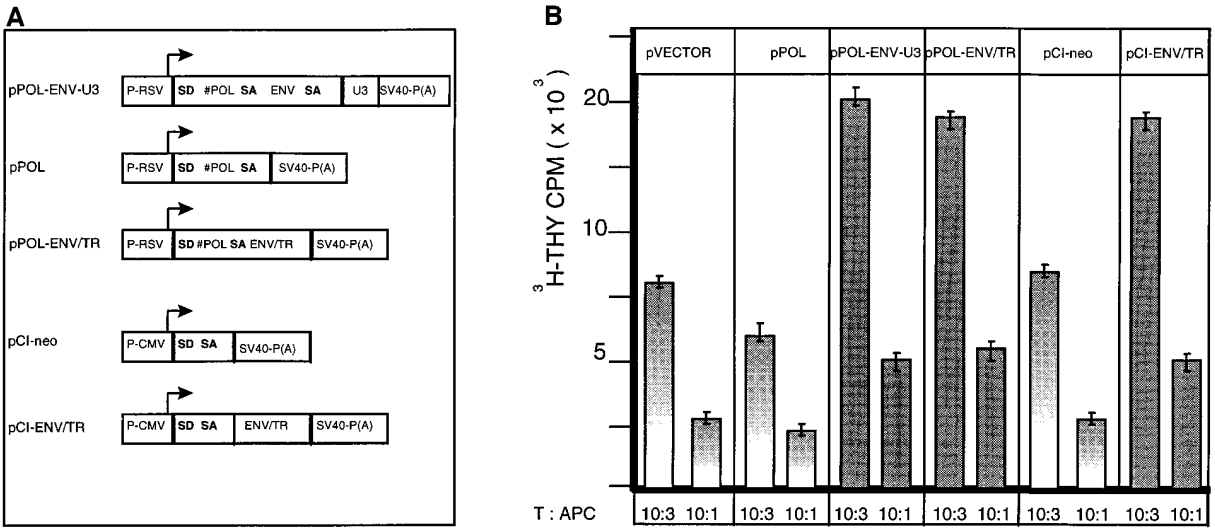

Figure 6. The N-Terminal *env* Moiety of IDDMK$_{1,2}$22 Mediates the SAG Effect

(A) Based on the construct pPOL-ENV-U3, different deletional mutants were generated that comprised (1) pPOL: the *pol* gene; (2) pPOL-ENV/TR: the *pol*- and the N-terminal moiety of the *env* gene; and (3) pCI-ENV/TR: the N-terminal moiety of the *env* gene alone.

(B) PBL from MHC-unrelated donors were cocultured with mitomycin C–treated THP1 cells as described in Figure 4. The individual transfectants are indicated with the names of the constructs above the bars: (1) pVECTOR, (2) pPOL, (3) pPOL-ENV-U3, (4) pPOL-ENV/TR, (5) pCI-neo, and (6) pCI-ENV/TR). One of at least three independent ^{3}H-thymidine incorporation experiments with allogeneic T cells stimulated by the individual transfectants is shown. The ratio between T cells and transfectants is indicated below the bars (T:APC).

variations in the level of MHC class II expression or by the individual MHC haplotypes present. On the other hand, it may be due to differential expression of costimulatory molecules or secretion of cytokines. In conclusion, by all criteria known to date, IDDMK$_{1,2}$22 encodes a mitogenic activity having all features of a Vβ7-specific SAG.

The SAG Function Is Mediated by the N-Terminal Moiety of the IDDMK$_{1,2}$22 ENV Protein

A series of deletional mutants were generated that contained either the truncated *pol-env*-U3 region (pPOL-ENV-U3), the truncated *pol* gene alone (pPOL), or the truncated *pol* gene followed by the *env* gene truncated downstream of the premature stop codon found in all clones (pPOL-ENV/TR) (Figure 6A). In addition, a C-terminally truncated *env* gene was generated as an individual expression unit (pCI-ENV/TR). As shown in Figure 6B, by excluding the *env* coding region, the SAG function is selectively lost (pPOL). If, however, the truncated *env* gene is included (pPOL-ENV/TR), the stimulatory capacity is restored to levels comparable to pPOL-ENV-U3. In addition, expression of the truncated *env* gene alone (pCI-ENV/TR) is sufficient for function. These findings demonstrate that the SAG function is mediated by the N-terminal moiety of the ENV protein comprising 153 amino acids. The nucleotide and predicted amino acid sequences of the minimal stimulatory region are shown in Figure 7. As shown in Figure 3A, this predicted protein resembles the N-terminal ENV proteins of related HERVs (HERV-K10), and those of the B-type retroviruses (MMTV, JSRV). However, there is no significant sequence homology with either MMTV-SAG, other SAGs, or autoantigens known to be important in IDDM.

Discussion

Here, we provide evidence showing that a human endogenous retrovirus, IDDMK$_{1,2}$22, is released from leukocytes in patients with acute-onset type I diabetes. In preliminary experiments, IDDMK$_{1,2}$22 RNA sequences were detectable in the plasma of IDDM patients at disease onset but not in the plasma of age-matched healthy controls. This novel human retrovirus is related to MMTV and encodes a SAG with functional characteristics similar to the one encoded by MMTV. In contrast to MMTV, however, the IDDM-associated SAG is encoded within the retroviral *env* gene rather than within the 3′ LTR. It has the same TCR Vβ7 specificity as the SAG originally identified in the IDDM patients. This SAG is thus likely to be the cause of the Vβ7-enriched repertoire of islet-infiltrating T lymphocytes.

IDDMK$_{1,2}$22 as a Member of the HERV-K Class of Endogenous Retroviruses

HERV-K genomes exist in two different forms, type I genomes that are largely splice deficient and type II genomes that generate three subgenomic mRNAs (Ono, 1986; Tönjes et al., 1996). A 292 bp insert at the *pol-env* boundary with clustered nucleotide changes downstream of the splice acceptor site are present in type II but not in type I genomes (Tönjes et al., 1996). The insert affects both the *env* gene and the *pol* gene: type II genomes have a stop codon between *env* and *pol* that is missing in type I genomes, and they have a considerably longer N-terminal *env* region. The 292 bp insert and the clustered nucleotide changes have been proposed to be responsible for the efficient splicing of type II genomes (Tönjes et al., 1996). IDDMK$_{1,2}$22 is missing the 292 bp

```
ACA TTT GAA GTT CTA CAA TGA ACC CAT CAG AGA TGC AAA GAA AAG CGC CTC CAC GGA 57

GAT GGT AAC ACC AGT CAC ATG GAT GGA TAA TCC TAT AGA AGT ATA TGT TAA TGA TAG 114
  M   V   T   P   V   T   W   M   D   N   P   I   E   V   Y   V   N   D   S 19

TGT ATG GGT ACC TGG CCC CAC AGA TGA TCG CTG CCC TGC CAA ACC TGA GGA AGA AGG 171
  V   W   V   P   G   P   T   D   D   R   C   P   A   K   P   E   E   E   G 38

GAT GAT GAT AAA TAT TTC CAT TGG GTA TCA TTA TCC TCC TAT TTG CCT AGG GAG AGC 228
  M   M   I   N   I   S   I   G   Y   H   Y   P   P   I   C   L   G   R   A 57

ACC AGG ATG TTT AAT GCC TGC AGT CCA AAA TTG GTT GGT AGA AGT ACC TAC TGT CAG 285
  P   G   C   L   M   P   A   V   Q   N   W   L   V   E   V   P   T   V   S 76

TCC TAA CAG TAG ATT CAC TTA TCA CAT GGT AAG CGG GAT GTC ACT CAG GCC ACG GGT 342
  P   N   S   R   F   T   Y   H   M   V   S   G   M   S   L   R   P   R   V 95

AAA TTA TTT ACA AGA CTT TTC TTA TCA AAG ATC ATT AAA ATT TAG ACC TAA AGG GAA 399
  N   Y   L   Q   D   F   S   Y   Q   R   S   L   K   F   R   P   K   G   K 114

AAC TTG CCC CAA GGA AAT TCC TAA AGG ATC AAA GAA TAC AGA AGT TTT AGT TTG GGA 456
  T   C   P   K   E   I   P   K   G   S   K   N   T   E   V   L   V   W   E 133

AGA ATG TGT GGC CAA TAG TGT GGT GAT ATT ACA AAA CAA TGA ATT CGG AAC TAT TAT 513
  E   C   V   A   N   S   V   V   I   L   Q   N   N   E   F   G   T   I   I 152

AGA TTA G 520
  D   *   153
```

Figure 7. The Nucleotide and Deduced Amino Acid Sequence of IDDMK$_{1,2}$22-SAG

The minimal stimulatory sequence corresponding to the insert of pCI-ENV/TR comprises a C-terminally truncated protein of 153 amino acids. There is only one ORF with a stop codon at position 518. The first potential start codon in a favorable context is at position 59. Two potential N-linked glycosylation sites are present at positions 106 and 182, respectively. The degree of homology with other retroviral ENV proteins is shown in Figure 3A. No significant homology was detected with the SAG of MMTV or with autoantigens known to be important in IDDM.

insert but has two in-frame stop codons between *env* and *pol* and the clustered nucleotide changes downstream of the SA typical of those found in type II genomes. In terms of splice efficiency, IDDMK$_{1,2}$22 may be in an intermediate position between type I and II genomes. This and the altered N-terminal sequences in IDDMK$_{1,2}$22 with respect to type II genomes may affect SAG expression in vivo. However, as shown in Figure 4, the 3′-terminal moiety (POL-ENV-U3) of the IDDMK$_{1,2}$22 genome mediates the SAG function in vitro. Moreover, it is known from MMTV that the SAG function in vivo may be present at levels where the respective protein remains undetectable (Winslow et al., 1992; reviewed by Acha-Orbea and MacDonald, 1995).

The Model: Human Self SAGs as Activators of Autoreactive T Cells in Type I Diabetes

We propose a model according to which induction of self SAGs in systemic and professional APCs, outside the pancreas, leads to autoimmunity in genetically susceptible individuals. The model implies two steps: the first is systemic, the second organ-specific. The initial event is a systemic, polyclonal activation of a Vβ-restricted T cell subset, triggered by the expression of an endogenous retroviral SAG in professional MHC class II$^+$ APCs. In a second step, autoreactive T cells within the subset of SAG-activated T lymphocytes initiate organ-specific tissue destruction. The evidence presented here, however, does not rule out that the release of the IDDMK$_{1,2}$22 RNA sequences in vivo and the SAG function associated with IDDM in these patients are the consequences rather than the causes of the inflammation.

The expression of self SAGs can in principle be modulated by two variables: physiological endogenous stimuli or environmental stimuli. A possible physiological stimulus might be steroid hormones. HERV-K10 expression is steroid inducible in vitro, and this is possibly the result of hormone response elements (HRE) present in its LTR (Ono et al., 1987). IDDMK$_{1,2}$22 and HERV-K10 share the same putative HRE in their respective LTRs (Ono et al., 1987), (Figure 3). One is thus tempted to speculate that steroid inducibility of IDDMK$_{1,2}$22 could also occur in vivo, in analogy to the well documented example of the transcriptional control by steroid hormones of the MMTV promoter (reviewed by Acha-Orbea and MacDonald, 1995). Infectious agents are of major importance when considering environmental factors. Examples include the cellular SAGs that are expressed by herpesvirus-infected monocytes and B lymphocytes (Dobrescu et al., 1995; Sutkowski et al., 1996). In both cases, HERVs have not been excluded as a potential source of the SAG activity. It is thus conceivable that SAGs are being selectively expressed in response to ubiquitous pathogens such as herpesviridae (reviewed by Roizman, 1996). In fact, HERVs are induced by a variety of environmental stresses, and some of them behave as hepatic acute-phase genes (reviewed by Wilkinson et al., 1994).

The experimental evidence presented suggests that the RT activity, the IDDMK$_{1,2}$22 RNA sequences, and, in consequence, the SAG may derive from leukocytes rather than from the pancreatic β cells. This may indicate that expression of the retroviral SAG is induced preferentially in systemically circulating professional MHC class II$^+$ APCs. The highest rate of IDDM coincides with puberty (10–14 years) in both sexes (Bruno et al., 1993). Infections with ubiquitous viruses (reviewed by Roizman, 1996) may act synergistically with an increase in the circulating levels of steroids to enhance expression of the SAG in professional APCs. Autoreactive T cells can be readily demonstrated in the mature repertoire of healthy individuals (Pette et al., 1990). However, to be able to migrate to the target tissue, these T cells have to be activated (reviewed by Steinman, 1995). These considerations lead us to the hypothesis that among the Vβ7$^+$ T cells activated by IDDMK$_{1,2}$22-SAG, some are autoreactive and migrate to the target tissue where β cell–specific death ensues. Once β cells die, cellular antigens are liberated and the immune response is perpetuated through determinant spreading (reviewed by Tisch and McDevitt, 1996).

The Concept of IDDMK$_{1,2}$22-*sag* as an Autoimmune Gene

Known genes conferring susceptibility to autoimmune diseases are host-derived, stably inherited Mendelian traits and contribute in a cumulative fashion to the familial clustering of the disease without causing disease per se (reviewed by Vyse and Todd, 1996). IDDMK$_{1,2}$22

should be viewed as a mobile genetic element with the potential to move within the host genome due to multiple mechanisms, including retrotransposition, homologous recombination, gene conversion, and capture, resulting in multiple copies of individual HERVs (reviewed by Preston and Dougherty, 1996; Wain-Hobson, 1996). This renders family studies dealing with searches for HERV disease association difficult. It should be noted, however, that there is little or no plus/minus genetic polymorphism in different humans at the HERV-K loci and as yet no evidence for mobility. Interestingly, an IDDMK$_{1,2}$22-related HLA-DQ-LTR is associated with susceptibility to IDDM, possibly due to cosegregation with the HLA (Figure 3C) (Badenhoop et al., 1996). In addition, infectious transmission cannot be excluded, as is the case for two closely related virus groups containing endogenous and exogenous variants: MMTV and JSRV (Figures 4A and 4B) (reviewed by Acha-Orbea and MacDonald, 1995; York et al., 1992).

In summary, this candidate autoimmune gene has distinctly different features from classical, disease-associated susceptibility genes. It has the potential of being transmitted as either an inherited trait or as an infectious agent. Moreover, this gene has no apparent essential function for the host, but it may have instead an inducible and intriguing potential to directly cause disease whenever expressed in genetically susceptible individuals.

Experimental Procedures

Patients

The islets and spleens from patients with acute-onset and chronic IDDM and nondiabetic organ donors were provided by the Pittsburgh Transplant Institute (Conrad et al., 1994). The plasma and genomic DNA from patients and controls for the epidemiological study were isolated by the Diabetes Register in Torino, Italy (Bruno et al., 1993). The samples were collected within 1 month after the clinical diagnosis from patients, aged 0–29 years (Bruno et al., 1993).

RT Assays

RT assays were performed as described (Pyra et al., 1994).

Isolation of Full-Length Retroviral Genomes

A description of the criteria used to identify unknown retroviral 5′ R-U5s and 3′ R-poly(A)s has been published (Weissmahr et al., 1997).

(I) Primer sequences for the 3′ moiety of the putative retroviral genomes (abbreviations are according to the recommendations of the Nomenclature Committee of the International Union of Biochemistry [1985]):

(A) RT region: RT 1a, 5′YAAATggMgWAYgYTAACAgACT3′; RT 1b, 5′YAAATggMgWAYgYTAACTgACT3′; RT 2a-nested, 5′CgTCTA gAgCCYTCTCCggCYATgATCCCg3′; RT 2b-nested, 5′CgTCTAgAg CCYTCTCCggCYATgATCCCA3′.

(B) 3′ U3-R-Poly(A)s (all primers have an identical 5′ anchor):

(i) 5′TgCgCCAgCAATgTATCCATg3′+ sequence-specific part: #1K1,2-1, 5′gggTggCAgTgCATCATAggT3′; #4K1,2-4, 5′gggAgAgg gTCAgCAgCAgACA3′; #K1,2-10, 5′gACAgCAAgCCAgTgATAAgC A3′; #K1,2-16, 5′ggAACAgggACTCTCTgCA3′; #K1,2-17, 5′gggAAg ggTAAggAAgTgTg3′; #K1,2-22, 5′ggTgTTTCTCCTgAgggAg3′; #K1,2-26, 5′gAAgAATggCCAACAgAAgCT3′; #K1,2-27, 5′gggAAACAAggA gTgTgAgT3′.

(ii) Common, secondary anchor primer:

3′ U3-R-poly(A)s common, 5′CATgTATATgCggCCgCTgCgCCAg CAATgTATCCATgg3′.

(II) Primer sequences for the 5′ moiety of the genome:

(A) RT region: RT 1, 5′TATCTTTCgTTTCTgCAgCAC3′; RT 2, 5′TAACTggTTgAAgAgCTCgACC3′.

(B) 5′-R-U5: R-U5-1, 5′ATACTAAggggACTCAgAggC3′; R-U5-2, 5′CAgAggCTggTgggATCCTCCATATgC3′.

The PCR conditions were as follows: 1 × 94°C for 2 min; 45°C for 5 min; 68°C for 30 min; 10 × 94°C for 15 sec; 45°C for 30 sec + 1°C/cycle; 68°C for 3 min, 30 sec; 25 × 94°C for 15 sec; 55°C for 30 sec; 68°C for 3 min 30 sec + 20 sec/cycle. Primers were used at 300 nM final concentration, dNTPs at 200 mM, with 52 U/ml of Taq-Pwo polymerase mix (Boehringer Mannheim). One vol% of first-round PCR was subjected to a nested PCR. Size-selected and purified amplification products were blunted, and EcoRI was adapted and subcloned into EcoRI-digested λZAPII arms. After two rounds of hybridization, 20 individual clones were rescued as plasmids. Eleven clones were selected for further analysis based on a conserved restriction pattern. An equivalent procedure was followed for the 5′ moiety of the genome. Sequencing was performed on an automatic sequencer (ABI, Perkin Elmer) using subgenomic clones.

Epidemiological Study: RNA-PCR

Three ml of blood was collected in EDTA tubes (Vacutainer) and further processed within 6 hours. Samples were subjected twice to centrifugation, for 4 × 10^3 G, 10 min at 4°C. Total RNA was extracted from 560 μl of plasma (QIAamp; Qiagen). Four vol% of total RNA was used for a single tube RT–PCR using thermostable AMV, Taq, and Pwo (Boehringer Mannheim). Reactions contained at a final concentration: di-Na salts of dNTPs at 0.2 mM; DTT at 5 mM; 10 U recombinant RNAsin (Promega); 1.5 mM $MgCl_2$; R-poly(A) primer 5′ TTT TTg AgT CCC CTT AgT ATT TAT T 3′; U3 primer 5′ Agg TAT TgT CCA Agg TTT CTC C 3′, both at 0.3 μM. RT was performed at 50°C for 30 min directly followed by 94°C for 2 min; 94°C for 30 sec, 68°C for 30 sec, −1.3°C each cycle, 68°C for 45 sec for a total of 10 cycles; 94°C for 30 sec, 55°C for 30 sec, 68°C for 45 sec for a total of 25 cycles. The amplified material (487 bp) was subjected to agarose gel electrophoresis followed by alkaline transfer and hybridization with probes generated from the IDDMK$_{1,2}$22 U3-R-region.

Genomic PCR

Genomic DNA (100 ng) was subjected to PCR. Reactions contained at a final concentration: dNTPs at 200 μM; 1.5 mM $MgCl_2$; 2.6 U of Taq-Pwo (Boehringer Mannheim); U3 primer 5′ Agg TAT TgT CCA Agg TTT CTC C 3′; R primers either 5′ CTT TAC AAA gCA gTA TTg CTg C 3, or 5′ gTA AAg gAT CAA gTg CTg TgC 3′ at 300 nM. The amplified products were 300 and 395 bp in size, respectively. The cycling profile was as follows: 94°C for 2 min; 94°C for 15 sec, 68°C for 30 sec, −1.3°C each cycle, 72°C for 45 sec for a total of 10 cycles; 94°C for 15 sec, 55°C for 30 sec, 72°C for 45 sec for a total of 25 cycles.

Sequence Alignment and Phylogenetic Trees

Sequences were aligned with CLUSTAL W (Thompson et al., 1994). Alignments were checked and manually corrected with the SEA VIEW multiple sequence alignment editor (Galtier et al., 1996). Phylogenetic trees were computed from multiple alignments using the "neighbor joining" method (Saitou and Nei, 1987).

Expression

Constructs

pPOL-ENV-U3: a SacI–NotI fragment derived from 11 IDDMK$_{1,2}$22 clones was ligated with a BamHI–SacI adapter containing a consensus SD and with a NotI–XbaI adapter, was subcloned into BamHI–XbaI-digested pλDR2 arms, and selected for by two rounds of screening, and plasmids were rescued. At least five independent clones were used for transfections. pPOL: pPOL-ENV-U3 was digested with KpnI–NotI, blunted, and religated. pPOL-ENV/TR: a stimulatory clone was digested with XbaI and religated. pCI-ENV/TR: 1 ng of pPOL-ENV-U3 was amplified with the primers 5′ gAC TAA gCT TAA gAA CCC ATC AgA gAT gC 3′ and 5′ AgA CTg gAT CCg TTA AgT CgC TAT CgA CAg C 3′. The amplified products were subcloned into pCI-neo (Promega).

Cells and Cell Lines

Monocytic cell lines: THP1, U937. B-lymphoblastoid cell lines: Raji, BOLETH, SCHU and WT 51. T cells of molecularly MHC-typed blood donors were generated by positive selection with anti-CD3-coated immunomagnetic beads (Milan-Analytika).

Transfections

Transient transfectants were used for functional assays 48 hours after transfection; stable transfectants were selected for 2 weeks in progressive concentration of Hygromycin B to a final concentration of 250 μg/ml for lymphoblastoid lines, and 50 μg/ml for monocytic cell lines.

Functional Assays

Transfectants were treated with Mitomycin C (Calbiochem) at 100 μg/ml per 10^7 cells for 1 hour at 37°C and washed extensively. Proliferation assays. Then 10^5 CD3 bead–selected, MHC-compatible T cells or Ficoll-Paque-isolated allogeneic PBL were cultured with transfectants at stimulator:responder ratios of 1:1, 1:3, and 1:10 for 48 and 72 hours in 96 round-bottom wells at 37°C. ^{3}H-thymidine was then added at 1 μmCi/well and incorporation measured after 18 hours incubation at 37°C. FACS analysis and antibodies used were as described; after 3 days of specific stimulation, at T:non-T ratios of 1:1 for syngeneic and 10:3 for allogeneic stimulations, the T cells were further expanded in 20 U/ml recombinant IL-2 for 6 days before flow cytometric analysis (Conrad et al., 1994).

Acknowledgments

Correspondence should be addressed to B. C. This work was supported by a Swiss National Science Foundation grant and the Department of Surgery, University of Geneva Medical School. We thank Dr. Behboo for patient and control samples; Drs. Callan and Romagné for supplying anti-Vβ7 antibodies 3G5 and 20E; Dr. Duret for sequence alignments and phylogenetic trees; S. Bontron for RNA analysis; M. Zufferey for THP1-CIITA transfectants; Drs. Acha-Orbea, Clarkson, Garcin, Reith, Steimle, and Strubin for helpful suggestions; and L. Otten and Dr. Reith for critical reading of the manuscript.

Received December 23, 1996; revised June 13, 1997.

References

Acha-Orbea, H., Shakow, A.N., Scarpellino, L., Kolb, E., Müller, V., Vessaz-Shaw, A., Fuchs, R., Blöchlinger, K., Rollini, P., Billotte, J., et al. (1991). Clonal deletion of Vβ14-bearing T cells in mice transgenic for mouse mammary tumor virus. Nature *350*, 207–211.

Acha-Orbea, H., and MacDonald, H.R. (1995). Superantigens of mouse mammary tumor virus. Annu. Rev. Immunol. *13*, 459–486.

Badenhoop, K., Tönjes, R.R., Rau, H., Donner, H., Rieker, W., Braun, J., Herwig, J., Mytilineos, J., Kurth, R., and Usadel, K.H. (1996). Endogenous retroviral long terminal repeats of the HLA-DQ region are associated with susceptibility to insulin-dependent diabetes mellitus. Hum. Immunol. *50*, 103–110.

Brocke, S., Gaur, A., Piercy, C., Gautam, A., Gijbels, K., Fathman, C.G., and Steinman, L. (1993). Induction of relapsing paralysis in experimental autoimmune encephalomyelitis by bacterial superantigen. Nature *365*, 642–644.

Bruno, G., Merletti, F., Vuolo, A., Pisu, E., Giorio, M., and Pagano, G.F. (1993). Sex differences in incidence of IDDM in age-group 15–29 yr. Diabetes Care *16*, 133–136.

Choi, Y., Kotzin, B., Herron, L., Callahan, J., Marrack, P., and Kappler, J. (1989). Interaction of staphylococcus aureus toxin "superantigens" with human T cells. Science *86*, 8941–8945.

Choi, Y., Kappler, J.W., and Marrack, P. (1991). A superantigen encoded in the open reading frame of the 3′ long terminal repeat of mouse mammary tumor virus. Nature *350*, 203–207.

Cole, B.C., and Griffiths, M.M. (1993). Triggering and exacerbation of autoimmune arthritis by the mycoplasma arthritidis superantigen MAM. Arthritis Rheum. *36*, 994–1002.

Conrad, B., Weidmann, E., Trucco, G., Rudert, W.A., Behboo, R., Ricordi, C., Rodriquez-Rilo, H., Finegold, D., and Trucco, M. (1994). Evidence for superantigen involvement in insulin-dependent diabetes mellitus aetiology. Nature *371*, 351–355.

Dobrescu, D., Ursea, B., Pope, M., Asch, A.S., and Posnett, D.N. (1995). Enhanced HIV-1 replication in Vβ12 T cells due to human cytomegalovirus in monocytes: evidence for a putative herpesvirus superantigen. Cell *82*, 753–763.

Fleischer, B., and Schrezenmeier, H. (1988). T cell stimulation by staphylococcal enterotoxins. J. Exp. Med. *167*, 1697–1707.

Galtier, N., Gouy, M., and Gautier, C. (1996). SeaView and Phylo-win, two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. *12*, 543–548.

Heidmann, O., and Heidmann, T. (1991). Retrotransposition of a mouse IAP sequence tagged with an indicator gene. Cell *64*, 159–170.

Howell, M.D., Diveley, J.P., Lundeen, K.A., Esty, A., Winters, S.T., Carlo, D.J., and Brostoff, S.W. (1991). Limited T-cell β-chain heterogeneity among interleukin 2 receptor-positive synovial T cells suggests a role for superantigen in rheumatoid arthritis. Proc. Natl. Acad. Sci. USA *88*, 10921–10925.

Karvonen, M., Tuomilehto, J., Libman, I., and LaPorte, R. (1993). A review of the recent epidemiological data on the worldwide incidence of type 1 (insulin-dependent) diabetes mellitus. Diabetologia *36*, 883–892.

Lo, D., Reilly, C.R., Scott, B., Liblau, R., McDevitt, H.O., and Burkly, L.C. (1993). Antigen-presenting cells in adoptively transferred and spontaneous autoimmune diabetes. Eur. J. Immunol. *23*, 1693–1698.

Mackay, C.R. (1993). Homing of naive, memory and effector lymphocytes. Curr. Opin. Immunol. *5*, 423–427.

Mach, B., Steimle, V., Martinez-Soria, E., and Reith, W. (1996). Regulation of MHC class II genes: lessons from disease. Annu. Rev. Immunol. *14*, 301–331.

McClure, M.A., Johnson, M.S., and Doolittle, R.F. (1988). Sequence comparisons of retroviral proteins: relative rates of change and general phylogeny. Proc. Natl. Acad. Sci. USA *85*, 2469–2473.

Medstrand, P., and Blomberg, J. (1993). Characterization of novel reverse transcriptase encoding human endogenous retroviral sequences similar to type A and type B retroviruses: differential transcription in normal human tissues. J. Virol. *67*, 6778–6787.

Nomenclature Committee of the International Union of Biochemistry (NC-IUB). (1985). Nomenclature for incompletely specified bases in nucleic acid sequences. Recommendations 1984. Eur. J. Biochem. *150*, 1–5.

Oldstone, M.B.A. (1990). Molecular mimicry and autoimmune disease. Cell *50*, 819–820.

Ono, M. (1986). Molecular cloning and long terminal repeat sequences of human endogenous retrovirus genes related to types A and B retrovirus genes. J. Virol. *58*, 937–944.

Ono, M., Yasunaga, T., Miyata, T., and Ushikubo, H. (1986). Nucleotide sequence of human endogenous retrovirus genome related to mouse mammary tumor virus genome. J. Virol. *60*, 589–598.

Ono, M., Kawakami, M., and Ushikubo, H. (1987). Stimulation of expression of the human endogenous retrovirus genome by female steroid hormones in human breast cancer cell line T47D. J. Virol. *61*, 2059–2062.

Paliard, X., West, S.G., Lafferty, J.A., Clements, J.R., Kappler, J.W., Marrack, P., and Kotzin, B.L. (1991). Evidence for the effects of a superantigen in rheumatoid arthritis. Science *253*, 325–329.

Pette, M., Fujiita, K., Wilkinson, D., Altmann, D.M., Trowsdale, J., Giegrich, G., Hinkkanen, A., Epplen, J.T., Kappos, L., and Weckerle, H. (1990). Myelin autoreactivity in multiple sclerosis: recognition of myelin basic protein in the context of HLA-DR2 products by T lymphocytes of multiple sclerosis patients and healthy donors. Proc. Natl. Acad. Sci. USA *87*, 7968–7972.

Preston, B.D., and Dougherty, J.P. (1996). Mechanisms of retroviral mutation. Trends Microbiol. *4*, 16–21.

Pyra, H., Böni, J., and Schüpbach, J. (1994). Ultrasensitive retrovirus detection by a reverse transcriptase assay based on product enhancement. Proc. Natl. Acad. Sci. USA *91*, 1544–1548.

Roizman, B. (1996). Herpesviridae. In Fields Virology, B.N. Fields et al., eds. (Philadelphia: Lippincott-Raven Publishers), pp. 2221–2230.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. *4*, 406–425.

Stegall, M.D., Lafferty, K.J., Kam, I., and Gill, R.G. (1996). Evidence of recurrent autoimmunity in human allogeneic islet transplantation. Transplantation *61*, 1272–1274.

Steinman, L. (1995). Escape from "horror autotoxicus": pathogenesis and treatment of autoimmune disease. Cell *80*, 7–10.

Sutkowski, N., Palkama, T., Ciurli, C., Sekaly, R.P., Thorley-Lawson, D.A., and Huber, B.T. (1996). An Epstein-Barr virus-associated superantigen. J. Exp. Med. *184*, 971–980.

Tassabehji, M., Strachan, T., Anderson, M., Campbell, R.D., Collier, S., and Lako, M. (1994). Identification of a novel family of human endogenous retroviruses and characterization of one family member, HERV-K(C4), located in the complement C4 gene cluster. Nucleic Acids. Res. *22*, 5211–5217.

Temin, H.M. (1981). Structure, variation and synthesis of retrovirus long terminal repeat. Cell *27*, 1–3.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matric choice. Nucleic Acids. Res. *22*, 4673–4680.

Tisch, R., and McDevitt, H. (1996). Insulin-dependent diabetes mellitus. Cell. *85*, 291–297.

Tydén, G., Finn, P.R., Sundkvist, G., and Bolinder, J. (1996). Recurrence of autoimmune diabetes mellitus in recipients of cadaveric pancreatic grafts. N. Engl. J. Med. *335*, 860–863.

Tönjes, R.R., Löwer, R., Boller, K., Denner, K., Hasenmaier, B., Kirsch, H., König, H., Korbmacher, C., Limbach, C., Lugert, R., et al. (1996). HERV-K: the biologically most active human endogenous retrovirus family. J. AIDS Hum. Retrovirol. *13*, 261–267.

Vyse, T.J., and Todd, J.A. (1996). Genetic analysis of autoimmune diseases. Cell *85*, 311–318.

Wahle, E., and Keller, W. (1996). The biochemistry of polyadenylation. Trends Biochem. Sci. *21*, 247–250.

Wain-Hobson, S. (1996). Running the gamut of retroviral variation. Trends Microbiol. *4*, 135–141.

Weissmahr, R.N., Böni, J., and Schüpbach, J. (1997). Reverse transcriptase activity in chicken embryo fibroblast culture supernatants is associated with particles containing endogenous avian retrovirus EAV-O RNA. J. Virol. *71*, 3005–3012.

Whitcomb, J.M., and Hughes, S.H. (1992). Retroviral reverse transcription and integration: progress and problems. Annu. Rev. Cell. Biol. *8*, 275–306.

White, J., Herman, A., Pullen, A.M., Kubo, R., Kappler, J.W., and Marrack, P. (1989). The Vβ-specific superantigen staphylococcal enterotoxin B: stimulation of mature T cells and clonal deletion in neonatal mice. Cell *56*, 27–35.

Wilkinson, D.A., et al. (1994). Endogenous human retroviruses. In The Retroviridae, J.A. Levy, ed. (New York: Plenum Press), pp. 465–535.

Winslow, G.M., Scherer, M.T., Kappler, J.W., and Marrack, P. (1992). Detection and biochemical characterization of the mouse mammary tumor virus 7 superantigen (Mls-1^a). Cell *71*, 719–730.

Wucherpfennig, K.W., and Strominger, J.L. (1995a). Selective binding of self peptides to disease-associated major histocompatibility complex (MHC) molecules: a mechanism for MHC-linked susceptibility to human autoimmune diseases. J. Exp. Med. *181*, 1597–1601.

Wucherpfennig, K.W., and Strominger, J.L. (1995b). Molecular mimicry in T cell–mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. Cell *80*, 695–705.

Xiong, Y., and Eickbusch, T. (1990). Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. *9*, 3353–3362.

York, D.F., Vigne, R., Verwoerd, D.W., and Guerat, G. (1992). Nucleotide sequence of the Jaagsiekte retrovirus, an exogenous and endogenous type D and B retrovirus of sheep and goats. J. Virol. *66*, 4930–4939.

GenBank Accession Numbers

The accession numbers for the sequences reported in this paper are AFO12327–AFO12337.