Protein Science (1996), *5*:240–247. Cambridge University Press. Printed in the USA.

Structure-based modeling of the ligand binding domain of the human cell surface receptor CD23 and comparison of two independently derived molecular models

JÜRGEN BAJORATH AND ALEJANDRO ARUFFO
Bristol-Myers Squibb Pharmaceutical Research Institute, Seattle, Washington 98121

(RECEIVED September 27, 1995; ACCEPTED November 9, 1995)

Abstract

CD23, a type II membrane receptor protein, recognizes four different ligands via its extracellular C-type lectin domain: immunoglobulin E (IgE), CD21, and the β2-integrins CD11b and CD11c. CD23 specifically interacts in a calcium-dependent manner, "lectin-like" with carbohydrate moieties expressed on CD21 and CD11b/c, but also "lectin-unlike" with protein epitopes on IgE. As a first step in analyzing the multiple binding specificities associated with CD23 in more detail, we report a detailed molecular model of the lectin-like domain of human CD23 (hCD23). The model was built based on information provided by X-ray structures of mannose binding protein (MBP) and E-selectin, both of which are members of the calcium-dependent (C-type) lectin superfamily. Sequence-structure comparisons suggest that hCD23 is structurally more similar to MBP than to E-selectin. The hCD23 model is compared to an independently derived model. Although the CD23–carbohydrate and CD23–protein interactions are both calcium dependent, analysis of the model suggests the presence of distinct binding sites for these ligands.

Keywords: CD23; C-type lectins; comparative modeling; protein structure prediction; receptor–ligand interactions; structure comparison

CD23 is a member of the C-type lectin superfamily (Drickamer, 1988, 1993). Its extracellular region includes a C-type lectin domain and three short consensus repeats that are followed by a single transmembrane region and a short cytoplasmatic tail (Kikutani et al., 1986). The CD23 consensus repeats are thought to form α-helical coiled-coils (Beavil et al., 1992), resulting in oligomerization of the extracellular region. On the basis of sequence similarities, the CD23 lectin domain belongs to subgroup II of the C-type lectins (Drickamer, 1993), which includes several other type II transmembrane receptor proteins. CD23 is expressed on many haematopoietic cell types, including T- and B-cells, monocytes, platelets, dendritic cells, and natural killer cells (Richards & Katz, 1991; Delespesse et al., 1992). CD23 specifically recognizes at least four different ligands. It was originally described as a low-affinity receptor for IgE (Yukawa et al., 1987). The IgE binding site was mapped to the CD23 lectin domain (Bettler et al., 1991). CD23 also recognizes CD21, a cell-surface receptor for several complement proteins and interferon-α (Aubry et al., 1994). Specific CD23–CD21 interactions can account for many, but not all of CD23s biological effects, including its involvement in the regulation of IgE synthesis and cellular adhesion (Richards & Katz, 1991; Bonnefoy et al., 1993). Recently, it was shown that CD23 binds to the CD18/CD11b and CD18/CD11c β2 integrins expressed by monocytes (Lecoanet-Henchoz et al., 1995). This interaction is thought to trigger the production of inflammatory cytokines such as interleukin-1β and -6 and TNF-α. These findings implicate CD23 in a number of chronic inflammatory diseases such as rheumatoid arthritis (Hellen et al., 1991). Indeed, antibodies against CD23 significantly improve established type II collagen-induced arthritis in a murine model (Plater-Zyberg & Bonnefoy, 1995). Thus, CD23 may be an attractive target for the treatment of chronic inflammatory diseases.

Taken together, the studies discussed above (Yukawa et al., 1987; Bettler et al., 1991; Aubry et al., 1994; Lecoanet-Henchoz

Reprint requests to: Jürgen Bajorath, Bristol-Myers Squibb Pharmaceutical Research Institute, 3005 First Avenue, Seattle, Washington 98121.

et al., 1995) suggest that the binding sites for all four ligands reside in the CD23 C-type lectin domain and that the CD23-ligand interactions are generally calcium dependent and of relatively low affinity. The CD23–CD21 and CD23–β2 integrin interactions are "lectin-like" in that CD23 recognizes specific carbohydrate structures expressed on CD21 and CD11b/c, respectively. In contrast, CD23 binds deglycosylated IgE (Richards & Katz, 1991), suggesting the "lectin-unlike" recognition of a protein (rather than a carbohydrate) epitope. Thus, the C-type lectin domain of CD23 displays a remarkable specificity profile whose molecular origin is yet to be determined. As a first step in analyzing the differences in CD23 specificity in detail, we have attempted to construct a reliable molecular model of the lectin-like domain of human CD23 (hCD23). Molecular models of several members of the C-type lectin family have been reported previously. These include E-selectin (Erbe et al., 1992; Mills, 1994; Bajorath et al., 1995), P-selectin (Bajorath et al., 1993; Erbe et al., 1993; Hollenbaugh et al., 1993), CD69 (Bajorath & Aruffo, 1994), and also a model of CD23 (Padlan & Helm, 1993). Although the details of the respective modeling protocols differ considerably, all of these molecular models were constructed based on the X-ray structure of the rat mannose binding protein (MBP) (Weis et al., 1991, 1992), which was the first structure of a C-type lectin to be determined. The modeling studies were complicated by the unusual finding that MBP consists, to ~50%, of nonstandard (non-α or -β) secondary structure elements (~20% loops and ~30% extended regions of nonstandard secondary structure) (Weis et al., 1991). In the presence of limited sequence identity between MBP and other C-type lectins, often less than 30%, it was difficult to decide whether these regions are structurally conserved in other C-type lectins and how to model regions corresponding to unusual secondary structure elements. In the case of E-selectin, a comparison of a model with the later determined X-ray structure (Graves et al., 1994) was possible, and it was shown that the availability of only MBP as structural template limited modeling accuracy in regions of unusual secondary structure (Bajorath et al., 1995).

Other X-ray structures, including C-type lectin domains, trimers of human (Sheriff et al., 1995) and rat (Weis & Drickamer, 1994) MBP and, in addition, E-selectin (Graves et al., 1994), have recently become available. The lectin domain of human MBP is, in sequence and structure, very similar to its rat homologue (Sheriff et al., 1995). On the other hand, MBP and E-selectin display less than 30% sequence homology and, although the carbohydrate recognition domains of these proteins adopt the same fold, their structures differ considerably in some regions (Graves et al., 1994). Therefore, comparison of these two C-type lectin structures allows a better definition of conserved C-type lectin core regions and of variable structural elements, and provides an improved basis for comparative modeling (Bajorath et al., 1993) of other members of the C-type lectin superfamily. We have applied this knowledge to the model building of the hCD23 lectin-like domain (Kinemage 1). The model was assessed using stereochemical and inverse folding analysis. A detailed comparison with an independently derived model of CD23 (Padlan & Helm, 1993) illustrates the opportunities and also the "resolution limits" of comparative modeling. In the absence of an experimentally determined structure, these models provide a structural basis for the analysis of the different specificities associated with the lectin-like domain of CD23.

Results and discussion

Structurally conserved regions in the lectin domains of MBP and E-selectin

We have first compared the X-ray structures of the C-type lectin domains of MBP at 1.7 Å (Weis et al., 1992) and of E-selectin at 2.0 Å resolution (Graves et al., 1994) in detail to identify regions that are structurally invariant despite the presence of limited sequence similarity. The X-ray structures were initially superposed using the backbone segments of corresponding α-helices and β-strands. Structurally conserved regions were then identified by sequential least-squares superposition of backbone segments of increasing length followed by RMS comparison. Nine backbone segments in MBP and E-selectin, including two regions of extended unusual secondary structure, superpose with RMS deviations (RMSDs) of less than 0.7 Å (Table 1). Figure 1 illustrates the spatial arrangement of the structurally conserved regions. The backbone RMSD for simultaneous superposition of the structurally conserved regions is ~1.2 Å. Excluding helix α2, the cumulative backbone RMSD for the superposition is ~0.8 Å. The low RMSDs for helix α2 and its preceding segment, when superposed separately (Table 1), are increased to ~1.4 Å when compared in combination (residues 25–41 in E-selectin). Although helix α2 is conserved as a single secondary structure element in MBP and E-selectin, it is rigid-body shifted in the context of the three-dimensional structures and adopts a different relative spatial position. Therefore, understanding the putative spatial position of helix α2 in hCD23 is a critical aspect of the modeling (see below).

Sequence comparison in light of X-ray structures

Following structural comparison, amino acid sequences of the C-type lectin domains of murine and human CD23 were aligned against a template that included the topological alignment of the MBP and E-selectin X-ray structures (Graves et al., 1994) and the structurally highly conserved regions as identified above.

Table 1. *Structurally conserved regions in MBP and E-selectin*[a]

Secondary structure	MBP	E-selectin	RMSD (Å)
β1	111–114	2–5	0.37
α1	121–131	12–22	0.30
NR	134–139	25–30	0.57
α2	142–150	33–41	0.34
β2	157–160	51–54	0.40
NR	180–188	75–83	0.65
β3	195–199	90–94	0.48
β4	202–209	102–109	0.28
β5	210–219	110–119	0.64

[a] Residues in the C-type lectin domains of the mannose binding protein (MBP) and E-selectin that form conserved secondary structure elements are listed. Names of secondary structure elements are taken from MBP (Weis et al., 1991), except "NR," which indicates extended regions of unusual (non-α, non-β) secondary structure. RMSDs are calculated for pairwise direct least-squares superposition of the backbone atoms of corresponding segments in MBP and E-selectin.

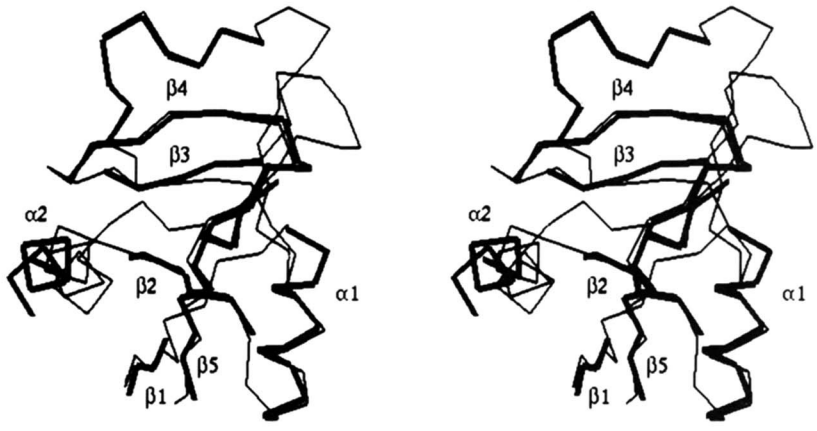

Fig. 1. Schematic representation of the C-type lectin fold and structurally conserved regions in the X-ray structures of rat MBP and E-selectin. The α-carbon trace of MBP is shown (thin line) with structurally conserved regions in E-selectin superposed (thick line). Major secondary structure elements in MBP are labeled (as in Fig. 2). Structurally conserved regions listed in Table 1 were included in the cumulative superposition (except helix α2, which is conserved as a single secondary structure element, but shifted in the context of the three-dimensional structures).

The structure-oriented sequence alignment is shown in Figure 2. Although the sequence of the hCD23 lectin domain is overall less than 30% identical to MBP and E-selectin, hCD23 shares many C-type lectin consensus residues with these proteins (Weis et al., 1991). These residues are thought to be determinants or characteristics of the C-type lectin fold. Of the 36 consensus positions within the aligned region, 19 residues are identical in rat MBP and hCD23 and 12 are conservatively replaced. The conservation or conservative replacements include 4 cysteine residues involved in the formation of disulfide bonds and, in addition, 18 hydrophobic core residues. Thus, the three-dimensional structures of these proteins are expected to be more similar than indicated by their level of sequence identity. The conservation of consensus residues, boxed in Figure 2, allows a high confidence sequence to structure alignment of hCD23 to MBP and E-selectin, including the locations of insertions and deletions. When compared to MBP, hCD23 displays two one-residue insertions that are predicted to extend the loops connecting α2/β2

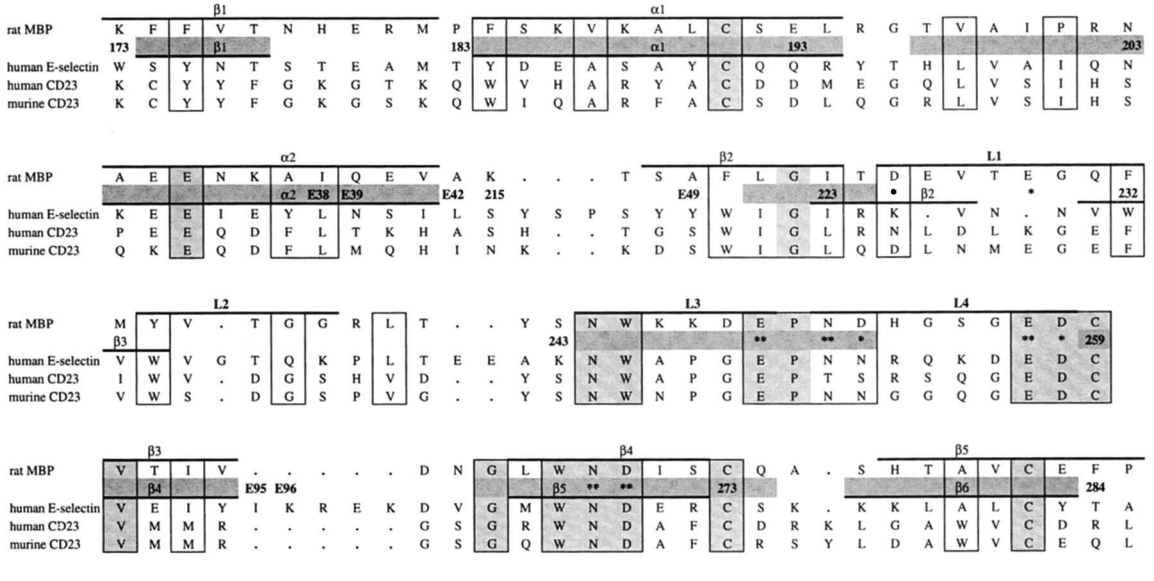

Fig. 2. Alignment of C-type lectin domain sequences of murine and human CD23, human E-selectin, and rat MBP. The alignment is combined with a topological alignment of MBP and E-selectin (Graves et al., 1994) and the analysis of structurally invariant regions in MBP and E-selectin presented herein. Major secondary structure elements in MBP (Weis et al., 1991) and E-selectin (Graves et al., 1994) are overlined and labeled. L1–L4 denote extended regions of nonclassical secondary structure in MBP. Residue numbers are given for human CD23. The numbers of E-selectin residues specifically discussed in the text are also given and designated E (for example, E38). Residues forming two calcium-binding sites in MBP are labeled with asterisks. Residues in MBP that form a calcium-binding site conserved in E-selectin are double-labeled (**). A calcium-binding site only present in ligand-bound MBP (Weis et al., 1992), was not considered. Consensus residues that are thought to be critical for the C-type lectin fold (Weis et al., 1991) are boxed. Residues invariant in the compared sequences are shaded. Horizontal (darker) shading shows the regions in MBP and E-selectin, defined here as structurally invariant (i.e., whose backbone segments superpose with less than 0.7 Å RMSD). The sequences human E-selectin and MBP are ~28% identical. Sequence identity for MBP and the modeled region of human CD23 is ~24% and the corresponding sequence segments of human and murine CD23 are ~67% identical.

and $\beta 4/\beta 5$, respectively. In contrast, four regions in the selectins show insertions and one region deletions relative to MBP and hCD23. Based on this observation and the analysis of helix positions, as discussed below, we have started the modeling of hCD23 from MBP, the highest resolution C-type lectin structure currently available (Weis et al., 1992).

Modeling of structurally conserved regions

The analysis of backbone segments in the C-type lectin domains of MBP and E-selectin that are, despite sequence variation, highly conserved, aids in the identification and modeling of structurally conserved and variable regions in hCD23. The conservation of consensus residues (Fig. 2) suggests that the structurally conserved segments (Table 1) are also invariant in hCD23. Therefore, the respective backbone segments of MBP were included in the hCD23 model. As noted above, the classification of helix $\alpha 2$ as a structurally conserved element must be considered with caution. In E-selectin, this helix moves away from the protein core when compared to MBP and adopts a different spatial position. We have analyzed the interactions of helix $\alpha 2$ in E-selectin and MBP with the remainder of the proteins to decide how to model the position of this helix in hCD23. In E-selectin, two pairwise interactions appear to cause the rigid-body shift of $\alpha 2$ relative to MBP. Residues 95 and 96 (I-K) in E-selectin, part of an insertion between $\beta 4$ and $\beta 5$ that is absent in MBP and CD23 (Fig. 2), contact residues 38 and 39 (L-N) of $\alpha 2$. Likewise, residue 42 (L) at the C-terminus of $\alpha 2$ in E-selectin packs against residue 49 (Y). These large residues (L/Y), which form this interaction in E-selectin, correspond to a pair of small residues in both MBP (A/A) and hCD23 (A/S). Thus, the conservation of these residues and the absence of the insertion in MBP and hCD23 suggested that the position of helix $\alpha 2$ in hCD23 should be similar to MBP.

Modeling of structurally variable regions

In contrast to the structurally conserved regions discussed above, the conformation of other segments in MBP and E-selectin, including regions of insertions or deletions, are significantly different. These segments were regarded as structurally variable in hCD23 and modeled by systematic conformational search. For example, the segment including CD23 residues 253–256 corresponds to a structurally variable region in MBP and E-selectin and the corresponding region in hCD23 shows no sequence conservation. This region in hCD23 was therefore modeled by conformational search. Backbone segments of structurally variable regions in MBP and E-selectin were only included in the model if one of the sequences was very similar to hCD23 and, in addition, if interactions with the residual protein appeared to be also conserved. For example, CD23 residues 241–243 (T-Y-S) are similar to corresponding residues in MBP (D-Y-S), but different from the selectins that show no conservation in this region and, in addition, a two-residue insertion. Thus, although not conserved in MBP and E-selectin, this segment should be conserved in MBP and CD23 and its backbone conformation was included the hCD23 model. Following these criteria, six regions were, after initial side-chain modeling (see the Materials and methods), considered to be structurally variable and subjected to systematic conformational search in the following order: 236–239, 214–218, 226–230, 253–256, 178–182, 274–277. Conformational searches resulted in acceptable conformations (on the basis of selection criteria given in the Materials and methods) for five of these regions (except 178–182), including the two regions with insertions relative to MBP (214–218, 274–277). These five conformations were included in the hCD23 model.

Putative calcium binding sites

On the basis of our structure-oriented sequence alignment, only one of four residues that form calcium binding site 1 (Ca1) in MBP is conserved in hCD23 (D258), and one Ca1 residue in MBP is conservatively replaced (N225). However, at position 252 in hCD23, a serine replaces aspartic acid in MBP and, in addition, the positively charged residue K229 replaces a negatively charged calcium ligand (glutamic acid) in MBP. In accordance with these observations, Ca1 in MBP is predicted to be absent in hCD23 and was not included in the model. In contrast, four of five ligands (E249, E257, N269, D270) that form calcium binding site 2 (Ca2) in MBP (Weis et al., 1991), which is also present in E-selectin (Graves et al., 1994), are conserved in hCD23. One Ca2-binding residue is not conserved in hCD23: T251 replaces an asparagine at the corresponding position in MBP. This change may be expected to weaken the Ca2 coordination sphere, but should retain calcium binding. Thus, a calcium position approximately corresponding to Ca2 in MBP was included in the hCD23 model. Based on its conservation, we predict that this calcium represents a functionally important calcium-binding site in hCD23.

Refinement and assessment of the model

The initially assembled model of the hCD23 lectin domain, including residues 173–285 and one calcium ion, was subjected to energy refinement until the RMS derivative of the energy function was ~1.2 kcal/mol Å. At this stage of the minimization, the backbone (all atom) RMSD relative to the initially assembled model was ~0.7 (~1.0) Å. Prior to energy minimization, interatomic distances within the predicted calcium site were slightly adjusted to accommodate the N/T substitution at position 251. During energy minimization, the residues that participate in the formation of the predicted calcium coordination sphere were constrained to their adjusted positions. As a result, the initial geometry of this approximate calcium site remained largely conserved. The calcium ligating side-chain atoms are in less than 3 Å distance to the calcium in the final model. The modeled calcium coordination sphere is incomplete due to the absence of water positions. When assessed with PROCHECK (Laskowski et al., 1993), the model showed good stereochemistry, no residues with disallowed main torsion angles, and no short intramolecular contacts. Energy profiles of the model were calculated using ProsaII (Sippl, 1993) to assess its sequence-structure compatibility. Figure 3 shows the results of the energy profile analysis of the model compared to the profiles of the MBP and E-selectin X-ray structures. As to be expected, the calculated pairwise residue interactions are more favorable in the crystallographic parent structure than in the model. However, the low average energy values of pairwise residue interactions and the shape similarity of the hCD23 model compared to the profile of the MBP X-ray structure indicate that the hCD23 model is overall correctly folded (Sippl, 1993) and that significant errors in the core regions of the model are absent. The energy

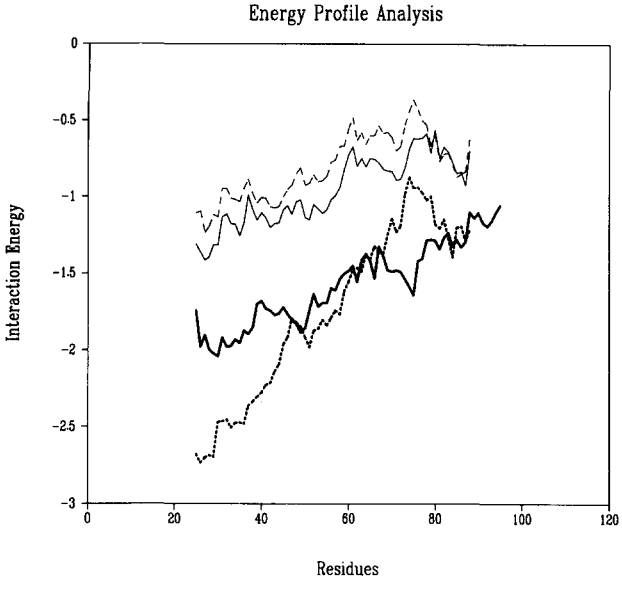

Fig. 3. Comparison of energy profiles. Profiles were calculated using ProsaII (Sippl, 1993) for our hCD23 model (solid line), the CD23 model of Padlan and Helm (1993) (dashed line), and the X-ray structures of MBP (Weis et al., 1992) (thick dotted line) and E-selectin (Graves et al., 1994) (thick solid line). Pairwise residue interaction energy is given in units of E/kT (E, interaction energy in kcal/mol; k, Boltzmann constant; T, temperature in K), calculated at each residue position, and averaged using a 50-residue window (Sippl, 1993).

profiles do not necessarily allow the detection of errors in modeled loop conformations.

Comparison with an independently derived CD23 model

Padlan and Helm (1993) have reported models of human and murine CD23 using rat MBP at 1.7 Å resolution as template structure. These models were generated prior to the availability of the E-selectin crystal structure. The CD23/MBP sequence alignment was based on global sequence similarities and the presence of defined secondary structure elements in MBP where the introduction of gaps was avoided. Padlan and Helm (1993) used options of the FRODO program (Jones, 1985) for their modeling. Side-chain conformations were modeled manually to provide, if possible, maximum overlap with the original side chain in MBP. Backbone conformations were only modified to accommodate the two insertions in hCD23 relative to MBP. The overall backbone conformation of the model is therefore more similar to MBP than expected at the level of ~30% sequence identity (Chothia & Lesk, 1986). The model was subjected to energy minimization. No calcium position was included in the final model (PDB code "1HLI"). Energy profiles were calculated for the previously derived and the present model (Fig. 3). The profiles are similar and suggest essentially equivalent sequence-structure compatibility. Figure 4 shows a superposition of the two hCD23 models. Despite the expected global similarity of the models, conformational differences are observed in several loops and in regions with insertions relative to MBP. Regions in the two models with significant differences in backbone conformation are given in Table 2. The backbone RMSD for superposition of predicted structurally conserved regions (Table 1; Fig. 2) is ~0.7 Å, and the RMSD for superposition of all residues (173–285) is ~0.9 Å. Thus, both differences in the model refinement procedures and in the modeling of backbone conformations of regions considered to be structurally variable in our study contribute to overall deviations. When taking side-chain atoms into account, the RMSD increases to ~2.0 Å due to some substantial differences in the modeling of side-chain conformations. Packing interactions in both models are overall similar because the conformations of residues in the core regions were modeled in a similar way. Figure 5 shows the superposition of the models focused on the predicted calcium-binding site and its vicinity and illustrates a number of considerable differences in the modeling of surface residues. For example, the previously modeled conformation of residue R253 partially blocks access to the putative calcium coordination sphere, whereas the other does not. In the FRODO-modeled insertion of the $\beta 4/\beta 5$-loop in hCD23, residue L277 interacts with M261. In contrast, in the CONGEN-generated conformation of this region, this interaction is absent. As a result of this difference, the conformation of M261 is less restricted and the shape of this region can be modulated by selection of alternative rotamer conformations. Thus, due to the limitations of comparative model building, many details such as side-chain conformations or specific interactions between residues may not be predicted with confidence and alternatives need to be considered.

Model analysis

We have used the hCD23 model to outline potential binding sites for the different (carbohydrate and protein) ligands, a prereq-

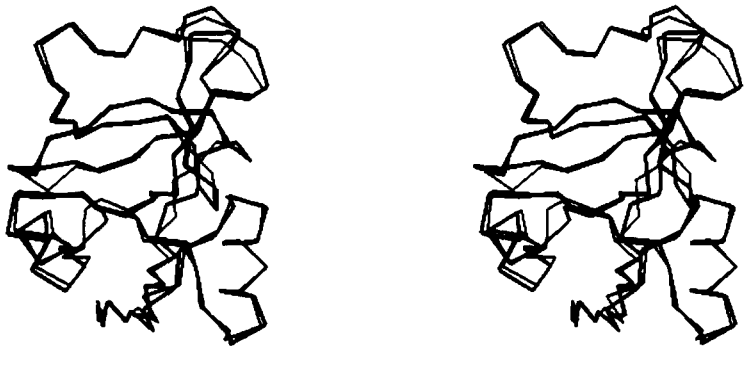

Fig. 4. Comparison of hCD23 models. The α-carbon traces of the present hCD23 model (thick line) and the model of Padlan and Helm (1993) (thin line) are shown. Orientation is similar to Figure 1. The stereo view was obtained by superposition of the backbone segments of the invariant regions highlighted in Figure 2.

Table 2. *Regions in two models of CD23 with significant differences in backbone conformation*

Secondary structure	Residues	Variable	RMSD (Å)
α2/loop	214–218	+	1.41
NR	227–229	+	1.01
NR	236–238	+	1.95
β3/loop	263–266	−	1.00
Loop	274–278	+	1.71

[a] Segments in the present hCD23 model and the model of Padlan and Helm (1993), which, when least-squares superposed, show significant differences in backbone conformation (RMSD of at least 1 Å) are listed. Segments are designated according to secondary structures elements in MBP; NR, extended regions of nonregular (non-α, non-β) secondary structure. Residue numbers are given for CD23. Variable (+/−) indicates whether the segments correspond to, or overlap with, a region considered to be structurally variable (178–182, 214–218, 226–230, 236–239, 253–256, 274–277) in our study.

uisite for targeted mutagenesis. Bettler et al. (1992) have carried out CD23 homologue-scanning mutagenesis and assessed mutant proteins in IgE binding assays. These researchers also used a panel of conformation sensitive anti-CD23 monoclonal antibodies to assess the structural integrity of their mutant proteins. In this study, four segments in hCD23 (185–190, 224–233, 238–243, 251–256) were identified as critical for the interaction with IgE. When mapped on the hCD23 model, these four segments cluster on a region of its surface (Fig. 6 and Kinemage 1). Although the individual residues that form the IgE binding site are not known, the spatial relation of these segments in the model implicates a relatively large and continuous surface area in IgE binding, consistent with the presence of protein–protein interactions. That residues in segment 251–256, which is adjacent to the putative calcium-binding site, are critical for the hCD23/IgE interaction may explain its calcium dependence. Considering the mode of mannose binding to MBP (Weis et al., 1992), it is likely that carbohydrate binding to hCD23 directly involves the calcium coordination sphere. Because the carbohydrate epitopes recognized by hCD23 are yet to be determined, carbohydrate binding may also extend to other regions of the molecule such as, for example, observed for E- and P-selectin (Erbe et al., 1992, 1993; Bajorath et al., 1993; Hollenbaugh et al., 1993; Graves et al., 1994), which recognize sialylated Lewis X tetrasaccharide structures. In the hCD23 model, a hydrophobic surface patch flanked by positively charged residues (Fig. 6 and Kinemage 1) represents an attractive region that approximately corresponds to the location of residues critical for ligand binding to the selectins. In hCD23, this region is proximal to the putative calcium site, but does not include segments critical for the hCD23/IgE interaction. Thus, the predicted arrangement of the residues depicted in Figure 6 suggests the possibility that the hCD23 binding sites for protein and carbohydrate ligands are spatially distinct.

Conclusions

The analysis of structurally conserved and variable regions in MBP and E-selectin, C-type lectins with limited sequence similarity and known three-dimensional structure, provides an improved basis for modeling of other members of the C-type lectin family. Knowledge-based modeling and contemporary structure assessment techniques have been applied to generate a detailed three-dimensional model of the C-type lectin domain in hCD23. The comparison of two independently derived CD23 models shows that, in this case, differences in the applied modeling procedures result in global RMSDs at the 2-Å level, although the models agree overall quite well. This illustrates the limits of comparative model building. The availability of two structural models allows the identification of regions modeled with higher and lower confidence and to assess alternative residue arrangements and interactions. In the absence of an experimental structure, these considerations aid in the design of mutagenesis experiments and the outline of potential binding sites. A conclusion drawn from model analysis is that binding sites for protein and carbohydrate ligands in hCD23 may be distinct, although all CD23–ligand interactions appear to be calcium dependent. This hypothesis can be experimentally tested in hCD23 model-based mutagenesis studies.

Materials and methods

For structure comparison and model building, the X-ray structures of E-selectin at 2.0 Å resolution (Graves et al., 1994) and of MBP at 1.7 Å resolution (monomer B of the asymmetric unit) (Weis et al., 1992), the highest-resolution MBP structure currently available, were selected. Initial sequence alignments were generated using GCG programs (Genetics Computer Group, Inc., Madison, Wisconsin) and manually modified to obtain structure-oriented alignments. Computer graphics, structural

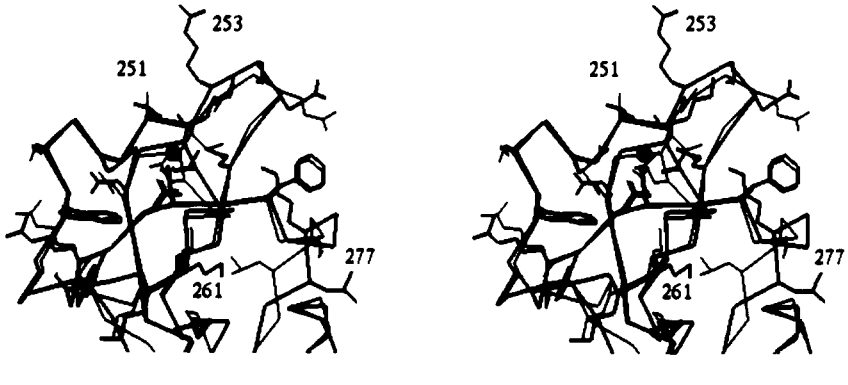

Fig. 5. Close-up stereo view of the region surrounding the predicted calcium-binding site in hCD23. Models are superposed and represented as in Figure 4. The putative calcium position is shown as a black sphere. Side chains of residues forming the calcium-binding site and of other residues in the vicinity are included in the comparison. Residues discussed in the text are numbered.

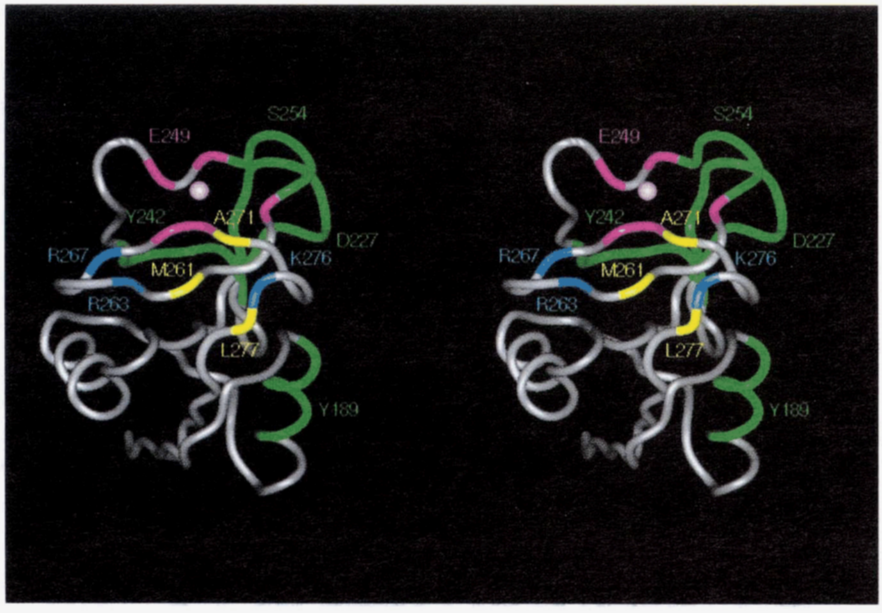

Fig. 6. Outline of potential ligand binding sites in human CD23. The hCD23 model is represented as a solid ribbon (silver) in an orientation similar to Figure 4. Backbone segments that include residues critical for IgE binding (Bettler et al., 1992) are colored in green (185–190, 224–233, 238–243, 251–256). Single residues in these segments are labeled to indicate their positions in the model. The proposed calcium ion is depicted as a pink sphere, and residues that participate in the formation of the calcium-binding site (for example, E249) are shown in magenta. Charged (R263, R267, K276) and hydrophobic (A271, M261, L277) residues discussed in the text are colored in blue and yellow, respectively, and labeled.

manipulations, and energy minimization calculations were carried out using the InsightII/Discover program package (Biosym Technologies, Inc., San Diego, California). For energy refinement with Discover, hydrogen atoms were added to the model according to pH 7 and partial charges were assigned to all atoms. No water positions were included. The calculations were carried out using CVFF force field parameters, a distance-dependent dielectric constant ($4r$), and a 15-Å cutoff distance for the treatment of nonbonded interactions. Conservative side-chain replacements (such as Y/F or L/I) were modeled in conformations as similar as possible to the original conformation. Other residue replacements were modeled using low-energy rotamer conformations (Ponder & Richards, 1987) as described (Bajorath & Fine, 1992) and implemented in InsightII. Low-energy combinations of side-chain rotamers were calculated for clusters of spatially adjacent residues using an 8-Å cutoff distance for nonbonded interactions. The conformations of loops (or other segments considered to be conformationally variable) were modeled by systematic conformational search using CONGEN (Bruccoleri & Novotny, 1992) essentially following a previously described protocol (Bruccoleri et al., 1988). Conformations with negative potential energy and smallest solvent-accessible surface area within 3 kcal/mol of the energy minimum were selected as the best conformations. Because conformational search results are environment dependent, these calculations were carried out after all other parts of the model were completed, but prior to global energy refinement. Loops were modeled in order of increasing interloop interactions. The stereochemistry of the model was analyzed using PROCHECK (Laskowski et al., 1993). Sequence-structure compatibility was assessed by energy profile analysis using the PROSAII program as described (Sippl, 1993). The pairwise residue interactions were calculated using β-carbon potentials and, for graphical representation of the profiles, a 50-residue window was used for energy averaging at each residue position. If not stated otherwise, stereo figures were generated using InsightII.

Supplementary material in Electronic Appendix

Coordinates of the hCD23 model are provided on diskette to accompany the manuscript. Coordinates have also been submitted to the Brookhaven Protein Data Bank (IDCODE "IKJE").

Acknowledgments

We thank Debby Baxter for her help in the preparation of the manuscript. We also thank Drs. Padlan and Helm for making coordinates of their CD23 models publicly available.

References

Aubry JP, Pochon S, Gauchat JF, Nueda-Marin A, Holers VM, Graber P, Siegfried C, Bonnefoy JY. 1994. CD23 interacts with a new functional extracytoplasmatic domain involving N-linked oligosaccharides on CD21. *J Immunol 152*:5806–5813.

Bajorath J, Aruffo A. 1994. Molecular model of the extracellular lectin-like domain in CD69. *J Biol Chem 269*:32457–32463.

Bajorath J, Fine RM. 1992. On the use of minimization from many randomly generated loop structures in modeling antibody combining sites. *Immunomethods 1*:137–146.

Bajorath J, Stenkamp R, Aruffo A. 1993. Knowledge-based model building of proteins: Concepts and examples. *Protein Sci 2*:317–334.

Bajorath J, Stenkamp R, Aruffo A. 1995. Comparison of a protein model with its X-ray structure: The ligand binding domain of E-selectin. *Bioconjugate Chem 6*:3–6.

Beavil AJ, Edmeades RL, Gould HJ, Sutton BJ. 1992. α-Helical coiled-coil stalks in the low-affinity receptor for IgE (FcϵPII/CD23) and related C-type lectins. *Proc Natl Acad Sci USA 89*:753–757.

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.

Bettler B, Texido G, Raggini S, Rüegg D, Hofstetter H. 1992. Immunoglobulin E-binding site in Fcϵ receptor (FcϵRII/CD23) identified by homolog-scanning mutagenesis. *J Biol Chem 267*:185–191.

Bonnefoy JY, Aubry JP, Gauchat JF, Graber P, Life P, Flore-Romo L, Mazzei G. 1993. Receptors for IgE. *Curr Opin Immunol 5*:944–947.

Bruccoleri RE, Haber E, Novotny J. 1988. Structure of antibody hypervariable loops reproduced by a conformational search algorithm. *Nature 335*:564–568.

Bruccoleri RE, Novotny J. 1992. Antibody modeling using the conformational search program CONGEN. *Immunomethods 1*:96–106.

Chothia C, Lesk AM. 1986. The relation between the divergence of sequence and structure in proteins. *EMBO J 5*:823–826.

Delespesse G, Sarfati M, Wu CY, Fournier S, Letellier M. 1992. The low affinity receptor for IgE. *Immunol Rev 125*:77–97.

Drickamer K. 1988. Two distinct classes of carbohydrate recognition domains in animal lectins. *J Biol Chem 263*:9557–9560.

Drickamer K. 1993. Ca^{2+}-dependent carbohydrate recognition domains in animal proteins. *Curr Opin Struct Biol 3*:393–400.

Erbe DV, Watson SR, Presta LG, Wolitzky BA, Foxall C, Brandley BK, Lasky LA. 1993. P- and E-selectin use common sites for carbohydrate ligand recognition and cell adhesion. *J Cell Biol 120*:1227–1235.

Erbe DV, Wolitzky BA, Presta LG, Norton CR, Ramos RJ, Burns DK, Rumberger JM, Rao NBN, Foxall C, Brandley BK, Lasky LA. 1992. Identification of an E-selectin region critical for carbohydrate recognition and cell adhesion. *J Cell Biol 119*:215–227.

Graves BJ, Crowther RL, Chandran C, Rumberger JM, Li S, Huang KS, Presky DH, Familetti PC, Wolitzky BA, Burns DK. 1994. Insight into E-selectin/ligand interaction from the crystal structure and mutagenesis of the lec/EGF domains. *Nature 367*:532–538.

Hellen EA, Rowlands DC, Hansel TT, Kitas GD, Crocker J. 1991. Immunohistochemical demonstration of CD23 expression on lymphocytes in rheumatoid synovitis. *J Clin Pathol 44*:293–296.

Hollenbaugh D, Bajorath J, Stenkamp R, Aruffo A. 1993. Interaction of P-selectin (CD62) and its cellular ligand: Analysis of critical residues. *Biochemistry 32*:2960–2966.

Jones TA. 1985. Interactive computer graphics: FRODO. *Methods Enzymol 115*:157–171.

Kikutani H, Inui S, Sato R, Barsumian EL, Owaki H, Yamasaki K, Kaisho T, Uchibayashi N, Hardy RR, Hirano T, Tsunasawa S, Sakiyama F, Suemura M, Kishimoto T. 1986. Molecular structure of human lymphocyte receptor for immunoglobulin E. *Cell 47*:867–885.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr 26*:283–291.

Lecoanet-Henchoz S, Gauchat JF, Aubry JP, Graber P, Life P, Paul-Eugene N, Ferrua B, Corbi AL, Dugas B, Plater-Zyberk C, Bonnefoy JY. 1995. CD23 regulates monocyte activation through a novel interaction with the adhesion molecules CD11b-CD18 and CD11c-CD18. *Immunity 3*:119–125.

Mills A. 1994. Modelling of the carbohydrate recognition domain of human E-selectin. *FEBS Lett 319*:5–11.

Padlan EA, Helm BA. 1993. Modeling of the lectin-homology domains of the human and murine low-affinity Fcε receptor (FcεRII/CD23). *Receptor 3*:325–341.

Plater-Zyberk C, Bonnefoy JY. 1995. Marked amelioration of established collagen-induced arthritis by treatment with antibodies to CD23 in vivo. *Nature Medicine 1*:781–785.

Ponder JW, Richards FM. 1987. Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J Mol Biol 193*:775–791.

Richards ML, Katz DH. 1991. Biology and chemistry of the low affinity IgE receptor (FcεRII/CD23). *Crit Rev Immunol 11*:65–86.

Sheriff S, Chang CY, Ezekowitz RAB. 1994. Human mannose-binding protein carbohydrate recognition domain trimerizes through a triple α-helical coiled-coil. *Nature Struct Biol 1*:789–793.

Sippl MJ. 1993. Recognition of errors in three-dimensional structures of proteins. *Proteins Struct Funct Genet 17*:355–362.

Weis WI, Drickamer K. 1994. Trimeric structure of a C-type mannose-binding protein. *Structure 2*:1227–1240.

Weis WI, Drickamer K, Hendrickson WA. 1992. Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature 360*:127–134.

Weis WI, Kahn R, Fourme R, Drickamer K, Hendrickson WA. 1991. Structure of the calcium-dependent lectin domain from a rat mannose-binding protein determined by MAD phasing. *Science 254*:1608–1615.

Yukawa K, Kikutani H, Howaki H, Yamasaki K, Yokota A, Nakamura H, Barsumian EL, Hardy RR, Suemura M, Kishimoto T. 1987. A B cell specific differentiation antigen, CD23, is a receptor for IgE (FcεR) on lymphocytes. *J Immunol 38*:2576–2580.