PROTEINS: Structure, Function, and Genetics 7:227–233 (1990)

Modeling of the Human Intercellular Adhesion Molecule-1, the Human Rhinovirus Major Group Receptor

Vincent L. Giranda, Michael S. Chapman, and Michael G. Rossmann
Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

ABSTRACT **A model has been built of the amino-terminal domain of the intercellular adhesion molecule-1 (ICAM-1), the receptor for most human rhinovirus serotypes. The model was based on sequence and presumed structural homology to immunoglobulin constant domains. It fits well into the putative receptor attachment site, the canyon, on the human rhinovirus-14 (HRV14) surface in a manner consistent with most of the mutational data for ICAM-1 (Staunton, D. E., Dustin, M. L., Erickson, H. P., Springer, T. A. Cell, in press, 1989) and HRV14 (Colonno, R. J., Condra, J. H., Mizutani, S., Callahan, P. L., Davies, M. E., Murcko, M. A. Proc. Natl. Acad. Sci. U.S.A. 85: 5449–5453, 1988).**

Key words: rhinovirus receptor, ICAM-1, structure prediction, immunoglobulin superfamily, docking receptor to virus

INTRODUCTION

Human rhinoviruses are picornaviruses and are one of the major causes of the common cold. Picornaviruses are icosahedral with each icosahedron comprised of 60 protomers. Each protomer is composed of four polypeptides, VP1–VP4 (Fig. 1). There are 100 characterized serotypes of rhinoviruses which have been classified into two groups with respect to their cellular receptor. There are at least 78 human rhinovirus serotypes which bind to the major group receptor and at least 10 serotypes which bind to the minor group receptor.[1–3] Structural studies of both major and minor group viruses have shown that a canyon encircles each icosahedral 5-fold axis.[4,5] The canyon hypothesis proposes that the residues inside the canyon are important in virus to cell binding, and that these residues can escape immunological surveillance because the canyon is sufficiently narrow to exclude immunoglobulins (Fig. 2).[6] This theory is supported by the observation that amino acids in the canyon are more conserved than other surface amino acids.[7] Also, mutation of some canyon floor amino acids results in altered viral to cell binding.[8] Furthermore, antiviral compounds which bind beneath and distort the canyon floor inhibit attachment of human rhinovirus-14 (HRV14) to HeLa cells.[9]

The human intercellular adhesion molecule-1 (ICAM-1) has been shown to be the cellular receptor for the major group of human rhinoviruses.[10,11] The ICAM-1 molecule shows sequence homology with immunoglobulin domains and is thought to be a member of the immunoglobulin superfamily.[12,13] It has five immunoglobulin-like domains, termed D1–D5 (numbered sequentially from the amino end), connected to a transmembrane and a cytosolic domain. Unlike immunoglobulins, ICAM-1 appears to remain mostly monomeric.[11] Immunoglobulin constant domains consist of seven β-strands (βA–βG), forming two β-sheets arranged in a β-sandwich. Strands βB and βF are linked by a disulfide bond.[14]

Mutational analysis of ICAM-1 has indicated that the unglycosylated N-terminal domain (D1) is important in viral–cell binding. The D2 domain is also thought to be important in viral binding, but D2-viral interactions are not as well characterized.[15,35] ICAM-1 cleavage products, which are devoid of all but domains D1 and D2, show near normal binding to cells.[15] Other immunoglobulin superfamily surface antigens utilized as viral receptors include CD4 for human immunodeficiency virus-1[16] and the poliovirus receptor.[17,18]

Structural comparisons of evolutionarily divergent proteins which perform similar functions have demonstrated that three-dimensional structure is conserved in spite of low sequence homology.[19–21] Using this principle, it is possible to construct a model of the N-terminal domain of ICAM-1 based on the presumed structural similarity between it and domains of immunoglobulin superfamily proteins with known structures.

This study was undertaken to determine whether a single N-terminal domain of ICAM-1 could fit into the rhinovirus-14 canyon as would be predicted by the canyon hypothesis. In contrast, the dimeric vari-

Received December 12, 1989; accepted December 27, 1989.
Address reprint requests to Dr. Michael G. Rossmann, Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907.

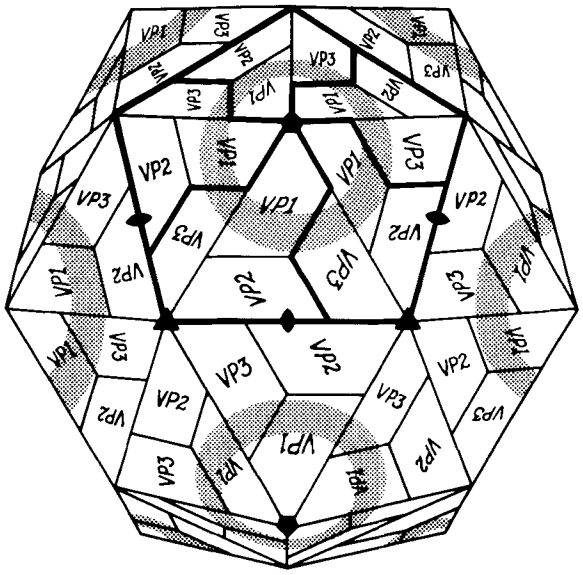

Fig. 1. A schematic diagram of HRV14 showing the icosahedral symmetry, subunit organization, and the canyon position (shaded). VP4 is located on the interior of the viral capsid.

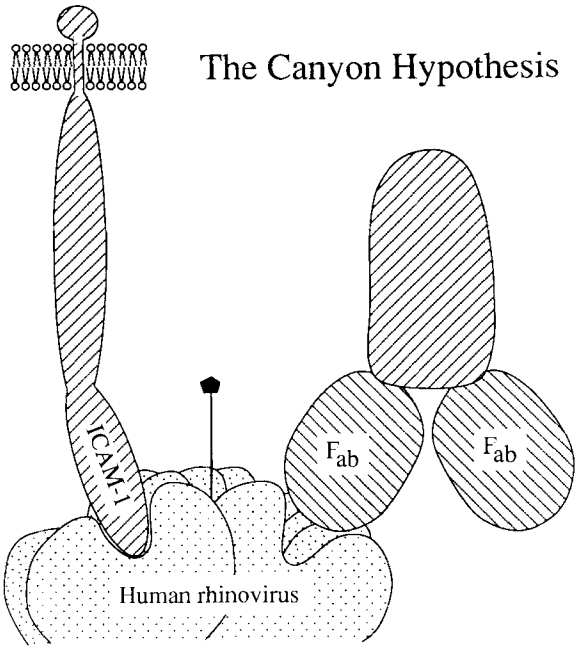

Fig. 2. The canyon hypothesis suggests that the large variable end of an immunoglobulin cannot enter the canyon, while a narrower receptor molecule would be able to enter the canyon and interact with residues on the canyon floor. The shape of the ICAM-1 molecule is loosely based on the observations by Staunton et al.[15]

able domains of an antibody do not fit into the canyon. The model also provides a basis for forming hypotheses about specific interactions between rhinovirus-14 and ICAM-1.

MATERIALS AND METHODS

Alignment of the ICAM-1 Sequence

Several immunoglobulin constant domains were structurally aligned in order to obtain a sequence profile and a consensus structure on which to base the ICAM-1 model. Examples of refined immunoglobulin light (CL) and heavy (C1, C2, C3) constant domains were obtained from the Brookhaven Protein Data Bank.[22] Domains CL and C1 were obtained from entry r1fb4,[23] and domains C2 and C3 were obtained from entry r1fc1.[24] Sequences of the IgG domains were aligned to each other on a structural basis using the computer program HOMO.[25,26] This program looks for topological equivalences between two molecules and, based on these structural equivalences, is able to align their sequences.

The ICAM-1 sequence was aligned to a profile of the four immunoglobulin domains, rather than individual sequences. Profiles provide a broader target, where sequence similarity is slight, and give more importance to regions of high similarity. The alignment was accomplished with a slightly adapted version of the program ProfileGap.[27] The method was modified to facilitate changing of weights. For example, large residue matching weights were used to force each of the cysteines of the immunoglobulin disulfides to align with any of the ICAM-1 cysteines. Furthermore, with knowledge of the three-dimensional structure of the profile, weights could be set to inhibit residue insertions or deletions within secondary structure.

Modeling of the ICAM-1 Amino-Terminal Domain

Modeling was performed using the FRODO computer program.[28] Residues of ICAM-1, which had been aligned with β-strands of the immunoglobulin domains, were built into the immunoglobulin backbone structure. Loop regions were built using the "bones" option of the FRODO program.[29] This option searches a structural database for peptides with the same number of amino acids that have starting and ending positions similar to those required by the model loop. Loops used as templates were chosen based on structural and sequence criteria. Template loops which interfered sterically with the remainder of the structure were first excluded. From those remaining, the loop with sequence most similar to ICAM-1 was chosen. Special consideration was given to those sequences with prolines or glycines at the same loop positions as required by the ICAM-1 alignment. The ICAM-1 N-terminal domain model was idealized using the "refi" option in the FRODO program.[30]

The model was visually docked into the HRV14 canyon to allow for maximum interaction with those residues at the base of the canyon that were shown

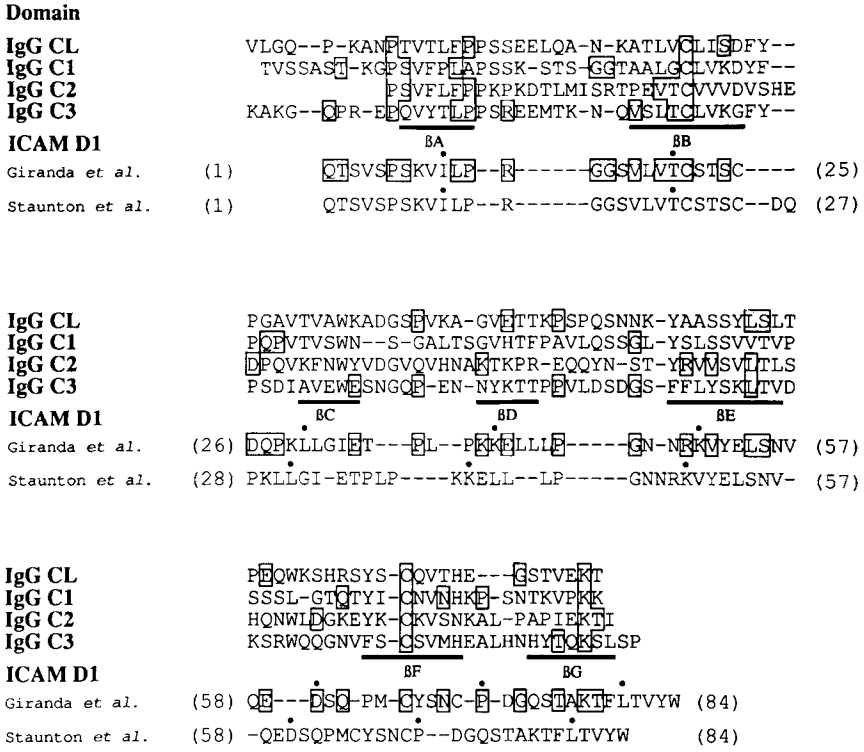

Fig. 3. The sequence alignment used in the formation of the proposed ICAM-1 domain 1 model is compared to the four IgG domains to which it was aligned. Also shown for comparison is the alignment presented by Staunton et al.[13] Boxed residues show where an ICAM-1 residue is aligned to an identical residue in one of the IgG domains. Regions of β-strand in the IgG domains are shown as heavy bars.

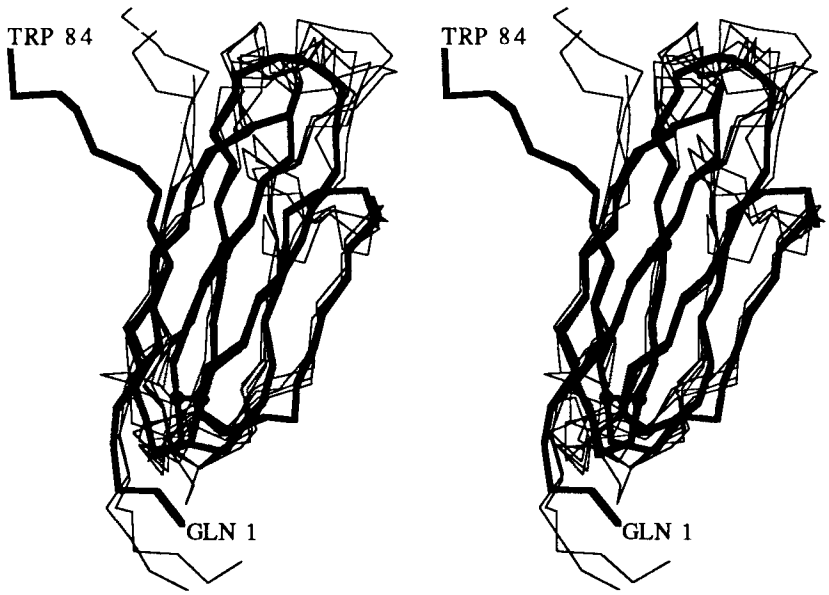

Fig. 4. The C_α backbone structure of the ICAM-1 model (thick) is compared to that of four IgG domains (thin lines).

to be important by mutational analysis[8] and WIN drug binding studies.[9,31,32] The conformations of some amino acid side chains for both the model and the canyon were changed to improve shape and charge complementarity during the docking procedure.

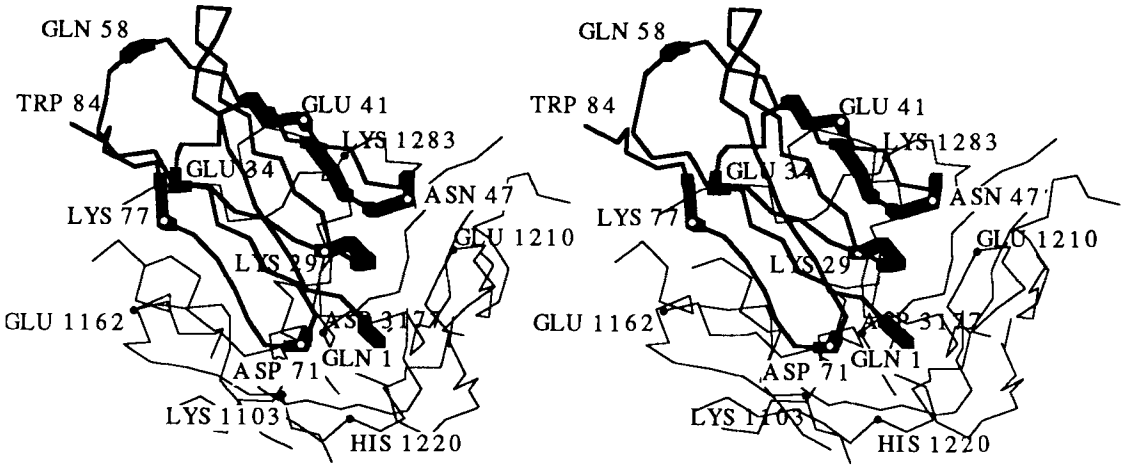

Fig. 5. C_α diagram of the ICAM-1 model docked into the HRV14 canyon shown in cross-section. The virus interior is toward the bottom of the page. The "north" canyon rim (Fig. 6) is to the left and the "south" rim is toward the right. The HRV14 backbone is shown in thin lines, ICAM-1 in medium lines, ICAM-1 mutated regions (Table I) in heavy lines. Open circles (ICAM-1) and filled circles (HRV14) represent specific interacting residues.

RESULTS AND DISCUSSION

Sequence Alignment and Model of the ICAM-1 N-Terminal Domain

An initial sequence alignment between ICAM-1 and the profile showed the weakest sequence homology in the regions of βC and βD. This region also proved to be structurally implausible. This initial sequence alignment predicted an extremely short loop between the βC and βD strands consisting of only two amino acids. This would have required elimination of some regions of β-strand in order to link the amino acids in βC and βD and would also have placed lysine 29* in the hydrophobic interior of the β-sandwich. The final alignment shown here is shifted by two residues in βC and four residues in βD toward the amino terminus of the immunoglobulin sequence profile (an "N-terminal shift"). This does not greatly worsen the alignment, but does lead to a structurally realistic model. No other adjustment was made to the alignment derived from the program ProfileGap.

The final sequence alignment of ICAM-1 with immunoglobulin constant domains is shown in Figure 3. This alignment, based on sequence and structural considerations, allows the construction of a structurally sound model and differs slightly from the alignment of Staunton et al.[13] that was based primarily on sequence considerations. Some of the differences reflect the use of structural information in the sequence alignment. Other differences probably reflect the inherent uncertainties of aligning dissimilar sequences. However, the secondary structural prediction implied by the present results shows reasonable agreement with the secondary structural predictions of Staunton et al.[33]

The alignments, the one presented here and that of Staunton et al.,[13] are identical in strands βA and βB. The latter has a deletion in the βC strand, and requires the carboxylate of Glu-34 to reside inside the β-sandwich. In βC the alignment presented here

TABLE I. ICAM-1 D1 Mutations Important in Binding to HRV14[15]*

Mutation number	Mutation	On ICAM-1/HRV14 interface
1	Q1T/KA	Yes
2	D26QPK/ALPE	Yes
3	E34/A	Yes
4	K39KE/ERQ	Yes
5	G46NN/ASI	Yes
6	R49KV/EKL	No
7	Q58/H	No
8	D71/N	Yes
9	K77T/ES	Yes

*The first column gives the number assigned to a mutation corresponding to Figure 6. The numbering system used in the second column is that used by Staunton et al.[15]

*Residues in the ICAM-1 D1 domain are numbered sequentially from 1–84. Residues in HRV14 are numbered sequentially in each of the viral proteins VP1, VP2, VP3, and VP4 starting at 1000, 2000, 3000, and 4000, respectively.

Fig. 6. "Road map" of HRV14[7] illustrating the predicted binding position of ICAM-1. The triangle represents one asymmetric unit of the icosahedral virus, bounded by a 5-fold vertex at the top and 3-fold vertices on either side of the base (compare with Fig. 1). The surface residues of HRV14 using the one letter amino acid code are shown. Canyon rims are indicated by a heavy line. Residues within 5 Å of the predicted ICAM-1 structure are shaded. Residues which form specific interactions with mutated regions of the ICAM-1 model are darkly shaded. The circled numbers correspond to positions of ICAM-1 mutations (Table I).

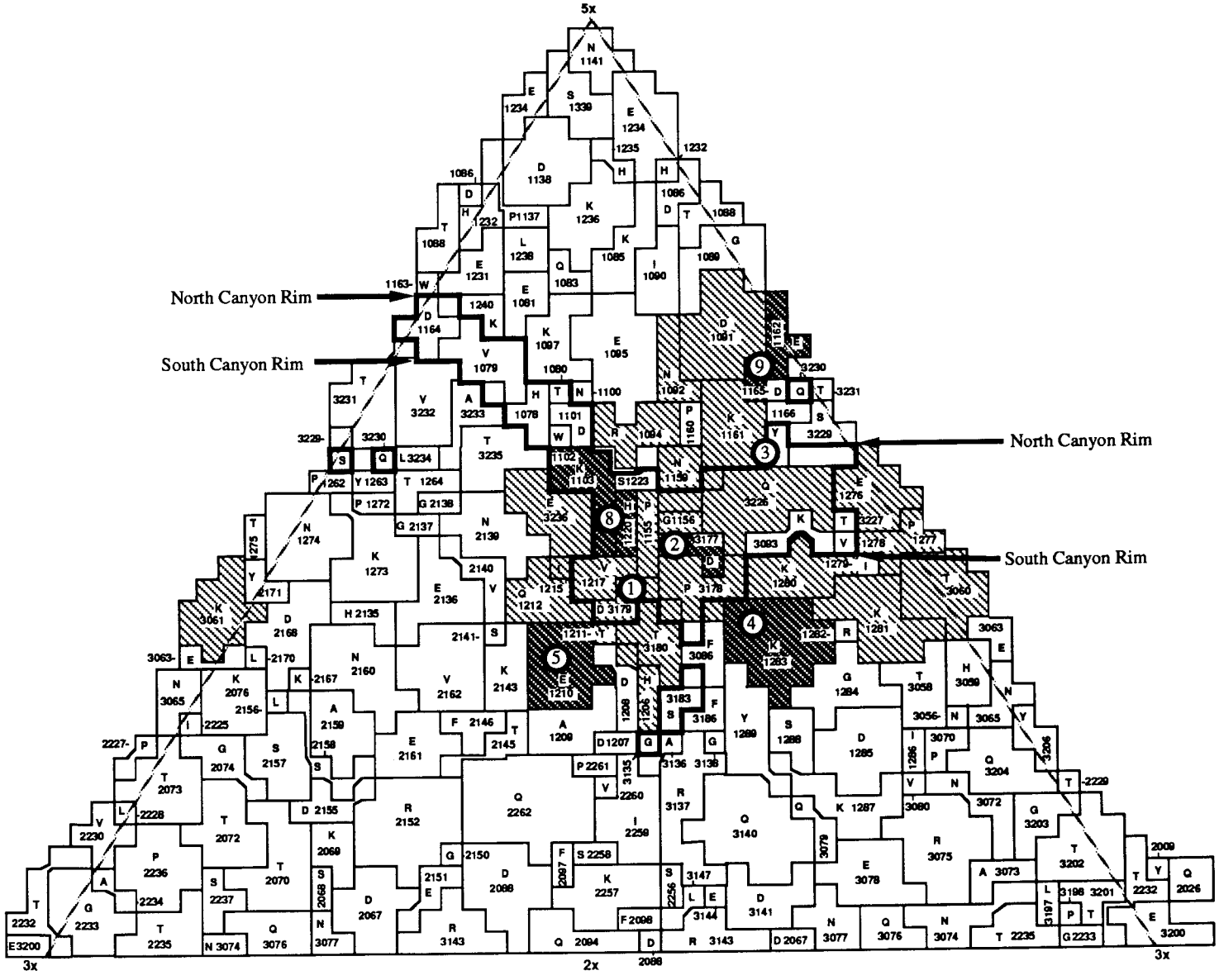
North Canyon Rim
South Canyon Rim
North Canyon Rim
South Canyon Rim
5x
3x
2x
3x

Fig. 6.

is shifted 1–2 residues C-terminally relative to that of Staunton et al.[13]

Because the βD strand lies on the edge of the β-sandwich, residues on either side of the strand are exposed to solvent. Thus, charged residues can be present on either side of the strand. The homology of the βC and βD strands is weak. The current alignment avoids a deletion in βD. Nevertheless, there is considerable uncertainty of alignments in this region. Strands βE and βG are shifted 1 and 4 residues C-terminally with respect to Staunton's alignment. Structural evidence does not support one sequence alignment over the other.

Staunton et al.[13] align Cys-69, in strand βF, to a position corresponding to a conserved disulfide bond. Their alignment puts Asp-71 and Gln-73 into the interior of the β-sandwich and requires a proline to be in the middle of βF. In the alignment presented here it is Cys-65 that is aligned to the conserved disulfide bond. It does not put any large polar or charged side chains into the interior of the β-sandwich, and allows for the formation of an additional disulfide linkage between Cys-25 and Cys-69. This additional disulfide linkage is consistent with results obtained from the alignment of ICAM-1 with ICAM-2.[33]

The C_α positions (Fig. 4) of the ICAM-1 N-terminal domain (D1) model are similar to those of immunoglobulin constant domains in the regions of predicted β-sheet structure. The loops of the ICAM-1 are shorter than the corresponding loops for immunoglobulin constant domains. All of the proline residues in the model are either in loops or at the end of β-strands. Charged residue side chains are in contact with solvent, although Lys-50 is in an atypical conformation.

Docking of ICAM-1 Into the HRV14 Canyon

Model building of ICAM-1 was completed independently of information concerning the shape and charge distrubition of the HRV14 canyon. Thus, the model is not biased by the canyon hypothesis. Two considerations were made in docking the model into the canyon. First, the shapes of the receptor model and the canyon should be complementary. The model should fit into the canyon sufficiently well so that residues on the canyon floor, especially those shown experimentally to be important in rhinovirus binding, are near receptor model residues. Second, the charges between the model and the canyon should be complementary.

The model of ICAM-1 has an ovoid shape approximately twice as long as wide. The model fits well into the deepest part of the canyon (Fig. 5). In this orientation, based solely on steric criteria, the canyon does not constrain rotation of the receptor model about its long axis. This constraint was added by considering charge complementarity. Of several possible orientations about the long axis, the preferred docking arrangement would allow the formation of a number of charged or polar interactions: Gln-27 with Thr-3180, Lys-29 with Asp-3177, Glu-41 with Lys-1283, Asn-47 with Glu-1210, Asp-71 with His-1220 and Lys-1105, and Lys-77 with Glu-1162. The docked model is in good agreement with the ICAM-1 mutational analysis.[15] Of the many mutants tested by Staunton et al.,[15] there were nine (Table I) that altered ICAM-1 to HRV14 binding without showing large conformational changes when probed by monoclonal antibody binding. In the docked ICAM-1 model, 7 of the 9 regions were in contact with residues of the rhinovirus (Table I, Fig. 6).

Some of the observed behavior of rhinovirus canyon mutants[8] and virus bound by WIN compounds[9,32] can be rationalized with the model presented here. Canyon residues His-1220 and Lys-1103 have been implicated in binding by both mutational analysis and drug studies. In the current model His-1220 and Lys-1103 interact with ICAM-1 Asp-71. Thus, changing the His-1220 to less basic residues or perturbing its position on the base of the canyon floor would be expected to decrease viral to cell binding, in accordance with observation.[8] The behavior of the Lys-1103 mutants is complicated and suggests that both charge and steric factors are important. Colonno et al.[8] found that when Pro-1155 was changed to a smaller Gly residue, binding of the virus to its receptor is increased. In the model presented here Pro-1155 forms a shelf near His-1220 and can sterically inhibit the receptor from moving deeper into the canyon, thus binding more tightly. Ser-1223 has been implicated in the binding of ICAM-1, but in this model it is 5 Å from the binding interface.

The proposed three-dimensional model accounts only for the interaction between the N-terminal domain, D1, of the ICAM-1 and the HRV14 canyon. A large region of the canyon is not interacting with the D1 domain and could interact with the D2 domain. Specifically, the "north" rim of the canyon (Fig. 6), near residues Glu-1081 and His-1232, could interact with the D2 domain. It has been suggested that this region may determine specificity between the major and minor group rhinoviruses.[5]

The human rhinovirus receptor, ICAM-1, as modeled is able to fit into the canyon and interact with residues on its floor and is, therefore, consistent with the canyon hypothesis. This model also suggests the importance of specific interactions between rhinovirus of the major receptor group and ICAM-1. Mutational and structural analyses have been initiated to test this model.

ACKNOWLEDGMENTS

We are most grateful to Don Staunton and Tim Springer for access to their data in advance of publication as well as many helpful discussions. Fig-

ures 4 and 5 were drawn with the MACINPLOT program.[34] We thank Sharon Wilder and Helene Prongay for help in preparation of the manuscript. The work was supported by grants from the National Institutes of Health and the Sterling Research Group to MGR.

REFERENCES

1. Colonno, R.J. Cell surface receptors for picornaviruses. BioEssays 5:270–274, 1986.
2. Colonno, R. J., Callahan, P. L., Long, W. J. Isolation of a monoclonal antibody that blocks attachment of the major group of human rhinoviruses. J. Virol 57:7–12, 1986.
3. Tomassini, J. E., Maxson, T. R., Colonno, R. J. Biochemical characterization of a glycoprotein required for rhinovirus attachment. J. Biol. Chem. 264:1656–1662, 1989.
4. Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H.J., Johnson, J. E., Kamer, G., Luo, M., Mosser, A. G., Rueckert, R. R., Sherry B., Vriend, G. Structure of a human common cold virus and functional relationship to other picornaviruses. Nature (London) 317:145–153, 1985.
5. Kim, S., Smith, T. J., Chapman, M. S., Rossmann, M. G., Pevear, D. C., Dutko, F. J., Felock, P. J., Diana, G. D., McKinlay, M. A. Crystal structure of human rhinovirus serotype 1A (HRV1A). J. Mol. Biol. 210:91–111, 1989.
6. Rossmann, M. G., Arnold, E., Griffith, J. P., Kamer, G., Luo, M., Smith, T. J., Vriend, G., Rueckert, R. R., Sherry, B., McKinlay, M. A., Diana, G., Otto, M. Common cold viruses. Trends Biochem. Sci. 12:313–318, 1987.
7. Rossmann, M. G., Palmenberg, A. C. Conservation of the putative receptor attachment site in picornaviruses. Virology 164:373–382, 1988.
8. Colonno, R. J., Condra, J. H., Mizutani, S., Callahan, P. L., Davies, M. E., Murcko, M. A. Evidence for the direct involvement of the rhinovirus canyon in receptor binding. Proc. Natl. Acad. Sci. U.S.A. 85:5449–5453, 1988.
9. Pevear, D. C., Fancher, M. J., Felock, P. J., Rossmann, M. G., Miller, M. S., Diana, G., Treasurywala, A. M., McKinlay, M. A., Dutko, F. J. Conformational change in the floor of the human rhinovirus canyon blocks adsorption to HeLa cell receptors. J. Virol. 63:2002–2007, 1989.
10. Greve, J. M., Davis, G., Meyer, A. M., Forte, C. P., Yost, S. C., Marlor, C. W., Kamarck, M. E., McClelland, A. The major human rhinovirus receptor is ICAM-1. Cell 56:839–847, 1989.
11. Staunton, D. E., Merluzzi, V. J., Rothlein, R., Barton, R., Marlin, S. D., Springer, T. A. A cell adhesion molecule, ICAM-1, is the major surface receptor for rhinoviruses. Cell 56:849–853, 1989.
12. Simmons, D., Makgoba, M. W., Seed, B. ICAM, an adhesion ligand of LFA-1, is homologous to the neutral cell adhesion molecule NCAM. Nature (London) 331:624–627, 1988.
13. Staunton, D. E., Marlin, S. D., Stratowa, C., Dustin, M. L., Springer, T. A. Primary structure of ICAM-1 demonstrates interaction between members of the immunoglobulin and integrin supergene families. Cell 52:925–933, 1988.
14. Williams, A. F., Barclay, A. N. The immunoglobulin superfamily-domains for cell surface recognition. Annu. Rev. Immunol. 6:381–405, 1988.
15. Staunton, D. E., Dustin, M. L., Erickson, H. P., Springer, T. A. The LFA-1 and rhinovirus binding sites of ICAM-1 and arrangement of its Ig-like domains. Cell, in press, 1990.
16. Maddon, P. J., Dalgleish, A. G., McDougal, J. S., Clapham, P. R., Weiss, R. A., Axel, R. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. Cell 47:333–348, 1986.
17. Mendelsohn, C. L., Wimmer, E., Racaniello, V. R. Cellular receptors for poliovirus: Molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. Cell 56:855–865, 1989.
18. White, J. M., Littman, D. R. Viral receptors of the immunoglobulin superfamily. Cell 56:725–728, 1989.
19. Rossmann, M. G., Moras, D., Olsen, K. W. Chemical and biological evolution of a nucleotide-binding protein. Nature (London) 250:194–199, 1974.
20. Remington, S. J., Matthews, B. W. A general method to assess similarity of protein structures, with applications to T4 bacteriophage lysozyme. Proc. Natl. Acad. Sci. U.S.A. 75:2180–2184, 1978.
21. Rossmann, M. G., Rueckert, R. R. What does the molecular structure of viruses tell us about viral functions? Microbiol. Sci. 4:206–214, 1987.
22. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., Tasumi, M. The Protein Data Bank: A computer based archival file for macromolecular structures. J. Mol. Biol. 112:535–542, 1977.
23. Marquart, M., Deisenhofer, J., Huber, R., Palm, W. Crystallographic refinement and atomic models of the intact immunoglobulin molecule Kol and its antigen-binding fragment at 3.0 Å and 1.9 Å resolution. J. Mol. Biol. 141: 369–391, 1980.
24. Deisenhofer, J. Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A *Staphylococcus aureus* at 2.9- and 2.8-Å resolution. Biochemistry 20:2361–2370, 1981.
25. Rossmann, M. G., Argos, P. Exploring structural homology of proteins. J. Mol. Biol. 105:75–95, 1976.
26. Matthews, B. W., Rossmann, M. G. Comparison of protein structures. Meth. Enzymol. 115:397–420, 1985.
27. Gribskov, M., McLachlan, A. D., Eisenberg, D. Profile analysis: Detection of distantly related proteins. Proc. Natl. Acad. Sci. U.S.A. 84:4355–4358, 1987.
28. Jones, T. A. A graphics model building and refinement system for macromolecules. J. Appl. Crystallogr. 11:268–272, 1978.
29. Jones, T. A., Thirup, S. Using known substructures in protein model building and crystallography. EMBO J. 5:819–822, 1986.
30. Hermans, J., Jr., McQueen, J. E., Jr. Computer manipulation of (macro)molecules with the method of local change. Acta Crystallogr. A30:730–739, 1974.
31. Smith, T. J., Kremer, M. J., Luo, M., Vriend, G., Arnold, E., Kamer, G., Rossmann, M. G., McKinlay, M. A., Diana, G. D., Otto, M. J. The site of attachment in human rhinovirus 14 for antiviral agents that inhibit uncoating. Science 233:1286–1293, 1986.
32. Badger, J., Minor, I., Kremer, M. J., Oliveira, M. A., Smith, T. J., Griffith, J. P., Guerin, D. M. A., Krishnaswamy, S., Luo, M., Rossmann, M. G., McKinlay, M. A., Diana, G. D., Dutko, F. J., Fancher, M., Rueckert, R. R., Heinz, B. A. Structural analysis of a series of antiviral agents complexed with human rhinovirus 14. Proc. Natl. Acad. Sci. U.S.A. 85:3304–3308, 1988.
33. Staunton, D. E., Dustin, M. L., Springer, T. A. Functional cloning of ICAM-2, a cell adhesion ligand for LFA-1 homologous to ICAM-1. Nature (London) 339:61–64, 1989.
34. Smith, T. J. MACINPLOT, a program to display density and atomic models on the Macintosh personal computer. J. Appl. Crystallogr., in press, 1990.
35. Marlin, S.D., Staunton, D.E., Springer, T.A., Stratowa, C., Sommergruber, W., Merluzzi, V.J. A soluble form of intercellular adhesion molecule-1 inhibits rhinovirus infection. Nature (London) 344:70–72, 1990.