Published in final edited form as:

Nat Immunol. 2006 August ; 7(8): 875–882. doi:10.1038/ni1359.

Ancient evolutionary origin of diversified variable regions demonstrated by crystal structures of an immune-type receptor in amphioxus

José A Hernández Prada¹, Robert N Haire², Marc Allaire³, Jean Jakoncic³, Vivian Stojanoff³, John P Cannon^{2,4}, Gary W Litman^{2,4,5}, and David A Ostrov¹

¹Department of Pathology, Immunology and Laboratory Medicine, University of Florida College of Medicine, Gainesville, Florida 32610, USA

²Department of Pediatrics, University of South Florida College of Medicine, Children's Research Institute, St. Petersburg, Florida 33701, USA

³Brookhaven National Laboratory National Synchrotron Light Source, Upton, NY 11973, USA

⁴Immunology Program, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida 33612, USA

⁵Department of Molecular Genetics, All Children's Hospital, St. Petersburg, Florida 33701, USA

Abstract

Although the origins of genes encoding the rearranging binding receptors remain obscure, it is predicted that their ancestral forms were nonrearranging immunoglobulin-type domains. Variable region-containing chitin-binding proteins (VCBPs) are diversified immune-type molecules found in amphioxus (*Branchiostoma floridae*), an invertebrate that diverged early in deuterostome phylogeny. To study the potential evolutionary relationships between VCBPs and vertebrate adaptive immune receptors, we solved the structures of both a single V-type domain (to 1.15 Å) and a pair of V-type domains (to 1.8 Å) from VCBP3. The deduced structures show integral features of the ancestral variable-region fold as well as unique features of variable-region pairing in molecules that may reflect characteristics of ancestral forms of diversified immune receptors found in modern-day vertebrates.

The immunoglobulin 'superfamily' is one of the most prevalent families of protein domains in the animal kingdom^{1–4}. The functions of the various members of this superfamily are disparate and provide a notable example of selection acting on a modular structure, the immunoglobulin domain fold, to form molecules with specialized functions in processes

© 2006 Nature Publishing Group

Correspondence should be addressed to D.A.O. (ostroda@pathology.ufl.edu) or G.W.L. (litmang@allkids.org).

Author Contributions: J.A.H.P., crystallization, data collection, data reduction, structure determination, structural refinement, structural analysis and composition of the manuscript; R.N.H., protein expression, refolding and purification of recombinant V1 and V1-V2 VCBPs and contributions to the manuscript; M.A., space-group determination and phase determination and cryoprotection of native and selenomethionine crystals for the high-resolution structure of VCBP3 V1 (1XT5); J.J., data collection, space-group determination and phase determination of native and selenomethionine crystals of VCBP3 V1V2 (2FBO); V.S., data collection supervision and mounting of native and selenomethionine crystals of VCBP3 V1(1XT5) and V1V2 (2FBO); J.P.C., generation of constructs for VCBP3 V1 and VCBP3 V1V2 expression and contributions to the manuscript; G.W.L., data analysis, interpretation, discussions and fundamental contributions to the manuscript; D.A.O., crystal mounting, cryoprotection, supervision of MAD and native data collection of VCBP3 V1 and SAD and native collection of VCBP3 V1V2 (2FBO) and fundamental contributions to the manuscript.

Competing Interests Statement: The authors declare that they have no competing financial interests.

such as cell-cell adhesion and immune recognition⁵. Immunoglobulin domains are composed structurally of a β -sandwich fold; a prominent feature of this fold is a disulfide bridge connecting two β -sheets with one invariant tryptophan residue packing against the disulfide bond⁶. Immunoglobulin domains can be classified as variable (V), constant (C), strand-switched (S) and hybrid (H) sets based on β -strand topology². The V set of domains range from being nonpolymorphic or minimally polymorphic to having high genetic complexity because of germline and somatic diversification. T cell antigen receptors (TCRs) and immunoglobulins, which constitute the rearranging antigen-binding receptors, are the most extensively characterized V-set molecules, although other molecules such as CD4 (ref. 7), CD8 (ref. 8) and at least two classes of natural killer cell receptors⁹ also have this structure. It is likely that the ancestral element that diversified into the rearranging antigen receptors was an immunoglobulin-like sequence with characteristics of the V set, but its properties and patterns of evolutionary diversification remain unknown^{10,11}.

Although certain exceptions exist^{12,13}, most antigen receptors used in adaptive immune responses consist of heterodimers in which V-set immunoglobulin domains are packed to juxtapose the protein segments with the highest degree of germline and somatically induced variation¹⁴. This feature of tertiary-quaternary structure forms a vast range of binding surfaces that accounts for highly diverse antigen receptor repertoires. The unusual genetics that account for the hyperdiverse structures of immunoglobulins and TCRs seem to be limited to the jawed vertebrates^{10,11}.

Alternative forms of immune-type receptors are found throughout invertebrate and protochordate phylogeny, but so far only the variable region-containing chitin-binding proteins (VCBPs) in amphioxus (*Branchiostoma floridae*) have shown evidence of extensive regionalized hypervariation in predicted V domains^{11,15,16}. VCBPs consist of two N-terminal predicted V regions and a C-terminal chitin-binding domain. The V-family distribution, high degree of germline polymorphism, tissue-specific expression and use of a chimeric immunoglobulin-lectin structure of VCBPs are all consistent with involvement of these molecules in immune recognition^{11,15,16}. The predicted primary structure of at least one VCBP family shows that the regionalized hypervariation is substantially displaced from that found in immunoglobulins and TCRs; that is, it does not occur in the sequence regions corresponding to the complementarity-determining regions (CDRs) of these rearranging receptors. Furthermore, a contiguous joining (J) region, which is a ubiquitous feature of antigen-binding receptors, is absent in the sequence^{11,15–17}. VCBPs may well reflect structural characteristics of an important transition between nonrearranging innate pattern-recognition molecules and the conventional adaptive immune receptors¹⁸. To determine how structural features of the V domains found in this putative receptor family potentially relate to those of the rearranging antigen receptors of modern jawed vertebrates as well as to hypothetical ancestral V-set molecules, we solved the structure of the N-terminal V regions of VCBP3 from amphioxus by X-ray crystallography.

Results

Both immunoglobulin folds of VCBP3 are of the V type

Immunoglobulin domains are structurally classified as V type if they have a conserved nine-strand secondary structure topology forming two β -sheets packed tightly in the aligned mode consisting of a front sheet (strands A'GFCC'C'') and back sheet (strands ABED)². Like several nonvertebrate immune-type receptors^{11,19,20}, VCBPs are chimeric structures composed of immunoglobulin domains and lectin domains¹⁵. Extensive sequence variation in the two tandem immunoglobulin domains has been described among the five known VCBP subfamilies; in addition, regionalized germline hypervariation has been characterized most extensively in the N-terminal predicted V domains of members of the VCBP2 family.

To assess how the three-dimensional structure of VCBP V domains might relate to those of other immune-type receptors, we expressed and crystallized both the single N-terminal V domain (V1) and the two tandem V domains (V1 and V2) of VCBP3 and determined their structures by the multiple-wavelength and single-wavelength anomalous dispersion methods²¹, respectively.

Native and selenomethionine crystals of the V1 domain of VCBP3 were isomorphous to each other ($59.2 \times 59.2 \times 79.3$ Å; $P3_121$) and diffracted to extremely high resolution (1.2 Å or higher; Fig. 1 and Supplementary Tables 1 and 2 online). Using native data, we refined the VCBP3 V1 structure to a resolution of 1.15 Å (ref 22), the highest level of refinement reported so far to our knowledge for a V-type immunoglobulin domain. We assessed secondary structure and hydrogen (H)–bonding patterns between strands²³ and found that VCBP3 V1 adopted the V-type immunoglobulin fold²⁴. The C' and C'' strands of VCBP3 V1 were part of a main-chain H-bond network stabilizing the front sheet (Fig. 1a), in contrast to C-type, S-type and H-type immunoglobulin domains, which do not contain C'' strands². The classification of VCBP domains in an invertebrate (*B. floridae*) as V type suggests that the V domain had already been adopted for use in diversified families of putative receptor molecules by the time the vertebrate lineage separated from invertebrates in evolution.

To our knowledge, VCBP3 V1 is the first crystal structure of a V-type immunoglobulin domain solved to a resolution limit at which it is possible to find ordered H atoms²⁵ (1.2 Å or higher; Fig. 2). The level of refinement achieved allowed us to infer previously unrecognized structural features that probably stabilize the core of the VCBP3 V1, providing further insight into the functional relevance of evolutionarily conserved architectural features of the immunoglobulin fold. Specifically, H atoms apparent in difference electron-density maps (found by subtracting the calculated diffraction data of the final structure from the observed data) support the idea of a function for noncanonical interactions between key conserved residues at the core of the VCBP3 V1 V-type immunoglobulin domain. Those interactions include molecular interactions involving marginally polarized bonds, such as a carbon–hydrogen (CH) bond. Most noncanonical interactions²⁶ consist of aliphatic hydrogen–carbonyl oxygen (CH–OC) H bonds and π – π interactions. The H-atom donors in noncanonical interactions often involve α - or β -carbons of an amino acid. In other noncanonical interactions, π electrons from aromatic side chains can function as H-atom acceptors, such as C α H– π (Fig. 2). There were several noncanonical interactions in the core of the VCBP3 V1, demonstrated by electron density noted at the positions representing participating H atoms. These noncanonical interactions were at the invariant tryptophan residue (Trp42 in VCBP3 V1) packed against the conserved intrachain disulfide bond (Cys27–Cys110; Fig. 2), a hyperconserved structural hallmark of immunoglobulin domains. Trp42 was involved in four noncanonical interactions: one main-chain–main-chain interaction (CH–OC), two main-chain–side-chain interactions (CH– π) and one side-chain–side-chain (π – π) π stacking interaction. Cys27 and Cys110, which form an intrachain disulfide bond, also demonstrated critical noncanonical interactions. For example, we identified a main-chain–side-chain interaction (CH– π) between Cys110 and Trp42. Moreover, electron density representing the ordered H atoms participating in these key stabilizing interactions was apparent in the core of VCBP3 V1 (Fig. 2).

In addition to those structural features common to all immunoglobulin domain subtypes (an invariant tryptophan residue packed against the intrachain disulfide bond), we found that most invariant conserved positions in V domains (seven of eight) had a highly distinctive pattern of noncanonical H-bond interactions (Table 1). The expanded definitions of canonical and noncanonical interactions between invariant core residues presented here provide a new basis for understanding the patterns of strong evolutionary conservation of the

V-type immunoglobulin fold. However, there are no solved structures of V-type immunoglobulin domains in the Protein Data Bank (PDB) of atomic resolution (1.2 Å or higher) other than VCBP3 V1 (1.15 Å). The combined stabilizing effect of canonical and noncanonical interactions in the core of vertebrate adaptive immune receptors remains to be determined, although such effects are strongly predicted by our study, as the hydrophobic V-type immunoglobulin core is highly conserved⁶.

Despite the insight gained from the atomic-resolution crystal structure of VCBP3 V1, it is most likely that VCBP function is mediated through V1-V2 interactions. We expressed, purified and crystallized VCBP3 (paired domains: V1·V2) to determine structural relationships between the tandem immunoglobulin domains of VCBP3. Native and selenomethionine substituted forms of the V1·V2 crystal were isomorphous to each other and belonged to the space group P6₁, with unit cell dimensions of $a = b = 109.4$ Å and $c = 48.3$ Å. We collected X-ray data from native and selenomethionyl protein crystals and refined those to 1.85 Å (Supplementary Tables 1 and 2).

The two tandem immunoglobulin domains of VCBPs had been predicted to have V-set immunoglobulin folds based on the presence of conserved residues that are canonical in antigen receptor V domains⁶, despite a relatively low degree of sequence similarity to V region-containing antigen-binding receptors (24%)¹⁵. The crystal structure of the V1·V2 dimer demonstrated one molecule in the asymmetric unit; both VCBP3 domains had strand topologies characteristic of the V-set immunoglobulin domains used by rearranging antigen receptors. Structural analysis with the Search for Structural Motifs program²⁷ showed that both V-set immunoglobulin domains of VCBP3 adopt the V-type immunoglobulin fold and were extremely similar to each other at the structural level (Fig. 1b). The degree of similarity between VCBP3 V1 and V2 was higher than any solved structure in the PDB with an r.m.s.d. of 1.4 Å for Cα atoms. The exceptionally high degree of structural similarity between V1 and V2, despite a relatively low degree of sequence identity (about 22%), suggests that both domains are under strong selective pressure to maintain a highly specific topography that promotes the central feature of quaternary structure (that is, paired V-type immunoglobulin domains), which also is a dominant feature of antigen receptors used in adaptive immunity.

Both V domains of VCBP3 had conserved intrachain disulfide bonds linking the back and front sheets by the B and F strands, respectively (Cys27-Cys110 in V1, and Cys160-Cys232 in V2). V2 also contained a disulfide bond linking the B and C strands (Cys162-Cys165) stabilizing the BC loop, which is analogous to CDR1 in antigen receptors. Notwithstanding the high degree of similarity in the folding of V1 and V2, among solved structures in the PDB aligned by the Search for Structural Motifs program²⁷, V1 was most similar to the TCR V_δ domain of PDB accession number 1TVD:B²⁸ (Fig. 1a), followed by the central nervous system—specific autoantigen myelin oligodendrocyte glycoprotein (PDB accession number 1PY9A)²⁹ and the human coxsackie and adenovirus receptor D1 (PDB accession number 1EAJ:B³⁰; r.m.s.d., 1.7–1.8 Å; sequence identity, 22–26%). The front sheet of VCBP3 V1 (A'GFCC'C'') was more similar to that of V_β, V_γ, V_δ, V_H and V_L than to that of V_α (A'GFCC'), in which the C'' strand of V_α is connected by H bonds to the back sheet (ABEDC'')^{14,31}. Again, it is notable that despite the considerable structural identity, the actual predicted peptide sequences had relatively low overall relatedness.

Three-layer packing and J-region structure in VCBP3

The dimer interface between V domains in antigen receptors consists of a specialized three-layer packing mode in which side chains from residues in the highly twisted edge strands, C' and G, interact to comprise the inner layer, whereas the two outer layers are formed by main and side chains of the central β-strands, F and C. In antigen receptor V domains, the

highly twisted edge strands fold over the inner strands of the front sheet to mediate critical contacts that form the inner core layer of the three-domain packing interface⁵ (Fig. 3a). A three-layer interface was formed between V1 and V2 of VCBP3, and similar interface residues in the C' and G strands interacted (Fig. 3b). Main-chain-to-side-chain H-bond interactions that comprise the outer layer were present between V1 and V2 of VCBP3 (Fig. 4a), in which Arg106 from V1 formed four H-bond interactions across the V1-V2 interface to Glu240, Leu241 and Ala243 in V2. The specialized three-layer packing mode used by VCBP3 is found in antigen receptor heterodimers and differs from the frequently noted two-layer interface that occurs in most protein-protein interactions in which β -sheets interact.

The structural importance of J-region gene segments in antigen receptor dimerization and combinatorial diversity arises from a conserved sequence motif, FGXGTXLXV (where 'X' indicates any amino acid). The amino acid sequences of VCBPs do not contain recognizable consensus J regions, which has been regarded as a substantial variation from the characteristic features of known antigen-binding receptors³². However, the packing of V domains in the structure of VCBP3 was likewise very similar to that of antigen receptors in that both use the G strands encoded by J region-like sequence elements. Notably, the crystal structure of VCBP3 showed that J region-related sequences were present in the G strands at the V1-V2 front-sheet-to-front-sheet interface (Fig. 4b). V-domain G-strand residues, which are encoded by J gene segments, participate in the unique front-sheet-to-front-sheet dimerization mode⁶ used exclusively by antigen receptors. In VCBP3, the J region-analogous sequence region is likewise intimately involved in V-domain packing (Fig. 3).

Opposing orientation of CDR-analogous loops

Heterodimeric antigen receptors in jawed vertebrates pack in a head-to-head way, which positions the hypervariable CDR loops in a way that forms a binding site composed of two V domains, each of which are present on and form an individual subunit; for example, heavy and light chains in immunoglobulins. The V regions in VCBP3 packed in a head-to-tail relationship (Fig. 5a,b) similar to those found in crystal structures of a V_{δ} domain (PDB accession number 1TVD)²⁸ and cell adhesion molecules such as myelin oligodendrocyte glycoprotein (PDB accession number 1PY9:A)²⁹. In this configuration, the VCBP3 residues that correspond to jawed vertebrate CDRs are opposed 180° and cannot cooperatively define a paired V combining site analogous to that seen in the heterodimeric immunoglobulins and TCRs (such as V_H , V_L , V_{α} and V_{β}). If the regions of VCBPs that correspond to the CDRs of the rearranging antigen-binding receptors are involved in binding, it would be through an entirely unrecognized mechanism.

VCBP hypervariation is on a contiguous surface

The hypervariability in antigen-binding receptors is regionalized to the sequences that correspond to the CDR loops, which are combined in paired V domains (combinatorial diversity) to form antigen receptor repertoires with diverse specificities. As described above, the hypervariation noted in VCBPs did not occur in the CDR analogous loops, nor did the relative orientation of those loops in the solved structure of VCBP3 defined here support the idea that those positions are involved in the cooperative function of a single binding surface. At least five sequence-related families of VCBPs (36–46% identity) are found in amphioxus and their structural folds can be modeled with high confidence against the solved structure presented here. Hyper-variation in VCBP2, the most extensively characterized VCBP family in terms of regionalized genetic variation^{15,16}, was located mostly in the N-terminal portions (A, A' and B strands and their connecting loops) of both the V1 and V2 domains, in contrast to the characteristic placement of hypervariable residues in immunoglobulins and TCRs (BC loop-CDR1, C'C'' loop-CDR2, FG loop-CDR3). We mapped the hypervariable positions of VCBP2 onto the solvent-accessible surface of VCBP3 to glean information about

potential ligand-binding sites. As with all VCBPs, VCBP2 family V domains were predicted to pack like antigen-binding receptors, which would position their hypervariable residues on a contiguous patch of solvent-accessible surface defined by both V domains. We determined the putative interaction site at a hypervariable region composed of two separate domains (Fig. 5c,d) with a solvent-accessible surface of approximately 800–1,200 Å², which is consistent with the range of surface area buried between antibody-antigen and in complexes of TCR–peptide–major histocompatibility complex in solved crystal structures^{33–35}.

The solved structure of VCBP3 also shows how different V domains use different surfaces for intermolecular binding interactions. TCR and immunoglobulin V domains use the membrane-distal CDR loops to bind antigen or complexes of peptide and major histocompatibility complex, respectively (Fig. 6). The human coxsackie and adenovirus receptor D1 (PDB accession number 1EAJ:B)³⁰ binds the adenovirus knob fiber protein through residues located on the front sheet (A''GFCC'C''). The high degree of contiguous hypervariation reported before for VCBP2, modeled on the solved structure of VCBP3, is located on the edge of the V domains (A, A', B and connecting loops; Fig. 6) and is consistent with involvement of these residues in specific binding processes. Alternative surfaces of the conserved V domain have specific functions and/or high sequence variation, such as the front sheet binding mode of recognition seen in viral receptors, the edge of the V domain in VCBPs and sites formed by combined (paired) CDR1-CDR3 found in the rearranging adaptive immune receptors.

Discussion

A variety of solutions to the general problem of immune protection has emerged during the evolution of invertebrates and vertebrates. Specifically, many different immunoglobulin domain and leucine-rich repeat-containing molecules, as well as several different immunoglobulin domain–lectin chimeras, have extensive germline and somatic variation, which in at least one case may approach the variation found in rearranging immune receptors in mammals¹¹. The V1 and V2 domains of VCBP3 are the first examples to our knowledge of highly diversified, V-set immunoglobulin domains that have been identified in invertebrates.

Of the various molecules being investigated as alternative mediators of innate and anticipatory immune responses, the VCBPs are of particular interest from several standpoints. Like immunoglobulins and TCRs, they can be categorized into distinct V families. VCBPs also have a very high degree of sequence polymorphism^{15,16}, which ongoing studies suggest is even greater than that reported before (N. Schnitker, L. Dishaw and G.W.L., unpublished observations). Furthermore, all of the other known diversified families of V region-containing molecules, such as immunoglobulins, TCRs, Siglec, TREM–TREM-like transcript and CMRF-35-like molecule, as well as products of genes identified in jawless vertebrates^{36,37}, are expressed by immune cells and seem to mediate leukocyte regulation or function. Novel immune-type receptors, a particularly diversified family of V region-containing genes, show allogeneic binding specificity¹¹ (J.P.C., D.D. Eason and G.W.L., unpublished observations). Other single V region-containing molecules function as natural killer cell activating receptors, such as NKp30 and NKp44, whereas others are involved in immune recognition (reactivity) or serve as coreceptors. There is at least one other case of a V region-containing molecule in an invertebrate; however, it seems to function in neuronal development³⁸. All diversified families of V regions described so far seem to mediate immune recognition or are expressed by cells involved in immune-type functions.

The studies reported here demonstrate that the VCBPs are related closely on a structural basis to the rearranging antigen-binding receptors of jawed vertebrates. Not only are the constituent immunoglobulin domains of VCBPs prototypic V-set domains, but they also show structural similarities to the more recently derived antigen receptors in folding and packing interactions. Specifically, V1 and V2 form an intrachain dimer that resembles the interchain dimer formation seen in the rearranging antigen-binding receptors. Like their more recently derived counterparts, VCBPs also have three-layer packing. The particularly long evolutionary history of these unique modes of packing and dimerization, established in the solved VCBP structure, emphasizes their function in creating advantageous binding properties of paired V domains, even though the mode of dimerization of VCBPs and TCRs is very different and the putative site that may be involved in specific VCBP interactions is displaced.

The presence of a J region-like structure in VCBP3 provides another indication that the VCBPs are in the same structural family as the antigen receptors and other immunologically relevant molecules. However, the J region-like segment in VCBP3 was not found in a contiguous structure, which distinguishes it from V domains of rearranging antigen-binding receptors and from other J region-containing structures, such as CD8 β , which has a J region-like sequence motif and has the characteristic β -bulge in the middle of the G strand, as in antigen receptors. If VCBPs better reflect the structure of an ancient form of V-type immune recognition molecule than can be inferred from V region-containing molecules found in vertebrates, it is not unreasonable to predict that a recombination event between short, separated gene segments, equivalent to FGXD in V1 and TXLXV in V2 of VCBP3, could have given rise to a single exon corresponding to the exons encoding J regions in the more modern rearranging antigen-binding receptors¹⁵. Notably, in VCBP4, a proto-typic FGXG segment is found in V1 in the position corresponding to the FGXD of VCBP3 shown here. Although the association defined in the J region-homologous stretches of V1 and V2 could be fortuitous, it is apparent that the J region-like sequences function in domain packing of V regions, a likely characteristic of the ancestral immune-type receptors that predated the receptors that function as mediators of recombination-activating gene-dependent adaptive immunity.

One theory of primordial antigen receptors is that they would have characteristics of cell adhesion molecules³⁹, some of which share features with contemporary antigen receptors, such as incorporation of the V-set immunoglobulin fold⁴⁰. For reasons based mainly on their function and the endogenous character of their ligands, cell adhesion molecules tend to show limited polymorphism in a species. In contrast, postulated changes from a self to a self-nonselving-distinguishing mode of function in the progenitor of an immune-type receptor would involve the emergence of highly variable combining sites, especially given the propensity of many pathogens to use genetically sophisticated mechanisms to avoid host defense⁴¹ and a corresponding increase in polymorphism of those molecules in a given species.

In addition to identifying a unique V domain-packing mode in VCBP3, our findings strongly suggest that molecules such as VCBPs or their forerunners are adapted to processes such as effecting immune recognition through not only the unique nature of their dimerization but also the clustering of hypervariable residues at the solvent-accessible surface comprising both V domains. Natural selection for the latter effect, which is loosely analogous to combining-site formation in heterodimeric antigen receptors, could have accounted for stabilization of diversified genes encoding V regions into the germline before the innovation of a basic mechanism of somatic reorganization in vertebrates. The mechanisms by which such dynamic changes in receptor diversification came about are becoming increasingly more apparent^{42,43}. The VCBPs in amphioxus may represent

contemporary reflections of a transitional step from a monomorphic germline V-type receptor to a somatically variable CDR-based mechanism of antigen recognition. The suggestion that amphioxus could share features with the ancestor of all deuterostomes⁴⁴ further emphasizes the relevance of both this species and the VCBPs in the evolution of immune-type recognition.

Methods

Constructs and expression

The cDNA construct and expression methods for VCBP3 V1 have been described⁴⁵. The cDNA sequences encoding VCBP3 V1V2 were amplified by PCR and were cloned into a bacterial expression vector for expression and refolding. Oligonucleotides were designed based on the VCBP3 cDNA sequence (GenBank accession number AF520474): VCBP3V1-XC-F1, 5'-ATGCAGTCCATCATGACCGTCCGCA-3' (ATG start codon plus VCBP3 nucleotides 49–70) and VCBP3V2-XC-R1, 5'-CTATCAGACCTTGAGAATGGTTGAGGAC-3' (antisense stop codons CTA and TCA plus VCBP3 antisense nucleotides 801–787). The final peptide encoded by the PCR amplicon represents the two tandem V domains of VCBP3, beginning at the first residue after the predicted signal peptide and extending through the end of the C-terminal immunoglobulin-like domain (amino acids 17–267 of GenBank accession number AAN62850). PCR products were ligated into pETBlue-1 (EMD Biosciences) and were sequenced for confirmation. The construct was then transformed into the *Escherichia coli* Tuner strain (EMD Biosciences) for expression induced by isopropyl β -D-1-thiogalactopyranoside. Cultures were grown at 37 °C to an absorbance at 600 nm of 0.5–0.9, and 100 mM isopropyl β -D-1-thiogalactopyranoside was added to a final concentration of 1 mM. Cultures were grown an additional 5 h at 37 °C. Induced bacterial cultures were centrifuged and were stored overnight in 20% sucrose and 10 mM EDTA.

Refolding and purification

The methods used for VCBP3 V1 and tandem V1-V2 domains are similar to those reported before⁴⁵. Thawed bacteria were brought to a volume of 200 ml in sucrose and EDTA, and egg-white lysozyme (Sigma-Aldrich) was added at a concentration of 1 mg/ml to the bacterial slurry, followed by processing in an EmulsiFlex C5 high-pressure homogenizer (Avestin) at 10,000 p.s.i. for two cycles. Phenylmethanesulfonyl fluoride was added to a final concentration of 0.1 mM, and 60 μ l of lysonase recombinant lysozyme and 60 μ l of benzonase nuclease (EMD Biosciences) were added. The homogenate was incubated at 23 °C for 20 min. Centrifugation for 25 min at 15,000g separated inclusion bodies from soluble components. Inclusion bodies were washed with 10 mM Tris, pH 8.0, 5 mM EDTA and 0.1% Triton X-100, pH 8.0, followed by four to six alternating washes in 10 mM Tris, pH 8.0, and deionized water. SDS-PAGE confirmed that the inclusion bodies contained inducible protein in a major band of the predicted size. Inclusion bodies were then solubilized overnight at 4 °C in 10 ml of 7.8 M guanidinium HCl and 50 mM Tris, pH 8.0. Solubilized protein (about 5–10 mg/ml) was centrifuged at 100,000g for 30 min to remove nucleated aggregates and then was added immediately to a Tris[2-carboxyethyl]phosphine hydrochloride (Pierce) disulfide reducing agarose gel column with a bed volume of 2.5 ml. The column eluate was slowly dripped over 4 h into about 300 ml of cold, stirred refolding buffer (0.55 M guanidine, 0.44 M L-arginine, 55 mM Tris, pH 8.2, 21 mM NaCl, 0.88 mM KCl, 1 mM EDTA, 1 mM reduced glutathione and 1 mM oxidized glutathione), was brought to a temperature of 23 °C over 1 h in a refrigerated water bath and was dialyzed overnight at 4 °C against 10 mM Tris and 50 mM NaCl, pH 8.0. The dialysate was centrifuged at 15,000g for 20 min and was concentrated 10- to 25-fold in an Amicon ultrafilter with a PM10 membrane (Millipore). Clarified concentrate was then separated by fast-performance

liquid chromatography on a Superdex 75 column (GE Healthcare) and the appropriate size fraction was collected as the final purification step. Final recovery of purified VCBP3 protein was approximately 10 mg per liter of starting culture.

Crystallization

For VCBP3 V1, hanging-drop vapor-diffusion crystallization experiments were done at 18 °C with drops consisting of 2 μ l of protein at a concentration of 4 mg/ml plus 2 μ l of reservoir solution. Drops were equilibrated on siliconized slides against 1 ml of reservoir solution (1.2–1.4 M ammonium sulfate, 0.1 M NaCl, 0.1 M HEPES, pH 8.5, and 12% glycerol). For VCBP3 V1-V2, sparse matrix crystallization⁴⁶ screens from Hampton Research were used to identify crystallization conditions: Crystal Screen 1 and Crystal Screen 2. Hanging drops equilibrated against 1 ml of reservoir solution yielded crystals of VCBP3 by the vapor-diffusion method, in which 2 μ l of protein solution (10 mg/ml in 10 mM Tris, pH 8.0, and 50 mM NaCl) and an equal volume of ‘mother liquor’ were mixed on siliconized slides. The data reported here were obtained from crystals grown in 1.5–2.0 M sodium formate, 0.2 M MgCl and 0.1 M citrate at a pH of 6.6 or 4.6 for native or selenomethionine derivative data, respectively. Crystals were successfully cryoprotected by being soaked in either glycerol, for VCBP3 V1, or 20% ethylene glycol, for VCBP3 V1-V2.

Data collection and processing

For VCBP3 V1, three separate multiple-wavelength anomalous dispersion experiments were done in which data sets were collected at three wavelengths (selenium inflection point, peak and remote) from three single crystals (diffracting to 1.7 Å, 1.4 Å and 1.17 Å) of VCBP3 D1 selenomethionyl derivative at beamline X6A of the National Synchrotron Light Source (Brookhaven National Lab, Upton, New York). Short exposures (5 s) were collected at each wavelength to reduce the effects of radiation damage, yielding three data sets from a single crystal complete from 30 Å to 1.3 Å. Longer exposures (30–60 s) were subsequently collected for higher-resolution reflections (to 1.17 Å). A complete native data set with a resolution of 1.15 Å was collected at beamline X6A. All data sets were integrated and scaled with the DENZO and SCALEPACK programs⁴⁷. The refined image scale factors in SCALEPACK for each frame were evaluated for evidence of radiation damage. Data collection and reduction statistics for the native crystal and the multiple-wavelength anomalous dispersion experiment of highest resolution are in Supplementary Table 1. VCBP3 V1V2 crystals were cryocooled by being immersed in liquid nitrogen and data were collected at beamline X6A of the National Synchrotron Light Source. A single-wavelength anomalous dispersion experiment was done by collection of diffraction data from a single selenomethionine derivative crystal at the peak of its absorption edge. Derivative data were collected with oscillation angles of 0.5° and 20-second exposures. Although pH affected the morphology of the crystals, diffraction data at both high pH (6.6) and low pH (4.6) corresponded to the space group P6₁-P6₅.

Intensities were indexed and integrated with DENZO and reduced with SCALEPACK⁴⁷. The calculated molecular weight of the crystallized polypeptide and post-refined unit-cell dimensions were submitted to the Matthews Probability Calculator⁴⁸ to obtain information about the contents of the asymmetric unit. Given a cell volume of 507,762.4 Å³ and a molecular weight of 27,373 Da, a single molecule of VCBP3 V1V2 is found at the asymmetric unit. The corresponding Matthews coefficient is 3.09 Å²/Da and the solvent content is 60.2%. Systematic absences were consistent with screw axes of 1/6 and 5/6 unit cell translations.

Native data reduction statistics reflect the quality of two different data sets each collected from independent crystals that were reduced and scaled together before post-refinement

iterations and outlier rejections in SCALE-PACK. A total of 600 frames of native diffraction data collected at 15-second exposures and 0.4° oscillations were merged in this way (Supplementary Table 1, collection parameters).

Phasing and refinement

For VCBP3 V1 crystals grown in sodium formate, the space group and unit cell dimensions are consistent with those reported for crystals grown in sodium acetate, P3₁21 (ref. 45). The SOLVE/RESOLVE software package⁴⁹ was used to identify a single selenium atom in the asymmetric unit, calculate multiple-wavelength anomalous dispersion phases, electron-density maps, apply density modification and automatically trace the chain. Experimental phases were calculated to a resolution of 1.3 Å, and model building of 127 (of 135) residues proceeded unambiguously; every amino acid side chain was visible in experimental electron-density maps. Building was done with the O9.0 program⁵⁰. Native data were refined initially with the Crystallography & NMR system⁵¹ and were confirmed with PROCHECK⁵² to a resolution of 1.5 Å. Higher-resolution refinement to 1.15 Å was done with SHELXL⁴⁶ using anisotropic thermal-displacement refinement. Refinement statistics are in Supplementary Table 2. Xtalview⁵³ was used to monitor refinement of solvent structure and split side-chain conformations at atomic resolution. Identical test sets of reflections were used to calculate R_{free} statistics and the test set of reflections comprise 11.3% of the data sets.

For VCBP3 V1-V2, the selenomethionine derivative diffraction data was successfully phased with the software package SOLVE⁴⁹ by the single-wavelength anomalous dispersion method. Initial experimental electron-density maps were evaluated with the Coot tool⁵⁴. Statistical density modification was done with RESOLVE. For VCBP3 V1-V2, model building of the main chain was done with Coot during rounds of refinement against derivative diffraction data. VCBP3 was refined with Shelx97 program²² and the Crystallography & NMR system⁵¹. The Xtalview program⁵³ was also used for molecular graphics. Most of the main chains and side chains were built before being refined against native diffraction data, in which the same test set of reflections was used to calculate R_{free} . Several rounds of refinement were done before water molecules began to be assigned through multiple cycles of refinement until most peaks with signal/noise ratios greater than 2 were occupied. The Crystallography & NMR system was used to assign the first solvent shell with the automatic water-picking software, but in later stages of refinement, SHELXL was also used to assign solvent molecules to $F_o - F_c$ sigma-weighted electron-density maps. Refinement continued until convergence for the model was achieved (PDB accession number 2FBO). Statistics are summarized in Supplementary Table 2.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank B. Pryor for editorial assistance. Supported by the National Institutes of Health (R01 AI23338 to G.W.L. and R01 DE013883 and R21 HL080222 to D.A.O.), the Cure Autism Now Foundation (2908051-12 to D.A.O.), the US Department of Energy (DE-AC02-98CH10886 for beamline X6A) and the National Institutes of Health, National Institute of General Medical Sciences (GM-0080 for beamline X6A).

References

1. Orengo CA, Thornton JM. Protein families and their evolution—a structural perspective. *Annu Rev Biochem.* 2005; 74:867–900. [PubMed: 15954844]

2. Bork P, Holm L, Sander C. The immunoglobulin fold. Structural classification, sequence patterns and common core. *J Mol Biol.* 1994; 242:309–320. [PubMed: 7932691]
3. Barclay AN. Membrane proteins with immunoglobulin-like domains—a master super-family of interaction molecules. *Semin Immunol.* 2003; 15:215–223. [PubMed: 14690046]
4. Hunkapiller T, Hood L. Diversity of the immunoglobulin gene superfamily. *Adv Immunol.* 1989; 44:1–62. [PubMed: 2646860]
5. Eason DD, et al. Mechanisms of antigen receptor evolution. *Semin Immunol.* 2004; 16:215–226. [PubMed: 15522620]
6. Chothia C, Gelfand I, Kister A. Structural determinants in the sequences of immunoglobulin variable domains. *J Mol Biol.* 1998; 278:457–479. [PubMed: 9571064]
7. Garrett TP, Wang J, Yan Y, Liu J, Harrison SC. Refinement and analysis of the structure of the first two domains of human CD4. *J Mol Biol.* 1993; 234:763–778. [PubMed: 8254672]
8. Leahy DJ, Axel R, Hendrickson WA. Crystal structure of a soluble form of the human T cell coreceptor CD8 at 2.6 Å resolution. *Cell.* 1992; 68:1145–1162. [PubMed: 1547508]
9. Cantoni C, et al. The three-dimensional structure of the human NK cell receptor NKp44, a triggering partner in natural cytotoxicity. *Structure (Camb).* 2003; 11:725–734. [PubMed: 12791260]
10. Laird DJ, De Tomaso AW, Cooper MD, Weissman IL. 50 million years of chordate evolution: seeking the origins of adaptive immunity. *Proc Natl Acad Sci USA.* 2000; 97:6924–6926. [PubMed: 10860947]
11. Litman GW, Cannon JP, Dishaw LJ. Reconstructing immune phylogeny: new perspectives. *Nat Rev Immunol.* 2005; 5:866–879. [PubMed: 16261174]
12. Desmyter A, et al. Crystal structure of a camel single-domain V_H antibody fragment in complex with lysozyme. *Nat Struct Biol.* 1996; 3:803–811. [PubMed: 8784355]
13. Stanfield RL, Dooley H, Flajnik MF, Wilson IA. Crystal structure of a shark single-domain antibody V region in complex with lysozyme. *Science.* 2004; 305:1770–1773. [PubMed: 15319492]
14. Garcia KC, et al. Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. *Science.* 1998; 279:1166–1172. [PubMed: 9469799]
15. Cannon JP, Haire RN, Litman GW. Identification of diversified genes that contain immunoglobulin-like variable regions in a protochordate. *Nat Immunol.* 2002; 3:1200–1207. [PubMed: 12415263]
16. Cannon JP, Haire RN, Schnitker N, Mueller MG, Litman GW. Individual protochordates possess unique immune-type receptor repertoires. *Curr Biol.* 2004; 14:R465–R466. [PubMed: 15203016]
17. Davis MM, Bjorkman PJ. T-cell antigen receptor genes and T-cell recognition. *Nature.* 1988; 334:395–402. [PubMed: 3043226]
18. Medzhitov R, Janeway CA Jr. Advances in immunology: innate immunity. *N Engl J Med.* 2000; 343:338–344. [PubMed: 10922424]
19. Zhang SM, Adema CM, Kepler TB, Loker ES. Diversification of Ig superfamily genes in an invertebrate. *Science.* 2004; 305:251–254. [PubMed: 15247481]
20. Watson FL, et al. Extensive diversity of Ig-superfamily proteins in the immune system of insects. *Science.* 2005; 309:1874–1878. [PubMed: 16109846]
21. Wang JW, et al. SAD phasing by combination of direct methods with the SOLVE/RESOLVE procedure. *Acta Crystallogr D Biol Crystallogr.* 2004; 60:1244–1253. [PubMed: 15213386]
22. Schneider TR, Sheldrick GM. Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr.* 2002; 58:1772–1779. [PubMed: 12351820]
23. Carter P, Andersen CA, Rost B. DSSPcont: Continuous secondary structure assignments for proteins. *Nucl Acids Res.* 2003; 31:3293–3295. [PubMed: 12824310]
24. Holm L, Sander C. Protein folds and families: sequence and structure alignments. *Nucl Acids Res.* 1999; 27:244–247. [PubMed: 9847191]
25. Dauter Z, Lamzin VS, Wilson KS. The benefits of atomic resolution. *Curr Opin Struct Biol.* 1997; 7:681–688. [PubMed: 9345627]
26. Babu MM. NCI: A server to identify non-canonical interactions in protein structures. *Nucl Acids Res.* 2003; 31:3345–3348. [PubMed: 12824323]

27. Kolodny R, Koehl P, Levitt M. Comprehensive evaluation of protein structure alignment methods: scoring by geometric measures. *J Mol Biol.* 2005; 346:1173–1188. [PubMed: 15701525]
28. Li H, et al. Structure of the V δ domain of a human $\gamma\delta$ T-cell antigen receptor. *Nature.* 1998; 391:502–506. [PubMed: 9461220]
29. Clements CS, et al. The crystal structure of myelin oligodendrocyte glycoprotein, a key autoantigen in multiple sclerosis. *Proc Natl Acad Sci USA.* 2003; 100:11059–11064. [PubMed: 12960396]
30. van Raaij MJ, Chouin E, van der Zandt H, Bergelson JM, Cusack S. Dimeric structure of the coxsackievirus and adenovirus receptor D1 domain at 1.7 Å resolution. *Structure.* 2000; 8:1147–1155. [PubMed: 11080637]
31. Ostrov DA, Shi W, Schwartz JC, Almo SC, Nathenson SG. Structure of murine CTLA-4 and its role in modulating T cell responsiveness. *Science.* 2000; 290:816–819. [PubMed: 11052947]
32. Kaufman J. The origins of the adaptive immune system: whatever next? *Nat Immunol.* 2002; 3:1124–1125. [PubMed: 12447367]
33. Rudolph MG, Luz JG, Wilson IA. Structural and thermodynamic correlates of T cell signaling. *Annu Rev Biophys Biomol Struct.* 2002; 31:121–149. [PubMed: 11988465]
34. Luz JG, et al. Structural comparison of allogeneic and syngeneic T cell receptor-peptide-major histocompatibility complex complexes: a buried alloreactive mutation subtly alters peptide presentation substantially increasing V(β) Interactions. *J Exp Med.* 2002; 195:1175–1186. [PubMed: 11994422]
35. Lo Conte L, Chothia C, Janin J. The atomic structure of protein-protein recognition sites. *J Mol Biol.* 1999; 285:2177–2198. [PubMed: 9925793]
36. Suzuki T, Shin I, Fujiyama A, Kohara Y, Kasahara M. Hagfish leukocytes express a paired receptor family with a variable domain resembling those of antigen receptors. *J Immunol.* 2005; 174:2885–2891. [PubMed: 15728499]
37. Haruta C, Suzuki T, Kasahara M. Variable domains in hagfish: NICIR is a polymorphic multigene family expressed preferentially in leukocytes and is related to lamprey TCR-like. *Immunogenetics.* 2006; 58:216–225. [PubMed: 16541254]
38. Seeger MA, Kaufman TC. Characterization of amalgam: a member of the immunoglobulin superfamily from *Drosophila*. *Cell.* 1988; 55:589–600. [PubMed: 3141062]
39. Streltsov VA, et al. Structural evidence for evolution of shark Ig new antigen receptor variable domain antibodies from a cell-surface receptor. *Proc Nat Acad Sci USA.* 2004; 101:12444–12449. [PubMed: 15304650]
40. Cannon JP, Haire RN, Rast JP, Litman GW. The phylogenetic origins of the antigen binding receptors and somatic diversification mechanisms. *Immunol Rev.* 2004; 200:12–22. [PubMed: 15242392]
41. Raskin DM, Seshadri R, Pukatzki SU, Mekalanos JJ. Bacterial genomics and pathogen evolution. *Cell.* 2006; 124:703–714. [PubMed: 16497582]
42. Chatterji M, Tsai CL, Schatz DG. New concepts in the regulation of an ancient reaction: transposition by RAG1/RAG2. *Immunol Rev.* 2004; 200:261–271. [PubMed: 15242411]
43. Fugmann SD, Messier C, Novack LA, Cameron RA, Rast JP. An ancient evolutionary origin of the *Rag1/2* gene locus. *Proc Natl Acad Sci USA.* 2006; 103:3728–3733. [PubMed: 16505374]
44. Delsuc F, Brinkmann H, Chourrout D, Philippe H. Tunicates and not cephalo-chordates are the closest living relatives of vertebrates. *Nature.* 2006; 439:965–968. [PubMed: 16495997]
45. Hernandez Prada JA, Haire RN, Cannon JP, Litman GW, Ostrov DA. Crystallization and preliminary x-ray analysis of VCBP3 from *Branchiostoma floridae*. *Acta Crystallogr D Biol Crystallogr.* 2004; 60:2022–2024. [PubMed: 15502315]
46. Weeks CM, et al. Automatic solution of heavy-atom substructures. *Methods Enzymol.* 2003; 374:37–83. [PubMed: 14696368]
47. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 1997; 276:307–326.
48. Kantardjieff KA, Rupp B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.* 2003; 12:1865–1871. [PubMed: 12930986]

49. Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr*. 1999; 55:849–861. [PubMed: 10089316]
50. Jones TA, Zou JY, Cowan SW, Kjeldgaard. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A*. 1991; 47:110–119. [PubMed: 2025413]
51. Brunger AT, et al. Crystallography & NMR system: A new software suite for macro-molecular structure determination. *Acta Crystallogr D Biol Crystallogr*. 1998; 54:905–921. [PubMed: 9757107]
52. Laskowski RA, Moss DS, Thornton JM. Main-chain bond lengths and bond angles in protein structures. *J Mol Biol*. 1993; 231:1049–1067. [PubMed: 8515464]
53. McRee DE. XtalView/Xfit—A versatile program for manipulating atomic coordinates and electron density. *J Struct Biol*. 1999; 125:156–165. [PubMed: 10222271]
54. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*. 2004; 60:2126–2132. [PubMed: 15572765]
55. Lefranc MP, et al. IMGT-ONTOLOGY for immunogenetics and immunoinformatics. *In Silico Biol*. 2004; 4:17–29. [PubMed: 15089751]

Figure 1.

Structural comparison of the VCBP3 domain fold and packing interactions with antigen receptors. **(a)** Among V-set immunoglobulin domains, VCBP3 V1 is most similar to a TCR V δ domain (in salmon superimposed on V1, in cyan); the CC' loop is curled over the front sheet (A'GFCC'C'') in V1. **(b)** VCBP3 V1 (cyan) superimposed on V2 (gold). The FG loops and G strands, encoded by J region-like elements in VCBPs, are nearly identical, whereas the BC and CC' loops of V1 are longer than those in V2.

Figure 2.

Noncanonical interactions in the core of VCBP3 V1 solved by multiple-wavelength anomalous dispersion and refined to 1.15 Å. The side chain of Trp46 and the intrachain disulfide bond (Cys27-Cys110) are in the core of VCBP3 V1. Noncanonical interactions mediated by π electrons and H atoms apparent in the $F_o - F_c$ electron density are presented at the level of 3σ (red) or 2σ (white) and are superimposed onto the final model (carbon, yellow; sulfur, orange; nitrogen, blue; oxygen, red). The $2F_o - F_c$ electron density calculated from phases to 1.15 Å is contoured at the level of 2σ (green-blue).

Figure 3.

Packing interactions in antigen receptors and paired V-type immunoglobulin domains in VCBP3. **(a)** The three-layer packing interaction characteristic of the interface between V domains of antigen receptors. TCR V α (gold and gray) interacts with the corresponding V β chain (magenta) by a front-sheet-to-front-sheet interaction in which V α side chains (cyan) from the highly twisted edge strands C' and G fold over the central strands of the V β front sheet to form an inner layer at the core of the three-layer interface between the domains. **(b)** A three-layer packing interaction at the interface between VCBP3 V1 and V2 in which residues form an inner layer and the C' loop folds over the central strands of the front sheet. VCBP3 V1 is gold (β-sheet), gray (loop regions) and red (helical regions); VCBP3 V2 is magenta; and side chains of Pro43, Leu104 and Phe106 **(a)** or Leu42, Leu122 and Ala124 **(b)** are cyan. View is rotated 180° about a vertical axis in the plane of the page compared with Figure 1.

Figure 4.

Structural elements at the V-domain interface. **(a)** V domain–packing interactions in VCBP3, including the H-bond network formed between residues across the interface between V1 and V2 of VCBP3. Interactions between side chains of one V domain and main-chain atoms of the opposing V domain form the outer layer of the three-layer packing arrangement noted in antigen receptors and VCBP3. The guanidinium group of Arg106 from VCBP3 V1 forms H-bond interactions with main-chain atoms from V2. Carbon atoms are magenta for V1 and green for V2; nitrogen atoms are blue; oxygen atoms are red. **(b)** Conformational similarities to J regions of antigen receptors: structural comparison of the G strands in antigen receptors (cyan) with VCBP3 V1 (gold). The sequence motif FGXGTXLxV is highly conserved in J gene segments of antigen receptors and corresponds to G-strand residues critical in V-domain packing. In VCBP3, the G strands of V1 and V2 are in an extended β -strand geometry, whereas antigen receptors have a conserved β -bulge at the position corresponding to FGX-(β -bulge)-G.

Figure 5.

Crystal structure of VCBP3 V1-V2 solved by single-wavelength anomalous dispersion and refined to 1.85 Å. **(a,b)** Secondary structure (β -strands, gold; loop regions, gray; helices, red). The G strand and FG loop encoded by a J gene segment-like element is cyan. Loops corresponding to CDR regions in TCR and immunoglobulin: BC loop, CDR1; C'Cx02033, CDR2; FG, CDR3. **(c,d)** The molecular surface of VCBP3 V1-V2. The molecular surface of V1 is magenta and that of V2 is violet. Polymorphic residues in VCBP sequences (gold) form a contiguous patch of solvent-exposed hypervariable residues (outlined by a green dashed rectangle). The views in **b,d** are rotated 180° about a vertical axis in plane of the page with respect to **a,c**.

Figure 6.

The distribution of known binding sites in V domains occupying alternative surfaces in the V-type fold in the TCR, a V-type viral receptor and the contiguous region of hypervariation in VCBPs. The main-chain atoms of a V_δ domain (PDB accession number 1TVD:B²⁸; red), the human coxsackie and adenovirus receptor D1 (PDB accession number 1EAJ:B³⁰; blue) and VCBP3 V1 (gold) are superimposed. The main-chain and side-chain atoms of the CDR loops in V_δ and the front-sheet residues involved in binding to the adenovirus knob fiber protein (PDB accession number 1KAC) are presented. The residues corresponding to the VCBP hypervariable edge strand residues are spheres. **(a)** V domains are oriented with the CDRs on the top surface and the front sheets, used in viral receptor interactions, facing right. The edge of VCBP3 V1 is oriented toward the viewer. **(b)** The orientation of V domains is rotated 90° around a horizontal axis on the plane of the page compared with **a**. The highly conserved V-type domain supports a particularly extensive range of different recognition-interaction functions.

Table 1
Noncanonical interactions in the evolutionarily conserved core of VCBP3

	VCBP3	Chothia ⁶	IMGT immunoglobulin ⁵⁵	Noncanonical interaction ²⁶	Electron donor	Acceptor
Gly	20	16/15	16	0	–	–
	27	23/22	23	CH-OC	0	1
	42	35/36	41	CH-PI	2	3
Trp				CH-OC		
				PI-PI		
Arg	84	61/66	75	CH-OC	0	1
Leu	95	73/80	89	CH-OC	0	1
Asp	104	82/86	98	CH-OC	0	2
Tyr	108	86/90	102	CH-PI	1	3
				CH-OC		
Cys				PI-PI		
	110	88/92	104	CH-OC	1	1
				CH-PI		

Chothia designations are light chain/heavy chain. IMGT, ImMunoGeneTics information system.