Molecular Immunology 47 (2010) 2119–2128

Contents lists available at ScienceDirect

Molecular Immunology

journal homepage: www.elsevier.com/locate/molimm

Novel atypical nucleotide insertions specifically at V_H–D_H junction generate exceptionally long CDR3H in cattle antibodies

Madhuri Koti [a], Galina Kataeva [b], Azad Kumar Kaushik [a,*]

[a] Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada N1G 2W1
[b] Mobix Lab, McMaster University, Hamilton, Ontario, Canada

ARTICLE INFO

Article history:
Received 21 December 2009
Accepted 21 February 2010

Keywords:
CDR3H
Antibody diversity
Bovine D_H gene segment

ABSTRACT

Some IgM cattle antibodies are amongst the largest known to exist in jawed vertebrates where CDR3H size may extend up to 61 amino acids. To understand the origin of such an exceptionally long CDR3H, bovine D_H gene locus was completely characterized from Holstein cattle that revealed the presence of a total of eight D_H genes, including D_HQ52, with a distinct organization in sub-clusters. However, a total of 10 D_H genes are identified at the polymorphic D-gene locus in cattle that are classified into four families, designated as BovD_HA, BovD_HB, BovD_HC and BovD_HD. In fetal B-cells, VDJ recombinations encoding long CDR3H (>50 codons) are directly encoded by the single germline V_H gl.110.20, the longest D_H2 and the J_H1 genes, apart from few N- and P-nucleotide additions at the junctions. Further, non J-proximal D_H7 gene is preferentially expressed in fetal B cells. The adult VDJ recombinations, however, are distinctly remarkable for 'conserved short nucleotide sequence' ('CSNS'; 13–18 nucleotides), of non-V_H or D_H gene origin, inserted specifically at V_H–D_H junctions resulting in extension of CDR3H size up to 61 codons. Together with P-nucleotides, N-additions (1–9 nucleotides) are noted at both the V_H–D_H and D_H–J_H junctions. Such 'CSNS' insertions at V_H–D_H junction of adult VDJ recombinations encoding exceptionally long CDR3H provide novel mechanism of antibody diversification in cattle, not yet observed in other species. Further, analysis of V_H–D_H–J_H recombinations originating from fetal B-cells reveals the presence of substitution, deletion or addition mutations without prior exposure to external antigen. Thus, somatic hypermutations may contribute towards diversification of the developing nascent antibody repertoire in cattle. In conclusion, the outlined experiments provide novel antibody diversification mechanism via 'CSNS' insertions, specifically at the V_H–D_H junction, in generating exceptionally long CDR3H extending up to 61 codons in cattle antibodies.

1. Introduction

The antigen recognition by the variable region of antibodies is mediated by six hypervariable polypeptide loops, three each from the heavy (CDR1H, CDR2H and CDR3H) and light (CDR1L, CDR2L and CDR3L) chain variable domains. The CDR3H constitutes floor of the antigen combining site, makes maximum contact with an epitope (Wilson and Stanfield, 1994) and is, thus, critical to determining antibody specificity. Three germline genetic elements, V_H, D_H, and J_H (Tonegawa, 1983) encode the CDR3H region resulting in an extensive diversity apart from junctional variations introduced during recombination. The D_H gene segments may enhance diversity of the CDR3H via D–D fusions (VanDyk and Meek, 1992) as well. The nucleotide insertions may either be non-templated (N) introduced by TdT enzyme or template dependent originating from palindromic (P) bases (Alt and Baltimore, 1982). In humans, GC-rich DIR elements up to 126 bp located between the DM and DN genes (Ichihara et al., 1988; Sanz, 1991) contribute to diversity of the CDR3H region. As such, CDR3H loop varies considerably both in sequence composition and size as a result of mechanistic factors involved in recombination and germline D_H gene segment utilization. Upon encounter with an antigen in the periphery, further antibody diversification occurs via somatic mutations involving point mutations (Wagner and Neuberger, 1996) and insertions or deletions (Wilson et al., 1998).

While the process of V_H–D_H–J_H recombination (Tonegawa, 1983) encoding antibodies is common to all jawed vertebrates, significant differences exist in germline V_H sequence divergence across species. A relatively large pool of divergent germline V, D and J genes are capable of generating significant combinatorial diversity in humans (Kabat and Wu, 1971; Matsuda et al., 1998). However, availability of a relatively small number of functional germline genes with limited sequence divergence in species, such as, chicken (Reynaud et al., 1987), rabbit (Becker and Knight, 1990),

Abbreviations: CDR3H, complementarity determining region 3 of heavy chain; RSS, recombination signal sequence; DH gene, diversity gene; RF, reading frame.

* Corresponding author. Tel.: +1 519 824 4120x54389; fax: +1 519 837 1802.
E-mail address: akaushik@uoguelph.ca (A.K. Kaushik).

0161-5890/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.molimm.2010.02.014

pig (Sun et al., 1994), sheep (Dufour et al., 1996; Reynaud et al., 1989) and cattle (Berens et al., 1997; Saini et al., 1997; Sinclair et al., 1997) restricts generation of comparable combinatorial diversity. Obviously, these species use alternate strategies for antibody diversification, for example, gene conversion in bursal B cells of chicken (Reynaud et al., 1989) and appendix B cells of rabbit (Becker and Knight, 1990; Crane et al., 1996; Weinstein et al., 1994). The somatic hypermutations are involved in diversifying the developing B cell repertoire in Ileal Peyer's Patches of sheep without prior exposure to exogenous antigens (Reynaud et al., 1991, 1995).

Our laboratory demonstrated a unique characteristic of cattle IgM antibodies where CDR3H loop extended up to 61 amino acids (Kaushik et al., 2002; Saini et al., 1999). This is apart from the fact that bovine antibodies, similar to camel (10–24 aa) (Muyldermans et al., 1994), humans (2–26 aa) (Kabat and Wu, 1971) and chicken (15–30 aa) (Reynaud et al., 1989) tend to have a relatively long CDR3H, in general. Approximately, 8–10% of peripheral B cells express IgM with exceptionally long (>50 aa) CDR3H (Kaushik et al., 2002; Saini et al., 1999) with multiple cysteine residues (Kaushik et al., 2009; Saini et al., 1999). The CDR3H loop in cattle antibodies shows extensive size heterogeneity ranging from 3 to 61 codons that becomes evident during fetal B cell stage (Saini and Kaushik, 2002). The cattle antibodies with exceptionally long CDR3H are functional as some of these recognize multiple conserved natural antigens and, therefore, exist as natural autoantibodies (Saini et al., 1999; Saini and Kaushik, 2001). Since CDR3H loop composition and size influences the three-dimensional configuration of antigen combining site that dictates antigen specificity, it is essential to determine the origin of such an exceptionally long CDR3H in cattle antibodies. Partial characterization of a 2.3 kb bovine genomic DNA fragment revealed the presence of two short (14–19 codons) and one long (49 codons) germline D_H gene segment (Shojaei et al., 2003). However, these observations, did not explain the origin of an exceptionally long CDR3H that extended up to 61 codons. The outlined experiments demonstrate that specific insertion of novel 'conserved short nucleotide sequences' (CSNS) at V_H–D_H junction, apart from direct contribution of V_H, D_H and J_H elements, provides the molecular basis of the origin of exceptional CDR3H loop extending up to 61 amino acids.

2. Materials and methods

2.1. Bovine D_H gene segment probe

The 2.3 kb D_H gene specific DNA probe (Shojaei et al., 2003) was prepared by digesting the plasmid pBGD14s with EcoRI and radiolabeled (Saini et al., 1997; Shojaei et al., 2003) using [α-^{32}P]dCTP by random priming (Roche, Canada).

2.2. Bovine genomic library screening

Bovine primary genomic DNA library (titer 1.36×10^{10} plaque forming units/ml), constructed in Lambda FIX II vector with an insert size of 9–23 kb, from the liver of a 2-year-old Holstein cow (Stratagene, LaJolla, USA), was screened with bovine D_H-gene specific DNA probe (Shojaei et al., 2003). Briefly, a total 1×10^6 phage clones (99% probability of DNA sequence represented in the library of bovine genome) were screened by transferring and UV crosslinking (Stratalinker) plaques onto nitrocellulose membrane (Boehringer Mannheim, Germany). The membranes were pre-hybridized (50% formamide, 5× Denhardt's solution, 5× SSPE, 0.5% SDS, 100 μg/ml sonicated and denatured salmon sperm DNA) at 42 °C followed by hybridization with the [α-^{32}P]dCTP labeled 2.3-kb D_H gene specific probe (Shojaei et al., 2003). The membranes were washed and exposed to XAR-5 film (Eastman Kodak Company, Rochester, NY, USA). The identified positive phage clones were stored at −70 °C.

2.3. Southern hybridization

Fifteen microgram of genomic DNA isolated from liver (Stratagene, LaJolla, USA) from two pure bred (Holstein, Jersey) and two cross-bred (Red Angus × Simmental and Red Angus × Simmental × Charolais) cows was digested with BamHI, fractionated in 0.8% agarose gel and depurinated in 0.2N HCl prior to transfer onto nitrocellulose (Schliecher and Schuell Inc., USA). The transferred DNA was UV crosslinked (Stratalinker) and membranes were pre-hybridized (50% formamide, 5× Denhardt's solution, 5× SSPE, 0.5% SDS, 100 μg/ml salmon sperm DNA) at 42 °C followed by hybridization (50% formamide, 5× Denhardt's solution, 5× SSPE, 0.5% SDS, 100 μg/ml salmon sperm DNA and $(2.5–3.0) \times 10^6$ dpm of the [α-^{32}P]dCTP labeled 2.3-kb D_H gene specific probes) (Saini et al., 1997; Shojaei et al., 2003). The membranes were washed and exposed to XAR-5 films (Eastman Kodak Company, Rochester, NY, USA) at −70 °C.

2.4. DNA sequencing and analysis

Three clones, BDG1-2, BDG2-4 and BDG5-3(2-11-C) that showed distinct restriction enzyme digestion pattern were selected for nucleotide sequencing. The phage DNA was isolated (Promega, USA) and subjected to automated DNA sequencing using a series of forward and reverse internal primers (MOBIX LAB, McMaster University, Hamilton, Canada). The DNA sequences were analyzed using DNA Star (South Park St., Madison, USA) and Geneious (Biomatters Ltd.) programs. The phylogenetic relationships were determined using Megalign program (Lasergene Software, South Park St., Madison, USA). For somatic mutation analysis, germline and fetal rearranged VDJ sequences (Berens et al., 1997; Saini and Kaushik, 2002) were compared from Holstein cattle in the conserved region to avoid allelic differences as described (Kaushik et al., 2009). D_H gene segments were identified in fetal (Berens et al., 1997; Kaushik et al., 2009; Saini and Kaushik, 2002) and adult (Berens et al., 1997; Saini et al., 1997, 1999) VDJ recombinations to determine their expression.

2.5. Statistical analysis

The data were analyzed using Fisher's exact Chi-squared test (SIGMA STAT 3.2 software package). Two-way ANOVA was performed using GraphPad InStat version, 3.05 (Graph pad Software, San Diego, CA, USA).

3. Results

3.1. Ten bovine D_H gene segments are classified into four families

Complete screening of bovine genomic library from a Holstein cattle ($p = 0.99$) by hybridizing with bovine D_H genespecific DNA probe led to isolation of three phage clones, BDG2-4, BDG1-2 and BDG5-3(2-11-C). Upon sequencing, BDG1-2 and BDG5-3(2-11-C) were noted to be overlapping clones resulting in assembly of 16.123 kb nucleotide sequence (Fig. 1) with a subcluster of three D_H gene segments, D_H1, D_H2 and D_H3 earlier identified (Shojaei et al., 2003). Analysis of BDG2-4 clone spanning 18,374 kb (Fig. 1) led to identification of four new germline D_H gene segments designated as D_H4, D_H5, D_H6 and D_H7, each flanked by a recombination signal sequence (RSS) with a 13 and 12 bp spacer on 5′ and 3′ ends, respectively (Fig. 1). The D_H4, D_H5, D_H6 and D_H7 genes are relatively short comprising, 36, 67, 42 and 58 bp, respectively (Figs. 1 and 2). Thus, a total of seven D_H gene segments are identified upon screening of