The Structure of the Macrophage Signal Regulatory Protein α (SIRPα) Inhibitory Receptor Reveals a Binding Face Reminiscent of That Used by T Cell Receptors*

Deborah Hatherley†1, Karl Harlos†1, D. Cameron Dunlop2, David I. Stuart‡1,3, and A. Neil Barclay†1,3

From the †Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE and §Division of Structural Biology and the Oxford Protein Production Facility, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, United Kingdom

Signal regulatory protein (SIRP) α is a membrane receptor that sends inhibitory signals to myeloid cells by engagement of CD47. The high resolution x-ray structure of the N-terminal ligand binding domain shows it to have a distinctive immunoglobulin superfamily V-like fold. Site-directed mutagenesis suggests that CD47 is bound at a surface involving the BC, FG, and DE loops, which distinguishes it from other immunoglobulin superfamily surface proteins that use the faces of the fold, but resembles antigen receptors. The SIRP interaction is confined to a single domain, and its use of an extended DE loop strengthens the similarity with T cell receptor binding and the suggestion that they are closely related in evolution. The employment of loops to form the CD47-binding surface provides a mechanism for small sequence changes to modulate binding specificity, explaining the different binding properties of SIRP family members.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact. The atomic coordinates and structure factors (code 2uv3) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).
1 Supported by the Medical Research Council.
2 Supported by a Wellcome Trust studentship.
3 To whom correspondence should be addressed. Tel.: 441865275591; E-mail: neil.barclay@path.ox.ac.uk.

The abbreviations used are: SIRP, signal regulatory protein; SIRPα d1, signal regulatory protein α N-terminal domain; IgSF, immunoglobulin superfam-

†1,3

ily; MAD, multielement anomalous dispersion; RU, response units; MHC, major histocompatibility complex; BisTris, 2-(bis(2-hydroxyethyl)

amino)-2-(hydroxymethyl)propane-1,3-diol; CHO, Chinese hamster ovary; SeMet, selenomethionine; TCR, T cell receptor; MES, 4-morpholinoethane-
sulfonic acid; mAb, monoclonal antibody; Endo Hf, endoglycosidase Hf.

proteins with three extracellular IgSF domains (reviewed in Refs. 1 and 2). SIRPα (CD172a or SHPS-1 (4)) interacts with a ligand CD47 expressed on many cells (5–7) and gives an inhibitory signal through immunoreceptor tyrosine-based inhibition motifs in the cytoplasmic region that interact with phosphatases SHP-1 and SHP-2 (8) (Fig. 1). In contrast, SIRPβ has a short cytoplasmic region and associates with a transmembrane adapter protein DAP12 containing immunoreceptor tyrosine-based activation motifs to give an activating signal (9–12). SIRPγ contains a very short cytoplasmic region lacking obvious signaling motifs but also binds CD47, albeit 10 times weaker than SIRPα (13, 14).

A number of features distinguishes the SIRPs from other cell surface protein interacting pairs. First, although the sequences of SIRPs and SIRPβ are very similar (~90% identity (Fig. 2)), they show very different binding to CD47. Second, SIRPα, SIRPβ, and SIRPγ all have three IgSF domains, and SIRPα and SIRPγ interact with CD47 that contains only one IgSF domain. This three-plus-one domain topology is unique in interactions between leukocyte membrane proteins where the most common arrangement of interactions is end-on between proteins each with two IgSF domains; however, it could maintain the 14 nm distance between membranes that is a very common spacing in cell contacts involving immune cells, the immunological synapse (15–17) (Fig. 1). Third, CD47 has a single IgSF domain linked to a region that spans the membrane five times, a very unusual topology (18). Fourth, the topology of the protein at the membrane surface may be important. There is a disulfide bridge that links between the IgSF domain of CD47 and one of the loops between the transmembrane regions, which is required for optimal binding (19). Fifth, SIRPβ but not SIRPα is present on the cell surface as a dimer because of a Cys bridge in IgSF domain 3 (20) suggesting that cis interactions may be important in this family. Sixth, the SIRPs are unusual in that they have immunoglobulin C1-set domains that are normally only present in proteins involved in antigen recognition, MHC antigens, TCR, and antibody (21). It has been suggested that SIRPs are closely related in evolution to TCR and primitive antigen recognition proteins (22).

We describe here the x-ray crystallographic analysis of the ligand binding domain of SIRPα, and we show by site-directed mutagenesis that it interacts with its ligand CD47 in a novel way by involving the loops at the end of the domain. This may explain the unusual sensitivity of the protein specificitiy to amino acid changes.
EXPERIMENTAL PROCEDURES

Production of the SIRPα N-terminal V Domain (SIRPα d1)—A construct for the production of SIRPα N-terminal V domain (SIRPα d1) consisted of the SIRPα signal sequence and the N-terminal IgSF domain (residues 1–148; GenBank™ accession number CAA71403) followed by the sequence TRHHH-HHH. The amino acid terminus is predicted to be residue 30 of the precursor, giving the sequence EEEL. The numbering in this paper is that of this mature form. This construct was expressed using the pEE14 vector in the Lec3.2.8.1 variant of CHO cells as described (23). The recombinant protein was purified by nickel affinity chromatography, eluted with imidazole. N-Linked sugars were removed by incubation with Endo Hf (New England Biolabs) at 1 unit/µg for 90 min at 37 °C. The protein was purified by gel filtration in 10 mM Hepes, pH 7.4, 140 mM NaCl, 0.02% NaN₃ (HBS) and concentrated for crystallization trials. Selenium-labeled SIRPα was similarly prepared, but once the transfected CHO cells were confluent, the media were removed and the cells washed with PBS and incubated with methionine-free Dulbecco’s modification of Eagle’s media containing 30 mg/liter L-selenomethionine, 50 mg/liter L-cystine, and 2 mM sodium butyrate. It was not possible to estimate the level of selenomethionine incorporation by mass spectrometry; however, subsequent crystallographic analysis indicated it to be ∼60%, sufficient for structure determination.

Recombinant extracellular SIRPα consisting of all three domains was produced in a similar manner (residues 1–349; accession number CAA71403) followed by the sequence TRHHHHHH). Recombinant extracellular region of CD47 with a biotinylation site and a polyhistidine tag (CD47) was prepared by transient expression in 293T cells using the pEFBOS vector (24) and biotinylated using the Bir enzyme as described (25). The CD47 fragment was subcloned into pEE14 and expressed by CHO K1 cells. Recombinant protein was purified by nickel affinity chromatography and gel filtration in HBS.

Surface Plasmon Resonance Analysis of the Interaction of SIRPα d1 and d1–3 with CD47—The interactions were analyzed using a BIAcore™ 2000 at 37 °C as described (26). Briefly, ~8000 RU of streptavidin were coupled to a CM5 research grade chip using amine coupling, and biotinylated CD47 and CD4d3 were bound (659.7 and 724.6 RU, respectively). For kinetic analysis serially diluted monomeric SIRPα purified proteins were injected at the indicated active concentrations over the flow cells. KD values were obtained by both nonlinear curve fitting and Scatchard transformations to the binding data. Extinction coefficients of 9770, 36,850, and 22,550 M⁻¹ cm⁻¹ for SIRPα d1, SIRPα d1–3, and CD47, respectively, were calculated by Vector NTI (Invitrogen).
Crystallization, Data Collection, Structure Determination, and Refinement—Recombinant SIRPα d1 protein was concentrated to 45 mg/ml. Crystallization screening experiments were set up at 21 °C in the Oxford Protein Production Facility crystallization facility as vapor diffusion experiments using 200-nl plus 100-nl droplets of protein and precipitant. These experiments yielded crystals under several conditions, the best being with ammonium sulfate, sodium/potassium phosphate, and lithium-sulfate. At these conditions three-row optimization experiments were set up with protein/precipitant volumes of 100:100, 200:100, and 300:100 nl (27). All crystals had a platelike morphology. Some crystals grew to their full size within 24 h, whereas others took well over 1 month to appear. The crystal used for native data collection was grown at protein/precipitant volumes of 200:100 nl from 2.4 M ammonium sulfate, 0.1 M MES, pH 6.0 (Hampton Grid Screen, C3); the SeMet crystal was grown from protein at 24 mg/ml from 1.6 M ammonium sulfate, 0.1 M Hepes, pH 7.0 (Hampton Grid Screen, B4), at a protein/precipitant ratio of 300:100 nl.

Diffraction data of native and SeMet crystals were collected at 100 K at the European Synchrotron Radiation Facility (Grenoble, France) on BM14, the UK CRG beamline (Table 1), using a “marmosaic-225” CCD-detector (Marresearch, Norderstedt, Germany). Crystals were dipped into perfluoropolyether PFO-X125/03 (Lancaster Synthesis) before freezing in a cold nitrogen gas stream. Diffraction data were processed with HKL2000 (28).

The native Patterson map of SeMet peak data showed a strong peak (0.5, 0.2, 0.5) indicating non-crystallographic, translational symmetry. Anomalous Patterson maps of the peak data set showed good peaks on all of the Harker planes of space group P2₁,2,2 in addition to this translation peak. All Patterson peaks could be explained by two selenium sites, one per molecule as expected from the sequence for two molecules per asymmetric unit. The selenium positions were confirmed using the hkl2map (29), and phases were calculated with Solve and Resolve (30). Applying noncrystallographic symmetry information led to a clear electron density map.

The initial model was built using Arp/Warp (31) followed by manual model building and several refinement cycles with refmac5 (32), leading to a model with R-factor of 28.7% (R-free = 33.1%) in the resolution range 50 to 2.25 Å. A region of close contact between molecule A and B was apparent. This dimer was used as a starting model for molecular replacement to solve the structure of the native crystal because the self-rotation function of the native data had indicated the presence of a 2-fold noncrystallographic symmetry. The self-rotation function calculated with CNS (33) using data from 15 to 4 Å had a peak at ψ = 90.0°, φ = 64.3°, and κ = 180° with a peak height of 4.47 σ. Starting with the SeMet dimer model, a
clear molecular replacement solution was found by the program Phaser (34) for the native data. The same dimer was found in the SeMet and native crystal. The initial model of the native structure was again built with Arp/Warp followed by several cycles of refinement and model building using REFMAC, including TLS modeling (32) and Coot (35). This led to the refined structure of SIRPα d1 (Table 1).

Some native as well as SeMet crystals were found to possess considerably larger unit cell dimensions (space group C222, unit cell dimensions a = 115.4 Å, b = 139.4 Å, and c = 80.9 Å). These crystals were highly twinned and unsuitable for structural analysis. Twinned and untwinned crystals grew under identical conditions. Twinned and untwinned crystals could not be distinguished by their morphology.

**Mutagenesis of SIRPα**—Mutants were introduced by PCR into the pEFBOS vector (24) containing SIRPα (d1–3) linked to rat CD4d3 + 4 as an antigenic label (25). The proteins were expressed by transient transfection in 293T cells, concentrated, and immobilized on a BIAcore chip to which OX68 anti-rat CD4d3 + 4 mAb had been coupled in BIAcore TM 2000 at 25 °C (36). In each experiment using four flow cells, ~1200 response units (RU) of wild type, negative control (CD4), and two mutants were immobilized. Recombinant CD47 extracellular domain (0.2 μM) (purified and expressed as for SIRPα) was passed over the mutants to test for loss of ligand binding. The mutants were also tested with the SIRPα mAb Se5A5 that recognizes domain 1 and OX117 that recognizes domains 2 or 3 (13). The use of 0.2 μM CD47 (just below the KD value) provides a sensitive assay to detect changes in affinity. For each mutant the specific binding in response units was determined and compared with the binding obtained with wild type SIRPα (Fig. 3B). The binding values were adjusted for any differences in amounts of the SIRP proteins immobilized (36). Mutants were categorized as loss of binding site (less than 35% of wild type binding), intermediate (35–65%), and no effect (>65% of wild type binding).

**RESULTS**

**Expression of SIRPα d1**—The ligand binding, N-terminal domain of SIRPα (SIRPα d1) was expressed at high levels in the Lec3.2.8.1 variant of CHO cells and purified by nickel affinity chromatography. This cell line has defective glycosylation apparatus that renders glycoproteins sensitive to Endo Hf treatment (23). SIRPα d1 migrated as a major band of ~16 kDa compatible with the single IgSF domain and a minor band of higher molecular weight that was removed by Endo Hf (Fig. 3E). This is in line with expectations, because SIRPα d1 possesses a single potential N-linked glycosylation site, with sequence NITP, and it is established that a proline at that position can reduce the efficiency of glycosylation (37).

The Ligand Binding Potential Is Present in the N-terminal Domain of SIRPα—Previous studies had shown that the N-terminal domain of SIRPα could bind the ligand CD47 (5), but a

---

**Table 1**

| Data collection statistics | Native | SeMet peak | Remote | Inflection |
|----------------------------|--------|------------|--------|------------|
| Beamsline                  | BM14   | BM14       |        |            |
| Resolution limits (Å)      | 0.8856 | 0.9738     | 0.9079 | 0.9792     |
| Space group                | C2     | P21 2      |        |            |
| Unit cell dimensions (Å)   | 76.8, 56.3, 82.2, 114.9 | 76.7, 82.9, 55.2 |        |            |
| No. of observations        | 441122 | 981952     |        | 326916     |
| Unique reflections         | 29,544 | 14,665     | 16,725 | 13,420     |
| Completeness (%)           | 99.2 (98.0) | 96.2 (78.9) | 88.6 (61.1) | 88.3 (55.5) |
| Rmerge (%)                 | 22.1 (3.8) | 44.1 (7.3) | 22.3 (4.7) | 21.1 (4.2) |
| Rfree (%)                  | 8.7 (28.4) | 10.1 (29.9) | 7.6 (25.3) | 7.8 (25.3) |

**Phasing statistics**

|                         | Native | SeMet peak | Remote | Inflection |
|-------------------------|--------|------------|--------|------------|
| No. of selenium atoms   | 2 (1 per molecule) |            |        |            |
| Resolution (Å)           | 30-2.25 |            |        |            |
| Solvent content (%)      | 0.58   |            |        |            |
| Overall figure of merit  | 0.61   |            |        |            |

**Refinement statistics**

|                         | Native | SeMet peak | Remote | Inflection |
|-------------------------|--------|------------|--------|------------|
| Resolution limits (Å)   | 20-1.8 |            |        |            |
| No. of reflections in working set | 25,160 |            |        |            |
| No. of reflections in test set | 1497   |            |        |            |
| R factor (working set)  | 20.4   |            |        |            |
| R-free (%)              | 24.9   |            |        |            |
| No. of atoms (protein/other) | 1937/353 |        |        |            |
| r.m.s.d./bond lengths (Å) | 0.016  |            |        |            |
| Average R-factors (Å)² (overall/protein/other) | 28/27/30 |        |        |            |
| B-factor r.m.s.d. (Å)²: main chain bonded/main chain angles/side chain bonded/side chain angles | 2.9/3.1/4.0/5.4 |        |        |            |

**Note:**

- Numbers in parentheses refer to the appropriate outer shell.
- Rmerge = [Σ(hkli)Fcalc(hkli) − Σ(hkli)Fobs(hkli)]/[Σ(hkli)Fobs(hkli)], where Σ(hkli) is the intensity of an individual measurement of a reflection, and Σ(hkli) is the average intensity of that reflection.
- R-factor of test set (5% of the data removed prior to refinement).
- Other atoms include 299 water molecules, 6 sulfate ions, and 2 MES molecules.
- r.m.s.d. indicates root mean square deviation from ideal geometry.
Structure of the SIRPα Domain That Binds CD47

A ribbon diagram of the polypeptide showing the V-type Ig fold with strands labeled from A1 to G2. The protein is oriented such that the other SIRP domains would be below this domain and the extracellular space above it. Color varies smoothly from blue at the N terminus to red at the C terminus. The sequence at the C terminus includes part of the His tag added for purification (shown in gray). Secondary structure assignments are from program DSSP (64). Shown here is molecule B, B, close packing of the two independent molecules of SIRPα d1 (A and B) around a noncrystallographic 2-fold axis as seen in the crystal. An extensive region of close contact is formed by the DE loops as well as the B1/B2 strands and part of the A1 strands of both molecules. The inset shows a close up of this interface. These interactions are unlikely to represent a physiological dimer at the cell surface because molecules A and B point into opposite directions. This as well as Figs. 5 and 6 were drawn with Pymol.

Contribution of the other domains could not be excluded. This was tested by preparing recombinant SIRPα containing all three IgSF domains (d1–d3). A recombinant protein consisting of the single IgSF domain of CD47 together with a tag that could be biotinylated was prepared to enable immobilization of CD47 to streptavidin previously coupled to the BIAcore chip. Fig. 3 shows that SIRPα d1 and d1–3 bound CD47 with comparable affinities, $K_D$ of 0.3 and 1.3 μM at 37 °C and 0.2 and 0.6 μM at 25 °C (data not shown), respectively. The experiment was repeated in the opposite orientation (passing recombinant soluble CD47 over the chip). Despite using freshly prepared CD47, some aggregation occurred, but an estimate for the $K_D$ of 0.8 μM at 25 °C was obtained (Fig. 3C). The values are comparable with 2 μM at 37 °C from previous data (13). Thus the best estimate for the SIRPα/CD47 interaction is $K_D \approx 1.2$ μM with no significant contribution from domains 2 and 3.

Crystallization and Structure Determination—SIRPα d1 formed crystals rapidly that diffracted to ~1.8 Å (see “Experimental Procedures”). Molecular replacement phasing was unsuccessful, and the structure was solved by multiwavelength anomalous diffraction (MAD) analysis of a crystal (belonging to a different space group, see Table 1) of selenomethionated SIRPα d1 prepared by growing the CHO clone in the presence of selenomethionine. The analysis was complicated by the presence of pseudo symmetry (see “Experimental Procedures”); however, the structure was eventually refined satisfactorily at 1.8 Å resolution ($R$-free = 24.9%; for other details see Table 1) to yield a model for the protein consisting of 299 water molecules, 6 sulfates, and 2 molecules of MES buffer. The crystallographic unit cell contains two SIRPα d1s related by a noncrystallographic 2-fold axis, with structures that are similar (overall fold shown in Fig. 4A, root mean square deviation on 117 C-αs of 0.55 Å apart from a few differences, which are minor with the exception of a significant difference in the DE loop that makes up the core of the major crystal contact (~3.5 Å deviation at C-α of residue Thr-67). This contact is between the two noncrystallographically related molecules of SIRPα d1, which are in a trans-orientation, inconsistent with the formation of cis dimers at the cell surface (see Fig. 4 and “Discussion”).

Structural Characteristics of SIRPα d1—The structure consists of a typical V-like Ig domain with two β-sheets linked by a conventional Cys bridge between β-strands B and F (Fig. 4). The A strand switches from the GFC to BED faces as found in many Ig V domains. There is also a short $3_{10}$ helix (residues 83–85). The closest structures to SIRPα d1 were all IgSF domains, as judged by systematic comparison using the Dali and SSM servers (38, 39); however, none were especially close and indeed the ranking of closest structures differed between the two servers. A composite, representative list of the 12 most similar domains was therefore compiled, and by using this information and a general knowledge of the field, definitive superpositions were obtained using the program SHP (39). The result is presented in Fig. 5, with an overlay of the C-α backbones of these structures shown in Fig. 5B, whereas Fig. 5A presents the global fold comparison data in a systematic way, by using pairwise analysis of all structures to produce a phylogenetic tree (40). This confirms that SIRPα d1 is substantially different from all other known structures, although it is more similar to the V domains of immunoglobulins and TCRs than it is to the CD2 adhesion molecules. The closest similarities are to TCR-like molecules and CTLA4. The distinctive features of SIRPα d1 include, most notably, an unusually large DE loop but also protruding FG and CC’ loops and the absence of a C’’ strand (Figs. 4 and 5).

The presence of two molecules in the crystallographic asymmetric unit allows us to provide at least a partial assessment of...
the flexibility of the molecule. The linker region between d1 and d2 (residues 114–120) is notable
for the absence of flexibility (despite being involved in different interactions in the two mole-
cules). This suggests that these two domains are likely to be
locked together as a rather rigid unit, despite this sequence only
containing a single proline residue. In fact there are only two
regions of significant flexibility (excluding the C-terminal His6
tag); the C′D loop (which is significantly displaced from the
position characteristic of molecules containing a C′ strand, see
Fig. 5B) shows modest differences, but the greatest variation is
seen in the extended DE loop (up to ~4 Å). This loop is crucial
for the striking and extensive contact in the crystal between
the two SIRPα d1 molecules (Fig. 5B), and despite the deviations
between the two molecules their mutual interaction conspires
to produce an almost exact 2-fold relationship (179.6° rotation
and 0.3 Å translation). This interface is novel for interactions of
Ig-like domains and is substantial, occluding 750 Å2 of surface
area, compared with 650–690 Å2 for the CD2/CD2 interaction
observed in the crystal (41, 42) and about 600 Å2 for the actual
ligand interaction between CD2 and CD58 (43). Furthermore
the anti-parallel orientation of these domains (see Fig. 4B)
is such that the contact could potentially mimic interactions with
membrane proteins on other cells, either through ligand bind-
ing (CD47) or homophilic interactions. This surface also shows
some variation between SIRP family members (Fig. 2) making it
an attractive candidate for the CD47-binding site, but mutagenesis experiments (see below) allow us to discount this
possibility. It also seems unlikely that it represents a biologically
relevant homophilic interaction, as no interaction of SIRPα was
observed during biochemical analysis or by surface plasmon
resonance (data not shown).

Identification of the Binding Region on SIRPα for CD47 by
Site-directed Mutagenesis—On the basis of the SIRPα d1 struc-
ture, a panel of 19 single residue mutants was designed to map
the CD47-binding surface (shown in Figs. 2 and 6). There was
a particular focus on three regions as follows: (i) the GFCC′ face
first identified for CD2 (42) that is frequently involved in leuko-
cyte protein interactions involving IgSF domains (six muta-
tions); (ii) the interface observed in the crystal contact involving
the DE loop (four mutations); and (iii) the loops corresponding
to the complementarity determining regions found in antibod-
ies and TCRs (the BC, C′C′, and FG loops, seven mutations).
The strategy was to select outpointing side chains and to intro-
duce major changes, the aim being to disrupt the interface,
enabling its extent to be mapped, rather than finding the ener-
getically most important residues in the interaction, a strategy
we developed with the CD2/CD48 and CD200/CD200R sys-
tems (36, 44). Mutants were introduced into a chimeric protein
that contained CD4 domains 3 as an antigenic tag, pro-
duced in a transient mammalian expression system, immobi-
lized in the BIAcore™ using a mAb against the CD4d3
tag, and tested for binding to CD47 and mAb recognizing SIRPα
(Fig. 3). The concentration of CD47 used was close to the KD
value; this allows for sensitive detection of mutants. All the
mutants were expressed at satisfactory levels, could be immo-
bilized through the CD4 mAb, and reacted with the OX117
mAb that recognizes a determinant in domains 2 or 3 of SIRPα,
indicating that the proteins were well folded. The mutants were
also tested with a mAb against the N-terminal domain SE5A5.
This mAb blocks ligand binding (6), and three mutants that
destroyed ligand binding activity also lost SE5A5 activity
(V33E, K96E, and I31K); three mutants (F94R, D100K, and
Structure of the SIRPα Domain That Binds CD47

S66D) were fully active for mAb binding but lost ligand binding activity, and two (M72R, R69E) had partial SE5A5 activity. This indicates that there is overlap between the antibody and ligand-binding sites. CD47 binding data for the mutants are summarized in Figs. 2 and 6.

Mutants in that region of the DE loop (Fig. 2) involved in the extensive crystal contact (Fig. 4B) had no effect on CD47 binding, indicating that this interface is not mimicking the CD47 interaction. Residues in the FG and BC loops caused dramatic loss of activity, whereas mutants in the main faces of the β-sheets had no effect. Some mutants had partial effects, and most of these were in the putative binding surface involving the BC, C'D, and FG loops together with the top of the DE loop. One mutant that disrupts CD47' binding (S66D) is somewhat distant (almost 10 Å) from the putative binding face, but this side chain forms a hydrogen bond with the main chain so it may have an indirect effect by stabilizing the conformation of the DE loop. Overall, the mutants affecting ligand binding clearly suggest that CD47 binds through the ends of the domain rather than the face. This is illustrated in Fig. 6B, top view, that shows that the binding face involves almost all the top of the SIRPα d1 molecule, covering, at a minimum, a triangular area of ~17 Å in width and ~10 Å in height. Fig. 6C shows that the binding face is strongly polarized electrostatically, comprising an extensive basic patch.

DISCUSSION

The CD47 binding domain of SIRPα is a classic Ig V-type domain. IgSF domains are particularly common in surface membrane proteins of leukocytes, and in many cases N-terminal V-type domains mediate interactions with membrane proteins on opposing cells (18). In this respect SIRPα d1 is fairly typical; however, the structure shows a number of distinctive features, and the mutagenesis analysis shows that the binding face is unusual in that it involves the loops that are the equivalent of the hypervariable loops of antibodies and TCRs. This is in line with a very recent study (45) using mutational analysis of SIRPβ and SIRPα to find residues crucial in binding CD47. Residues Val-27 and Gln-37 had effects on binding in agreement with the involvement of the BC loop.

The involvement of the loops is unusual in that IgSF domains involved in cell adhesion tend to associate via the faces of the β-sheets. This face to face interaction was first proposed on the basis of pseudo-homophilic interactions seen in crystals of CD2 that involved the GFCC face (41, 42) and subsequently observed in, for instance, B7/CTLA4 (46), the Coxsackie and adenovirus receptor (47), and CD2/CD58 (43). This mode of association is probably common to many interactions in this family of proteins, as illustrated by the recent structure of the homophilic interacting NTB-A (48). The homodimerization in Necl-1 involves the CC'CD strands and intervening loops (49),...
Structure of the SIRPα Domain That Binds CD47

and the interaction with CD4 and MHC class II involves mainly the face of the fold (50). CD8 does interact with MHC class I through its loops, but it interacts as a dimer and thus resembles TCR and Ig antigen recognition (51, 52). The recognition of MHC class I by the leukocyte LIR and KIR receptor family also involves the loops of two domains, but the interaction is centered around the junction of the two domains (53–55). In contrast in SIRPα the binding involves the complementarity-determining region-like loops of only one domain and in addition the top of the unusually large DE loop (Arg-69). The use of loops presented at the ends of IgSF domains for ligand binding is, however, a recurrent theme in the rapid generation of different fine binding specificities; it is key to antigen recognition in the context of both antibodies and TCRs and is used by the LIR and KIR receptors. There is, however, also a parallel among the adhesive interactions, the interaction of CTLA4 with B7-1 and B7-2 (56, 57). In that system the B7s contribute a classical “side” face of the β-sandwich; however, CTLA4 approaches this face in such a way that the corner of the molecule forms the point of contact, i.e., in a more “end-on” fashion than in the classical adhesive interactions, but not so squarely onto the loops as is seen in the immune receptors and, apparently, in SIRPα. Nevertheless, the residues involved have some similarity with those implicated in SIRPα (58). This is illustrated in Fig. 6D where the CTLA4 of the CTLA4/B7-2 complex (56) has been superposed on SIRPα d1 to show how the region occluded by B7-2 overlaps the critical binding residues of position of SIRPα d1. Note that there are major differences in the binding loops between CTLA4 and SIRPα, which is reflected in the steric clashes between SIRPα and B7-2, largely because of the prominent DE loop of SIRPα and the shortening of the FG loop in CTLA4, so the B7-2 slips down and is disconnected from the DE loop of CTLA4 (this is even more pronounced in the complex of CTLA4 with B7-1 (57)). In this respect SIRPα instead resembles TCRs; it has an extended DE loop, which appears to be directly involved in ligand binding. Some TCRs use the DE loop to recognize MHC peptide, and this region is termed the HV4 or 4th hypervariable region (59, 60). In contrast the DE loop is not used for antigen recognition by antibodies.

In the case of the SIRPs, the use of loops for binding may provide a molecular explanation for the different recognition properties of the three SIRPs despite close sequence similarity (Fig. 2). Examination of the sequence differences between the SIRPs, which can bind CD47 (SIRPα and SIRPγ), and SIRPβ, which does not, shows few candidates for residues essential for CD47 binding especially when one looks at other SIRPα sequences (61) that presumably correspond to polymorphisms as only one gene has been found in the human genome (2). The use of loops for binding provides a mechanism whereby small chemical differences in side chains can have large effects on the shape of the loops without disrupting the core β fold of the molecule, as exemplified by the variation observed in the complementarity-determining regions of antibodies and TCRs.

The SIRPs show additional similarities to TCRs and antibodies, in that their other two Ig-like domains have sequence similarity to the class of IgSF C domains found only in proteins associated with antigen recognition such as TCR, Ig itself, MHC antigens, and β2-microglobulin. One theory proposed to explain this was that they represent the precursors of the antigen receptors (22); however, because no SIRP homologues have been recognized in species such as amphibians and fish, it may be more likely that SIRP evolved from an antigen receptor (1, 2). The similarities in the binding face of SIRPα and the antigen receptors, together with the overall fold similarity, provides further support for a close evolutionary link between these molecules.

The data presented here indicate that the top of the Ig-like domain of SIRPα interacts with the single Ig-like domain of CD47. Assuming that the three domains of SIRPα are in a linear array (and the rigidity of the d1–d2 linker is consistent with this), then this interaction would span four Ig domains, as in the immunological synapse as illustrated in the scheme in Fig. 1. The orientation of CD47 may be important because it contains an unusual disulfide bond linking the extracellular domain of CD47 to one of the small extracellular loops of the 5-transmembrane segment (19).

Another aspect of proteins mediating cell-cell interactions is the use of the multimeric state (dimeric or higher order) of the proteins on the cell surface to modulate signaling outcome. This has been shown to be important in the B7 CD28/CTLA-4 system (46). We found no evidence of dimerization of the recombinant SIRPα N-terminal domain, and the crystal contacts observed are not compatible with cis interactions at the cell surface. However, it is still possible that the full extracellular domain can associate.

The finding that SIRPα recognition uses the loops at the top of the domain rather than the faces of the fold provides a mechanism to manipulate ligand binding specificity by modest chemical changes in amino acid side chains. In addition to antigen receptors, this mode of recognition is used by NK paired receptors LIR and KIRs, where there is evidence for rapid changes in specificity driven by pathogens. There are similarities between the NK self-recognition system and the highly polymorphic SIRPs in that CD47 is a marker of self, because cells deficient in CD47 are more susceptible to lysis by phagocytes (62). The main difference between the two systems is that, unlike the MHC antigen targets of the NK receptors, CD47 is not polymorphic. Nevertheless, the SIRP CD47 system of paired receptors recognition has evolved, presumably in response to pathogen load, to provide a highly sensitive recognition system regulating myeloid cell activity.

Acknowledgments—We are grateful to H. Belrhali and M. Walsh of BM14, the UK MAD Beamline at the European Synchrotron Radiation Facility (Grenoble, France) for assistance with data collection, and T. S. Walter, J. Ren, N. Abescia, M. Koch, S. Graham, C. Siebold, J. Kaufman, M. H. Brown, and J. Grimes for help and advice. The Oxford Protein Production Facility is funded by the Medical Research Council with additional finance from the Biotechnology and Biological Sciences Research Council and is part of the Structural Proteomics IN Europe consortium QLG2-CT-2002-00988.

REFERENCES
1. Barclay, A. N., and Brown, M. H. (2006) Nat. Rev. Immunol. 6, 457–464
2. van Beek, E. M., Cochrane, F., Barclay, A. N., and van den Berg, T. K. (2005) J. Immunol. 175, 7781–7787
