and deoxyguanosine 5′-triphosphate (dGTP), 0.14 mM deoxythymidine 5′-triphosphate (dTTP), 4.3 μM dUTP-11-digoxigenin, 100 pmol each of forward and reverse primer, 1.0 μl (5 U) of *Taq* polymerase (Amplitaq, Perkin-Elmer, Norwalk, CT), and 0.001% gelatin (w/v)]. The reaction mixture was amplified in a 48-well thermocycler (Perkin-Elmer Cetus, Norwalk, CT) programmed for 40 cycles of thermal denaturation (94°C for 1 min), reannealing (58°C for 2 min), and extension (74°C for 1.5 min), with 5 s added for each of the 40 cycles.

15. Cycled cells were centrifuged onto microscope slides that had been treated with Denhardt's solution, washed with PBS, and incubated with a 1:1000 dilution of antibody to digoxigenin conjugated with alkaline phosphatase for 2 hours at 37°C. Slides were washed with PBS and incubated with NBT/X-phosphate substrate for 10 min at ambient temperature. Cells were counterstained with Fast green, mounted on cover slips, and visualized under light microscopy.
16. Sequence-specific oligonucleotide probes containing either single or multiple fluorescein-tagged nucleotides were synthesized on a 380B DNA synthesizer (Applied Biosystems, Foster City, CA) with carboxyfluorescein phosphoramidites. Fluoresceinated probes were as follows: MFA-1 (5′-XTTCTCTATCAAAGCAACCCACCTC-CCAATC-3′; nucleotides 6069 to 6083 and 8431 to 8445, numbered according to the HIV-1 HXB2 strain); human lymphocyte antigen (HLA) GH64 (5′-XTGGACCTGGAGAGGAAGGAGACTG-3′), where X represents the position of a 6-carboxyfluorescein amidite (Applied Biosystems); and SK19-3 (5′-ATCCYGGGATTAAATAAAATYGTAA-GAATGTATAGYCCTAC-3′), where Y represents the position of a 5-carboxyfluorescein phosphoramidite [P. Theisen *et al.*, *Tetrahedron Lett.* **33**, 5033 (1992)]. The synthesized material was alkaline-deprotected and purified by elution through an oligonucleotide purification cartridge. 5-Carboxyfluorescein phosphoramidite incorporation was verified by ultraviolet spectroscopy.
17. A 400-pM sample of the viral or HLA-DQα sequence–specific fluorescein-tagged oligonucleotide probe and sonicated herring sperm DNA (10 μg/ml; Sigma) was added to the PCR reaction tube; the sample was denatured at 95°C for 2 min and was then hybridized at 56°C for 2 hours. The cells were washed for 30 min with 2× SSC–50% formamide–bovine serum albumin (BSA; 500 μg/ml) at 42°C, 30 min with 1× SSC–50% formamide–BSA (500 μg/ml) at 42°C, 30 min with 1× SSC-BSA (500 μg/ml) at ambient temperature, and briefly with PBS at ambient temperature. Cells were resuspended in PBS (pH 8.3) and counterstained with propidium iodide (0.01 μg/ml) for flow cytometric analysis.
18. PBMCs were isolated from fresh heparinized blood layered on a Histopaque 1077 (Sigma) discontinuous density gradient and centrifuged for 30 min at 1600 rpm at ambient temperature. The turbid layer was removed and was washed twice with 3 volumes of RPMI and once with PBS. Cells were divided into portions and treated with STF and proteinase K as described (*14*).
19. The cell suspension was filtered through a 37-μm nylon mesh and analyzed by flow cytometry with an EPICS PROFILE II flow cytometer. Laser excitation was 15 mW at 488 nm. Instrument sensitivity was standardized before each experiment with Immuno-Bright calibration beads (Coulter Source, Marietta, GA). The percentage of fluorescence-positive cells was determined by integration over a range of 0.1% positive counts on the identically treated negative sample (100% uninfected PBMCs). At least 2000 events within the propidium iodide–positive window gate were counted for each sample.
20. The established acute lymphoblastic leukemia cell line CEM was obtained from T. Minowada through the AIDS Research and Reagent Program, Division of AIDS, NIAID, NIH. The cells were maintained in RPMI 1640 medium containing 2.5 mM Hepes buffer (pH 7.4), L-glutamine, 10% fetal bovine serum, penicillin G (100 U/ml), streptomycin sulfate (100 μg/ml), and amphotericin B (0.25 μg/ml) in a humidified incubator with a 5% CO_2 atmosphere.
21. B. K. Patterson *et al.*, unpublished data.
22. Southern blot hybridization was performed as described previously (*25*).
23. After proteinase K treatment (*14*), 40 μl of a reaction mixture containing 10.0 U of thermostable rTth reverse transcriptase (Perkin-Elmer Cetus), 90 mM KCL, 100 mM tris-HCl (pH 8.3), 1.0 mM $MnCl_2$, 200 μM each of dGTP, dATP, and dCTP, 125 μM dithiothreitol, 4 μM dUTP-11-digoxigenin (Boehringer Mannheim), 40 U of RNasin (Promega, Madison, WI) RNase inhibitor, and 100 pmol of the appropriate downstream primer was added to each sample. Samples were incubated for 15 min at 70°C and then placed on ice. Cells were resuspended in 160 μl of a PCR reaction mixture containing 100 mM KCl, 10 mM tris-HCl (pH 8.3), 0.75 mM EDTA, 0.05% Tween-20, 5.0% (v/v) glycerol-chelating buffer (Perkin-Elmer Cetus), 2 mM MgCl, 100 pmol of upstream primer, and 0.5 μl of *Taq* polymerase. Samples were then cycled in an automated thermal cycler. All solutions were made with diethyl pyrocarbonate–treated analytical reagent water (Mallinkrodt).
24. The pNL4-3–infectious molecular clone was obtained from M. Martin through the AIDS Research and Reagent Program, Division of AIDS, NIAID, NIH. Transfections were performed as described previously (*25*).
25. M. R. Furtado, R. Balachandran, P. Gupta, S. M. Wolinsky, *Virology* **185**, 258 (1990).
26. HIV-1–positive and –negative cell populations from the HIV-1–infected patients were sorted (EPICS C, Coulter), and the cells were lysed with proteinase K (100 μg/ml) at 56°C for 45 min. The probe:product DNA heteroduplex was resolved on a 6% nondenaturing polyacrylamide gel and analyzed by a 373 DNA sequencer with GeneScan 672 software (Applied Biosystems).
27. S. M. Schnittman *et al.*, *Ann. Intern. Med.* **113**, 438 (1991); G. Pantaleo *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9838 (1991).
28. M. E. Harper, L. M. Marselle, R. C. Gallo, F. Wong-Staal, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 772 (1986).
29. G. Pantaleo, C. Graziosi, A. S. Fauci, *N. Engl. J. Med.* **328**, 327 (1993).
30. C. Janeway, *Nature* **349**, 459 (1991).
31. A. Amadori *et al.*, *J. Immunol.* **148**, 2709 (1992).
32. L. Meyaard *et al.*, *Science* **257**, 217 (1992).
33. K. MacDonell *et al.*, *Am. J. Med.* **89**, 708 (1990).
34. Supported by NIH grants HD-26619-01 and AI-32535 and contract PO1-25569 (to S.M.W.). We thank N. Bouck, G. K. Lewis, L. Chow, J. Georgi, D. Bolognesi, M. Greenberg, A. T. Haase, H. Temin, D. Scarpelli, and J. Phair for advice and discussion and B. Giusti, J. Drew, S. Wu, and T. Meiers for technical support.

23 November 1992; accepted 23 February 1993

Crystal Structure of Domains 3 and 4 of Rat CD4: Relation to the NH_2-Terminal Domains

R. L. Brady,* E. J. Dodson, G. G. Dodson, G. Lange, S. J. Davis, A. F. Williams, A. N. Barclay

The CD4 antigen is a membrane glycoprotein of T lymphocytes that interacts with major histocompatibility complex class II antigens and is also a receptor for the human immunodeficiency virus. The extracellular portion of CD4 is predicted to fold into four immunoglobulin-like domains. The crystal structure of the third and fourth domains of rat CD4 was solved at 2.8 angstrom resolution and shows that both domains have immunoglobulin folds. Domain 3, however, lacks the disulfide between the beta sheets; this results in an expansion of the domain. There is a difference of 30 degrees in the orientation between domains 3 and 4 when compared with domains 1 and 2. The two CD4 fragment structures provide a basis from which models of the overall receptor can be proposed. These models suggest an extended structure comprising two rigid portions joined by a short and possibly flexible linker region.

The CD4 antigen is present on the subset of T lymphocytes that recognize foreign antigens in the presence of major histocompatibility complex (MHC) class II antigens [reviewed in (*1*)]. It interacts with MHC class II molecules (*2*) and is also the receptor by which the human immunodeficiency virus (HIV) binds its host (*3*, *4*). The extracellular portion of CD4 has been predicted, on the basis of sequence analysis, to fold into four immunoglobulin (Ig)-like domains, although the evidence for this organization in domain 2 (D2), D3, and D4 is weaker than in D1 (*5*–*8*).

Attempts to determine the structure of the full extracellular portion by protein crystallography have so far been unsuccessful (*9*, *10*); but the structures for the two NH_2-terminal domains in which the HIV-binding site is located have been determined, and both have an Ig fold (*11*, *12*). We report the crystal structure, determined at 2.8 Å resolution, of the third and fourth domains (D3D4) of rat CD4. Rat D3D4 shows considerable sequence homology to the equivalent domains from human CD4; however, crystallization of human D3D4, which contains an additional glycosylation site, has not been reported.

A protein consisting of CD4 D3 and D4

R. L. Brady, E. J. Dodson, G. G. Dodson, G. Lange, Department of Chemistry, University of York, York Y01 5DD, United Kingdom.
S. J. Davis, A. F. Williams, A. N. Barclay, Medical Research Council (MRC) Cellular Immunology Unit, Sir William Dunn School of Pathology, Oxford OX1 3RE, United Kingdom. A. F. Williams died of cancer in April 1992.

*To whom correspondence should be addressed.

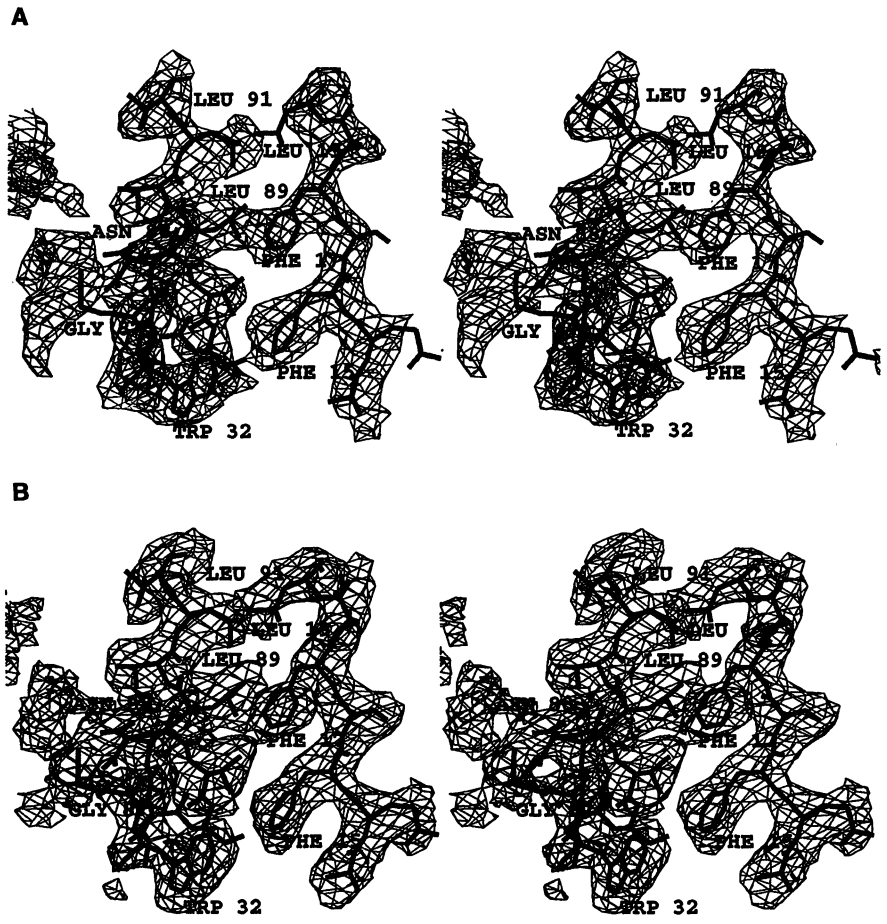

Fig. 1. Stereo pairs of (**A**) section from the MIR electron density map (8.0 to 4.0 Å) with the corresponding residues of the present model and (**B**) the equivalent section from a $2F_o - F_c$ map from which the labeled residues have been omitted from the phasing. Both contoured at 1 sigma. Poorly ordered density into which carbohydrate can be fitted extends to the left from Asn[88]; this density is broken at the contour level shown but is continuous at 0.5 sigma. Comparison of the original MIR map with an F_c map derived from the final model gives correlation coefficients of 0.68 for main chain and 0.45 for side chain density. The region shown is the disulfide equivalent site in D3. The side chains of Phe[17] and Leu[89] form almost continuous density reminiscent of a disulfide bridge, although the two β sheets are considerably further apart than is normally observed in Ig domains. The invariant tryptophan in strand C (Trp[32]), which normally lies adjacent to the disulfide bond, is here most proximal to the Phe[15]-Gly[87] pair. The absence of a side chain at position 87 (strand F) creates a pocket partially occupied by the Val[102] side chain (strand G). Residue Asn[88] is an N-linked glycosylation site (conserved in human CD4); density for the branched carbohydrate stem was visible in the original MIR and later electron density maps.

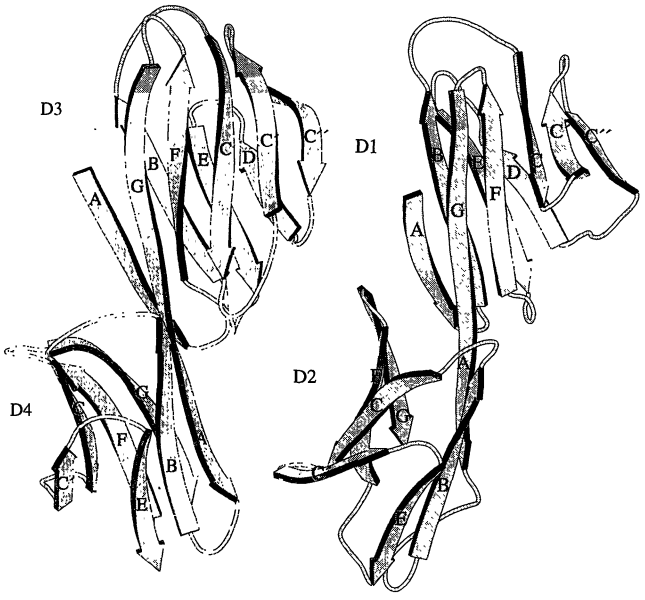

Fig. 2. Schematic diagram showing the structure of rat CD4 D3D4 (**left**) and, in a similar orientation, human CD4 D1D2 (**right**) from (*11*, *12*). Drawn with the program MOLSCRIPT (*26*). Strands labeled A through G.

was expressed in Chinese hamster ovary cell lines (*13*). Crystals of the purified protein were obtained from ammonium sulfate precipitants (*14*). The crystal structure was solved by multiple isomorphous replacement (MIR) (Table 1) after attempts with molecular replacement with use of the structure of the NH_2-terminal fragment were unsuccessful. The model consists of 1368 non-hydrogen protein atoms and 15 solvent molecules, giving an *R* factor of 23.3% (using all 7133 reflections 10 to 2.8 Å), with root-mean-square (rms) deviations of bond and angle lengths of 0.022 and 0.081 Å, respectively. This model does not include partially ordered carbohydrate (about 10% of the atoms present), which cannot be included reliably.

Only limited and restrained temperature factor (*B* value) refinement has been performed on protein atoms; the overall *B* value for all atoms refines to 40 Å^2, which reflects the high overall temperature factor evident for data collected from these crystals. Three residues had conformations just outside energetically unfavorable positions in a Ramachandran plot: Asn[20], Glu[24], and Gln[41]. These all fall within poorly defined and presumably disordered loop regions: the BC loop (residues 20 to 25), CC′ loop (37 to 42), and FG loop (95 to 98), all in D3. Further evidence for the correctness of the structure includes (i) identification of the disulfide in D4 (Cys[119] to Cys[161]) and disulfide-equivalent residues in D3 (Phe[17] to Leu[89]); (ii) density into which sugar units can be fitted is evident at the N-linked glycosylation site (Asn[88]) in the original MIR and all subsequent maps (Fig. 1); (iii) sequence fit to the core strands is good (average *B* value for main chain atoms of typical core strand residues is 28 Å^2) and is structurally sensible, whereas (iv) interpretation of the loop regions is more problematic (average *B* value for main chain atoms of 82 loop residues is 50 Å^2) and at present represents the best fit to the available data. Solvent-exposed residue side chains in general have high *B* values.

The crystal structure shows that the two domains form a closely associated rod-like structure (about 60 Å by 35 Å by 25 Å); both domains are antiparallel β barrels characteristic of the Ig fold and remarkably similar to the NH_2-terminal domains of CD4 (Figs. 2 through 4). This establishes the assignment of these two domains as Ig superfamily (IgSF) domains and thus shows that the extracellular portion of CD4 consists of four IgSF domains. The finding that all four domains of CD4 have Ig folds as predicted from the analysis of their amino acid sequences strengthens the argument that all members of the IgSF that were predicted in this way, about 40% of leucocyte surface proteins (*15*), will have Ig folds.

The arrangement of the nine strands in D3 is similar to that reported for D1 of CD4 and other IgSF variable (V set) domains (Figs. 2 and 3). However, the equivalent residues to the putative HIV gp120 binding site on the C″ strand (in D1) adopt a considerably different conformation. Domain 3 lacks the disulfide bond between the two β sheets that is normally found in Ig domains. The result is an increased separation of the two β sheets, accompanied by a generally higher proportion of bulkier, particularly leucine, internal hydrophobic side chains. The closest equivalent residues to the disulfide positions are Phe^{17} and Leu^{89}, whose Cα-Cα distance is 9 Å (Fig. 1). This is a greater separation than the range (5.6 to 7.4 Å) normally observed in Ig domains (*16*).

The structures of two other unusual IgSF domains lacking this disulfide have been determined. In CD2 D1, in which the cysteines are replaced by an isoleucine and a valine, there is no expansion; the reported separation (*17*) of 7.0 Å is within the normal range. In CD4 D2, in which the equivalent residues Leu^{116} and Cys^{159} are approximately 9 Å apart (*11*, *12*), there is an expansion similar to that observed in D3. As a net effect, the widening of D2 and D3 may provide greater surface area in the contact area between the two domains. The widening in D3 produces a slightly concave surface at the top of the domain into which loops at the base of D2 can be fitted (Fig. 5).

Domain 4, with seven strands, is similar to but even more compact than D2 of human CD4 (Figs. 2 and 3). Domain 4 is a truncated C2 set IgSF domain. The C2 set (*8*) was distinguished on the basis of sequence alignments from the archetypal C-type domain known from antibody structures (termed C1 set). In D4, as in D2, β strand D of antibody C-domains is replaced by a linking strand C′, which forms part of the same sheet as strand C. In D4, two prominent loops (between strands C and D and between F and G) protrude to the side (Fig. 2), although both are involved in crystal contacts. Domain 4 lacks many of the conserved residues normally present in Ig folds. Even though there are significant changes in the sequences of D2 and D4 (for example, the disulfide bond arrangements), the domains superimpose readily (Fig. 3).

Domains 3 and 4 are closely and presumably relatively rigidly associated, burying residues in the central, continuous GA strand. The buried residues in this interdomain region are predominantly hydrophobic, as are the residues between D1 and D2. Although the orientation of D1 to D2 and that of D3 to D4 are roughly similar, the positions of D1 and D3 are rotated 30° (Fig. 3C).

There are three less residues in the GA strand of D3D4 when compared with D1D2, although the overall length of the strand is only increased in the latter by about 4 Å because the additional residues

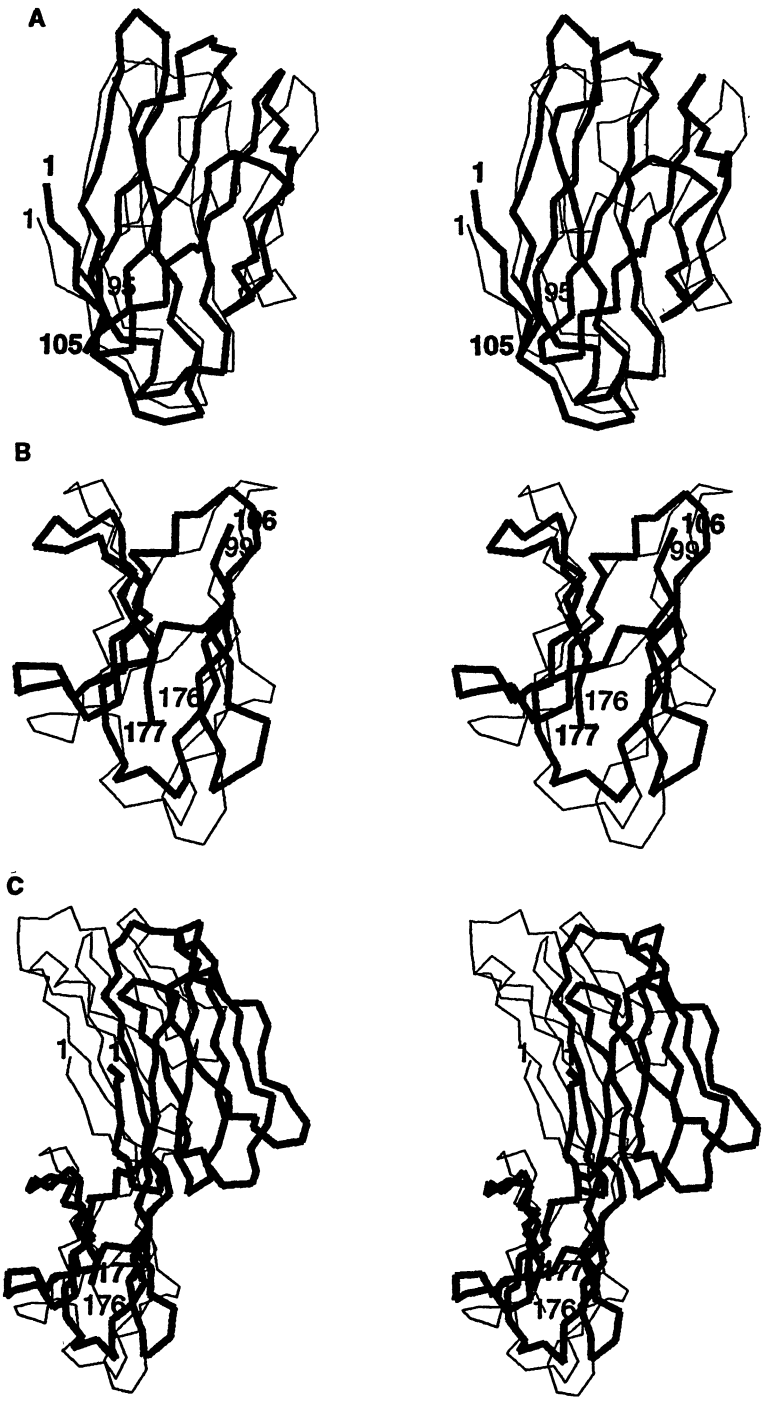

Fig. 3. Stereo pairs of (**A**) rat CD4 D3 (bold) overlapped on D1 (thin) of human CD4 (rms difference of 47 equivalent core Cα atoms = 2.7 Å); (**B**) rat CD4 D4 (bold) overlapped on D2 (thin) of human CD4 (rms difference of 36 equivalent core Cα positions = 2.1 Å); and (**C**) the D3D4 fragment (bold) overlaid on the D1D2 fragment (thin) by overlapping D2 and D4.

form a bulge in the A strand of D2. In spite of the differences in chain length and domain orientation, the buried solvent-accessible surface area in the fragments is the same [950 Å^2, as calculated with the use of QUANTA (Molecular Structures Inc., Burlington, Massachusetts) with hydrogens included]. The similarity of this junction suggests D3 and D4 associate to form a rigid rod-like unit, as has been shown for the D1D2 fragment through the study of different crystal forms (*12*).

Initial analyses of non-isomorphous crystal forms of D3D4 also show no overall rearrangement of the domains. Thus, this tightly abutting domain junction can generate not only rigidity but also overall changes in domain orientation. In the recent structure of CD2 (*18*), the linker region and relative orientations of D1 and D2 were considerably different from the arrangements seen in both D1D2 and D3D4 of CD4. Domain 1 of CD2 is tilted 45° from the rod axis, and there is flexibility in the linker region that is not seen in either of the CD4 structures. These differences profoundly affect the geometry of CD2 and, presumably, its function (*18*). Variation in interdomain organization that alters the exposure of protein surfaces mediating different functions is likely to be a characteristic of the family of molecules consisting of concatenated IgSF domains.

An analysis of various crystal forms of CD4 containing all four extracellular domains (*9*) has suggested that CD4 can dimerize, an effect attributed (*11*) to proposed dimerization loops in D3. The CD8 molecule, present on the T cell subset that interacts with the MHC class I molecule, exists as a dimer in crystals (*19*). In the crystals of D3D4, a dimeric interaction was generated by the crystallographic symmetry (Fig. 5). This crystallographic dimer is formed by antiparallel β sheet interactions between residues Ile^{99}, Leu^{100}, and Tyr^{101} in strand G of D3. Only the leucine is conserved in CD4 molecules from different species. This variation in the dimer-forming residues along with the relatively small contact area involved in the dimer interface (530 Å^2) suggest that the dimer in the crystal may only be the result of the lattice contacts. Recombinant soluble rat CD4 did not dimerize in solution (*20*).

The overall similarity of D3D4 to D1D2 is consistent with the suggestion that full-length CD4 has evolved by gene duplication of a two-domain structure (*7*). The structures of the D1D2 and D3D4 fragments can be used to model an overall structure of the extracellular

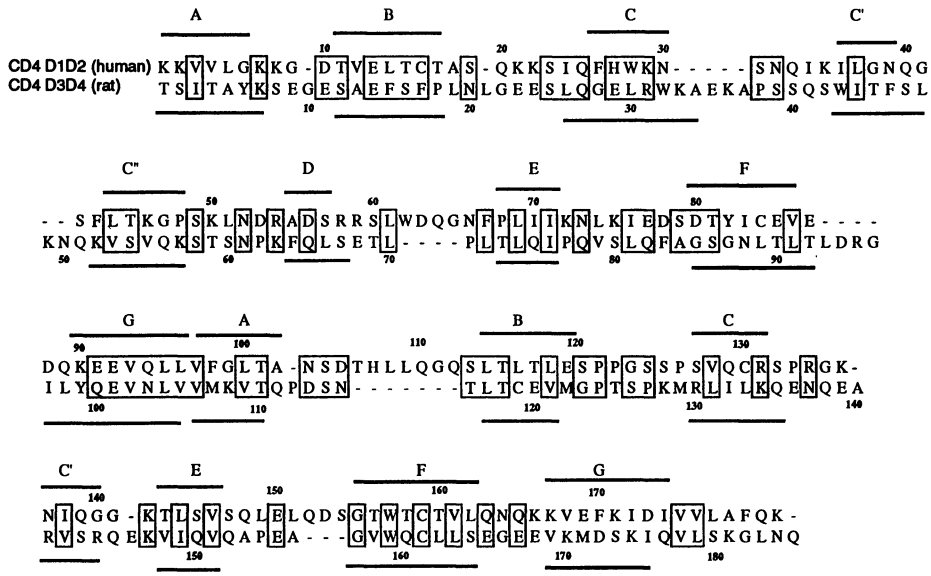

Fig. 4. Alignment of human CD4 D1D2 with rat CD4 D3D4. The alignment is based on structurally equivalent residues. The positions of the β strands in the two structures are indicated above and below the sequences. Data for human CD4 domains are from (*11*). Residues that are either identical or conservative substitutions are boxed. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

Table 1. Summary of native and heavy atom derivative data. Crystals: $P3_121$ with a = 77.78 Å, c = 82.48 Å. Each data set was collected from a single crystal at a wavelength of 1.54 Å (CuKα from a Rigaku RU-200 rotating anode), except for native (Mar), which was collected on the European Molecular Biology Laboratory (EMBL) beamline X11 at the Hamburg synchrotron operating at 0.96 Å. Intensity drops sharply beyond 4 Å, as is seen by average $\langle I\rangle$/sig$\langle I\rangle$ values for the synchrotron native data in resolution bins 25.0 to 6.0 Å, 30.5; 6.0 to 4.0 Å, 32.6; 4.0 to 3.0 Å, 11.5; and 3.0 to 2.8 Å, 3.5. Heavy atom parameters were refined with the program MLPHARE (*27*), resulting in a mean figure of merit of 0.71 (1984 measurements in range 8 to 4 Å). The protein chain was built into an electron density map calculated with unaltered MIR phases with the programs O (*28*) and FRODO (*29*). Most of the main chain and many side chains were traced directly into this initial map, giving an initial model with 171 residues, from which calculated structure factors gave a map that was correlated with the original MIR map (correlation coefficients of 0.63 and 0.34 for main chain and side chain regions, respectively). We further extended and refined the model in a series of cycles by using phase recombination with SIGMAA (*27*) to improve the MIR map and then cautious refinement with X-PLOR (*30*) (initially using the MIR phases as restraints). We initially used the R_{axis} native data, using the synchrotron native data only at a late stage of refinement. Only at the very final stages were $2F_o - F_c$ density maps used for rebuilding (*31*).

Type of Atom	Mode of data collection	Resolution (Å)	Measurements (total no.)	Reflections (total no.)	Multiplicity	Completeness (%)	R_{merge} (%)*	R_{deriv} (%)†	Sites (no.)	Relative occupancy	Phasing power‡	R_{Cullis}§
Native (R_{axis})	R_{axis}	3.0	32,891	6432	5.1	99.9	8.0					
Native (Mar)	EMBL Mar	2.8	41,516	7274	5.7	99.1	6.5	(11.2)				
Trans‖	Xentronics	3.7	30,780	3267	9.4	98.3	9.6	18.3	2	99, 35#	2.6, 1.4#	0.46, 0.53#
Trans¶	Xentronics	3.7	32,355	3263	9.9	98.3	10.5	18.1	2	100, 20#	2.4, 1.4#	0.49, 0.58#
K_2AuCl_4	R_{axis}	3.6	11,422	3518	3.2	99.9	10.4	21.4	2	84, 48#	1.5, 0.9#	0.69, 0.67#

*$R_{merge} = \Sigma\Sigma|I_i - \langle I\rangle|/\Sigma\Sigma I_i$ (double sum implies summation over all i intensities, I, of the measured reflections and subsequent summation over all the data). †$R_{deriv} = \Sigma||F_{PH}| - F_P|/\Sigma|F_P|$, where F_{PH} and F_P are the structure factors for native and derivative data. ‡Phasing power = $[\Sigma|F_H|^2/\Sigma(|F_{o,PH}| - |F_{c,PH}|)^2]^{1/2}$, where F_H is the structure factor for heavy atom data. §$R_{Cullis} = |(|F_P + F_H| - |F_{PH}|)|/|(|F_P| - |F_{PH}|)|$. ‖Crystals soaked in 2 mM trans-$Pt(NH_3)Cl_2$ for 3 weeks. ¶Crystals soaked in 2 mM trans-$Pt(NH_3)Cl_2$ for 2 days. #Values for acentric, centric measurements.

domains of the CD4 receptor, for which a crystallographic structure has not yet been determined (Fig. 5). The two possible models are based on the structures of human D1D2 and rat D3D4 with different interpolations of the five missing residues from the COOH-terminus of D2 and of the A strand of D3. For the model in Fig. 5B, the interactions between hydrophobic regions at the base of D2 and the top of D3 were maximized.

There may be considerable rotational freedom between the two fragments, particularly for the model illustrated in Fig. 5C. Both models assume that the five missing residues between the fragments adopt a fully extended β sheet conformation, resulting in a short, flexible "hinge"-type region between the fragments. This possibility appears to be supported by electron micrographs of four-domain CD4, in which a variety of conformations of the molecule are observed, each approximated by two domains joined end to end (*21*). Flexibility may be essential for the function of the receptor, particularly to allow fusion of HIV with the T cell membrane. The loops involved in the contacts between the two fragments in Fig. 5B, particularly loops AB and EF of D2, are highly conserved across the species. It is, of course, also possible that the residues joining the two fragments are not fully extended, in which case the fragments would be more intimately associated, consistent with sequence homology of this region with the junction of D1 and D2 (*12*).

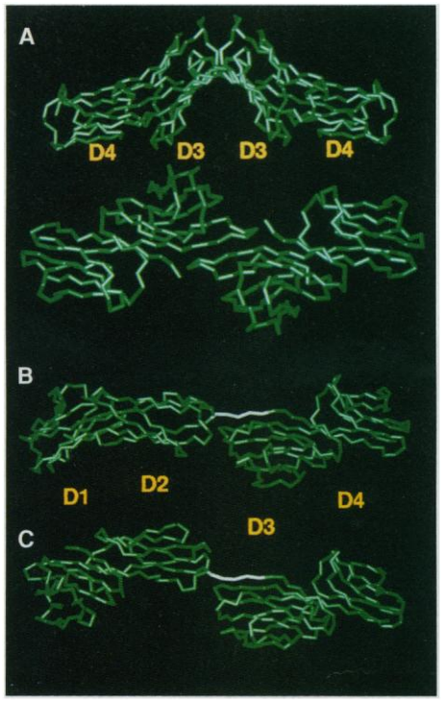

Fig. 5. (**A**) Two orthogonal views of the crystallographic dimer formed by antiparallel H bonding of residues Ile[99], Leu[100], and Tyr[101] of strand G of D3. (**B** and **C**) Possible models, shown as a Cα trace, for the complete extracellular portion of CD4 comprising human D1 and D2 joined to rat D3 and D4 by simple interpolation of the five missing residues to complete strand G of D2 and strand A of D3. Rat and human CD4 have 55% overall sequence identity or 70% sequence similarity with most core structurally important residues being highly conserved and hence are likely to adopt very similar folds. In (B), the connecting peptide (shown in pale green, with side chains in ideal conformations) joins the two fragments to bring conserved loop residues at the base of D2 into contact with the top of D3. The buried surface area between D2 and D3 in model (B) is approximately 600 Å^2. The loop residues involved in the contacts are 107 through 111 and 150 through 154 in D2, and 18 through 26 and 93 through 98 in D3. Both models are approximately 120 Å in length.

The MHC class II binding site on human CD4 may include regions in D1, D2, and D3 (*22–24*). Two sites in D3 have been implicated by mutagenesis studies (*22*). The equivalent residues in D3 of rat CD4 are located on the A strand (residues 1 through 5) and on the CC′ loop (37 through 42) and are not on the same face of the molecule. Rather than direct contacts, however, substitutions in the A strand may alter flexibility in the hinge region proposed in the models above and hence affect function. In contrast, the mutations in the CC′ loop are seen in the crystal structure to be in an exposed position where they are unlikely to affect the overall domain structure.

One study (*24*) has also implicated residues 165 through 172 in D4, which constitute the prominent FG loop protruding from the side of the domain (Fig. 2), although this mutant was reported to have disrupted antibody-binding characteristics, suggesting that the mutant was incorrectly folded. Interestingly, these residues are spatially close to the region implicated at the beginning of the A strand of D3. A concatenated rod-like structure for CD4 implies that interactions with the class II molecule might be restricted by the distance each molecule protrudes from the respective cell surfaces.

An alternative explanation for the mutation effects in D3 is that these regions exert an effect because they are involved in contacts with proteins that form the T cell receptor complex. Finally, the CD4-binding antibody Q425, which inhibits HIV fusion but not binding (*25*), has been mapped to residues that the rat CD4 structure shows to be in the D strand of D3. Binding of an antibody to this site could sterically hinder the proposed hinge region, which is in close proximity to the implicated residues.

The determination of crystal structures for the two fragments of CD4, each revealing a relatively rigid unit with tight internal organization, sets limits within which models of the overall receptor structure can be proposed. The structure of D3D4 and the models for all four domains form a basis for proposal of mutagenesis experiments to clarify the function of the CD4 molecule in HIV binding and fusion and its role in MHC class II recognition.

REFERENCES AND NOTES

1. J. R. Parnes, *Adv. Immunol.* **44**, 265 (1989).
2. G. Cammarota *et al.*, *Nature* **356**, 799 (1992).
3. A. G. Dalgleish *et al.*, *ibid.* **312**, 763 (1984).
4. D. Klatzmann *et al.*, *ibid.*, p. 767.
5. P. J. Maddon *et al.*, *Cell* **42**, 93 (1985).
6. S. J. Clark, W. A. Jefferies, A. N. Barclay, J. Gagnon, A. F. Williams, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1649 (1987).
7. A. F. Williams, S. J. Davis, Q. He, A. N. Barclay, *Cold Spring Harbor Symp. Quant. Biol.* **LIV**, 637 (1989).
8. A. F. Williams and A. N. Barclay, *Annu. Rev. Immunol.* **6**, 381 (1988).
9. P. D. Kwong *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6423 (1990).
10. S. J. Davis *et al.*, *J. Mol. Biol.* **213**, 7 (1990).
11. S.-E. Ryu *et al.*, *Nature* **348**, 419 (1990).
12. J. Wang *et al.*, *ibid.*, p. 411.
13. Recombinant rat CD4 D3 and D4 were expressed from a construct that included the sequence for the CD4 signal and for D3 and D4 with the use of the glutamine synthetase expression system (*20*) (R. L. Brady *et al.*, in preparation). The NH_2-terminal sequence was determined by protein sequencing to be TSITAYKSEG. Sialic acid was removed with neuraminidase digestion.
14. Crystals of the recombinant protein were obtained by vapor diffusion, in which 10-μl sitting drops containing the protein at 10 mg/ml in 20 mM tris buffer, pH 8, were equilibrated against a 1-ml reservoir containing the same buffer made 58 to 62% in saturated ammonium sulfate. The crystals, which diffract weakly, are hexagonal prisms with typical dimensions 0.2 mm by 0.15 mm by 0.15 mm.
15. A. N. Barclay *et al.*, *Leucocyte Antigens Factsbook* (Academic Press, New York, 1992), pp. 4–8.
16. J. S. Richardson, *Adv. Protein Chem.* **34**, 167 (1981).
17. P. C. Driscoll, J. G. Cyster, I. D. Campbell, A. F. Williams, *Nature* **353**, 762 (1991).
18. E. Y. Jones, S. J. Davis, A. F. Williams, K. Harlos, D. I. Stuart, *ibid.* **360**, 232 (1992).
19. D. J. Leahy, R. Axel, W. A. Hendrickson, *Cell* **68**, 1145 (1992).
20. S. J. Davis *et al.*, *J. Biol. Chem.* **265**, 10410 (1990).
21. J. T. Finch and R. L. Brady, personal communication.
22. D. Lamarre *et al.*, *EMBO J.* **8**, 3271 (1989).
23. L. K. Clayton, M. Sieh, D. A. Pious, E. L. Reinherz, *Nature* **339**, 548 (1989).
24. S. Fleury *et al.*, *Cell* **66**, 1037 (1991).
25. D. Healey *et al.*, *J. Exp. Med.* **172**, 1233 (1990).
26. P. J. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).
27. CCP4, The SERC (UK) Collaborative Computing Project No. 4, a Suite of Programs for Protein Crystallography (1979). Distributed from Daresbury Laboratory, Warrington, United Kingdom.
28. T. A. Jones, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr. Sect. A* **47**, 110 (1991).
29. T. A. Jones, *J. Appl. Crystallogr.* **11**, 268 (1978).
30. A. T. Brunger, *X-PLOR Manual* (Yale Univ. Press, New Haven, CT, 1990).
31. Further details of the analysis will be published elsewhere. R. L. Brady *et al.*, in preparation.

32. We thank P. C. E. Moody and Z. Dauter for assistance with x-ray data collection and processing; S. J. Lewis, J. Fennelly, and H. Sharma for technical assistance; A. C. Willis for peptide sequencing; J. Finch for electron microscopy studies; and H. Savage for general discussions. R.L.B., S.J.D., and G.L. were supported by the Medical Research Council (United Kingdom) AIDS-Directed Programme. Coordinates have been deposited in the Brookhaven Protein Databank (accession no. 1 CID).

7 December 1992; accepted 17 March 1993