Antigen Specificity and High Affinity Binding Provided by One Single Loop of a Camel Single-domain Antibody*

Received for publication, March 8, 2001, and in revised form, May 3, 2001
Published, JBC Papers in Press, May 7, 2001, DOI 10.1074/jbc.M102107200

Aline Desmyter, Klaas Decanniere, Serge Muyldermans, and Lode Wyns
From the Department Ultrastructure, Vlaams Interuniversitair Instituut voor Biotechnologie, Vrije Universiteit Brussel, Paardenstraat 65, B-1640 Sint Genesius Rode, Belgium

Detailed knowledge on antibody-antigen recognition is scarce given the unlimited antibody specificities of which only few have been investigated at an atomic level. We report the crystal structures of an antibody fragment derived from a camel heavy chain antibody against carbonic anhydrase, free and in complex with antigen. Surprisingly, this single-domain antibody interacts with nanomolar affinity with the antigen through its third hypervariable loop (19 amino acids long), providing a flat interacting surface of 620 Å². For the first time, a single-domain antibody is observed with its first hypervariable loop adopting a type-I canonical structure. The second hypervariable loop, of unique size due to a somatic mutation, reveals a regular tertiary structure determining region; cAb, camel single-domain antibody fragment; VHH, variable domain of heavy chain of conventional antibody; VH, variable domain of heavy chain antibody; VH, variable domain of heavy chain of conventional antibody (VH).

Conventional antibody IgG molecules consist of two light chains folded in two domains and two heavy chains folded in four domains (1). Surprisingly, the serum of Camelidae contains in addition a large proportion (~50%) of functional antibodies devoid of light chains and heavy chains possessing only three domains since the equivalent of the first constant domain is missing (2). The two C-terminal domains of the heavy chain homodimers within cameldins and conventional IgG molecules share large sequence identities and are responsible for the effector functions. Also, the N-terminal variable domain of the heavy chain antibody (referred to as VH) has an overall sequence and structure that is homologous to the variable domain (VH) of the heavy chain of a classical human antibody (4–8). Important amino acid differences occur between the VH and VHH in their framework 2 region. This region is hydrophilic in VHHs rendering the domain soluble in aqueous solution, whereas the region is hydrophobic in the VH, and its amino acids associate with the VL. The VHH domain represents the smallest naturally occurring, intact antigen-binding site (9), comprising only one single immunoglobulin domain with three antigen-binding loops (or complementarity determining regions, CD) heavy chain antibodies with high specificity and affinity can be generated against a wide variety of antigens (10). Their VHHs are readily cloned (11, 12) and expressed in bacteria and yeast (13) and are extremely stable (14).

In human and mouse, the first two antigen-binding loops of a VH domain, CDR1 and CDR2, can be assigned to a limited number of possible conformations referred to as canonical structures (15–18). The conformation of these loops depends both on their length and on the presence of specific residues at key positions. In contrast, the x-ray structure analysis of four VHH domains showed that their CDR1 and CDR2 deviate significantly from the canonical loop structures observed in human or mouse VHs (19).

The third antigen-binding loop (CDR3) of the VHH fragments is often constrained by an interloop disulfide bond and is, on average, longer than a human or mouse VH-CDR3 loop (4). This allows for a potentially larger antigen-binding surface (20).

About half of the dromedary single-domain binders to enzymes are potent inhibitors (12). This can be explained by their long CDR3 loop inserting into the active site cleft on the enzyme surface, as illustrated by the lysozyme binder cAb-Lys3. In this case, the N-terminal part of the 24-amino acid-long CDR3 loop protrudes from the remaining antigen-binding surface, penetrates deeply into the active site of the enzyme (6), and mimics the lysozyme natural substrate (21). However, several non-inhibiting antibody fragments with a long CDR3 loop were also isolated (12), and these fragments are not expected to interact with the active site of their enzymes. Therefore, it was hypothesized that these non-inhibiting VHH molecules would interact with other clefts present on the protein surface.

Here, we present the crystal structures of a camel VHH fragment, cAb-CA05, both as free antibody and in complex with its antigen. This specific non-inhibiting enzyme binder recognizes the bovine erythrocyte carbonic anhydrase with an affinity of 72 nM (Kd), which is in the same range as other VHHs or...
Single-domain Antibody Structure and Antigen Recognition

RESULTS

cAb-CA05 Sequence and General Structure—cAb-CA05 is a VH antibody fragment of 135 residues (M, 15,000) that binds specifically to bovine erythrocyte carbonic anhydrase (Kd 5 72 µM) but does not inhibit the enzymatic activity of the antigen. It was selected by panning from a VHH library of an immunized dromedary (12). The cAb-CA05 was extracted from the periplasm and purified by chromatography on Ni-NTA (Qiagen) and Superdex 75 (Amersham Pharmacia Biotech) and gel filtration (12). The cAb-CA05-carbonic anhydrase complex was prepared by mixing cAb-CA05 in phosphate buffered saline with bovine erythrocyte carbonic anhydrase (Sigma) in a molar ratio of 1:2.1 and applied on Superdex 75 (Amersham Pharmacia Biotech). Crystals from the antigen-free cAb-CA05 (1.7 mg/ml) and from the complex cAb-CA05-carbonic anhydrase (3.5 mg/ml) were grown in 25% (w/v) PEG8000 (Hampton), 0.1 M sodium citrate, pH 5.6, using the hanging drop vapor diffusion method. A data set to 2.1 Å for cAb-CA05 was collected using a Rigaku RU-H2R rotating anode generator (Kobe, Japan) and a MarResearch image plate (MarResearch, Hamburg, Germany). Data for the antigen-antibody crystal were collected on beam line BW7A at EMBL-Hamburg using a MarResearch image plate. Primary data processing was done with DENZO (22), scaling was done with SCALA, and further processing was done with the CCP4 program suite (23).

Structure Determination and Refinement—The structure of the free antigen was solved by molecular replacement as implemented in AMORE (24), using the VH H cAb-RN05 (Protein Data Bank entry code 1BZQ) as a search model. The search model was refined with X-PLOR (25) and REFMAC (26). The CDR loops were deleted from the search model and from the model constructed from the cAb-CA05, the structure was refined with ARP (27) and checked manually. Model and structure factors are deposited in the Protein Data Bank, entry 1F2X.

The antigen-antibody complex structure was solved by molecular replacement with the antigen-free cAb-CA05 structure and human carbonic anhydrase structure (1CA2) as search models. For refinement, we used the CNS program (28), and included a simulated-annealing step to reduce possible model bias. As the high resolution limit of the data set is 3.5 Å, a grouped B factor refinement scheme was used (only one B factor for main chain atoms and one B factor for side chain atoms/residue). Model and structure factors are accessible through the Protein Data Bank, entry 1G6V.

Superposition of structures or structure fragments and calculation of rmsd were done with LSQKAB (23). Inter-residue and inter-atom distances were calculated with CONTACT (23), and accessible surface areas were calculated with NACCESS (29). Figures were produced with MOLSCRIPT 2.1 (30) and rendered with RASTER3D (31).

The cAb-CA05 was crystallized both as free antibody and in complex with its antigen. Antigen-free cAb-CA05 crystallizes in space group P 21, with cell dimensions of a = 29.98 Å, b = 43.86 Å, c = 87.95 Å, β = 93.23° and two VH molecules in the asymmetric unit. The antigen-antibody complexes crystallize in space group P 4 2 2 2 with cell dimensions of a = 83.86 Å and c = 224.05 Å and one antigen-antibody complex in the asymmetric unit.

The structure of the antigen-free cAb-CA05 was refined to 2.1 Å resolution (Table I). It adopts the standard fold of an immunoglobulin variable domain with nine conserved anti-parallel β-strands (Fig. 2) and three hypervariable regions (FRs) at one end of the domain (1, 33, 34). The Cys-22 and Cys-92 are oxidized into an intradomain disulfide bond, conserved in all immunoglobulin domains. Its general structure superimposes very well with a human VH reference structure (1gxm) and with all available VH structures of camel (1mle, 1hxq) and llama (1hcv, 1qdo). Root-mean-square deviations for the main chain atoms of the framework residues (residues 2–24, 32–52, 55–72, 77–92, and 103–112) ranged between 0.58 and 0.88 Å.

The table of amino acid sequence alignment of VHs and a human VH of known structure. The sequences of cAb-CA05 (this study), cAb-Lys3 (6), cAb-RN05 (5), llama VH hcg (7), RR6 (8), and human Pot VH (50) are shown. The framework, the CDRs and the amino acid numbering (bottom line) are as defined by Kabat et al. (32). The VH-specific amino acids of the framework 2 region are in bold and an asterisk indicates their location.
Structural adaptations, however, are expected to occur in the side of the domain that corresponds to the VL-interacting side of a VH domain. This area is hydrophilic in all VHs by the presence of Val-37, Leu-45, Trp-47, and Trp-103 side chains, conserved in sequence and structural position (Fig. 3A) (35). The L45R and W47I substitutions and the Trp-103 rotated over its Cβ-Cγ bond in cAb-Lys3 (Fig. 3C) make this VH region more hydrophilic. The V37F mutation fills a hydrophobic pocket created by the side chains of the Trp-103, Tyr-91, and the CDR3, where the conserved Tyr (three amino acids upstream of Trp-103) plays a central role (6). The W103R substitution found here in cAb-CA05 renders this ‘former VL side’ of the VH even more hydrophilic. It also allows a shift of the Phe-37 side chain toward the Tyr-91 and Arg-103 (Fig. 3D). As a result, the backbone of the long CDR3 approaches the former VL-side even more closely (Fig. 4D). All these modifications occur in the absence of distortions of the framework structure. In contrast, the partial camelization of a human VH in this area by L45R and W47I substitutions makes the isolated domain more soluble but induces backbone deformations at positions 37–38 and 45–47 (36). In addition, the side chain of Trp-103 takes a completely new position (Fig. 5B).

cAb-CA05 Hypervariable Regions—The conformation of the H1 loop (residues 26–32), the solvent-exposed loop around the CDR1) of cAb-CA05 fits with the canonical structure type-1, a conformation observed in all VH structures containing a 7-amino acid H1 loop (15, 16). This canonical loop structure is shaped by a sharp turn at Gly-26, clustering of the hydrophilic side chains of Ala-24, Phe-27, Phe-29, and Met-34 (Fig. 4A), and the hydrophobic part of the Arg-94 side chain (Cβ-Cε of Lys-94 in Pot VH) (16). The sequence composition of the cAbCA05 H1 loop harbors the key elements for a type-1 structure except for the R94G and conservative P27Y and P29V substitutions (Fig. 1). These substitutions lead to a slightly different organization of the side chains forming the hydrophobic core of the loop but do not influence the main chain conformation of the loop (Fig. 4A). In contrast, all previously solved camel or llama VH structures had their H1 loop folded into completely different main chain architectures.

Residues 52–56 form a hairpin loop (denoted H2) that constitutes the antigen-binding region of the second hypervariable region (15, 16). Canonical structures of the H2 loop are described for loops with sizes of five, six, or eight amino acids. We previously showed that VH H2 loops of six amino acids adopt conformations not yet observed in VH structures (19). Here, we are facing an H2 loop with only four amino acids (Fig. 1) that adopts a regular β-hairpin structure (Fig. 4B) known as type II (37). A comparison of four- and five-amino acid-long H2 loops indicates that the addition of a fifth amino acid introduces a bulge at position 55 and converts the loop to an H2 canonical structure type-1 (Fig. 4, B and C). The CDR3 (residues 95–102) of a VH is on average longer than that of a VH (17 versus 12 residues) (4), although a notable fraction of llama VHs were found with ‘short’ CDR3 loops (20, 38). Another remarkable feature of the CDR3 of VHs is the frequent presence of a cysteine forming a disulfide bond with a cysteine in the CDR1 (4, 6). In this respect, the cAb-CA05 with a 19-amino acid-long CDR3 loop and cysteines at positions 33 and 100c forming a disulfide bond is comparable with cAb-Lys3 (6) (Figs. 1 and 4D).

The cysteine at position 100c can be considered to divide the CDR3 region into an N-terminal and a C-terminal part. The C-terminal part of the CDR3 loop folds back onto the side of the VH domain corresponding to the side of the VH interacting with VL (1) (see Table II for a list of contacting residues) (39). Large parts of the former VL-side are apparently shielded from the solvent by the C-end of the CDR3. A similar location and function has been observed for the C-terminal part of the cAb-Lys3 CDR3-loop (Fig. 4D) (6) for the entire, much shorter CDR3 loop of cAb-RN05 (5) and that of RR6 llama VH (8).

The N-terminal part of the CDR3 of cAb-CA05 and cAb-Lys3 are in a different environment. In cAb-Lys3, it forms a protruding loop (Fig. 4D) inserting in the catalytic site of the lysozyme (21). In cAb-CA05, this part of the loop does not extend into the solvent but associates with the residues of the remaining hypervariable loops (Table II). Thus, the N-terminal and C-terminal half of the CDR3 of cAb-CA05 contact different parts of the domain. Furthermore, the entire (long) CDR3 of cAb-CA05 appears to be well fixed by these abundant contacts and by the covalent Cys33-Cys100c disulfide bond.
The Antigen-Antibody Complex—In addition to the antigen-free cAb-CA05 crystal, crystals of the complex of cAb-CA05 with its antigen, carbonic anhydrase, were obtained and diffracted to 3.5 Å using synchrotron radiation. The structure was solved by molecular replacement (see “Experimental Procedures”) and the data and refinement statistics for the antibody-antigen complex structure are shown in Table I.

Binding of the antibody has little influence on the overall structure of the carbonic anhydrase. The rmsd is 0.7 Å for the Cα atoms between the bovine carbonic anhydrase molecule found here and the human carbonic anhydrase used as search model. Also, the structure of the cAb-CA05 antibody in the complex is the same as the uncomplexed structure (Fig. 2). The interloop disulfide bond is shown as well as the conserved disulfide bond between the scaffold Cys-22 and Cys-92. The Kabat numbering of a few amino acids is given for reference.

The epitope consists of two separate continuous segments within the carbonic anhydrase. A first stretch includes residues 46–52 and the second stretch involves residues 180–187 (Fig. 2, Table II). The first segment uses mainly main chain atoms (21 of 29 contacts), whereas the second part of the epitope involves mainly side chain atoms (50 of 58 contacts). The para-tope comprises both, the N- and the C-terminal part of the CDR3 loop. In the N-terminal part of the CDR3, many main chain atoms participate in antigen binding; 32 of 60 contacts (i.e., antibody atoms within 4.0 Å of antigen atoms) involve an antibody main chain atom. In the C-terminal part of the CDR3, only antibody side chain atoms contact the antigen.

The antigen-interacting surface of cAb-CA05 is to a large extent planar (Fig. 5). A solvent-accessible surface area of 622 Å² becomes buried upon antigen complexation. Remarkably, the CDR3 region provides this surface entirely. The antigen makes no contacts with the CDR1 and CDR2 loops apart from Tyr-32, which is at 3.7 Å from the carbonic anhydrase molecule. This is the first antibody fragment that shows high affinity and high binding specificity using one CDR loop for direct interaction.

**DISCUSSION**

The crystal structures of cAb-CA05, both antigen-free and in complex with carbonic anhydrase, reveals the structural adaptations in the VH chain that explain its solubility and antigen-
Antibody residues contacting the CDR3 residues\(^a\) | Interacting CDR3 residues\(^b\) | Antigen residues contacting the CDR3 residues\(^c\)
---|---|---
Tyr-27, Tyr-32, Cys-33 | Ser-95 | Leu-47, Ser-48, Val-49
Tyr-32 | Thr-96 | Val-49, Leu-189
Cys-33 | Val-97 | Val-49, Ser-50, Tyr-51, Asp-52, Arg-182
Thr-31 | Ala-98 | Asp-180, Arg-182
Leu-52 | Ser-99 | Asp-180, Arg-182
Leu-52, Tyr-58 | Thr-100 | Asp-180
Gly-100b | Gly-100a | Arg-182, Pro-186, Glu-187
Trp-100b | Arg-182, Pro-186, Glu-187
Cys-33, Thr-50, Ile-51, Tyr-58 | Cys-100c\(^d\), Ser-100d | Arg-182, Gly-183, Leu-185, Glu-187
Gly-47, Val-48, Tyr-59, Leu-60 | Gly-100f | Glu-187
Gly-35, Trp-36, Phe-37, Thr-50, Phe-37, Arg-45, Gly-47 | Arg-100g, Pro-100h | Glu-187
Cys-33, Met-34, Thr-50 | Tyr-100i | Glu-187
Val-2, Leu-4, Leu-4, Phe-37 | Asp-100j, Tyr-100k, His-101, Arg-102 | Pro-46

\(^a\) Residues of the antibody having atoms within 4.0 Å to CDR3, the residues in bold belong to CDR.
\(^b\) CDR3 residues contacting the remaining VHH, the antigen, or both are aligned in the column on the left, right, or middle, respectively.
\(^c\) Residues of the antigen having atoms within 4.0 Å to the antibody.
\(^d\) Cys residues connected by a disulfide bond.

The H2 loop of the cAb-CA05 is special because it contains only four amino acids instead of the conventional five, six or eight residues. Structurally, the H2 loop of four amino acids forms a regular \(\beta\)-turn of type II\(^d\) (37). Because H2 loops of this length are absent in germline VHHs (43), they are generated by a somatic mutation.

The long CDR3 loops of cAb-CA05 are divided based on distinct functions into an N-terminal and a C-terminal part, with the Cys-100c residue as midpoint making an interloop disulfide bond. A similar loop division was introduced in the CDR3 of cAb-Lys3 (6). The C-terminal part of both CDR3-loops is involved in extensive interactions with residues that constitute the VL-interacting surface of a VH in normal VH-VL pairs. Thus, the CDR3 loop covers the remaining hydrophobic patches in this area of the domain and shields these from the aqueous solvent. With the exception of the llama hcg VHH with its short CDR3 of eight amino acids, all VHHs of known structure share this feature. The VHH-specific amino acids Phe-37, Arg-45, and Gly-47 are all contacted by the CDR3 residues (Table II), supporting the idea that it is the combined effect of these hallmark VHH substitutions and the (C-end of the) long CDR3 that provides the optimal single-domain properties of a VHH.
with the VL in a VH-VL heterodimer. Furthermore, this novel CDR3 positioning in VHVs strongly suggests that the mutagenesis of the framework 2 amino acids of a VH to mimic the VHH of camelds will be insufficient to convert a VH into a functional, soluble single-domain antibody fragment. Additional CDR3 mutagenesis in a camelized VH will be a prerequisite to restore the antigen binding characteristics of the parental VH domain. This is exactly what the group of Riechmann (47) encountered when they used a camelized human VH to generate single-domain molecular recognition units.

The N-terminal part of the CDR3 of cAb-CA05 folds back over the CDR1 and CDR2. Both the N-terminal and the C-terminal part of the CDR3 participate in carbonate anhydride binding (Table II). However, the residues of the N-terminal part that interact with the remaining CDR loops of the VH domain bind with the antigen as well, whereas the C-terminal part of the CDR3-loop shows an alternating pattern of residues contacting either the antigen or the remaining VH domain. The N and C-terminal part of the CDR3 of cAb-CA05 forms one large surface that is essentially flat. Two marginal notes can be made from this observation. First, although this paratope architecture is also provided by the six hypervariable loops in the VH-VL heterodimers to recognize large antigens such as proteins, the difference is that the cAb-CA05 paratope is composed by CDR3 residues only. The concentration of the paratope into one single loop opens opportunities to design smaller peptidomimetics (9) or to randomize these residues and to create a synthetic library from which binders with new specificities could be retrieved by the phage display technology (48) or ribosome display (49). Secondly, it seems that the paratope of VHVs bears a large structural diversity including the formation of a flat surface as in cAb-CA05, a protruding loop as seen for the N-terminal part of the CDR3 of cAb-Lys3, or a cavity between the CDRs as observed for the hapten binder RR6. These VHVs contain a long CDR3 of 19, 24, and 16 residues, respectively. The N-terminal part of the long CDR3 of cAb-Lys3 protrudes from the remaining antigen-binding site and provides ~70% of the antigen binding surface by insertion into a cavity harboring the catalytic site of the lysozyme (6, 21). In contrast, in cAb-CA05 we are confronted with a long CDR3 that forms a planar surface that does not insert into a cavity on the antigen surface. It explains the failure of this antibody fragment to inhibit the enzymatic activity of its antigen. Furthermore, it proves that different parts of the carbonic anhydrase with a planar or a concave surface are antigenic for heavy and light chains.
Antigen Specificity and High Affinity Binding Provided by One Single Loop of a Camel Single-domain Antibody
Aline Desmyter, Klaas Decanniere, Serge Muyldermans and Lode Wyns

J. Biol. Chem. 2001, 276:26285-26290.
doi: 10.1074/jbc.M102107200 originally published online May 7, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M102107200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 46 references, 3 of which can be accessed free at
http://www.jbc.org/content/276/28/26285.full.html#ref-list-1