Protein Science (1997), 6:588–597. Cambridge University Press. Printed in the USA.

A 3D model for the measles virus receptor CD46 based on homology modeling, Monte Carlo simulations, and hemagglutinin binding studies

CHRISTIAN MUMENTHALER,[1] URS SCHNEIDER,[2] CHRISTIAN J. BUCHHOLZ,[2] DANIEL KOLLER,[2] WERNER BRAUN,[3] AND ROBERTO CATTANEO[2]

[1]Institut für Molekularbiologie und Biophysik, ETH-Hönggerberg, CH-8093 Zürich, Switzerland
[2]Institut für Molekularbiologie der Universität Zürich, Abteilung I, Hönggerberg, CH-8093 Zürich, Switzerland
[3]Sealy Center for Structural Biology, University of Texas Medical Branch, Galveston, Texas 77555-1157

(RECEIVED October 7, 1996; ACCEPTED December 2, 1996)

Abstract

The two terminal complement control protein (CCP) modules of the CD46 glycoprotein mediate measles virus binding. Three-dimensional models for these two domains were derived based on the NMR structures of two CCP modules of factor H. Both CD46 CCP modules are about 35 Å long, and form a five-stranded antiparallel β-barrel structure. Monte Carlo simulations, sampling the backbone torsion angles of the linker peptide and selecting possible orientations on the basis of minimal solvent-exposed hydrophobic area, were used to predict the orientation of CCP-I relative to CCP-II. We tested this procedure successfully for factor H. For CD46, three clusters of structures differing in the tilt angle of the two domains were obtained. To test these models, we mutagenized the CCP modules. Four proteins, two without an oligosaccharide chain and two with mutated short amino acid segments, reached the cell surface efficiently. Only the protein without the CCP-I oligosaccharide chain maintained binding to the viral attachment protein hemagglutinin. These results are consistent with one of our models and suggest that the viral hemagglutinin does not bind at the membrane-distal tip of CD46, but near the concave CCP-I–II interface region.

Keywords: complement control protein module; distance geometry; FANTOM; hemagglutinin binding; hydrophobic solvent-accessible surface area; measles virus receptor CD46; module assembly; Monte Carlo simulations

CD46, also known as membrane cofactor protein, protects autologous cells from complement lysis. This type 1 transmembrane glycoprotein contains four CCP modules of about 60 residues each (Liszewski et al., 1991) that have been implicated in cell attachment of measles virus (MV), an enveloped negative-strand RNA virus (Dörig et al., 1993; Naniche et al., 1993). Cell fusion experiments using recombinant glycoproteins from MV and a related virus showed that the MV hemagglutinin specifically determines attachment to CD46 (Nussbaum et al., 1995), and a soluble form of the MV hemagglutinin (sH) binds to CD46 with high affinity (Devaux et al., 1996).

CCP modules in other proteins mediate the binding of structurally different viruses. Epstein-Barr virus, an enveloped DNA virus, attaches to the two external CCP modules of the complement regulatory protein CR2 (CD21) (Martin et al., 1991; Molina et al., 1991; Moore et al., 1991), and several picornaviruses, small icosahedral RNA viruses, attach to the complement decay accelerating factor (DAF or CD55) (Bergelson et al., 1994, 1995). Echovirus 7 binds to the second, third, and fourth (last) CCP modules of DAF (Clarkson et al., 1995). Thus, the study of MV hemagglutinin binding to the CD46 CCP modules could give insights into the mode of cell attachment of other viruses.

Previous studies, while establishing the essential role of the CCP modules in MV binding (Varior-Krishnan et al., 1994), indicated that deletion of either the first or second modules results in loss of MV receptor function (Iwata et al., 1995; Manchester et al., 1995), and demonstrated that mutant proteins composed only of duplicated CCP-I or -II, or with these domains in an inverted order, cannot bind MV (Buchholz et al., 1996b). Moreover, it is known that N-linked oligosaccharides, either of the complex type or mannose-rich, are necessary for virus binding (Maisner et al., 1994; Maisner & Herrler, 1995) and that CCP-III and -IV reinforce binding of MV particles (Buchholz et al., 1996b).

Each CCP module has four conserved cysteines that form the disulfide bonds C1-C3 and C2-C4 (Janatova et al., 1989). Solution-state 2D ^{1}H-NMR spectroscopy provided the tertiary structure of

Reprint requests to: W. Braun, Sealy Center for Structural Biology, 301 University Boulevard, University of Texas Medical Branch, Galveston, Texas 77555-1157; e-mail: werner@nmr.utmb.edu.

two CCP modules, the fifth and the 16th repeat of human complement factor H (Barlow et al., 1991, 1992), and of a pair of modules, the 15th and 16th repeats of the same protein (Barlow et al., 1993). We used the NMR structures as a base for generating 3D models of the individual CCP-I and -II modules of CD46 with the program DIAMOD (Hänggi & Braun, 1994; Mumenthaler & Braun, 1995b) and performed Monte Carlo simulations with the program FANTOM (Schaumann et al., 1990; von Freyberg & Braun, 1993; Mumenthaler & Braun, 1995a) to model the relative orientation of the domains.

We tested different model structures by site-directed mutagenesis and expression of the mutant proteins in mouse cells. Effects of the mutations on MV hemagglutinin binding, cell–cell fusion, and monoclonal antibody binding were examined. A model structure consistent with these experimental data was identified. This model will be useful in further characterization of the hemagglutinin binding site and in the design of therapeutic agents to interfere with the MV attachment.

Results

3D models of the individual CCP-I and -II modules and their validation

Figure 1 shows the predicted structure of the external CD46 modules CCP-I (Fig. 1A) and -II (Fig. 1B). Both modules are about 35 Å long and contain five extended strands. The polypeptide chain forms an antiparallel β-barrel with up and down topology, wrapping around a hydrophobic core. The disulfide bonded cysteines are part of the hydrophobic cores, together with the additional residues P4, M10, L12, Y20, V26, Y28, T44, and W52 for CCP-I and I68, P71, A76, P78, Y83, F91, L105, L109, W116, and P121 for CCP-II. The tryptophan and the tyrosine residues are highly conserved among homologous CCP modules, and most of the other core residues are conserved or substituted by other hydrophobic residues. All charged residues are on the surface except E24, which forms an internal salt bridge with the partially buried K15, and E103, which forms a hydrogen bond with the side chain of H90. The N-glycosylation site of CCP-I (N49) is near the N terminus of CD46, whereas the N-glycosylation site of CCP-II (N80) can be regarded as part of the CCP-I–II interface region (magenta spheres in Fig. 1).

The quality of the CCP-I and -II model structures was verified with the program PROCHECK v4.0 (Laskowski et al., 1993). The Ramachandran plot revealed that 96% of all residues excluding glycines and prolines lie within the most favored (68%) or additionally allowed regions (38%). The three residues A41, A79, and A114 are situated in generously allowed regions, and the asparagine, to which the CCP-I oligosaccharide is attached (N49), is in a disallowed region. The backbone dihedral angles of all G and P as well as all the χ^1 and χ^2 conformations of the side chains were in allowed regions. The peptide bond planarity of the model structures, with an average deviation of 9.8°, was slightly worse than the expected 6.0 ± 3.0° variation.

The ratio of the charged to polar solvent-accessible surface area (ASA) in both CCP is an independent test for structural quality. Miller et al. (1987) found that this ratio changed from 1:3.7 in extended polypeptide conformations to 1:1.3 in native proteins. With a corresponding ratio of 1:1.3, our model captures the characteristics of native proteins having most of their charged groups on the solvent-exposed surface.

Relative orientation of the modules

We cannot use the factor H structure to predict the relative orientation of the modules because the tetrapeptide linker (EGLP) is different from the YRET tetrapeptide linker in CD46. Moreover, the NMR data indicate flexibility between the two factor H modules. To obtain a nonbiased prediction, we sampled different orientations by Monte Carlo simulations with FANTOM (Schaumann et al., 1990). Structures were then selected on the basis of low energy and low hydrophobic accessible surface area (ASA_h) (von Freyberg & Braun, 1993; Mumenthaler & Braun, 1995a).

This method could successfully reproduce the experimentally determined relative orientation of the factor H CCPs (Table 1). The 84 structures obtained by the Monte Carlo simulation were classified into six distinct clusters. The cluster with the lowest ASA_h contained structures close to the factor H NMR structure. The 12 structures with lowest ASA_h of 4,925–4,945 Å^2 had an RMS deviation (RMSD) of 3.2–3.3 Å (backbone atoms) and 3.4–3.5 Å (all heavy atoms) to the NMR structure (Fig. 2A). This deviation is similar to the spread in the ensemble of NMR structures to the mean NMR structure with an RMSD of 3.3 Å for all heavy atoms (Barlow et al., 1993). The lowest ASA_h of a structure with a wrong orientation is 4,951 Å^2, i.e., 26 Å^2 above the minimum, and the average value was 5,041 ± 72 Å^2. The differences are relatively small, but they originate exclusively from the small interface region between the two modules and are therefore significant.

We then used the same procedure to search for possible orientations of CCP-I relative to -II in CD46. The analysis of the 75 structures identified a large cluster (33 structures) that contains the structure with the lowest ASA_h. However, three structure clusters with differently oriented modules also had similar hydrophobic areas (Table 1; Fig. 2B). Furthermore, the differences in the ASA_h between the structure clusters are very small, and the clusters M1–M3 have nearly the same minimal ASA_h.

In Figure 3, representative structures of the four clusters (M1–M4) with different linker peptide conformations are displayed. To characterize the relative orientation of both modules, we calculated the tilt and twist angles as defined in Bork et al. (1996). In a procedure similar to the one presented by Barlow et al. (1993), we assigned a rigid frame ($\mathbf{O}^i$,$\mathbf{x}^i$,$\mathbf{y}^i$,$\mathbf{z}^i$) to every module **i**, where the unitary vector $\mathbf{z}^i$ points from the C^α position of cysteine 1 to the C^α position of cysteine 4, $\mathbf{x}^i$ is a perpendicular vector passing through the Cα atom of cysteine 2, and $\mathbf{y}^i$ is perpendicular to $\mathbf{x}^i$ and $\mathbf{z}^i$. Vector $\mathbf{z}^i$ points from the N-terminus to the C-terminus and corresponds approximately to the major axis of the inertia tensor. The tilt angle is then defined as the angle between the **z** vectors, i.e., $a\cos(\mathbf{z}^1 \cdot \mathbf{z}^2)$, and the twist angle is the angle between $\mathbf{x}^1$ and $\mathbf{x}^{2*}$, where $\mathbf{x}^{2*}$ is the orientation of $\mathbf{x}^2$ after rotating the second module by the tilt angle such that $\mathbf{z}^1$ and $\mathbf{z}^2$ are colinear.

Using these definitions, we found (tilt, twist) angle pairs of (57°, 95°), (112°, −84°), (17°, 113°), and (51°, −41°) for M1, M2, M3, and M4, respectively. Only the structure M1 comes close to the angles of the factor H module pair (95°, 103°). A recent study (Bork et al., 1996) analyzing all known 3D structures of module pairs found that, whereas the tilt angle varied significantly between 20° and 120° (within a maximum range of [0°, 180°]), the twist angles were clustered in a region from 100° to 180° (range [−180°,

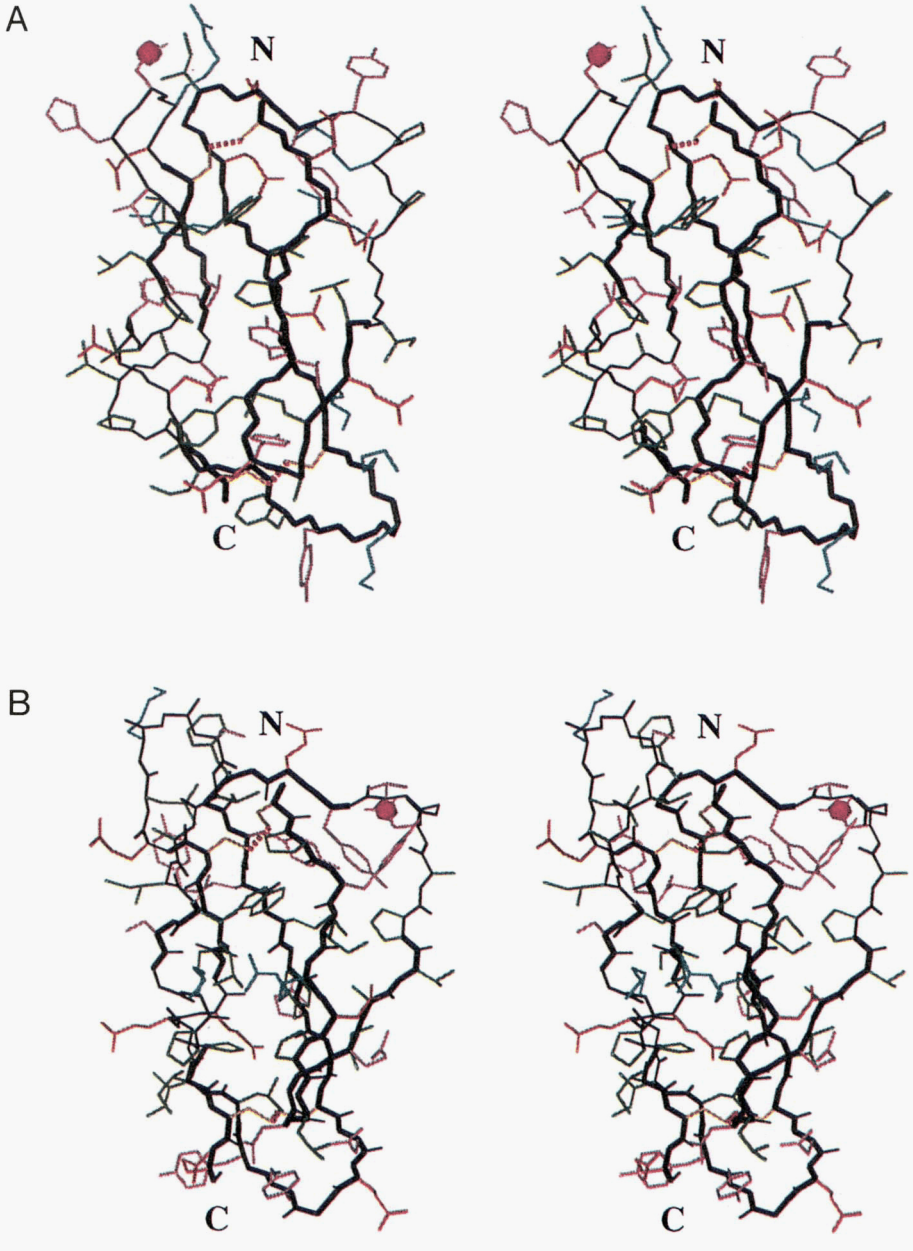

Fig. 1. Stereo view of the 3D models for the CD46 modules CCP-I (**A**) and -II (**B**). The amino acids are color coded: positively charged residues are blue (K, R), negatively charged residues are red (E, D), polar residues are pale magenta (S, T, Y, H, N, Q), and nonpolar residues are green (G, L, I, A, P, V, F, M, W) or yellow (C). The N-glycosylation sites, N49 and N80, are marked in magenta. The backbone regions of CCP-I and -II, which are assumed to be structurally similar to corresponding regions in the 15th and 16th module of factor H, are drawn with thick lines. This drawing, Figure 3, and Figure 5 were prepared with the program MOLMOL (Koradi et al., 1996).

180°]), with two exceptions at −39° and 32°. Only models M1 and M3 have twist angles in this most populated region.

To plan mutagenesis experiments (see below), we tentatively selected M1 as model structure. In this structure, Y61, which is part of the linker peptide between CCP-I and -II, is buried in the interface. Other less than 20% solvent-accessible residues in the interface region are I37, Y34, T64, T82, and E84. With 24% solvent accessibility, F85 may also play a role in the interface stabilization. It is also noteworthy that six charged residues, E8, K32, R62, E63, E84, and K110, are located near the interface, and might be involved in salt bridges. The pairs E8-R62, E84-R62, and E63-K110 are particularly close in the 3D structure.

Only the oligosaccharide situated near the CCP-I–II junction is essential for hemagglutinin binding

Two N-glycosylation motifs (N-X-S/T), one near the tip of the molecule and the other near the CCP-I–II interface in the predicted structure, were selected for mutation analysis. To obtain proteins lacking either one of the oligosaccharide chains, the N residues of the respective motifs were changed to Q. As expected, the apparent mobilities of the two mutants N49Q and N80Q on a denaturing protein gel were slightly increased (data not shown), indicating that both potential N-glycosylation sites are used. The level of cell surface expression of both mutant proteins was three times lower

Table 1. *Clusters of 3D structures obtained from the Monte Carlo simulations of the relative orientations of the CCP domains of factor H and CD46*

Protein[a]	$N_{structures}$[b]	ASA_h[c] (Å^2)	ΔASA_h[d] (Å^2)	N_{first}[e]	$RMSD_c$[f] (Å)
Factor H					
H1	15	4,926	0	32	3.2–4.8
H2	16	4,962	36	12	7.0–8.7
H3	5	5,007	81	3	6.3–7.6
H4	12	5,013	87	17	5.0–8.2
H5	5	5,035	109	18	9.9–10.8
H6	5	5,131	205	7	9.0–10.7
CD46					
M1	33	5,295	0	3	
M2	10	5,299	4	19	
M3	7	5,301	6	4	
M4	9	5,350	55	22	

[a] H1–H6 are the six clusters of 3D structures of factor H, and M1–M4 are the four clusters for CD46.
[b] Number of structures in the cluster.
[c] Lowest solvent-accessible hydrophobic surface area among all structures from this cluster.
[d] Difference to the lowest solvent-accessible hydrophobic surface area of the first cluster.
[e] Number of the first Monte Carlo structure that is contained in the cluster.
[f] RMSD to the two-domain NMR structure of CCP-15 and -16 of factor H.

than that of the standard (Fig. 4, proteins 1 and 2, column Se; the surface expression level of the N80Q mutant was marginally higher than that of the N49Q mutant). Antigenicity of both mutant proteins was similar to that of the standard, as measured by three CCP-I- (20.6, E4.3, and J4/48) and two CCP-II-specific antibodies (M75 and M177) (Fig. 4, MA columns). Thus, the omission of either oligosaccharide chain may slow down, but does not preclude, correct folding and intracellular protein transport.

To assay the function of proteins 1 and 2, we used a cell–cell fusion test based on mixing of cells expressing a CD46 protein variant with cells expressing the two viral proteins necessary for fusion, hemagglutinin and F (Buchholz et al., 1996b). After cocultivation, cells expressing functional CD46 proteins fuse and form syncytia. In the case of the standard CD46 protein, the percentage of nuclei in syncytia exceeds 50%. Figure 4, column Fu, presents the results of the cell–cell fusion tests: the N49Q mutant protein was almost as efficient as the standard in supporting fusion, whereas the N80Q mutant protein was considerably weaker (note that the fusion test is very efficient, and even with a moderate cell surface expression, a high percentage of nuclei in syncytia can be obtained). A protein with both N residues mutated to Q reached the cell surface inefficiently and completely lost fusion-support function (data not shown).

We then measured binding of a soluble form of MV hemagglutinin (sH, Devaux et al., 1996) to mutant and standard CD46 proteins. After normalization for surface expression, the N49Q protein was as efficient as the standard protein in binding (Fig. 4, column sHb), whereas binding of the N80Q protein was below detection level (10% of the standard). The fact that a weak activity of this protein was detected in the fusion assay is due to the higher sensitivity of this test.

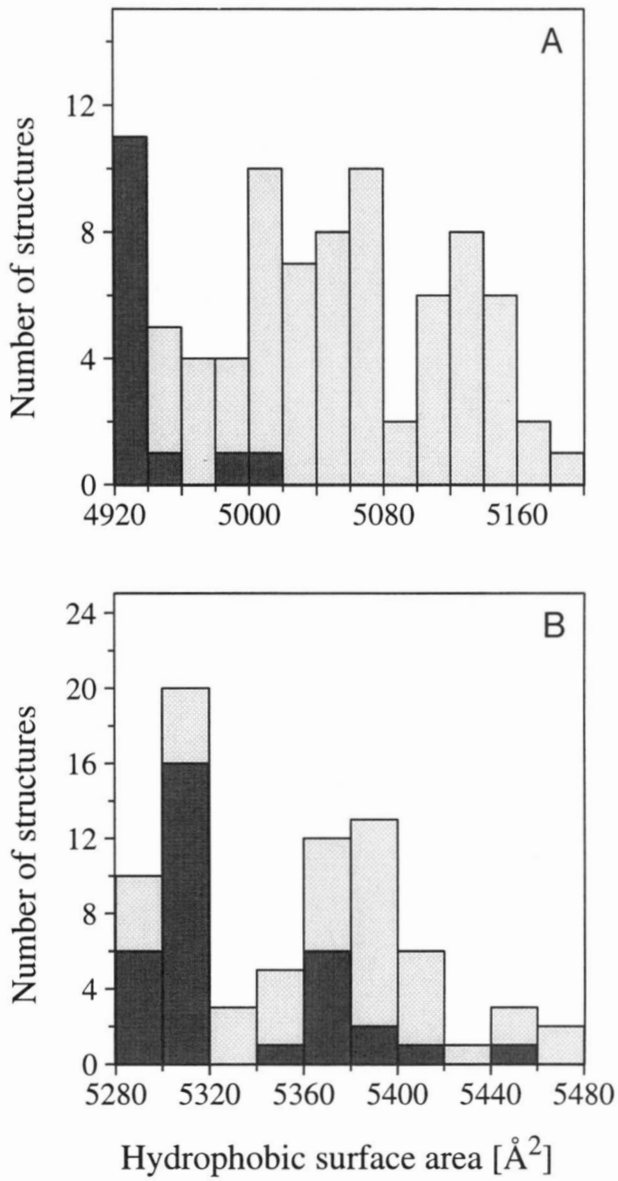

Fig. 2. Hydrophobic accessible surface area (ASA_h) of the structures calculated by Monte Carlo simulations. **A:** Distribution of ASA_h in the 84 structures obtained from the simulation with the factor H protein. Structures contained in the correct cluster, with a backbone RMSD of <3.5 Å to the NMR structure of factor H (Barlow et al., 1993), are highlighted in dark grey. **B:** Distribution of ASA_h in the 75 structures obtained from the simulation with CD46. Structures belonging to the largest cluster (M1) are shown in dark grey.

These results indicate that the oligosaccharide chain predicted to be near the CCP-I–II interface is important for binding and fusion-support function, whereas the one predicted to be at the membrane-distal protein end is not. Recently, the conclusion that the N-glycan of the CCP-II region is essential for the attachment of MV to CD46 has been reached independently by Maisner et al. (1996). Our data additionally demonstrate that the antigenicity of the modified proteins is maintained.

Protein segments essential for function

To verify if a systematic approach based on the exchange of progressively shorter protein segments could be used to narrow down

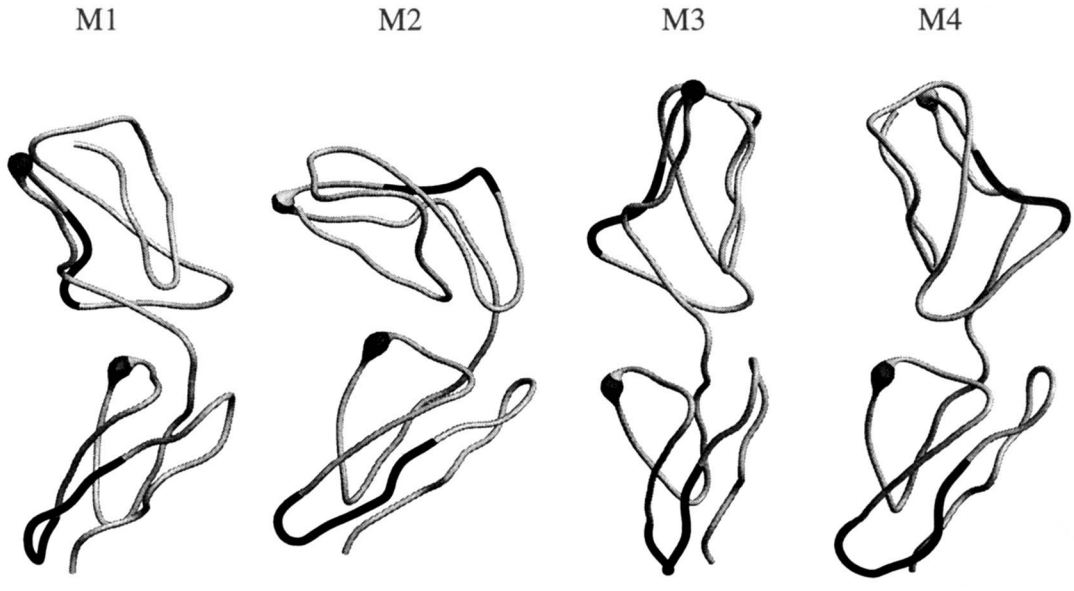

Fig. 3. Four possible orientations of CD46 CCP-I and -II modules as found by Monte Carlo simulations. Two protein segments (residues 39–44 and 94–105) that are important for soluble hemagglutinin binding are shown in black. The positions of the N-glycosylation sites, N49 and N80, are shown with grey spheres.

the MV binding regions in CCP-I and -II, we substituted the first half of CCP-I or the second half of CCP-II by the corresponding part of the other module (Fig. 4, proteins 3 and 4; the first and last residues of the exchanged segments were always conserved Cs).

Protein 4 reached the cell surface with an efficiency similar to the standard (Fig. 4, first column). This protein, as expected, bound three anti-CCP-I antibodies, but not two anti-CCP-II antibodies. It did not permit sH binding and MV-induced cell fusion. Because protein 3 and two other proteins with half-CCP substitutions were not expressed on the cell surface, we concluded that hybrid proteins with large exchanged segments may often fold poorly and thus reach the cell surface inefficiently.

Consequently, we constructed two protein hybrids with short exchange segments. First, we selected a 12-amino acid segment in CCP-II that was part of the long stretch exchanged in protein 4. These residues (94–105) are situated on the same face of CCP-II as the oligosaccharide attached to N80 (Fig. 3, M1–M4), but in a more membrane-proximal position. Protein 5 (Fig. 4) had the same functional characteristics as protein 4: it bound CCP-I antibodies, but not CCP-II antibodies, and neither bound sH nor supported fusion. These results indicate that, even in the context of an antigenically intact CCP-I module, a relatively small alteration of CCP-II can abolish MV binding.

The last hybrid was designed to test the four models of Figure 3. A short, structurally prominent six-amino acid loop (PLATHT, residues 39–44) situated below the CCP-I oligosaccharide was selected. This loop is predicted to be close to the CCP-II oligosaccharide in model M1, opposite to it in model M2, and at an intermediate distance in models M3 and M4. These residues were substituted by the corresponding six amino acids (GEKDSV) of the related protein DAF. The corresponding protein (6, Fig. 4) reached the cell surface with intermediate efficiency, lost, as expected, part of its antigenicity in CCP-I, did not completely maintain the CCP-II antigenicity, and completely lost sH binding and fusion-support function. These results are best accounted by model M1 (Fig. 5).

Discussion

Hemagglutinin binding and module orientation

Our data show that the omission of the CCP-II oligosaccharide chain, or the exchange of residues 39–44 in CCP-I or 94–105 in CCP-II result in a strong negative effect on MV hemagglutinin recognition. Are these three structural landmarks involved in MV binding directly? Because MV receptors with different oligosaccharide chains are functional (Maisner & Herrler, 1995), we favor the hypothesis that the N-glycan of CCP-II has an indirect role in stabilizing the protein fold, as observed using NMR for the oligosaccharide chain of the Ig-like protein CD2 (Wyss et al., 1995). Another CCP-II-altered protein with an exchange of residues 94–105, did not support MV binding and lost antigenicity to both CCP-II antibodies. This could be due to binding of both antibodies to the sH binding region or to a nonlocal change of the CCP-II conformation. The fact that both CCP-II antibodies inhibit virus infection (Iwata et al., 1995) is consistent with the first hypothesis.

The CCP-I mutant with an exchange of residues 39–44 retained part of its antigenicity to all CCP-I antibodies, which implies that the conformation of this domain was not altered drastically. Interestingly, this protein lost part of its antigenicity to the CCP-II antibodies. An explanation of this unexpected result becomes evident when model M1 (Fig. 5) is considered: in this structure, the exchanged CCP-I loop is particularly close to the interface, and to the CCP-II oligosaccharide. Thus, displacement of the oligosaccharide, or of neighboring residues buried in the interface (e.g., I37, Y34), may have altered the CCP-II antigenicity. Alternatively, the introduction of three additional charges (two negative, one positive) in this CCP-I mutant may influence the CCP-II antigenicity.

In summary, the model M1 accounts best for the hemagglutinin binding data. Model M2 cannot account for ligand binding because the two relevant amino acid segments would be situated on opposite sides of the structure. Model M3 presents the two segments on one side, but the low and—considering factor H—improbable tilt

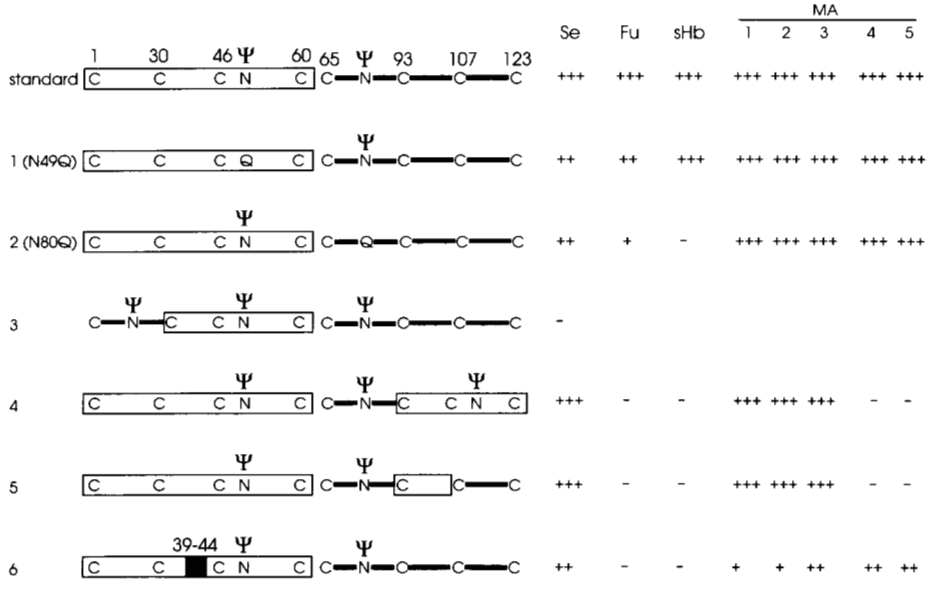

Fig. 4. Primary structure and functional characteristics of six CD46 mutants. A box denotes the amino acid sequence of CCP-I and a thick line indicates the sequence of CCP-II. In proteins 1–5, CCP-I–II are fused to the CD4 immunoglobulin-like domain 4 (Buchholz et al., 1996b); in protein 6, to the other half of the CD46 protein. The cysteine residues (C) and the oligosaccharide chains (Ψ) are indicated. All mutants have an additional oligosaccharide chain on the Ig-like domain 4 (proteins 1–5) or on CCP-IV (protein 6). Mutant proteins 1 and 2 lack a glycosylation site, proteins 3–5 have duplications of segments of CCP-I or -II, and a segment from the related protein DAF is substituted on positions 39–44 (black box) of protein 6. The degree of surface expression (Se), fusion-support activity (Fu), soluble hemagglutinin binding (sHb), and monoclonal antibody binding (MA 1–5) are indicated, the latter two normalized for surface expression. Monoclonal antibodies 20.6, E4.3, and J4/48 are directed against CCP-I (MA 1–3), M75 and M177 against CCP-II (MA 4–5). For proteins 1–5, values are standardized to those of protein I–II/4 (Buchholz et al., 1996b), and for protein 6, to those of the CD46 isoform C2 (Buchholz et al., 1996a). The number of plus signs indicates the level: +++ corresponds to 71–130% of the standard level; ++, 31–70%; +, 10–30%; −, less than 10%.

angle between CCP-I and -II spreads them out in space. Model M4 is not considered because of the higher hydrophobic solvent-accessible surface. Model M1 presents the two segments on the concave side of the CCP-I–II interface and, at the current state of knowledge, represents the experimental observations in the most consistent way.

Experimental data for the relative orientation of two CCP modules are available only for the 15th and 16th CCPs of factor H (Barlow et al., 1993), but recently a model of a CD21 CCP pair was presented (Molina et al., 1995). In this study, two separate binding regions to the complement factor iC3b were defined on the CD21 CCP-I and -II modules. To empirically form a contiguous binding surface for iC3b through the two binding sites, the first CD21 CCP had to be rotated by approximately 130° about the long axis of the second CCP domain, as compared to the relative orientation of both modules of factor H.

The approach used here to model the relative orientation of two CCPs is distinctly different because it is based on energetic parameters. It should be broadly applicable to predict the relative orientation of protein modules with known structure linked by peptides of unknown conformation.

Consequences of the 3D model for further mutational analysis

The 3D model can be used to map the hemagglutinin binding site more precisely. Our current mutation strategy involves substitutions of residues on the CD46 surface, including not only charged amino acids, but also some solvent-exposed hydrophobic residues. Indeed, if the expected regular inside/outside pattern of solvent-exposed β-sheets is fulfilled, some hydrophobic residues are predicted to be on the protein surface in model M1, for example, I45, L53, I92, and I104. They are all situated in β-sheets containing buried tryptophan or cysteine residues. The high number of solvent-exposed hydrophobic side chains reflects the fact that 44% of the residues of the CD46 sequence are hydrophobic, i.e., A, V, L, I, M, F, W, P or C, which is a higher fraction of hydrophobic residues than in factor H, with 35%. This suggests that hydrophobic interactions may be involved in CD46-MV hemagglutinin binding or in CD46 clustering following virus binding (Buchholz et al., 1996b).

Cell entry by viruses

Modular proteins have been selected as receptors by several viruses; e.g., the two amino-terminal CCP modules of CD21 are bound by the surface glycoprotein gp350/220 of the Epstein-Barr virus (Martin et al., 1991; Molina et al., 1991; Moore et al., 1991), and the outermost of the four Ig-like CD4 domains is bound by the HIV gp120 coat protein (Harrison et al., 1992; Hendrickson et al., 1992; Moebius et al., 1992). The modular architecture makes CCP proteins highly flexible, as documented by hydrodynamic analysis, transmission electron microscopy (DiScipio et al., 1992), and NMR spectroscopy (Barlow et al., 1993). Moreover, the crystals of the four Ig-like CD4 domains are not well ordered, limiting the determination of the 3D structure of CD4 to two-domain fragments. Thus, high flexibility might be a general requirement for virus receptors to support viral protein configuration changes that occur during cell entry.

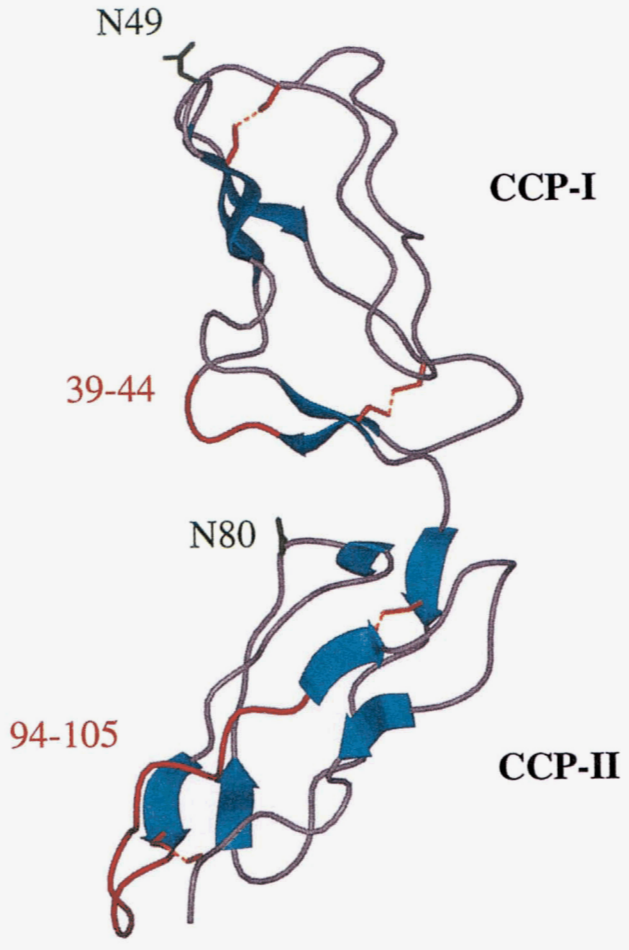

Fig. 5. Ribbon representation of CD46 model M1. Blue arrows indicate β-sheet regions. Side chains of cysteines are shown in orange and the side chains of asparagines at the N-glycosylation sites are green. Segments 39–44 and 94–105, important for hemagglutinin binding, are shown in red.

The gp120 binding site on CD4 has been mapped by site-directed mutagenesis studies and analysis of the 3D X-ray structure to a surface region of about 900 Å^2 near the C′-C″ ridge of the N-terminal domain of CD4 (Moebius et al., 1992). Although there is no similar detailed description of the virus binding sites for any CCP bearing receptor, our 3D model and the binding studies of the CD46 mutants in this paper indicate a quite distinct binding pattern. Whereas the HIV binding site is near the membrane-distal tip of the CD4 receptor protein, the MV hemagglutinin binding site might be located on the concave face formed by the CCP-I and -II modules.

Conclusions and perspectives

Our modeling approach integrates computational tools, including distance geometry, energy minimization, and Monte Carlo simulation with site-directed mutagenesis studies. This combination of methods should be broadly applicable to predict the relative orientation of other modules (Bork et al., 1996). The atomic coordinates of the four models, M1–M4, will be made generally available to allow testing of the accuracy of the modeling procedure once an experimental high-resolution X-ray or NMR structure is available. Currently, the 3D model is used to rationally plan mutagenesis experiments to map the hemagglutinin binding site more precisely. In addition, we have produced a set of peptides overlapping the two protein segments defined here, and are determining their ability to inhibit virus attachment to cells and replication. These peptides could be lead substances for the development of antiviral compounds controlling MV infection.

Materials and methods

Modeling of the individual CCP-I and -II modules

The 3D structure of the 15th and 16th modules of human factor H, entry 1hfh (Barlow et al., 1993) was obtained from the Brookhaven Protein Data Bank (Bernstein et al., 1977). The sequences of CD46 CCP-I and -II were aligned manually to the factor H sequence according to Figure 6. For the regions with reasonable similarity, upper and lower distance constraints between all Cα and Cβ atoms were extracted from the factor H structure with a tolerance of ± 1 Å. In the β-sheet regions, backbone angles were restrained to $-150° < \phi < -90°$ and $90° < \psi < 150°$. ϕ Angles of all other residues, with the exception of G, N, D, S, C, A, H and K, were restrained to $-180° < \phi < 0°$, following a recent statis-

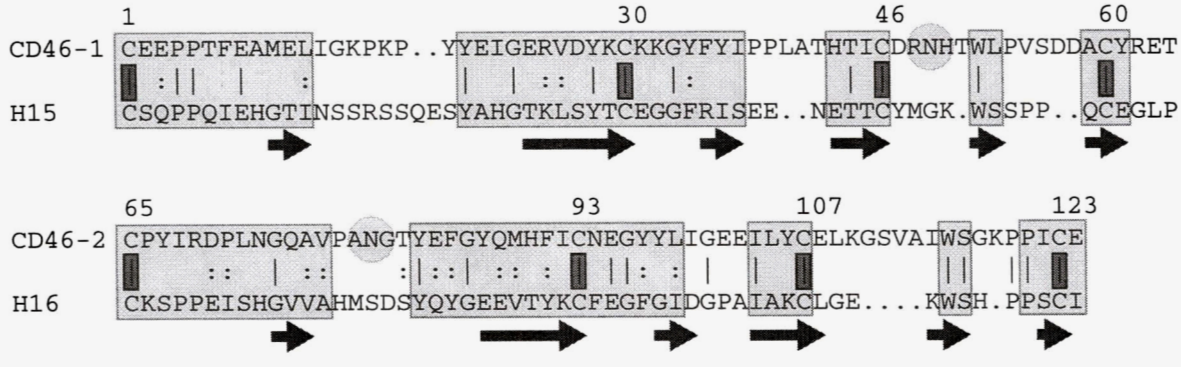

Fig. 6. Sequence alignment of the first two CD46 modules to the 15th and 16th factor H modules. Regions highlighted in light grey were considered to be structurally conserved. Aligned cysteines are joined by dark grey lines. Arrows below the sequences indicate regions where β-strand angle constraints were used in the modeling procedure. The two asparagine residues to which an oligosaccharide is attached are highlighted with a light grey circle.

tical study (Kamimura & Takahashi, 1994). Additionally, constraints for hydrogen bonds between the β-sheets and three other hydrogen bonds conserved in both modules of factor H were generated.

These data were used by the program DIAMOD, a modified version of the distance geometry program DIANA (Güntert et al., 1991), to generate model structures of the individual modules. The DIAMOD structure with lowest target function value was further refined by energy minimization using the program FANTOM (Schaumann et al., 1990; von Freyberg & Braun, 1993). The ECEPP/2 force field (Némethy et al., 1983) was extended by a term containing the solvent–protein interaction (Mumenthaler & Braun, 1995a). This hydrophobic interaction term was weighted highly for the first 600 minimization steps and weighted normally for the following 600 steps.

Determination of the relative orientation of two CCP modules

The ϕ and ψ backbone dihedral angles of the linker tetrapeptide were sampled by Monte Carlo simulations. Metropolis Monte Carlo simulations with 100 trial steps at constant temperature (T = 10,000 K) were performed with FANTOM. Initially, the backbone angles of the linker tetrapeptide between the two modules were set to an extended conformation ($\phi = -120°$ and $\psi = 120°$). All atoms within the two modules were kept fixed through upper and lower limit distance constraints. In every step, only one angle was changed to a random value, followed by 20 steps of Newton–Raphson minimization.

This procedure was first tested for factor H, where the relative orientation of two CCP modules is known. The NMR structure of factor H was regularized to standard ECEPP/2 geometry by extracting distance and angle constraints from the NMR structure. The RMSD of the regularized structure to the NMR structure was 1.1 Å for backbone atoms. In preliminary calculations, we used the full ECEPP/2 force field in vacuum, but the correct orientation was not found. In the absence of solvation terms, charged amino acids on the surface tended to form salt bridges with each other leading to large negative energies. To reduce such artifacts, the electrostatic energy term was omitted in all following calculations.

All structures reaching an energy below 5,000 kcal/mol were stored and subsequently minimized to a local minimum. The resulting structures were classified into different clusters, if the RMSD for backbone atoms was more than 4 Å (Mumenthaler & Braun, 1995b).

Viruses, antibodies, soluble MV hemagglutinin

The vaccinia virus vTF7-3, coding for the bacteriophage T7 RNA polymerase (Fuerst et al., 1986), and the VV-F/H virus, coding for the MV F and H proteins (Wild et al., 1992), were grown as described. The anti-CD46 monoclonal antibody 20.6 was described by Naniche et al. (1993), E4.3 and J4/48 were obtained from Serotec (Oxford, England), and M75 and M177 were kindly donated by Dr. T. Seya (Seya et al., 1990). Soluble MV hemagglutinin, purified as described by Devaux et al. (1996), was kindly donated by D. Gerlier.

Generation of plasmids coding for mutant CD46 proteins

The plasmid pI-II/4, producing the CD46–CD4 hybrid protein CCP-I–II/Ig4 (Buchholz et al., 1996b), was used to construct the plasmids coding for proteins 1–5. For the construction of plasmid pN49Q coding for protein 1, two PCR fragments were amplified using external primer A (CAGTACATCTACGTATTAGT, primers are always shown 5′-3′) and B (ATTTAGGTGACACTATAG) binding upstream and downstream of the coding region, respectively. Primer A was used in combination with primer CD46I$_{(+Pvu\ I)}$ (ATGATTC*CGATCG*CAAATAGT), and primer B with primer CD46I$_{(N\rightarrow Q,+Pvu\ I)}$ (TATTTG*CGATCG*GCAACATACATG, mutated residues are underlined) introducing the N49Q mutation and additionally creating a new, unique *Pvu* I restriction-site (italics) by a silent point mutation. After gel purification, the fragments were digested with *Hin*d III/*Pvu* I and *Pvu* I/*Xba* I and cloned into a *Hin*d III/*Xba* I digested pRc-CMV (Stratagene), in a three-way ligation.

For the construction of plasmid pN80Q coding for protein 2, only one fragment was amplified using primer B and the mutagenic primer CD46II$_{(N\rightarrow Q)}$ (TAAA*TGGCCA*AGCAGTCCCTGCA CAAGGGACTTACG), covering besides the N80 codon a *Msc* I site (italics), unique in the coding region. The PCR fragment was digested with *Msc* I/*Xba* I and ligated into pRc-CMV together with a *Hin*d III/*Msc* I fragment, obtained from restriction of pI-II/4.

Sequence duplications in constructs 3–5 were produced by overlap extension PCR (Ho et al., 1989). This strategy includes two rounds of PCR. In a first round, two fragments were produced. First, an upstream fragment, coding for the N-terminal part of the protein, using primer A in combination with a mutagenic primer. Second, a downstream fragment, coding for the C-terminal part, using primer B with another mutagenic primer. The primer combinations were: for constructs 3 and 4, primer CD46II;I(−) (CCTTTTTTACAAATAAAGTGCATCTGATAAC) with primer A; and primer CD46I;II(+) (GCACTTTATTTGTAAAAAAGGATA CTTCTA) with primer B. For construct 5, again primer CD46II;I(−) with primer A; and primer CD46II;I(+) (TAGATTATAAGTGT AATGAGGGTTATTACT) with primer B. pI-II/4 served as template for the amplification of the upstream fragments of constructs 4 and 5 and pII/4 (Buchholz et al., 1996b) for that of construct 3. For the amplification of the downstream fragments, the following templates were used: pI-II/4 for construct 3; pI/4 (Buchholz et al., 1996b) for construct 4; and pI;II/4, an intermediate construct harboring the C-terminal transition site from CCP-I to -II, for construct 5. Each fragment contained a stretch of complementary sequence, defining the transition site from CCP-I to -II, and the actual overlap used in the extension PCR connecting the upstream with the downstream fragment. The second round of PCR was based on the formation of the overlap in a fraction of the molecules. To reduce or eliminate background amplification products, the overlap extension product was amplified using a second pair of external primers, primer C (TCAAGCTTAATATGGAGCCTCC CGG) and D (GATCTAGATCAAATGGGGCTACATG), binding to the 5′ and the 3′ end of the coding region, respectively. After gel purification, the products were *Hin*d III/*Xba* I digested and cloned into the pRc-CMV vector in a two-way ligation.

All PCR reactions were conducted on the Stratagene robocycler, using the *Pyrococcus furiosus* DNA polymerase (Stratagene), harboring a 3′ → 5′ proofreading activity. In general, PCR reactions used 20 ng of template DNA, 50 pmol of each primer, 20 mmol dNTPs, and 2.5 U Pfu DNA polymerase in a final volume of 100 μL. For the overlap extension PCR, the reaction mixes were as following: 1 ng of each product from the first round of PCR, 100 pmol each of primer C and D, 20 mmol dNTPs, and 2.5 U Pfu DNA polymerase in a final volume of 100 μL. Thirty cycles of

denaturation (50 s, 94 °C), annealing (80 s, 43 °C or 45 °C in overlap extension PCR), and elongation (90 s, 72 °C) were performed. The nucleotide sequence corresponding to the CD46 part of the hybrid proteins was confirmed.

The plasmid encoding protein 6 was constructed by site-directed mutagenesis of plasmid pRc/CMV-CD46, which contains a cDNA coding for the C2 isoform of CD46 (clone 63 in Buchholz et al., 1996a). The site-directed mutagenesis system was based on double-stranded DNA (Chameleon kit, Stratagene) using a mutagenic primer (GGATACTTCTATATACCT**GGTGAAAAGGACTCTGTT**ATTT GTGATCGGAATCAT; mismatches in bold letters) and a selection primer (CAGATA*TACGTA*TTGACATTGAT; an *Sna*B 1 site (italics) was introduced to substitute a unique *Mlu* 1 site).

Transient expression of mutant proteins

Transient expression of standard and mutant CD46 proteins was in mouse Ltk$^-$ cells with the vTF7-3 system (Buchholz et al., 1996a). Protein analysis was as described (Buchholz et al., 1996b), by western blotting, fluorescence-activated cell sorter analysis using the appropriate antibodies, or by the cell–cell fusion assay. Binding assays of mutant CD46 proteins to soluble hemagglutinin were according to Devaux et al. (1996).

Acknowledgments

We acknowledge financial support to C.M. by the ETHZ. This work was supported by the Schweizerische Nationalfonds (grants 31-29343.90, 31-33746.92, and 31-45900.95) and the National Science Foundation (BIR-9632326). We thank P. Devaux, D. Gerlier, and T. Seya for reagents, C.H. Schein, H. Weber, and D. Gerlier for helpful comments on the manuscript, and M.A. Billeter for continuous support. The use of the computing facilities of the IPS of the ETHZ is gratefully acknowledged.

References

Barlow PN, Baron M, Norman DG, Day AJ, Willis AC, Sim RB, Campbell ID. 1991. Secondary structure of a complement control protein module by two-dimensional ^{1}H NMR. *Biochemistry 30*:997–1004.

Barlow PN, Norman DG, Steinkasserer A, Horne TJ, Pearce J, Driscoll PC, Sim RB, Campbell ID. 1992. Solution structure of the fifth repeat of factor H: A second example of the complement control protein module. *Biochemistry 31*:3626–3634.

Barlow PN, Steinkasserer A, Norman DG, Kieffer B, Wiles AP, Sim RB, Campbell ID. 1993. Solution structure of a pair of complement modules by nuclear magnetic resonance. *J Mol Biol 232*:268–284.

Bergelson JM, Chan M, Solomon K, St. John NF, Lin H, Finberg RW. 1994. Decay-accelerating factor, a glycosylphosphatidylinositol-anchored complement regulatory protein, is a receptor for several echoviruses. *Proc Natl Acad Sci USA 91*:6245–6248.

Bergelson JM, Mohanty JG, Crowell RL, St. John NF, Lublin DM, Finberg RW. 1995. Coxsackievirus 3B adapted to growth in RD cells binds to decay-accelerating factor (CD55). *J Virol 69*:1903–1906.

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.

Bork P, Downing AK, Kieffer B, Campbell ID. 1996. Structure and distribution of modules in extracellular proteins. *Quart Rev Biophys 29*:119–167.

Buchholz CJ, Gerlier D, Hu A, Cathomen T, Liszewski MK, Atkinson JP, Cattaneo R. 1996a. Selective expression of a subset of measles virus receptor competent CD46 isoforms in human brain. *Virology 217*:349–355.

Buchholz CJ, Schneider U, Devaux P, Gerlier D, Cattaneo R. 1996b. Cell entry by measles virus: Long hybrid receptors uncouple binding from membrane fusion. *J Virol 70*:3716–3723.

Clarkson NA, Kaufman R, Lublin DM, Ward T, Pipkin PA, Minor PD, Evans DJ, Almond JW. 1995. Characterization of the echovirus 7 receptor: Domains of CD55 critical for virus binding. *J Virol 69*:5497–5501.

Devaux P, Loveland B, Christiansen D, Milland J, Gerlier D. 1996. Interactions between the ectodomains of haemagglutinin and CD46 as a primary step in measles virus entry. *J Gen Virol 77*:1477–1481.

DiScipio RG. 1992. Ultrastructures and interactions of complement factors H and I. *J Immunol 149*:2592–2599.

Dörig RE, Marcil A, Chopra A, Richardson CD. 1993. The human CD46 molecule is a receptor for measles virus (Edmonston strain). *Cell 75*:295–305.

Fuerst TR, Niles EG, Studier FW, Moss B. 1986. Eukaryotic transient expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc Natl Acad Sci USA 83*:8122–8126.

Güntert P, Braun W, Wüthrich K. 1991. Efficient computation of three-dimensional protein structures in solution from NMR data using the program DIANA and the supporting programs CALIBA, HABAS, and GLOMSA. *J Mol Biol 217*:517–530.

Hänggi G, Braun W. 1994. Pattern recognition and self-correcting distance geometry calculations applied to myohemerythrin. *FEBS Lett 344*:147–153.

Harrison SC, Wang J, Yan Y, Garret T, Liu J, Moebius U, Reinherz, E. 1992. Structure and interactions of CD4. *Cold Spring Harbor Symp Quant Biol 57*:541–548.

Hendrickson WA, Kwong PD, Leahy DJ, Ryu SE, Yamaguchi H, Fleury S, Sekaly RP. 1992. Structural aspects of CD4 and CD8 involvement in the cellular immune response. *Cold Spring Harbor Symp Quant Biol 57*:549–556.

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site-directed mutagenesis by overlap-extension using the polymerase chain reaction. *Gene 77*:51–59.

Iwata K, Seya T, Yanagi Y, Pesando JM, Johnson PM, Okabe M, Ueda S, Ariga H, Nagasawa S. 1995. Diversity of sites for measles virus binding and for inactivation of complement C3b and C4b on membrane cofactor protein CD46. *J Biol Chem 270*:15148–15152.

Janatova J, Reid KBM, Willis AC. 1989. Disulfide bonds are localized within the short consensus repeat units of complement regulatory proteins: C4b-binding protein. *Biochemistry 28*:4754–4761.

Kamimura M, Takahashi Y. 1994. Phi-psi conformational pattern clustering of protein amino acid residues using the potential function method. *CABIOS 10*:163–169.

Koradi R, Billeter M, Wüthrich K. 1996. MOLMOL: A program for display and analysis of macromolecular structures. *J Mol Graphics 14*:51–55.

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.

Liszewski MK, Post TW, Atkinson JP. 1991. Membrane cofactor protein (MCP or CD46): Newest member of the regulators of complement activation gene cluster. *Annu Rev Immunol 9*:431–455.

Maisner A, Alvarez J, Liszewski MK, Atkinson DJ, Atkinson JP, Herrler G. 1996. The N-glycan of the SCR 2 region is essential for membrane cofactor protein (CD46) to function as a measles virus receptor. *J Virol 70*:4973–4977.

Maisner A, Herrler G. 1995. Membrane cofactor protein with different types of N-glycans can serve as measles virus receptors. *Virology 210*:479–481.

Maisner A, Schneider-Schaulies J, Liszewski MK, Atkinson JP, Herrler G. 1994. Binding of measles virus to membrane cofactor protein (CD46): Importance of disulfide bonds and N-glycans for the receptor function. *J Virol 68*:6299–6304.

Manchester M, Valsamakis A, Kaufman R, Liszewski MK, Alvarez J, Atkinson JP, Lublin DM, Oldstone MBA. 1995. Measles virus and C3 binding sites are distinct on membrane cofactor protein (CD46). *Proc Natl Acad Sci USA 92*:2303–2307.

Martin DR, Yuryev A, Kalli KR, Fearon DT, Ahearn JM. 1991. Determination of the structural basis for selective binding of Epstein-Barr virus to human complement receptor type 2. *J Exp Med 174*:1299–1311.

Miller S, Janin J, Lesk AM, Chothia C. 1987. Interior and surface of monomeric proteins. *J Mol Biol 196*:641–656.

Moebius U, Clayton LK, Abraham S, Harrison SC, Reinherz EL. 1992. The human immunodeficiency virus gp120 binding site on CD4: Delineation by quantitative equilibrium and kinetic binding studies of mutants in conjunction with a high-resolution CD4 atomic structure. *J Exp Med 176*:507–517.

Molina H, Brenner C, Jacobi S, Gorka J, Carel JC, Kinoshita T, Holers VM. 1991. Analysis of Epstein-Barr virus-binding site on complement receptor 2 (CR2/CD21) using human-mouse chimeras and peptides. *J Biol Chem 266*:121173–12179.

Molina H, Perkins SJ, Guthridge J, Gorka J, Kinoshita T, Holers VM. 1995. Characterization of a complement receptor 2 (CR2, CD21) ligand binding site for C3. An initial model of ligand interaction with two linked short consensus repeat modules. *J Immunol 154*:5426–5435.

Moore MD, Cannon MJ, Sewall A, Finlayson M, Okimoto M, Nemerow GR. 1991. Inhibition of Epstein-Barr virus infection in vitro and in vivo by soluble CR2 (CD21) containing two shorts consensus repeats. *J Virol 65*:3559–3565.

Mumenthaler C, Braun W. 1995a. Folding of globular proteins by energy minimization and Monte Carlo simulation with hydrophobic surface area potentials. *J Mol Mod 1*:1–10.

Mumenthaler C, Braun W. 1995b. Predicting the helix packing of globular proteins by self-correcting distance geometry. *Protein Sci 4*:863–871.

Naniche D, Varior-Krishnan G, Cervoni F, Wild TF, Rossi B, Rabourdin-Combe C, Gerlier D. 1993. Human membrane cofactor protein (CD46) acts as a cellular receptor for measles virus. *J Virol 67*:6025–6032.

Némethy G, Pottle MS, Scheraga HA. 1983. Energy parameters in polypeptides. 9. Updating of geometrical parameters, nonbonded interactions and hydrogen bond interactions for the naturally occurring amino acids. *J Phys Chem 87*:1883–1887.

Nussbaum O, Broder CC, Moss B, Stern LB, Rozenblatt S, Berger EA. 1995. Functional and structural interactions between measles virus hemagglutinin and CD46. *J Virol 69*:3341–3349.

Schaumann T, Braun W, Wüthrich K. 1990. The program FANTOM for energy refinement of polypeptides and proteins using a Newton-Raphson minimizer in torsion angle space. *Biopolymers 29*:679–694.

Seya T, Hara T, Matsumoto M, Akedo H. 1990. Quantitative analysis of membrane cofactor protein (MCP) of complement. High expression of MCP on human leukaemia cell lines, which is down-regulated during cell differentiation. *J Immunol 145*:238–245.

Varior-Krishnan G, Trescol-Biémont MC, Naniche D, Rabourdin-Combe C, Gerlier D. 1994. Glycosyl-phosphatidylinositol-anchored and transmembrane forms of CD46 display similar measles virus receptor properties: Virus binding, fusion, and replication; down-regulation by hemagglutinin; and virus uptake and endocytosis for antigen presentation by major histocompatibility complex class II molecules. *J Virol 68*:7891–7899.

von Freyberg B, Braun W. 1993. Minimization of empirical energy functions in proteins including hydrophobic surface area effects. *J Comp Chem 14*:510–521.

Wild TF, Bernard A, Spehner D, Drillien R. 1992. Construction of vaccinia virus recombinants expressing several measles virus proteins and analysis of their efficacy in vaccination of mice. *J Gen Virol 73*:359–367.

Wyss DF, Choi JS, Li j, Knoppers MH, Willis KJ, Arulanandam RN, Smolyar A, Reinherz EL, Wagner G. 1995. Conformation and function of the N-linked glycan in the adhesion domain of human CD2. *Science 269*:1273–1278.