FOR THE RECORD

Immunoglobulin fold characteristics of B7-1 (CD80) and B7-2 (CD86)

JÜRGEN BAJORATH, ROBERT J. PEACH, AND PETER S. LINSLEY

Bristol-Myers Squibb Pharmaceutical Research Institute, 3005 First Avenue, Seattle, Washington 98121

(RECEIVED August 15, 1994; ACCEPTED September 19, 1994)

Abstract: B7-1 and B7-2 are expressed on antigen-presenting cells and bind to the CD28 and CTLA-4 receptors on T cells. These interactions trigger a costimulatory pathway that is essential for T-cell activation. B7-1 and B7-2 are members of the immunoglobulin superfamily (IgSF) and, despite sharing common function, have only limited sequence similarity. The B7-1 extracellular region was previously subdivided into 2 IgSF domains, an N-terminal V(ariable)-like domain, followed by a C(onsant)-like domain. We recently reported that the V-like domains of B7-1 and B7-2 share some significant sequence similarities with 3 major histocompatibility complex (MHC)-encoded members of the IgSF. We have now applied inverse folding methodology to assess the compatibility of the B7-1 and B7-2 extracellular region sequences with currently available 3-dimensional structures. In these calculations, the sequences of the N-terminal (V-like) domains in B7-1 and B7-2 were not compatible with known structures, including the IgSF V-set. In contrast, the sequences of the C-like domains were compatible with IgSF C-set structures and were best recognized by the β 2-microglobulin (β 2m) domain of MHC Class I. A sequence comparison of the C-like domains in the B7 molecules showed that 11 of 17 rigorously conserved residues in B7-1 and B7-2 are not IgSF C-1 set consensus residues. When mapped onto the corresponding positions of the β 2m structure, the conserved residues in B7 cluster on the surface, where they may interact with the B7 V-like domain or other molecules.

Keywords: B7; immunoglobulin superfamily; inverse folding; sequence similarity; structural similarity; T-cell costimulation

T-cell costimulation via the interaction of CD28 and/or CTLA-4 on T cells with B7-related molecules on antigen-presenting cells is now established as an essential feature of T-cell-dependent immune responses (Linsley & Ledbetter, 1993). Two B7-related molecules with common functional properties have been identified, B7-1 (CD80) (Freeman et al., 1989) and B7-2 (CD86)

(Freeman et al., 1993) or B7-0 (Azuma et al., 1993). On the basis of the sequence of their extracellular region, the B7 molecules have been included in the immunoglobulin superfamily (IgSF) (Williams & Barklay, 1988). Despite their common receptor-binding properties, the extracellular regions of human B7-1 and B7-2 share only ~25% sequence identity and are thought to consist of 1 V(ariable)-like (~16% sequence identity) and 1 C(onsant)-like (~22% sequence identity) domain each. Previously we have reported (Linsley et al., 1994) that B7-1 and B7-2 share significant sequence similarities with 3 major histocompatibility complex (MHC)-encoded members of the IgSF including butyrophilin (Jack & Mather, 1990). The 3 MHC-encoded molecules and the B7 molecules share 40 consensus residues in their V-like domain, only 8 of which are known IgSF consensus residues. Much of the sequence identity beyond IgSF characteristics (Williams & Barklay, 1988) is found in regions that lack IgSF consensus residues. These regions approximately correspond to the A, G, C', and C'' β -strands of an IgSF V-fold.

We have now further assessed the IgSF characteristics of the B7 molecules by inverse folding calculations (for review, see Wodak & Rooman, 1993) using the Brookhaven Protein Data Bank (PDB) (Bernstein et al., 1977) structures. The compatibility of the B7-1 and B7-2 V- and C-like sequences with structures available in the PDB was assessed using the topological fingerprint method (Godzik et al., 1992, 1993) as implemented in the Matchmaker program (Tripos Assoc., Inc., St. Louis, Missouri). As shown in Table 1, these calculations did not support a strong relationship between the B7 V-like sequences and IgSF structure. In contrast, considerable sequence-structure compatibility was detected for the sequence of the C-like domains in B7-1 and B7-2 and IgSF constant domains, and the highest compatibility scores were calculated for the B7 sequences and the β 2-microglobulin (β 2m) structure (Table 1), which belongs to the IgSF C1-set. In addition, 3-dimensional outline models were generated for the first and the second domain of B7-1 based on the conjecture that the first and second extracellular domains of B7-1 adopt IgSF V- and C-folds, respectively. We have focused on B7-1 because the sequence-structure searches overall gave better results for B7-1 than for B7-2. The generated outline models (details are given in the caption of Fig. 1) were then assessed using the energy profiling method of Sippl (1993) as im-

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

Table 1. Sequence-structure compatibility calculations using the topology fingerprint method^a

Sequence	Structure matched	SDBM	Consistency
B71-C	β 2m	5.7	+
B72-C	β 2m	4.0	(+)
BT-C	β 2m	4.4	+
B71-V	—	—	—
B72-V	—	—	—
BT-V	IgV	3.8	+

^a Structure matched gives the highest scoring structure for each sequence (only in the case of overall consistent results). SDBM gives the statistical free energy in standard deviations below the mean of the Boltzmann distribution for all sequence-structure matches. Consistency summarizes whether the alignments obtained in different calculations (see below) were stable and equivalent. In the case of B72-C, (+) means that one of the calculations did not fit the above criteria. The butyrophilin V-like sequence (BT-V) matched several Ig variable folds (IgV) with comparable scores. Results shown were calculated using gap and gap extension penalties of 2.0 and 0.1, respectively. Calculations: The putative V- and C-like domains in human B7-1 and B7-2 were defined as (Azuma et al., 1993): B7-1 V-like, residues 19–137; B7-2 V-like, 1–125; B7-1 C-like, 138–234; B7-2 C-like 126–229. For each sequence several calculations with different gap (between 1.0 and 2.0) and gap extension (between 0.05 and 0.1) penalties were carried out to evaluate the consistency of the results, and frozen approximations and updated residue environments (thawed alignments) (Godzik et al., 1992) were calculated to assess the stability of the alignments. A topological fingerprint of telokin (Holden et al., 1992), the first IgSF I-set structure (Harpaz & Chothia, 1994), was added to the fingerprint database. As controls, calculations were carried out using the sequences of the CD8 Ig domain (Leahy et al., 1992) and of the β 2-microglobulin (β 2m) domain of MHC Class I (Garrett et al., 1989). Both test sequences were clearly recognized by their respective structures.

plemented in the ProsaII program (Version 2, Center of Applied Molecular Engineering, Techno-Z Salzburg, Salzburg, Austria) (Fig. 1). The “best” energy profile of the B7-1 V-domain is compared with the energy profile of the B7-1 C-domain model. The positive average energy of the V-like outline model indicates that this model is (at least partially) incorrectly folded. In contrast, the overall negative profile trace of the B7-1 C-like model indicates a correctly folded structure. The results of the energy profile analysis were therefore consistent with the results of the sequence-structure calculations.

The inverse folding analysis was unable to identify structures compatible with the sequences of the putative V-like domains in B7-1 and B7-2. In this analysis, “threading” the V-like sequences of B7-1 and B7-2 onto a typical IgSF V-fold results in compromised structural models. It is therefore possible that the putative V-like domain in B7-1 and B7-2 includes structural features which depart from known standard immunoglobulin folds. In contrast, the inverse folding analysis establishes that the sequences of the B7 C-like domains are compatible with IgSF C-domain structures and suggests that β 2m (which shares ~12% sequence identity with human B7-1) is a close structural equivalent of these domains. It has therefore been possible to map the position of residues that are conserved in the B7 C-domains to corresponding positions in the β 2m structure. The C-domains in the murine and human B7 molecules include 17 rigorously conserved residues, 11 of which are distinct from IgSF C1-set

Fig. 1. Energy profile analysis, generated using ProsaII, of outline models of the V-like domain (dashed line) and of the C-like domain (solid line) in B7-1 on the basis of an IgSF variable and constant (C1) fold, respectively. The energy profiles were calculated using β -carbon interaction energies and a 50-residue window for profile averaging. Outline models for the putative V-domain in B7-1 were based on an IgSF V-set structure (KOL; Marquardt et al., 1980), and the C-domain model was based on β 2-microglobulin of MHC Class I (Garrett et al., 1989). Due to the absence of IgSF consensus residues in parts of the B7-1 V-like sequence, several outline models with locally modified assignment of residues to β -strands or loop regions were built. B7-1 residues were “threaded” onto the Ig structures in similar positions via computer graphics. Insertions and deletions in loop regions were accommodated using CONGEN (Brucoleri et al., 1988).

consensus residues (Fig. 2A). Ten of these 11 residues (except B5 in β 2m) map to surface-accessible positions in β 2m (Fig. 2B). Eight of these exposed residues spatially cluster in the N-terminal part of the domain, whereas 2 (B8 and B10) are close to the C-terminus. All exposed residues are localized at the interface between β 2m and the α domains in MHC Class I. These observations suggest that this region in the B7 C-domain, like β 2m, is involved in protein-protein interactions. The conserved residues in the C-domain may either interact with the N-terminal B7 domain or, alternatively, with other molecules.

Acknowledgments

We thank Tripos Associates, St. Louis, Missouri, and Dr. W. Harte for providing us with a test version of the Matchmaker program and Dr. J. Novotny for making ProsaII available to us. We also thank D. Mosenkis for helpful discussions on Matchmaker.

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.
- Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol* 112:535–542.
- Brucoleri RE, Haber E, Novotny J. 1988. Structure of antibody hypervariable loops reproduced by a conformational search algorithm. *Nature* 335:564–568.