Protein Science (1994), *3*:1341–1343. Cambridge University Press. Printed in the USA.

FOR THE RECORD

Extending the B7 (CD80) gene family

PETER S. LINSLEY, ROBERT PEACH, PAUL GLADSTONE, AND JÜRGEN BAJORATH
Bristol-Myers Squibb Pharmaceutical Research Institute, Seattle, Washington 98121

(RECEIVED May 18, 1994; ACCEPTED June 1, 1994)

Abstract: B7-1 and B7-2 are members of the immunoglobulin superfamily (IgSF) and important regulators of T cell-mediated immune responses. Despite sharing only limited sequence identity, B7-1 and B7-2 bind common receptors, CD28 and CTLA-4, on T cells and have similar functional properties. We have found that the extracellular V(ariable)-like domains of B7-1 and B7-2 share significant sequence similarities with 3 major histocompatibility complex (MHC)-encoded members of the IgSF: butyrophilin, myelin/oligodendrocyte glycoprotein, and the chicken MHC molecule, B-G. This raises the question whether there is an evolutionary link between the MHC, which encodes molecules regulating the antigen specificity of T lymphocyte responses, and B7 molecules, which co-stimulate these responses in antigen-nonspecific fashion.

Keywords: B7; immunoglobulin superfamily; sequence similarity; structural similarity; T cell-mediated immunity

Binding of CD28 and/or CTLA-4 receptors on T cells to B7 (CD80)-related molecules on antigen-presenting cells triggers a T cell co-stimulatory pathway required for optimal immune responses (Linsley & Ledbetter, 1993). Two B7-related molecules have been cloned from man and mouse: B7-1 (Freeman et al., 1989, 1991) and B7-2 (Freeman et al., 1993a, 1993b) or B7-0 (Azuma et al., 1993). These molecules are members of the immunoglobulin superfamily (IgSF), comprising a V(ariable)-like and a C(onstant)-like domain in their extracellular regions. B7-1 and B7-2 share functional properties of binding to CTLA-4 and of co-stimulating T cell responses (see Linsley & Ledbetter, 1993, for review), despite having only ~25% amino acid sequence identity in their extracellular domains.

As part of ongoing efforts to elucidate the molecular basis of interactions between B7-related molecules and their receptors on T cells, we have analyzed the sequences of B7-related molecules relative to other members of the IgSF. We show that much of the sequence identity in the IgSF V-like domains of B7-1 and B7-2 beyond established IgSF consensus residues is shared with 3 other members of the IgSF, which are related: butyrophilin (BT), a component of the bovine milk fat globule membrane (Jack & Mather, 1990); myelin/oligodendrocyte glycoprotein (MOG), a component of the myelin sheath (Gardinier et al., 1992; Pham-Dinh et al., 1993); and the chicken major histocompatibility complex (MHC) molecule, B-G (Kaufman & Salomonsen, 1992; Steinman, 1993).

IgSF V-like domains were analyzed using the GCG package (Genetics Computer Corporation, Inc., Madison, Wisconsin, 1993). Two sequence comparison tables (Dayhoff et al., 1979; Risler et al., 1988) were used and gave similar results. Alignments were performed using PILEUP; position-specific sequence profiles were calculated using PROFILEMAKE and used to search the Swiss Protein Database with PROFILESEARCH. A profile calculated from the V-like domains of B7-1 and B7-2 showed higher sequence similarity with BT than with all other sequences in the Swiss Protein Database ($Z = 6.9$; Z scores >5 are considered statistically significant). Another profile was calculated from B7-1, B7-2, and the V-like domain of BT, and again used for database searching. This profile revealed higher similarity with MOG ($Z = 12.4$) than with all other sequences in the Swiss Protein Database; incorporation of MOG into the profile then gave highest similarity with B-G ($Z = 10.9$). A consensus sequence was calculated from these IgSF V-like domains using PRETTY for sequence identity in at least 4 of 7 sequences. Residues were assigned to framework β-strands of an IgSF variable fold by reference to IgSF consensus residues (Williams & Barclay, 1988). Alignments were adjusted manually to maximize conservation of IgSF determinants and to avoid gaps in predicted β-strands. These procedures revealed 40 consensus sequence residues from these 7 V-like sequences (Fig. 1). The significance of this consensus sequence was evaluated in 2 ways. The same procedures were used to calculate a consensus sequence from V-like domains of 18 other members of the IgSF, including murine Ig κ chains and VH domains; rat OX2; and human myelin P0 protein, T cell receptor α-chain, CD2, CD4, CD7, CD8, CD22, CD28, CD33, CD48, CD58, CTLA-4, and Thy-1. These V-like sequences shared only 8 consensus residues with the sequences shown in Figure 1 (magenta-shaded residues); these were the same as previously described consensus residues of IgSF V-like sequences (Williams & Barclay, 1988). As a second test, we conducted a PROFILESEARCH using the consensus sequence from Figure 1 modified such that the IgSF consensus residues from the B and F strands were made ambiguous. This

Reprint requests to: Jürgen Bajorath, Bristol-Myers Squibb Pharmaceutical Research Institute, 3005 First Avenue, Seattle, Washington 98121; e-mail: bajorath@protos.bms.com.

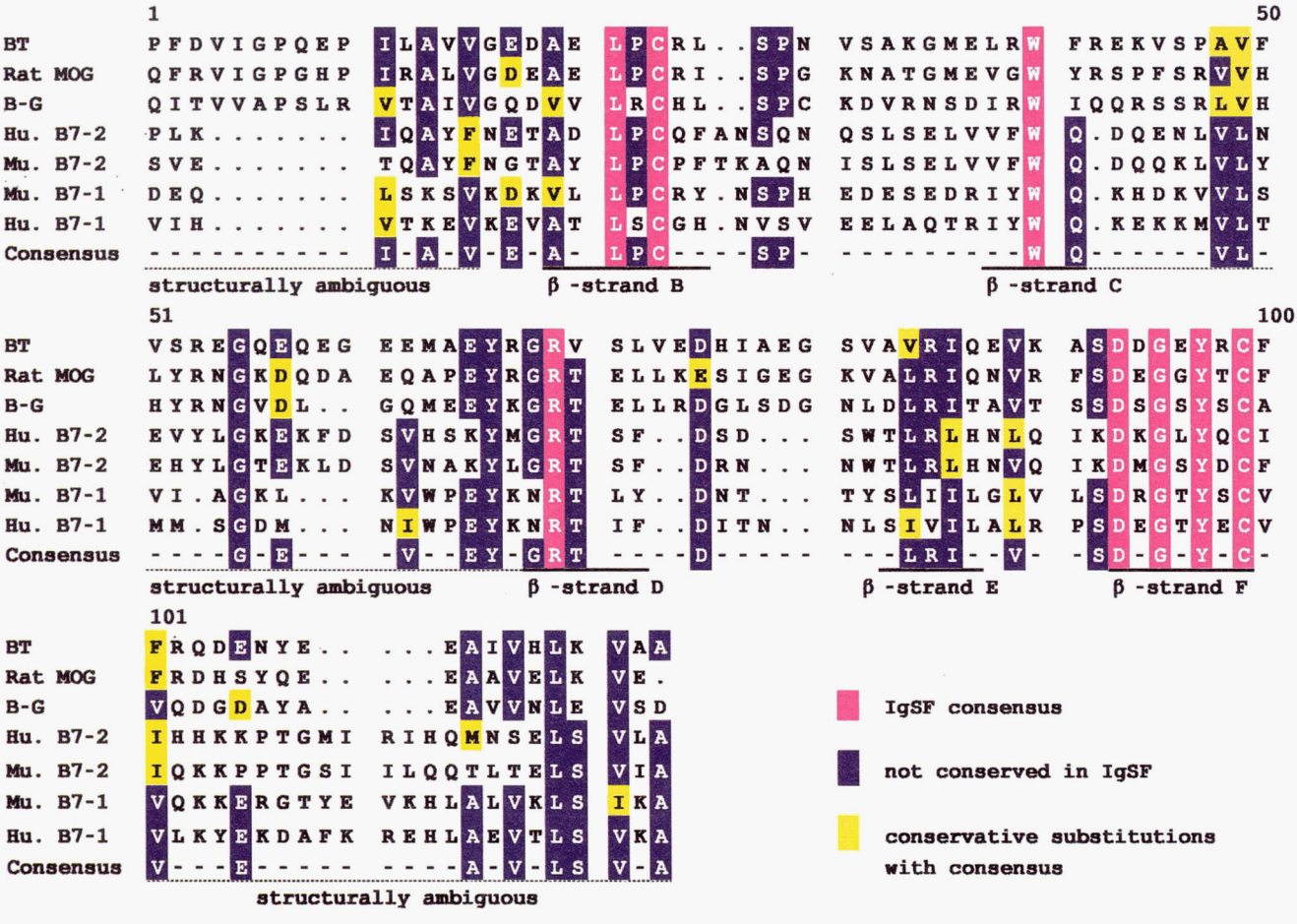

Fig. 1. Sequence comparison of the extended B7 family. Alignment and consensus sequence of V-like domains of BT, MOG, B-G, B7-1, and B7-2. Consensus residues in other IgV-like domains are colored magenta; consensus residues not shared in most IgV-like domains are colored blue; conservative substitutions with the consensus sequence are colored yellow. Tentative assignment of residues to β-strands of an IgSF V domain are indicated, as are positions that could not be assigned to an IgSF V domain (structurally ambiguous). Accession numbers of sequences are: BT, P188892; rat MOG, M99485; B-G, M61863; hu. (human) B7-1, M27533; mu. (murine) B7-1, X60958; hu. B7-2, L25259; mu. B7-2, L25606.

was done so as to remove their contribution to the consensus profile. This search gave higher similarity with the 7 sequences aligned in Figure 1 than with all other sequences in the Swiss Protein Database ($Z > 6$). These findings show that the consensus sequence shown in Figure 1 extends beyond IgSF characteristics.

The analysis reveals a significant sequence relationship of BT, MOG, BG, and the B7 molecules despite the absence of high overall sequence similarity; this is also a characteristic of the B7-1 and B7-2 homologues. The sequences of BT, MOG, and BG are more similar to one another than to B7-1 and B7-2 and vice versa, but the similarity of the extended B7 family, including BT, MOG, and BG, significantly exceeds the sequence similarities of these molecules to known IgSF members. This is illustrated by the fact that the extended B7 family exhibits 40 consensus residues in their V-like extracellular domain, only 8 of which are known IgSF consensus residues. Sequence conservation within the extended B7 family was highest in predicted IgSF β-strands and in regions that were not readily assigned to a V-domain fold due to the absence of IgSF consensus residues. Regions lacking IgSF consensus may or may not correspond structurally to the A, C′, C″, and G IgSF β-strands; we refer to these here as structurally ambiguous. Sequence–structure compatibility calculations (Lüthy et al., 1992) further support the existence of these structurally ambiguous regions in B7. Diversity between these sequences was highest in (loop) regions connecting predicted β-strands. Sequence similarity between B7-1, B7-2, and BT extended through their C-like domains (MOG and B-G do not contain a C-like domain), but the transmembrane and cytoplasmic regions of all of these molecules were distinct.

Based on these sequence similarities, we suggest that the V-like sequences of BT, MOG, B-G, B7-1, and B7-2 characterize a subfamily within the IgSF. The fact that B7-1 and B7-2 have similar functional properties but low sequence identity suggests that these molecules have diverged from a common ancestral Ig V-like domain. If BT, MOG, and B-G also diverged from the same ancestral Ig V-like domain, then divergence must have begun long ago, because expression of each of these molecules is restricted to very different tissues and possibly to different species (B-G expression has not been detected in mammals). Genes encoding these molecules are dispersed, because BT, MOG, and B-G are MHC-encoded (Kaufman & Salomonsen, 1992; Pham-Dinh et al., 1993), whereas human B7-1 maps outside the MHC

(Freeman et al., 1992). Like MHC class I and II genes, B-G genes are polymorphic. Others have suggested that, like MHC class I and class II molecules, B-G molecules may function in antigen-specific recognition (Kaufman & Salomonsen, 1992; Steinman, 1993). The sequence similarity between B-G antigens and B7 molecules strongly suggests an evolutionary link between antigen-specific and antigen-nonspecific regulators of lymphocyte function.

Given the case for evolutionary divergence, these molecules should have once shared, and perhaps still share, functional properties. There is little known of natural functions of BT, MOG, or B-G, although immunologic functions for the latter 2 molecules have been suggested (Kaufman & Salomonsen, 1992; Steinman, 1993). In addition to sequence similarity, B7 and B-G also share the ability to enhance immune responses; B7 costimulates T cell immunity (Linsley & Ledbetter, 1993) and B-G enhances immune responses to other antigens (the so-called "adjuvant effect") (Kaufman & Salomonsen, 1993). It may therefore be worthwhile to investigate whether B-G and MOG are capable of interacting with specific receptors on the T cell surface.

References

Azuma M, Ito D, Yagita H, Okumura K, Phillips JH, Lanier L, Somoza C. 1993. B70 antigen is a second ligand for CTLA-4 and CD28. *Nature 366*:76–79.

Dayhoff MO, Schwarz RM, Orcutt BC. 1978. In: Dayhoff MO, ed. *Atlas of protein sequence and structure*. Washington, D.C.: National Biomedical Research Foundation. pp 345–352.

Freeman GJ, Boriello F, Hodes RJ, Reiser H, Grippen JG, Ng JW, Kim J, Goldberg JM, Hathcock K, Laszlo G, Lombard LA, Wang S, Gray GS, Nadler LM, Sharpe AH. 1993a. Murine B7-2, an alternative s-4 counter-receptor that costimulates T cell proliferation and interleukin 2 production. *J Exp Med 178*:2185–2192.

Freeman GJ, Disteche CM, Gribben JG, Adler DA, Freedman AS, Dougery J, Nadler LM. 1992. The gene for B7, a costimulatory signal for T-cell activation, maps to chromosomal region 3q13.3–3q21. *Blood 79*:489–494.

Freeman GJ, Freedman AS, Segil JM, Lee G, Whitman JF, Nadler LM. 1989. B7, a new member of the Ig superfamily with unique expression on activated and neoplastic B cells. *J Immunol 143*:2714–2722.

Freeman GJ, Gray GS, Gimmi CD, Lombard DB, Zhou LJ, White M, Fingeroth JD, Gribben JG, Nadler LM. 1991. Structure, expression, and T cell costimulatory activity of the murine homologue of the human B lymphocyte activation antigen B7. *J Exp Med 174*:625–631.

Freeman G, Gribben JG, Boussiotis VA, Ng JW, Restivo VA Jr, Lombard LA, Gray GS, Nadler LM. 1993b. Cloning of B7-2: A CTLA-4 counter-receptor that costimulates human T cell proliferation. *Science 262*:909–911.

Gardinier MV, Amiguet P, Linington C, Matthieu JM. 1992. Myelin/oligodendrocyte glycoprotein is a unique member of the immunoglobulin superfamily. *J Neurosci Res 33*:177–187.

Jack LJ, Mather IH. 1990. Cloning and analysis of cDNA encoding bovine butyrophilin, an apical glycoprotein expressed in mammary tissue and secreted in association with the milk-fat globule membrane during lactation. *J Biol Chem 265*:14481–14486.

Kaufman J, Salomonsen J. 1992. B-G: We know what it is but what does it do? *Immunol Today 13*:1–3.

Linsley PS, Ledbetter JA. 1993. The role of the CD28 receptor during T cell responses to antigen. *Annu Rev Immunol 11*:191–212.

Lüthy R, Bowie JU, Eisenberg D. 1992. Assessment of protein models with three-dimensional profiles. *Nature 356*:83–85.

Pham-Dinh D, Mattei MG, Nussbaum JL, Roussel G, Pontarotti P, Roeckel N, Mather IH, Artzt K, Lindahl KF, Dautigny A. 1993. Myelin/oligodendrocyte glycoprotein is a member of a subset of the immunoglobulin superfamily encoded within the major histocompatibility complex. *Proc Natl Acad Sci USA 90*:7990–7994.

Risler JL, Delorme MO, Delacroix H, Henaut A. 1988. Amino acid substitutions in structurally related proteins. A pattern recognition approach. Determination of a new and efficient scoring matrix. *J Mol Biol 204*:1019–1029.

Steinman L. 1993. Connections between the immune system and the nervous system. *Proc Natl Acad Sci USA 90*:7912–7914.

Williams AF, Barclay AN. 1988. The immunoglobulin superfamily—Domains for cell surface recognition. *Annu Rev Immunol 6*:381–406.