UC Irvine
UC Irvine Previously Published Works

Title
The Structure of an Antitumor CH2-domain-deleted Humanized Antibody

Permalink
https://escholarship.org/uc/item/2qw674k9

Journal
Journal of Molecular Biology, 348(5)

ISSN
0022-2836

Authors
Larson, Steven B
Day, John S
Glaser, Scott
et al.

Publication Date
2005-05-01

DOI
10.1016/j.jmb.2005.03.036

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed
The Structure of an Antitumor CH2-domain-deleted Humanized Antibody

Steven B. Larson¹, John S. Day¹, Scott Glaser², Gary Braslawsky² and Alexander McPherson¹*

¹Department of Molecular Biology and Biochemistry
The University of California
Irvine, CA 92697-3900, USA

²Biogen Idec Corp., San Diego
CA 92122, USA

CH2-domain-deleted CC49 (HuCC49ΔCH2), a recombinant humanized antibody that recognizes the TAG-72 antigen expressed on a variety of human carcinomas, is secreted from cultured cells as a mixture of two homodimeric isoforms. Isoform A contains two covalent interchain disulfide bonds at heavy chain positions 239 and 242, while isoform B fails to develop any interchain disulfide bonds but has 239–242 intrachain disulfide bonds instead. Form A is currently in preclinical development as a therapeutic agent for treating colorectal carcinoma, though form B shows equal efficacy.

HuCC49ΔCH2 form B can be crystallized from sodium formate only in the presence of detergents. X-ray diffraction data were collected on a single cryo-cooled crystal grown with Triton X-100 and the structure was solved by molecular replacement. The model has refined to R = 0.246 (R_free = 0.297) for 2.8 Å data. The antibodies pack in the crystal around crystallographic 2-fold axes as tetramers with approximate 222 symmetry. Atomic force microscopy studies show that this tetrameric structure is the crystal building block and also exists free in the mother liquor. The tetramer is composed of two rings, back-to-back, with a thickness of ~83 Å. Each ring is composed of two antibodies with the complementarity-determining regions (CDR) of the two Fabs of one antibody interacting with the CDR regions of the second antibody in a head-to-head fashion. These rings are approximately 167 Å long and 112 Å wide. The CH3 domain is inverted with respect to the Fabs when compared to the usual orientation found in conventional antibodies. The polypeptides joining the CH3 domains to the Fab portions of the antibody are not seen and are almost certainly disordered. The antigen combining site of HuCC49ΔCH2 is very similar, but not identical, in topology and charge distribution to that of antibody B72.3, which binds a similar epitope on TAG-72. The combining site consists of a deep cleft, heavily lined with aromatic amino acid side-chains but bounded by numerous charged groups.

Introduction

CH2-domain-deleted antibodies are a class of genetically engineered antibody reagents being developed for radioimmunotherapy (RIT) of solid tumors. As radioimmuno therapeutics, these antibodies address the issue of alleviating certain dose-limiting toxicities, primarily bone marrow toxicity, found with full-length immunoglobulin G (IgG) radiolabeled antibodies in the circulatory system and vascular pools.¹ ³ HuCC49ΔCH2, a humanized CH2-domain-deleted CC49 monoclonal antibody (mAb), has high affinity for the TAG-72 antigen expressed on a majority of human carcinomas including colorectal, gastric, pancreatic, lung, and ovarian.⁴ In human tumor mouse xenograft models, treatment with HuCC49ΔCH2 resulted in tumors efficiently retaining the antibody, but rapid serum clearance was exhibited compared

Abbreviations used: RIT, radioimmunotherapy; mAb, monoclonal antibody; AFM, atomic force microscopy; CDR, complementarity-determining region; Ig, immunoglobulin.

E-mail address of the corresponding author: amcphers@uci.edu

© 2005 Elsevier Ltd. All rights reserved.

Keywords: quaternary structure; X-ray crystallography; antibody engineering; Tn antigen; antibody CC49

0022-2836/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
to full-length CC49 IgG, presumably due to the smaller size (~122 kDa) of the antibody and its inability to recycle through FcRn receptors. Recent clinical studies with low-dose $^{[131]}$I-HuCC49ΔCH2 in a small group of patients with metastatic colorectal carcinoma have shown the radioimmunotherapeutic to be well tolerated, and exhibit a demonstrably reduced serum half-life compared to earlier studies with the full-length CC49 IgG.

Biosynthesis of HuCC49ΔCH2 in mammalian cells produces two homodimeric isoforms present in approximately a 50:50 mixture. As illustrated in Figure 1, one isoform, referred to as form A, contains covalent interchain disulfide bonds at heavy chain hinge positions 239 and 242, Kabat numbering system. The second isoform, form B, is presumably held together by non-covalent interactions within the CH3 domains, and fails to develop interchain hinge disulfide bonds as evidenced by the formation of a 60 kDa product following non-reducing denaturing gel electrophoresis. Biodistribution and compound stability studies support that form A is the preferred molecule for therapeutic development and methods for the separation and purification of form A from form B have been developed.

Though derived independently, and from different sources, antibody CC49, from which HuCC49ΔCH2 was derived, is closely related to another murine antibody, B72.3. The latter antibody binds a similar epitope on TAG-72 and has similar, though not so effective, antitumor affinities to CC49. B72.3 is noteworthy because of the structural similarity to CC49 found here, and the fact that, because of its earlier discovery, it has been the subject of more extensive studies and characterization, including an X-ray structure analysis of its antigen binding fragment.

The engineered antibody studied here, HuCC49ΔCH2 form B, is a chimeric antibody in which the complementarity-determining regions (CDR) of the Fabs are of mouse origin, while the remainder of the antibody is derived from a human antibody sequence. The CH2 domain is entirely missing, and the CH3 domain is linked directly to the Fabs through polypeptides, which are 20 amino acid residues in length. With no interchain disulfide bonds between heavy chains, the three domains of the antibody might be expected to exhibit high degrees of flexibility and exist in a wide range of dispositions. Thus, it was with some surprise that we found that form B would crystallize, while the presumably more conformationally stable form A would not.

Details of the crystallization of HuCC49ΔCH2, the characterization of the crystals by X-ray diffraction, and an atomic force microscopy (AFM) investigation of the crystals were reported earlier. Those analyses indicated that the crystallization unit was composed of four HuCC49ΔCH2 antibodies arranged in a toroidal structure. They further showed that the tetrameric aggregates existed independent of the crystals and were present in solution. The analysis presented here is consistent with those conclusions and reveals the detailed structure of the antibodies and their oligomeric organization.

At this point in time, only four other intact monoclonal antibodies (mAbs) have been analyzed by X-ray crystallography, though two other intact hinge-deleted immunoglobulins have also been determined. At least two other intact antibodies have been crystallized but not solved. HuCC49ΔCH2 form B, therefore, represents the fifth intact monoclonal antibody whose structure has been solved by X-ray diffraction, though it does lack the CH3 domain. Furthermore, it is the first humanized antibody, and the first lacking disulfide bridges in the hinge region.

Results

The model

The atomic model of the asymmetric unit of the crystals of HuCC49ΔCH2 contains two antibodies, each with two light chains of 220 amino acid residues and two heavy chains of 324 amino acid residues for a total of eight chains, 2176 residues, and 16,740 atoms. The hinge regions of the heavy

![Figure 1. The CH2-domain-deleted antibody HuCC49ΔCH2 exists in two forms, A and B, differing only in the presence of interchain disulfide bonds (form A), or intrachain disulfide bonds (form B). The two isoforms occur in equal proportions and give rise in a Western blot to the bands shown here. The gel is a non-reducing, SDS-PAGE, followed by Western blot, of unprocessed culture supernatant from producer CHO cells. The gel was probed with rabbit anti-human κ-chain antibody. The cell line, in addition to producing the A and B forms of HuCC49ΔCH2, secretes an excess of light chains as both monomers and dimers. The two forms of HuCC49ΔCH2 have shown equal efficacy against tumor cells in clinical trials. In form B, the heavy chains are held together only by CH3 domain non-covalent interactions.](image-url)
chains, residues 219 through 238, which are not visible in electron density maps, are not included in the model. Antibody 1 is composed of light chains A and C, and heavy chains B and D; antibody 2 contains light chains E and G, and heavy chains F and H. Table 1 gives statistics regarding the model.

Figure 2 illustrates a representative region of electron density from a composite omit map. A backbone tracing of a single HuCC49 CH2 antibody is shown in Figure 3(a). Two of these antibodies are bound to each other in the asymmetric unit in a head-to-head fashion through the CDR regions to form a ring, as illustrated in Figure 3(b) and (c). Furthermore, two rings related by the crystallographic 2-fold axis along c form a toroid-like assembly, shown in Figure 4, in which the two rings are rotated with respect to each other by 8°.

Upon tetramer formation, a remarkable 7100 Å² of accessible surface area of the freestanding ring is buried, which is evenly distributed over all two-domain groups (see Table 2). The tetramer assembly has approximate 222 symmetry with a pseudo-2-fold axis parallel with each of the a and b-axes. This approximate 222 symmetry appears in the root-mean-square deviations (rmsds) between the equivalent chains; rmsds for the pseudo 2-fold-related light chains (A-E and C-G, Ca atoms of residues 1–215) are 0.47 Å and 0.51 Å, respectively, whereas rmsds for the other four pairings (A-C, A-G, C-E, E-G) range from 1.08–1.29 Å. A similar difference is seen in the rmsds (Ca atoms of residues 1–215) of the heavy chains; pairs B-E and D-H have rmsds of 0.35 Å and 0.39 Å, respectively, and pairs B-D, B-H, D-F, and F-H have rmsds in the range of 0.77–0.97 Å. The chains of the CH3 domains exhibit a narrow range in rmsds of 0.13–0.19 Å. The elbow angles of the Fab fragments are also consistent with pseudo 2-fold symmetry and account for the

| Table 1. Data collection and refinement statistics and parameters |
|---------------------------------------------------------------|
| **A. Data collection** |
| Wavelength λ (Å) 0.9194 |
| Resolution (Å) 2.8–45.1 |
| Rmerge 10.9 (5.0) |
| (I/σI) 99.87 (99.83) |
| Completeness 76.929 (7574) |
| Number of unique reflections 696,849 |
| Observations 76,929 |
| Redundancy 9.06 (7.62) |
| **B. Crystal data** |
| Space group P2_12_1 |
| a = 83; b = 224; c = 167 |
| Unit cell parameters (Å) |
| **C. Refinement and model statistics** |
| Resolution (Å) 2.8–45.1 |
| Rcryst 0.246/0.297/0.247 |
| Rfree 0.247/0.297/0.247 |
| Number of residues 2176 |
| Number of non-hydrogen atoms 16,740 |
| rmbsd from ideal geometry |
| Bond lengths (Å) 0.013 |
| Bond angles (deg.) 1.87 |
| Improper (deg.) 1.28 |
| Dihedrals (deg.) 27.3 |
| Ramachandran plot (%, mf,f,a,d) 80.7, 16.3, 2.1, 0.9 |

| NCS restraints |
|----------------|
| Equivalence groups (CA, N C atoms only) | NCS weight | Sigma B | (b) (Å²) |
| Light chain 1–114 | 210 | 6 | 31 |
| Light chain 115–215 | 150 | 12 | 45 |
| Light chain 216–220 | None | None | 90 |
| Heavy chain 1–39 | 210 | 6 | 32 |
| Heavy chain 40–45 | None | None | 35 |
| Heavy chain 46–114 | 210 | 6 | 26 |
| Heavy chain 115–128 | 150 | 12 | 45 |
| Heavy chain 129–136 | None | None | 99 |
| Heavy chain 137–213 | 150 | 12 | 43 |
| Heavy chain 214–218 | None | None | 90 |
| Heavy chain 219–238 (hinge region) | Not in model | Not in model | – |
| Heavy chain 239–315 | 210 | 6 | 47 |
| Heavy chain 316–319 | None | None | 40 |
| Heavy chain 320–344 | 210 | 6 | 57 |

* mf, most-favored; f, additionally allowed; g, generously allowed; d, disallowed.
disparity in the rmsds; the AB and EF Fab fragments have elbow angles of 144° and 145°, respectively, whereas the CD and GH fragments have angles of 137° and 139°, respectively.

There are several regions in the antibodies (besides the unseen hinges) that exhibit disorder, as indicated by large $B$ values. This is apparent in the average $B$ values of the NCS equivalent groups described in Table 1. In all chains the variable domains (residues 1–114) have lower $B$-factors than the constant domains. Although the variable domains of the two antibodies exhibit similar patterns in $B$-factors, with average $B$-factors in the narrow range of 22–36 Å², the constant domains of antibody 2 have significantly higher values than those of antibody 1 (61 Å² versus 34 Å²).

Several of these high-$B$-factor regions also exhibit poorly defined electron density in composite omit maps. Specifically, the loop at residues 129–136 of all heavy chains has broken, difficult-to-model density, consistent with the structures of many other antibodies and antibody fragments; the C-terminal ends of all $C_L$ and $C_H1$ domains are poorly defined despite the presence of the disulfide bond connecting the corresponding $C_L$ and $C_H1$ domains together; the maps in the hinge regions, residues 219–238, are totally devoid of recognizable structure; and, finally, the last two or three residues of the heavy chains have poorly defined density.

The missing hinge region is 20 amino acid residues in length and would span more than 60 Å in an extended conformation. The distance between the carbonyl carbon atom of residue 218 and the peptide nitrogen atom of residue 239 of the four heavy chains ranges from 15.3 Å in chain D to 20.4 Å in chain B. Thus, there is undoubtedly substantial slack in the hinges of both antibodies. The slack results from the absence of the interchain disulfide bonds at cysteine residues 224 and 227 (model numbering) that characterize form A of the antibody. It is the flexibility in the hinge that most likely favors crystallization of form B over form A. The hinge flexibility allows the $C_H3$ domains to invert the orientation from that found in previously analyzed intact antibodies.10–12 Furthermore, the lack of interchain disulfide bonds in the hinge removes constraints on the Fabs, permitting them to separate at the C-terminal end, thus, enabling each Fab moiety of one antibody to align in a colinear fashion with its Fab counterpart in the second antibody in a self-complementary manner (as seen in Figure 3(b)), similar to a previously described structure of an anti-antiidiotypic Fab fragment.13

The interactions between the Fabs of the two antibodies in the asymmetric unit are slightly different for the two interfaces (see Table 2), especially between chains D and E. Clearly, the dominant interactions are between the light chain of
one Fab and the heavy chain of its counterpart. Figure 5 illustrates the hydrogen bonding interactions between the heavy chain B and light chain G, which exhibit one additional hydrogen bond in contrast to the other three light chain–heavy chain interfaces. The residues from the light chain that are involved in these interactions are located on CDR1. Those involved from the heavy chain reside on
CDR2. Intermolecular contacts in a dimer bury approximately 700 Å² of surface area on each variable domain. In a tetramer, an additional approximately 700 Å² of surface is buried on each of the variable and constant domains. Thus, dimer interactions, on a domain basis, are almost equal to the average domain interactions in a tetramer. Figure 3 and the buried surface areas listed in Table 2 clearly show that in each antibody the CH3 domain interacts strongly with only one Fab. The CH3 domains of chains B and F have only seven and four interactions (interatomic distances < 4 Å), respectively, with one possible hydrogen bond for chain B. On the other hand, the CH3 regions of chains D and H have 88 and 82 interactions, predominantly with the light chains A and E, with 13 and ten possible hydrogen bonds, respectively. This suggests that the CH3 groups, loosely tethered to the Fabs, assume arbitrary dispositions (with respect to the Fabs) until they eventually bind firmly to one Fab, predominantly with its light chain. Because the two Fab fragments of one antibody must have the same orientation in order to form the tetramer, the CH3 domains of chains B and F do not have comparable regions of the light chain of the second Fab fragment in their respective antibodies with which to interact. Hence, the interactions involving chains B and F are limited and appear to be nothing more than packing interactions. This phenomenon could not occur in HuCC49ΔCH2 form A.

Packing

Figure 6 shows the packing of the toroid-like antibody assemblies. The thickness of the tetramer is approximately the same as the a axis of the unit cell (Figure 6(b)); the length is approximately equal to the length of the c axis (Figure 6(a)). There is evidence from AFM studies of HuCC49ΔCH2 and its crystals that, under the crystallization conditions noted above, tetrameric assemblies are present in solution, and that crystal growth proceeds through the incorporation of tetramers into the crystal lattice. In contrast to the ring-to-ring interface discussed above, there are only six contacts (interatomic distances < 4 Å) between tetramers along the a axis and a total of only 463 Å² of buried surface area per ring assembly.

Discussion

The crystallization of form B of HuCC49ΔCH2 and the failure of form A to do so was unanticipated because we naturally associate crystallizability with conformational stability. The puzzle, however, is solved by the association of two, and then four entire antibodies into structured oligomeric entities. The quaternary interactions create a conformationally stable crystallizing unit from a collection of highly flexible, loosely associated domains. Form A
of HuCC49DCH2 cannot form such oligomers because of the interchain disulfide bonds in its hinge. It may be significant that both forms A and B exhibit similar efficacies in clinical trials. Presumably this means that binding of antibody to antigen is strictly a function of the CDR regions of the Fabs, and is not a matter of overall antibody conformation, nor the degree to which various domains are associated.

In addition, B form assemblies likely exist only where the antibody is at high concentrations used for crystallization, though the large amount of surface area buried upon aggregation suggests that they might be present even at low concentrations.

### Table 2. Interactions in self-complementary Fab binding, in tetramer formation and between C\textsubscript{H}3 domains and the Fab groups

| Interactions | A···H | G···B | B···H | C···F | E···D | D···F |
|--------------|-------|-------|-------|-------|-------|-------|
| van der Waals contacts (<4 Å) | 39 | 39 | 7 | 37 | 27 | 7 |
| Number of contacts | | | | | | |
| Hydrogen bonds (in Å) | | | | | | |
| Y31-O···Y60-OH | 2.47 | 2.76 | 2.86 | 2.92 |
| Y31-O···L70-O | 2.55 | 2.55 | 2.57 | 2.55 |
| Y38-O···D56-O | 2.71 | 2.91 | 2.65 | 2.52 |
| N34-ND2···D56-OD2 | – | 3.29 | – | – |

### Table 2. Interactions in self-complementary Fab binding, in tetramer formation and between C\textsubscript{H}3 domains and the Fab groups

| Interaction | Group | Free-standing group ASA | Combined ASA | Buried surface area |
|-------------|-------|-------------------------|--------------|---------------------|
| Tetramer    | Ring  | 93,501                  | 86,400       | 7101 |
| Tetramer    | AB F\textsubscript{H-V} in ring | 9188 | 8408 | 781 |
| Tetramer    | AB F\textsubscript{H-Ch}1 in ring | 9088 | 8416 | 672 |
| Tetramer    | CD F\textsubscript{H-V} in ring | 9030 | 8199 | 831 |
| Tetramer    | CD F\textsubscript{H-Ch}1 in ring | 9623 | 9155 | 469 |
| Tetramer    | EF F\textsubscript{H-V} in ring | 9025 | 8202 | 823 |
| Tetramer    | EF F\textsubscript{H-Ch}1 in ring | 8996 | 8257 | 739 |
| Tetramer    | GH F\textsubscript{H-V} in ring | 8987 | 8127 | 860 |
| Tetramer    | GH F\textsubscript{H-Ch}1 in ring | 9559 | 9027 | 532 |
| Tetramer    | BD F\textsubscript{H-Ch}3 in ring | 10,008 | 9292 | 716 |
| Tetramer    | FH F\textsubscript{H-Ch}3 in ring | 9905 | 9318 | 677 |
| Fab-Fab     | AB F\textsubscript{H-V} in domain | 10,657 | 9950 | 707 |
| Fab-Fab     | GH F\textsubscript{H-V} in domain | 10,499 | 9760 | 739 |
| Fab-Fab     | CD F\textsubscript{H-V} in domain | 10,544 | 9783 | 761 |
| Fab-Fab     | EF F\textsubscript{H-V} in domain | 10,559 | 9826 | 733 |
| Fab-C\textsubscript{H}3 | AB F\textsubscript{H-Ch}1 in domain | 10,363 | 9804 | 558 |
| Fab-C\textsubscript{H}3 | CD F\textsubscript{H-Ch}1 in domain | 10,732 | 10,118 | 614 |
| Fab-C\textsubscript{H}3 | BD F\textsubscript{H-Ch}3 in domain | 10,389 | 10,325 | 64 |
| Fab-C\textsubscript{H}3 | EF F\textsubscript{H-Ch}3 in domain | 10,732 | 10,645 | 88 |
| Fab-C\textsubscript{H}3 | FH F\textsubscript{H-Ch}3 in domain | 10,288 | 9707 | 581 |
| Fab-C\textsubscript{H}3 | GH F\textsubscript{H-Ch}1 in domain | 10,726 | 10,092 | 634 |
| Fab-C\textsubscript{H}3 | FH F\textsubscript{H-Ch}3 in domain | 10,371 | 10,302 | 69 |

**Figure 5.** Stereo view of the hydrogen bonding interactions in the Fab–Fab interface between heavy chain B of antibody 1 and light chain G of antibody 2 (distances in Å). Except for the G-N34···B-D56 hydrogen bond, equivalent hydrogen bonding patterns exist between chains A and H, chains C and F, and chains D and E. Residue labels are composed of the chain name as the first letter, the one-letter amino acid code as the second letter and the residue number in the model as the last two characters.
concentration. Dilution in the serum of patients would, however, likely resolve the complexes into their individual components.

The quaternary structural arrangement that we observe here is quite intricate and involves a number of interactions. It seems unlikely that it is simply fortuitous and has no functional significance, though at present we can assign none. At the very least it illustrates the kinds of quaternary interactions of which an antibody is capable, given sufficient freedom. As with other intact mAbs that have been investigated, the C\textsubscript{4\text{\textbeta}}3 domains, like the intact Fc domains, appear relatively free to assume arbitrary orientations with respect to the Fabs. Here, the C\textsubscript{4\text{\textbeta}}3 domains are completely inverted (see Figure 3(a)) from their usual orientation and maintain a stable position in the crystal by virtue of contacts with the constant domains of the Fabs and the 2-fold related C\textsubscript{4\text{\textbeta}}3 domain in the tetramer. Also noteworthy, is the fact that the two Fabs in this structure are not related by a dyad axis of symmetry (true or pseudo) as they are in all other intact antibodies. Here, the orientations in space are the same and the Fabs parallel. There is no contact between Fabs within an individual antibody; contacts are between an Fab and two other Fabs within a tetramer.

The crystal structure of HuCC49ΔCH2 is in excellent agreement with the preliminary analysis of the crystals by AFM, and predictions based on that analysis. From AFM it was concluded that the antibodies first formed rings of two antibodies, and that these then associated back-to-back to form a tetrameric toroidal assembly. AFM further showed that this toroidal unit formed in the mother liquor prior to incorporation into a crystal. It cannot be argued, given the AFM results, that the tetrameric assembly, whose structure was solved here, is simply a fortuitous consequence of crystal packing interactions. It is a stable entity.

Although this is an intact antibody, it lacks the C\textsubscript{4\text{\textbeta}} domain and clearly, therefore, cannot be considered structurally in the same manner as those that preceded it.\textsuperscript{10-12} It undoubtedly lacks many of the effector functions normally present, as well as requisite sites at the C\textsubscript{4\text{\textbeta}}2-C\textsubscript{4\text{\textbeta}}3 elbow for the binding of proteins such as protein A and protein G. On the other hand, the long and unrestrained hinge polypeptides, each of which could extend as much as 60 Å, permit the antibody a reach of nearly 200 Å
between the CDRs of the two Fabs. This could significantly affect the extent of antibody–antigen interactions on, for example, a cell surface.

In this antibody assembly, having pseudo 222 symmetry, the CDRs of any Fab are symmetry-related to the CDRs of a second antibody, with which they hydrogen bond and otherwise interact, by a 2-fold axis of symmetry. Such an arrangement has been seen at least once before in the crystal structure of an anti-antiidiotypic antibody Fab. There too, Fabs coupled with one another through a 2-fold symmetrical interface. Similarly, Fabs in this crystal interact about other pseudo 2-fold axes creating interfaces employing lateral surfaces.

**Antigen combining site**

CC49 and B72.3 are two closely related antibodies that have been the focus of exceptional attention by cancer immunologists because of their remarkably high affinity and specificity for tumor cells. This observation of anti-tumor activity has withstood more than 20 years of extensive analysis and further characterization. Both antibodies, which recognize the mucin-like glycoprotein TAG-72, have been used in tumor marking and treatment. CC49, which has been termed a “second generation B72.3”, has been chosen for further development because of its significantly higher affinity for the tumor antigen. Engineered antibodies based on CC49, including the A form of the molecule described here, have now entered phase II clinical trials as reagents for diagnosis and treatment of human colorectal, breast, prostate, and ovarian carcinoma.

The antigen recognized by B72.3 appears to be the disaccharide N-acetyleneuraminic acid α(2→6)-N-acetylgalactosamine, O-linked to serine or threonine. This has been designated sialylated Tn antigen. It has been reported, using synthetic antigens, that both B72.3 and CC49 react strongly with a disaccharide cluster of three sialyl Tn epitopes-O-serine, but that they showed different reactivities to single sialyl Tn disaccharides, suggesting that CC49 recognizes an epitope distinct from that of B72.3.

Hansisch et al. used mass spectrometry to identify the epitope of CC49 and reported it to be the same as that for B72.3, but containing a galactose residue in addition, β(1→3)[N-acetyl neuraminic acid α(2→6)] N-acetyl galactosamine, O-linked to a polypeptide serine on TAG-72 has been identified as the epitope for CC49 based on mass spectrometry. The presence of the galactose residue apparently distinguishes this epitope from that for antibody B72.3. This core antigen likely arises from incomplete glycosylation and subsequent exposure on the surfaces of tumor cells.

In Figure 7 the trisaccharide shown here, galactose β (1→3)[N-acetyl neuraminic acid α (2→6)] N-acetyl galactosamine, O-linked to a polypeptide serine on TAG-72 has been identified as the epitope for CC49 based on X-ray crystallography and its combining site delineated. It is, therefore, available for comparison. No structures of Fab fragments from CC49 have been reported, but Tamura et al. have used homology rules to identify its six canonical hypervariable loops. The three-dimensional structure of a Fab fragment of B72.3 has been solved by X-ray crystallography and its combining site delineated. It is, therefore, available for comparison. No structures of Fab fragments from CC49 have been reported, but Tamura et al. have used homology rules to identify its six canonical hypervariable loops. The model for HuCC49ΔCH2 presented here reveals their relative dispositions in space. In other antibodies (Fab fragments) analyzed by X-ray crystallography, those having hapten-like antigens exhibited CDRs characterized by a deep pocket. Those binding to polypeptide and protein antigens had a flatter, less concave, sometimes even convex surface. Antibodies directed against oligosaccharides were generally found to feature deep grooves across the face of the combining site. In addition, oligosaccharide combining sites are usually characterized by the presence of a large number of aromatic amino acids. The antigen combining sites of CC49 and B72.3 conform well to the description of a typical oligosaccharide binding site.

In Figure 8 the alpha carbon backbone traces of CC49 and B72.3 are superimposed and seen both in profile and en face. In the HuCC49ΔCH2 toroidal complex, there are four crystallographically independent Fab groups. These have elbow angles ranging from 137° to 145°, an 8° variation. In the Fab for B72.3, the elbow angle was 137°, the lower end of the range for HuCC49ΔCH2. For the most part, the topologies of the two combining sites are almost the same, both characterized by a deep trench traversing the distal
The only major difference between the two lies in CC49 light chain CDR1, which contains an insertion of six amino acid residues compared with the corresponding CDR1 of B72.3 (see Table 3). This loop stands well above the cleft and emphasizes its depth. In the toroidal complex of HuCC49, this loop contributes hydrogen-bonding groups responsible for much of the association with a dyad-related antigen combining site of another antibody within the complex (see Figure 5). The close similarity of the conformations of the CDR loops in HuCC49 and the Fab fragment of B72.3 suggests that the formation of the toroidal complex by HuCC49 and B72.3 does not significantly perturb its structure. This is in spite of the fact that amino acid residues from these loops

| Chain | Sequencea |
|-------|-----------|
| LCDR1 | 24 Lys 25 Ser 26 Ser 27 Glu 28 a 29 Leu 30 Leu 31 Tyr 32 Ser 33 Gly 34 Asn |
| CC49  | Arg Ala Ser Ser Glu Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser |
| B72.3 | Ala Ala Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg |
| LCDR2 | 50 Tyr 51 Pro 52 Tyr 53 Tyr 54 Tyr 55 Tyr 56 Tyr |
| CC49  | Tyr Tyr Tyr Tyr Tyr Tyr Tyr Tyr |
| B72.3 | Tyr Tyr Tyr Tyr Tyr Tyr Tyr Tyr |
| LCDR3 | 89 Ser 90 Pro 91 Thr 92 Tyr 93 Ser 94 Tyr 95 Pro 96 Tyr 97 |
| CC49  | Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser |
| B72.3 | Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser |
| LCDR1 | 24 Lys 25 Ser 26 Ser 27 a 28 b 29 c 30 d 31 e 32 f 33 g 34 h |
| CC49  | Arg Ala Ser Ser Glu Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser |
| B72.3 | Ala Ala Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg |
| LCDR2 | 50 Tyr 51 Pro 52 Tyr 53 Tyr 54 Tyr 55 Tyr 56 Tyr 57 Tyr 58 Tyr 59 Tyr 60 Tyr 61 Tyr 62 Tyr 63 Tyr 64 Tyr 65 Tyr |
| CC49  | Tyr Tyr Tyr Tyr Tyr Tyr Tyr Tyr |
| B72.3 | Tyr Tyr Tyr Tyr Tyr Tyr Tyr Tyr |
| LCDR3 | 89 Ser 90 Pro 91 Thr 92 Tyr 93 Ser 94 Tyr 95 Pro 96 Tyr 97 |
| CC49  | Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser |
| B72.3 | Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser |

a The model numbering is sequential; therefore, the correspondence is: LCDR1, 24–40; LCDR2, 56–62; LCDR3, 95–103; HCDR1, 31–35; HCDR2, 50–54; HCDR3, 99–103.
are primarily responsible for dimerization of HuCC49ΔCH2.

Table 3 compares the amino acid sequences of the six CDRs of CC49 and B72.3. It shows that of 52 corresponding residues, 30 are identical between the two antibodies, but that many others represent only conservative changes. HCDR1 is identical in both antibodies, and HCDR2 has 14 of 17 residues the same. On the other hand, LCDR2 is identical at only one of seven positions. In spite of the amino acid variations, the conformations of the CDR loops are extremely similar. Minimization of differences in positions of 92 corresponding alpha carbon atoms of the variable regions of the two antibodies yields an rmsd of only 0.6 Å, with the CDR loops omitted. Based on this same superposition, the CDR loops alone have an rmsd of 2.6 Å for the 52 corresponding alpha carbon atoms.

The antigen binding site of CC49 displays a remarkable number of aromatic residues, as many as 23, in positions where they could conceivably contact the antigen. Involvement of some of these side-chains would perhaps require rearrangement of one or more loops to make contact with the antigen, but small loop movements would not be unreasonable.

We would not like to speculate as to which of the residues might be most important for antigen binding. Tamura et al. have, however, based on homology with other antibody–antigen complexes, as well as the binding affinities of CC49 mutants, made some predictions. Among those are that Tyr L94 and particularly His H32, which is common to both antibodies, would contact the antigen. In terms of the three dimensional structure, these appear reasonable.

Figure 9 is a surface representation of the combining sites for CC49 and B72.3, which also shows the electrostatic distributions. The deep groove crossing the face of the combining site is quite evident, though deepened by the protrusion of LCDR1 in CC49. There are broad, and even some
distinctive similarities in the electrostatic distributions, though they are by no means identical. In fact, the difference appears to be somewhat greater than we might have anticipated for two antibodies that have so similar an antigen. There is a fairly distinctive variation in charge over the combining site surfaces of both antibodies, which suggests that electrostatics may play an important role in recognition and association of the antigen.

It is noteworthy, we feel, that of all the Fab fragments used as probes in the molecular replacement search, one was particularly successful, and that was the Fab fragment derived from B72.3. That probe was also a human-murine chimeric Fab, was also specific for TAG-72, and it had an epitope closely related to that of CC49. In retrospect, it seems clear from the molecular replacement results that their structures were more similar overall than for two, otherwise arbitrary Fab fragments.

**Materials and Methods**

The antibody is produced by Biogen Idec of La Jolla, CA and stored in phosphate-buffered saline at pH 7.2. Details of the crystallization of HuCC49ΔCH2 form B have been described elsewhere, along with an atomic force microscopy investigation of the crystals. The crystals used for diffraction analysis were grown by vapor-diffusion at 17°C in the presence of 1.5 mM Triton X-100. Detergent was produced ordered crystals that diffract to adequate effective in yielding crystals, but, so far, only Triton X-100 was essential to crystallization, and many detergents were found in Table 1.

Data to 3.2 Å resolution were initially collected at 22°C from a crystal conventionally mounted in a quartz capillary on a Rigaku R-axis system (Molecular Structure Corp., Woodlands, TX) using an RU-200 generator fitted with Osmic mirrors and dual image plates. This data set was used for structure determination by molecular replacement. For refinement of the structure, data were collected from a single crystal at −173°C on beamline 5.0.2 at the Advanced Light Source at Berkeley, CA. Crystals were mounted on cryoloops (Hampton Research, Aliso Viejo, CA) and frozen directly in the cryostream at −173°C without any additional treatment with a cryo protectant. The mosaicity was 0.658 and the crystal-to-detector distance was 250.53 cm. The data were processed with D*TREK v9.029 to a resolution of 2.8 Å with statistics found in Table 1.

**Structure solution**

The structure was solved by molecular replacement using CNS and the room temperature data set in the resolution range of 5–15 Å. Sixteen molecular replacement probes (five of which were chimeric) were taken from the Protein Data Bank and used without modification to locate the Fab fragments. Similarly, 16 probes for the C1H3 domains were obtained from the Protein Data Bank, some of which had to be expanded by 2-fold symmetry. The PDB codes are used to identify the probes in the following discussion of the molecular replacement (MR) results.

A cross-rotation search was performed using each probe. The top 20 cross-rotation solutions for each probe were subsequently used in a translation search. Based on reasonable packing, only six of the Fab probes and only 13 of the C1H3 probes gave plausible translation solutions as their top solution. Based on the correlation coefficient (CC), the best Fab solution was obtained with probe 1BBJ (Fab fragment of antibody B72.3), which gave CC = 0.192; the second best Fab probe gave CC = 0.174 (probe 1I7ZB). The best C1H3 solution was obtained with probe 16X and gave CC = 0.166; the second best probe gave CC = 0.152 (probe 1H3W).

A summary of the molecular replacement results is found in Table 4. Starting with the best Fab solution obtained with probe 1BBJ, the second, third and fourth translation solutions were found from the second, fourth and third 1BBJ rotation function solutions, respectively. With the four Fab fragments of the asymmetric unit located, the top 30 rotation function solutions of each C1H3 probe were input to translation searches for ten translation trials. Out of the 4800 translation trials, only one solution (rotation solution 21 of probe 1FC1) produced a significant increase in the correlation coefficient. A search for the second C1H3 domain using this probe did not produce a translation solution that improved the correlation coefficient or the packing value. Therefore, the second C1H3 position was obtained by superposition of

---

**Table 4. Progress of structure solution by molecular replacement as each successive piece of the model was placed in the unit cell and refinement to the final model**

| Probe | Cross rotation solution | Correlation coefficient | Packing value |
|-------|-------------------------|-------------------------|---------------|
| 1BBJ  | 1                       | 0.192                   | 0.0851        |
| 1BBJ  | 2                       | 0.325                   | 0.1688        |
| 1BBJ  | 4                       | 0.420                   | 0.2539        |
| 1BBJ  | 3                       | 0.504                   | 0.3377        |
| 1FC1  | 21                      | 0.569                   | 0.3796        |
| 1FC1  | 1–30                    | 0.548–0.559             | 0.3871–0.4083 |

Refinement procedure | Resolution (Å) | $R$ | $R_{free}$ |
|--------------------|----------------|----|-----------|
| Rigid body          | 5.0            | 0.414 | 0.511 |
| Rigid body          | 4.0            | 0.407 | 0.462 |
| Rigid body          | 3.5            | 0.418 | 0.459 |
| Minimization        | 3.5            | 0.404 | 0.511 |
| Annealing, minimization, B-factor | 3.5 | 0.303 | 0.461 |
| Annealing, minimization, B-factor | 3.2 | 0.249 | 0.345 |
| Annealing, minimization, B-factor | 3.0 | 0.257 | 0.330 |
| Annealing, minimization, B-factor | 2.8 | 0.271 | 0.338 |
| Final model         | 2.8            | 0.246 | 0.297 |
the Fab fragments of the complete antibody onto the Fab fragments of the incomplete antibody.

Structure refinement

With all six molecular fragments in the asymmetric unit oriented and positioned in the unit cell, the amino acid sequences of the molecular probes were mutated to the HuCC49-CH2 sequence. This initial model was then subjected to a series of rigid body refinements against the cryo data beginning at 5 Å resolution and extending to 3.5 Å resolution, as noted in Table 4. After minimization, the model was rebuilt against a composite omit map. The model was then subjected to successive cycles of refinement by a simulated annealing protocol (starting at 3000 K with 25 K decrements, followed by 200 cycles of conjugate gradient minimization and ending with individual B-factor refinement) and model rebuilding with a gradual increase in resolution to 2.8 Å. At 2.8 Å resolution, various non-crystallographic symmetry restraints were investigated in search of those restraints that produced the best combination of R and Rfree. The restraints imposed on the final model are listed in Table 1.

Computer programs

Data were processed with D*TREK. Model refinement was performed with CNS. Model building was carried out with the program O. PROCHECK was used for evaluating the quality of the final and intermediate models. MAPMAN was used to manipulate maps for O. Figure 1 was made with Adobe Illustrator. Figures 2, 3, 4, 6, and 8 were rendered with PyMol. Figure 5 was created with MOLSCRIPT and rendered with Raster3D. GRASP was used to produce Figure 9.

Acknowledgements

We thank Aaron Greenwood for his assistance in preparing Figures. This research was supported by a grant from the National Institutes of Health (GM58868).

References

1. Mueller, B. M., Reisfeld, R. A. & Gillies, S. D. (1990). Serum half-life and tumor localization of a chimeric antibody deleted of the CH2 domain and directed against the disialoganglioside GD2. Proc. Natl Acad. Sci. USA, 87, 5702–5705.

2. Calvo, B., Kashmiri, S. V., Hutzell, P., Hand, P. H., Slavin-Chiorini, D. C., Schlom, J. et al. (1993). Construction and purification of domain-deleted immunoglobulin variants of the recombinant/chimeric B72.3 (y1) monoclonal antibody. Cancer Biother. Pharmacol., 8, 95–109.

3. Slavin-Chiorini, D. C., Kashmiri, S. V., Schlom, J., Calvo, B., Shu, L. M., Schott, M. E. et al. (1995). Biological properties of chimeric domain-deleted antitumor immunoglobulins. Cancer Res., 55, 5957S–5967S.

4. Slavin-Chiorini, D. C., Kashmiri, S. V., Lee, H. S., Milenic, D. E., Poole, D. J., Bernon, E. et al. (1997). A CDR-grafted (humanized) domain-deleted antitumor antibody. Cancer. Biol. Ther. Radiopharm., 12, 305–316.

5. Forero, A., Meredith, R. F., Khazaali, M. B., Carpenter, D. M., Shen, S., Thornton, J. et al. (2003). A novel monoclonal antibody design for radioimmunotherapy. Cancer. Biol. Ther. Radiopharm., 18, 751–759.

6. Kabat, E. (1991). Sequences of Proteins of Immunological Interest, US Department of Health and Human Services, NIH.

7. Noto, M., Teramoto, Y. A., Marianicostantini, R., Hand, P. H., Colcher, D. & Schlom, J. (1982). A monoclonal-antibody (B72.3) defines patterns of distribution of a novel tumor-associated antigen in human mammary-carcinoma cell-populations. Int. J. Cancer, 29, 539–545.

8. Brady, R. L., Edwards, D. J., Hubbard, R. E., Jiang, J. S., Lange, G., Roberts, S. M. et al. (1992). Crystal structure of a chimeric Fab’ fragment of an antibody binding tumour cells. J. Mol. Biol., 227, 253–264.

9. Larson, S. B., Kuznetsov, Y. G., Day, J., Zhou, J., Glaser, S., Braslawsky, G. et al. (2005). Combined use of AFM and X-ray diffraction to analyze crystals of an engineered, domain-deleted antibody. Acta Crystallogr. Sect. D, 61, 416–422.

10. Harris, L. J., Larson, S. B. & McPherson, A. (1999). Comparison of intact antibody structures and the implications for effector function. Advan. Immunol., 72, 191–208.

11. Saphire, E. O., Parren, P. W. H. I., Pantophlet, R., Zwick, M. B., Morris, G. M., Rudd, P. M. et al. (2001). Crystal structure of a neutralizing human IgG against HIV-1: a template for vaccine design. Science, 293, 1155–1159.

12. Harris, L. J., Larson, S. B., Skalesky, E. & McPherson, A. (1998). Comparison of the conformations of two intact monoclonal antibodies with hinges. Immunol. Rev., 163, 35–43.

13. Ban, N., Day, J., Wang, X., Ferrone, S. & McPherson, A. (1996). Crystal structure of an anti-anti-idiotypic antibody shows it to be self-complementary. J. Mol. Biol., 255, 617–627.

14. McPherson, A. (1999). Crystallization of Biological Macromolecules, Cold Spring Harbor Laboratroy Press, Cold Spring Harbor, NY.

15. Muraro, R., Kuroki, M., Wunderlich, D., Poole, D. J., Colcher, D., Thor, A. et al. (1988). Generation and characterization of B72.3 second generation monoclonal antibodies reactive with the tumor-associated glycoprotein 72 antigen. Cancer Res., 48, 4588–4596.

16. Molinolo, A., Simpson, J. F., Thor, A. & Schlom, J. (1990). Enhanced tumor binding using immunohistochemical analyses by second generation anti-tumor-associated glycoprotein 72 monoclonal antibodies versus monoclonal antibody B72.3 in human tissue. Cancer Res., 50, 1291–1298.

17. De Pascalis, R., Gonzalez, N. R., Padlan, E. A., Schuck, P., Batra, S. K., Schlom, J. et al. (2003). In vitro affinity maturation of a specificity-determining region grafted humanized anticarcinoma antibody: isolation and characterization of minimally immunogenic high-affinity variants. Clin. Cancer Res., 9, 3521–3531.

18. Gold, D. V. & Mattes, M. J. (1988). Monoclonal antibody B72.3 reacts with a core region structure of O-linked carbohydrates. Tumour Biol., 9, 137–144.

19. Zhang, S., Walberg, L. A., Ogata, S., Itzkowitz, S. H., Koganty, R. R., Reddish, M. et al. (1995). Immune sera

† www.adobe.com
‡ pymol.sourceforge.net
and monoclonal antibodies define two configurations for the sialyl Tn tumor antigen. Cancer Res. 55, 3364–3368.

20. O’Boyle, K. P., Markowitz, A. L., Khorshidi, M., Lalezari, P., Longenecker, B. M., Lloyd, K. O. et al. (1996). Specificity analysis of murine monoclonal antibodies reactive with Tn, sialylated Tn, T, and monosialylated (2→6) T antigens. Hybridoma, 15, 401–408.

21. Hanisch, F. G., Uhlenbruck, G., Egge, H. & Peter-Katalinic, J. (1989). A B72.3 second-generation-monoclonal antibody (CC49) defines the mucin-carried carbohydrate epitope Gal \(^\text{b}(1-3)\)NeuAc \(^\text{a}(2-6)\)GalNAc. Biol. Chem. Hoppe Seyler. 370, 21–26.

22. Feizi, T. (1985). Demonstration by monoclonal antibodies that carbohydrate structures of glycoproteins and glycolipids are onco-developmental antigens. Nature, 314, 53–57.

23. Tamura, M., Milenic, D. E., Iwahashi, M., Padlan, E., Schlam, J. & Kashmiri, S. V. (2000). Structural correlates of an anticarcinoma antibody: identification of specificity-determining residues (SDRs) and development of a minimally immunogenic antibody variant by retention of SDRs only. J. Immunol. 164, 1432–1441.

24. Cygler, M., Rose, D. R. & Bundle, D. R. (1991). Recognition of a cell-surface oligosaccharide of pathogenic salmonella by an antibody Fab fragment. Science, 253, 442–445.

25. Rees, A. R., Pedersen, J. T., Searle, S. M. J., Henry, A. H. & Webster, D. M. (1994). Antibody structure from X-ray crystallography and molecular modeling. In Immunochorimistry (van Oss, C. J. & van Regenmortel, M. H. V., eds), pp. 615–650, Marcel Dekker, New York, NY.

26. McCallum, R. M., Martin, A. C. R. & Thornton, J. M. (1996). Antibody–antigen interactions: contact analysis and binding site topography. J. Mol. Biol. 262, 732–745.

27. Cygler, M. (1994). Recognition of carbohydrates by antibodies. Res. Immunol. 145, 36–40.

28. Nicholson, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins: Struct. Funct. Genet. 11, 281–296.

29. Pfleghart, J. W. (1999). The finer things in X-ray diffraction data collection. Acta Crystallog. sect. D, 55, 1718–1725.

30. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallog. sect. D, 54, 905–921.

31. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H. et al. (2000). The Protein Data Bank. Nucl. Acids Res. 28, 235–242.

32. Jones, T. A. & Kjeldgaard, M. (1994). O–The Manual. 5.10, Uppsala University Press, Uppsala.

33. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1996). Procheck—a program to check the stereochemical quality of protein structures. J. Appl. Crystallog. 29, 905–921.

34. Kleywegt, G. G. & Jones, T. A. (1996). xdlMAPMAN and xdlDATAMAN—programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. Acta Crystallog. sect. D, 52, 826–828.

35. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallog. 24, 946–950.

36. Merritt, D. A. & Bacon, D. J. (1997). Raster3D: photorealistic molecular graphics. Methods Enzymol. 277, 505–524.

Edited by I. Wilson

(Received 25 January 2005; received in revised form 7 March 2005; accepted 14 March 2005)