Structure and Molecular Interactions of a Unique Antitumor
Antibody Specific for N-Glycolyl GM3*

Ute Krengel‡‡, Lise-Lotte Olsson§§, Carlos Martínez‡‡, Ariel Talavera§§, Gertrudis Rojas‡‡, Elin Mier§§, Jonas Ångström, and Ernesto Moreno§§§

The Journal of Biological Chemistry Vol. 279, No. 7, Issue of February 13, pp. 5597–5603, 2004

© 2004 by The American Society for Biochemistry and Molecular Biology, Inc.

This paper is available on line at http://www.jbc.org

N-Glycolyl GM3 ganglioside is an attractive target antigen for cancer immunotherapy, because this epitope is a molecular marker of certain tumor cells and not expressed in normal human tissues. The murine monoclonal antibody 14F7 specifically recognizes N-glycolyl GM3 and shows no cross-reactivity with the abundant N-acetyl GM3 ganglioside, a close structural homologue of N-glycolyl GM3. Here, we report the crystal structure of the 14F7 Fab fragment at 2.5 Å resolution and its molecular model with the saccharide moiety of N-glycolyl GM3, NeuGcGalaα4Glcβ. Fab 14F7 contains a very long CDR H3 loop, which divides the antigen-binding site of this antibody into two subsites. In the docking model, the saccharide ligand is bound to one of these subsites, formed solely by heavy chain residues. The discriminative feature of N-glycolyl GM3 versus N-acetyl GM3, its hydroxymethyl group, is positioned in a hydrophilic cavity, forming hydrogen bonds with the carbonyl group of Asp H52, the indole NH of Trp H33 and the hydroxyl group of Tyr H50. For the hydrophobic methyl group of N-acetyl GM3, this environment would not be favorable, explaining why the antibody specifically recognizes N-glycolyl GM3, but not N-acetyl GM3. Mutation of Asp H52 to hydrophobic residues of similar size completely abolished binding. Our model of the antibody–carbohydrate complex is consistent with binding data for several tested glycolipids as well as for a variety of 14F7 mutants with replaced VL domains.

Received for publication, October 24, 2003, and in revised form, November 19, 2003
Published, JBC Papers in Press, November 19, 2003, DOI 10.1074/jbc.M311693200

The interest in antibodies that specifically recognize tumor markers has revived during the last few years, boosted by the approval, in 1997, of the first monoclonal antibody for cancer treatment by the United States Food and Drug Administration (FDA). Today, a total of six antibodies have been approved by the FDA for the therapy of several types of malignancies, and dozens of different antibodies with potential therapeutic effects in cancer are in clinical trials.

One attractive target antigen for cancer immunotherapy is the N-glycolyl GM3 ganglioside (1), a membrane-associated glycosphingolipid with a terminal N-glycolylated sialic acid residue (NeuGa). This tumor marker is expressed in certain tumor cells, such as melanoma and breast tumors (2, 3), but is otherwise absent from normal human tissues (4), which opens up the possibility of using antibodies specific toward these molecules both for diagnosis and immunotherapy.

However, targeting N-glycolyl GM3 gangliosides is no easy task, as this glycosphingolipid shares common structural features with many other gangliosides expressed on the cell surface. In particular, N-glycolyl GM3 is highly similar to N-acetyl GM3, which is present in most human tissues. In fact, what should be targeted corresponds to an extremely subtle chemical modification: the difference between the N-glycolyl group in NeuGa and the N-acetyl group of NeuAc consists of the addition of a single oxygen atom to the N-acetyl moiety (CH2OH instead of CH3 in the context of a trisaccharide) (see Fig. 1).

A murine monoclonal antibody (mAb), termed 14F7, that has this singular property has recently been described (5, 6). It exhibits high specificity toward N-glycolyl GM3 and strongly recognizes human melanoma and breast cancer tissues. 14F7 is quite unique, not only because of its binding specificity and rather strong binding affinity (in the nanomolar range) for N-glycolyl GM3, but also because of several other properties. In particular, 14F7 is an IgG antibody, belonging to the IgG1 subclass, whereas most other anti-ganglioside antibodies are IgMs, as a result of only a primary immune response to carbohydrate antigens (7). Furthermore, in experiments carried out in mice, the 14F7 monoclonal antibody showed a remarkable inhibition of the growth of solid tumors (6). Recently, a pilot clinical trial with radiolabeled 14F7 has been completed, with very encouraging results showing that this antibody is able to specifically target breast tumors and their metastases in vivo.3

The abbreviations used are: NeuGa, 5'-N-glycolylneuraminidate; Fab, antigen-binding fragments of immunoglobulins; CDR, complementarity determining region; CL, constant light chain; CH, constant heavy chain; Fv, Fab variable domains; Gal, galactose; Glc, glucose; mAb, monoclonal antibody; NeuAc, 5'-N-acetylneuraminidate; PDB, Protein Data Bank; R.m.s.d., root mean square differences; r.s.c.c., real space correlation coefficient; VL, variable light chain; VH, variable heavy chain; FEG, polyethylene glycol; ELISA, enzyme-linked immunosorbent assay; PBS, phosphate-buffered saline.

1 The abbreviations used are: NeuGa, 5'-N-glycolylneuraminidate; Fab, antigen-binding fragments of immunoglobulins; CDR, complementarity determining region; CL, constant light chain; CH, constant heavy chain; Fv, Fab variable domains; Gal, galactose; Glc, glucose; mAb, monoclonal antibody; NeuAc, 5'-N-acetylneuraminidate; PDB, Protein Data Bank; R.m.s.d., root mean square differences; r.s.c.c., real space correlation coefficient; VL, variable light chain; VH, variable heavy chain; FEG, polyethylene glycol; ELISA, enzyme-linked immunosorbent assay; PBS, phosphate-buffered saline.

2 G. Rojas, A. Talavera, Y. Munoz, E. Rengifo, U. Krengel, J. Ångström, J. Gavilondo, and E. Moreno, submitted manuscript.

3 A. Casaco, personal communication.
To date, only very few structures of carbohydrate-binding antibodies have been reported (8–17). This is remarkable in the view of the more than 250 entries containing Fab structures that have been deposited in the Protein Data Bank (18). One of the possible reasons for this unbalance could be that carbohydrates in general are not very immunogenic, producing mostly low affinity antibody responses of the IgM isotype, as mentioned above. 14F7 is thus an exception to the rule, thanks to a successful immunization procedure (5). Here, we report the crystal structure of the unliganded Fab fragment of 14F7 at 2.5 Å resolution. This structure provided the starting point for computer docking studies with the saccharide moiety of N-glycolyl GM3, NeuGco4Galβ4GlcβCer. The resulting molecular model will provide a basis for further genetic manipulation of the binding site to improve tumor targeting.

EXPERIMENTAL PROCEDURES
Preparation and Purification of 14F7 Fab Fragments
Fab 14F7 was prepared from purified mAb using papain digestion. Digestion time, temperature, and papain: mAb ratio were optimized following the progressive degradation of the 50 kDa heavy chain band on a 10% SDS-PAGE under reducing conditions. A 3 mg/ml mAb solution in 100 mM potassium phosphate buffer, pH 7.0, 10 mM cysteine, 10 mM EDTA was incubated with papain (1:100 papain:mAb ratio) at 37 °C for 4 h. Enzymatic digestion was stopped with 20 mM iodoacetamide (Digita) and loaded on a protein A-Sepharose column (Amersham Biosciences) to retain the Fc and undigested mAb, while the flow-through contained the Fab. The Fab fraction was then dialyzed extensively against 25 mM Tris-HCl, pH 8.8, and 25 mM Tris-HCl, pH 7.5, over a period of 100 min at a flow rate of 0.5 ml/min. The main peak was collected and concentrated to 10 mg/ml using a 10 kDa cut-off disposable ultrafiltration device (Millipore).

Crystallization
Initial crystallization experiments from a mixture of Fab 14F7 isoforms were unsuccessful, but once the purification protocol was significantly improved, crystals could be obtained readily within 20 min to a few days, provided that the protein solution was freshly prepared. Crystals were grown by vapor diffusion using the hanging drop technique at room temperature. The crystallization conditions were rather broad, with the precipitant of choice being PEG (ranging from PEG 400 to PEG 8000), buffered at pH 7.0 to 9.0, with or without the addition of various salts as additives. The crystal shapes varied from plates to rods and needles.

Crystals from many conditions were tested for diffraction and all of them were isomorphous. The space group was P212121, with unit cell dimensions a = 52.5 Å, b = 79.9 Å, c = 121.9 Å, corresponding to a Matthews coefficient, V_m, of 2.5 Å³/dalton (40) with one molecule per asymmetric unit. Two crystals were used for final data collection. Both of them grew in droplets formed by a 1:1 mixture of the protein solution (10 mg/ml in 25 mM Tris, pH 8.8) and the reservoir solution, containing 15% PEG 400 and 100 mM Hepes buffer at pH 7.0.

Data Collection
Data collection was performed at BL711 at the MAX-lab II synchrotron in Lund, Sweden, on a MAR CCD detector. Since shock-freezing of the crystals in a nitrogen stream resulted in strongly anisotropic diffraction, we decided to collect data at room temperature, exposing five different spots on the same crystal, a long needle. To achieve higher completeness and redundancy, a second crystal was used (data collection at 10 °C). All available data to 2.5 Å resolution were processed and scaled with XDS (19–21), resulting in 17,786 unique reflections. Statistics for the final data set are summarized in Table I.

Structure Determination and Refinement
The structure was solved by Molecular Replacement with the program AMoRe (22), using the Fab coordinates PDB 1C5C as the search model. The sequence was corrected, and the model was subjected to rigid body refinement with four separate rigid domains (light chain constant (CL)/variable (VL) domains and heavy chain constant (CH)/variable (VH) domains) followed by simulated annealing (23). Six percent of all reflections were used for the test set (24). Thereafter, positional and temperature factor refinement with CNS (25) were alternated with molecular rebuilding (program O, (26)) in several rounds. Water molecules were placed at peak positions of σ|Fo|−|Fc| maps and 2|Fo|−|Fc| maps greater than 3σr or 1σr, respectively, and checked at the graphics display. 37 water molecules are included in the final model. Refinement results are summarized in Table I.

The structure was validated against composite annealed omit maps from CNS. In a Ramachandran plot (27) produced by PROCHECK (28), 87.0% of the residues are located in the most favored regions of the plot and 12.2% are in the additional allowed regions. Two residues are in generously allowed regions and, finally, one residue. Ala L51, is in a disallowed region. The residue, slightly outside a generously allowed region, L51 is positioned in a γ-turn in complementarity-determining region loop 2 (CDR L2) and is found in a similar place of the Ramachandran plot in many reported Fab structures. The real space correlation coefficient (r.s.c.c.) for this residue against a composite annealed omit map is 75%.

The sequences for the conserved domains have not been determined, but were taken from the Kabat data base (29) (www.kabatdatabase.com). The sequence for H187–H188 was first chosen as Pro-Arg, however, the electron density clearly indicated that H188 must be a tryptophan. An examination of the sequence data base revealed that Trp H188 is always accompanied by a threonine residue at H187. The real
space correlation coefficients for these two residues are 93% (Thr H187) and 87% (Try H188).

Here, we use the Kabat numbering scheme for the variable regions, which for this particular antibody introduced insertion letters in the heavy chain sequence after positions 52 (52A), 82 (82A to 82C), and 100 (100A to 100H).

Modeling of the 14F7-NeuGc-GM3 Complex
Preparation of Protein and Ligand—The crystal structure of the 14F7 Fab fragment was taken as the starting geometry for dynamics and docking simulations. The CL and CH1 constant domains and all water molecules were removed, thus leaving only the Fab variable domains (Fv) for calculations. Energy minimizations and dynamics simulations were performed using the program CHARMM, version 24 (39). Both docking and dynamics simulations were carried out using the Gasteiger–Marsili charges for the protein and, subsequently, an energy minimization was carried out in three steps, in order to refine the stereochemistry. First, harmonic constraints of 50 kcal/mol Å2 were applied to all a carbons during 100 steps of a steepest descent geometry relaxation. Thereafter, a similar cycle was performed with the constraints reduced to 10 kcal/mol Å2. Finally, the harmonic constraints were removed to perform 500 steps of conjugate gradient minimization. A distance dependent dielectric constant (ε = 8r) and a 12 Å cutoff distance were used for non-bonded interactions. The terminal disaccharide moiety of N-glycolyl GM3 (NeuGc3Galβ) used in the docking procedure was built using the program CHARMM, based on topology files included in the Quanta2000 package (Accelrys Inc.). As for the protein receptor, all hydrogens were explicitly included. The third carbohydrate residue, 4Glcβ1, was added after docking of the disaccharide since it does not participate in any protein interactions. Glycosidic dihedral angles are defined as described in Refs. 31–33.

Conformational Sampling in the Binding Site Region using Molecular Dynamics—Molecular dynamics was employed to explore the conformational space of the receptor binding site, using the energy-minimized structure as starting geometry. Most of the amino acids in the complementarity determining regions and several other neighboring residues were allowed to move. It should be noted that conformational sampling was not restricted to the side chains, but also included the backbone atoms. The structure was initially heated from 0 to 300 K during 10 ps and equilibrated at this final temperature during another 12 ps. Thereafter followed a 1000-ps simulation with a 0.001-ps integration step. The SHAKE algorithm as implemented in CHARMM was used to constrain bonds to hydrogens. The same non-bonding parameters as for energy minimization were applied. Coordinates were collected every 500 steps (0.5 ps), so that at the end, 2000 coordinate frames were stored in a trajectory file. These coordinate frames were subsequently clustered using an r.m.s.d. cutoff of 1 Å.

Docking Simulations of the 14F7-NeuGc-GM3 Complex—Docking simulations were performed employing a modified version of the program DOCK 3.5 (34, 35) called DOCKdyna. This program allows the screening of a much larger number (thousands) of receptor conformations for binding of a single ligand, by using a novel and automatic way of representing the receptor binding site (36) and re-calculating the force field grids used for energy evaluation for every receptor coordinate frame. Ligand flexibility was allowed. The top-ranking solutions, one from each cluster, were selected for visual inspection and analysis.

Analysis of 14F7 Mutants
Site-directed Mutagenesis—Modeling studies indicated that Asp H52 plays a crucial role in the binding of NeuGc-GM3. In order to probe the role of this residue for the molecular recognition of the saccharide ligand, two primers were designed, targeting the 14F7 residue Asp H52 for site-directed mutagenesis. One primer was used to mutate Asp H52 to Val (DH52V), the other one for mutation to Ile (DH52I). The substitutions were performed in the phage-displayed VL-shuffled single chain antibody cloned into the phagemid vector pM13K07 (37). Recombinant forms of all eight immunoreactive phage-displayed antibody fragments was determined by enzyme-linked immunosorbent assay (ELISA). Polyvinyl chloride microtiter plates (Costar, USA) were dried with N-glycolyl GM3 (1 μg/well) in methanol. Plates were washed with PBS and blocked with 2% skimmed milk powder in PBS (M-PBS) during 1 h at room temperature. Phage-displaying antibody fragments were diluted in M-PBS (10 μg/particles/ml) and incubated on the plates for 1.5 h at room temperature. After washing the plates with PBS plus Tween 20 0.1% (PBS-T), an anti-M13 mAb coupled to horseradish peroxidase (Amersham Biosciences), appropriately diluted in M-PBS, was added to the wells. The plates were washed again and the substrate solution (500 μg/ml orthophenylene diamine and 0.015% hydrogen peroxide in 100 mM citrate buffer, pH 5.5) was added. The reaction was stopped after 15 min with 2.0 M sulfuric acid. The absorption (492 nm) was determined with a microplate reader (Bio-Rad). Each sample was tested twice for the binding of N-glycolyl GM3. The 3Fm clone was included as a positive control for N-glycolyl GM3 recognition.

Figures and Tables—Fig. 1 was prepared using ChemDraw from CambridgeSoft and Fig. 2 using Quanta2000/CHARMM27 modeling package from Accelrys Inc.

RESULTS AND DISCUSSION
Crystal Structure of Fab 14F7—The crystal structure of the 14F7 Fab fragment was determined at 2.5 Å resolution and refined to crystallographic R-factors of 18.1% (Rmerge) and 23.8% (Rfree), respectively, with good stereochemistry. The electron density is well defined for a 2.5 Å structure (average r.s.c.c. for composite annealed omit map is 86%), due to the good data quality. Only a few regions in the structure are less well defined. In particular, the C-terminal residues L213-L214 and H128-H133, there is some discontinuous density, but not enough for this loop to be correctly traced. Many Fab structures deposited in the protein data base lack this particular loop. Further, due to weak electron density for their side chains, Lys H13 and Lys H64 were modeled as alamines and Cys H215 was substituted by a glycine residue.

Overall, the crystal structure of Fab 14F7 is highly similar to other Fab structures, with one noticeable variation. The complementarity determining region CDR H3 of 14F7 (residues 96–102 with eight letter insertions at 100), is among the longest found both in the Kabat data base and in the PDB, with 16 amino acid residues (see Table II). This long loop protrudes from the middle of the antibody-combining region, dividing the binding site into two well-separated zones that we will refer to as the VL and VH subsites (Fig. 2A).

4 E. Moreno, to be published.
In the crystal structure of 14F7, the long CDR H3 loop is well defined by electron density (r.s.c.c. /H11005 91%) and has rather low temperature factors (the average B-factor for this region is 41.2 Å², compared with 51.7 Å² for the whole protein), indicating that this part of the structure is quite rigid. However, a closer inspection of the crystallographic contacts revealed significant stabilizing interactions from symmetry related molecules. The strongest interactions involve three arginine side chains, Arg H98, H100, and H100A, two of which participate in bi-dentate salt bridges with the carboxylates from neighboring glutamate side chains (L79# and L81#), and one of them forming a hydrogen bond to a third glutamate residue (H191# in the constant region of the Fab). Additional stabilizing interactions involve Ile H100C and the hydrophobic parts of the Ser L 203#/Pro L204# side chains. All of these interactions are intermolecular interactions with symmetry-related molecules in the crystal (indicated by #).

Long loops that are not stabilized over long stretches by interactions with other residues are in general very flexible. It is thus suggestive that the rigidity of the CDR H3 loop is an artifact induced by crystallization and that this loop, which is only stabilized by interactions to symmetry related molecules in the crystal, is usually much more flexible, especially at its tip. Nevertheless, the molecular dynamics simulations we performed suggest that the conformation of CDR H3 is quite stable, and that no significant conformational changes of its backbone are to be expected upon binding of the antigen.

In this context, it is worthwhile noting the presence of three consecutive tyrosines at the base of CDR H3 (residues 100D, 100E, and 100F). The aromatic rings of these tyrosines are interacting with other neighboring aromatic residues, being embedded in the structure and having very favorable interactions with the surrounding amino acids. These three tyrosines, together with Pro H95, confer rigidity to the base and middle part of CDR H3.

Although both the VL and VH subsites are rather open, the tip of CDR H3 is bent toward the VH subsite. A small groove is created by the top part of this loop, especially by the side chains of Arg H98 and Arg H100A, together with several amino acids from CDR H1 and H2 and one single residue from the light chain (Trp L94).

Modeling of the 14F7-NeuGc-GM3 complex—Since 14F7 so

| PDB entry | Antibody name | CDR H3 sequence |
|------------|---------------|-----------------|
| 1RIH       | 14F7          | ESPRLRGIYIYAMDY|
| 1OM3, 1OP3, 1OP5 | 2G12         | KGSDRLSDNDP----FD4|
| 1UCB, 1CLY, 1CLZ | br96        | GLDGDG-------FAY|
| 1MPA-E     | L155-4        | GGGQDY--------GYD|
| 1M71, 1M7D, 1M7I | SyyaJ6     | GGAAGA--------MDY|
| 1FW, 1FX, 1FY | S-20-4       | HAY----------LDY|
| 1PSK       | ME36.1        | KS-------------FD4|

In the crystal structure of 14F7, the long CDR H3 loop is well defined by electron density (r.s.c.c. = 91%) and has rather low temperature factors (the average B-factor for this region is 41.2 Å², compared with 51.7 Å² for the whole protein), indicating that this part of the structure is quite rigid. However, a closer inspection of the crystallographic contacts revealed significant stabilizing interactions from symmetry related molecules. The strongest interactions involve three arginine side chains, Arg H98, H100, and H100A, two of which participate in bi-dentate salt bridges with the carboxylates from neighboring glutamate side chains (L79# and L81#), and one of them forming a hydrogen bond to a third glutamate residue (H191# in the constant region of the Fab). Additional stabilizing interactions involve Ile H100C and the hydrophobic parts of the Ser L 203#/Pro L204# side chains. All of these interactions are intermolecular interactions with symmetry-related molecules in the crystal (indicated by #).

Long loops that are not stabilized over long stretches by interactions with other residues are in general very flexible. It is thus suggestive that the rigidity of the CDR H3 loop is an artifact induced by crystallization and that this loop, which is only stabilized by interactions to symmetry related molecules in the crystal, is usually much more flexible, especially at its tip. Nevertheless, the molecular dynamics simulations we performed suggest that the conformation of CDR H3 is quite stable, and that no significant conformational changes of its backbone are to be expected upon binding of the antigen.

In this context, it is worthwhile noting the presence of three consecutive tyrosines at the base of CDR H3 (residues 100D, 100E, and 100F). The aromatic rings of these tyrosines are interacting with other neighboring aromatic residues, being embedded in the structure and having very favorable interactions with the surrounding amino acids. These three tyrosines, together with Pro H95, confer rigidity to the base and middle part of CDR H3.

Although both the VL and VH subsites are rather open, the tip of CDR H3 is bent toward the VH subsite. A small groove is created by the top part of this loop, especially by the side chains of Arg H98 and Arg H100A, together with several amino acids from CDR H1 and H2 and one single residue from the light chain (Trp L94).

Modeling of the 14F7-NeuGc-GM3 complex—Since 14F7 so
far resisted all attempts to crystallize in complex with its carbohydrate ligand, we resorted to construct a theoretical model of the complex of the antibody with its antigen, N-glycolyl GM3. The shape of the binding site strongly suggested that the binding site for N-glycolyl GM3 is located in the VH subsite, in the groove described above.

Starting from the crystal structure of the Fab fragment of 14F7 and using molecular dynamics, we generated an ensemble of 2000 receptor coordinate sets, which after clustering were reduced to 754 different binding site conformations. Subsequent docking simulations with the crystal disaccharide of NeuAc-GM3, NeuGc3Galβ3, resulted in a theoretical model of the antibody-ligand complex (Fig. 2B). Compared with the crystal structure of unliganded 14F7, no major changes were introduced in the computer model. Apart from Arg H100, docking merely affected the side chains of the groove residues. The protein backbone in CDR H3 superimposes with root mean square differences (r.m.s.d.) of 0.5 Å. The residue that is most affected by the docking simulations is Tyr H100D. Other minor conformational changes are observed for Arg H100A, Tyr H100F, Arg H98 and Trp L94 (in order of decreasing differences compared with the crystal structure). Trp H33 is totally superimposable, while Asp H52 exhibits a minimal shift of its side chain. The buried surface area for the trisaccharide ligand is 206 Å² (out of the total of 468 Å²). This number compares well with the buried areas for similar molecules reported in the PDB, such as sialyllactose ( NeuAc, as no NeuGc is present in the PDB), ranging from 139 to 310 Å².

In the model, the hydroxyl group (O11H) of the N-glycolyl moiety of the sialic acid is located in a highly hydrophilic pocket created by the carbohydrate ligand of Asp H52 and the indole NH of Trp H33, while the apolar side of the N-glycolyl hydroxymethyl group faces the aromatic ring of Tyr H56 forming a small hydrophobic patch (see Fig. 2 and Table III). Tyr H50, which is positioned somewhat more peripherally, also contributes a long-range hydrogen bond to the N-glycolyl OH. The carbonyl oxygen (O10) of the N-glycolyl moiety is hydrogen bonded to Thr H54. This residue may alternatively form a hydrogen bond to O7 of the glycerol tail.

Ligand interactions further involve the carboxylate group of the sialic acid residue (to Arg H98 εNH, and the main chain NH of Gly H100B) as well as all three hydroxyl groups of the glycerol tail as detailed in Fig. 2 and Table III. The Galβ4 residue also participates in hydrogen bonding interactions via the O4 oxygen to Arg H100A εNH and its glycosidic oxygen (formally belonging to NeuGc3) to Arg H100A εNH. The NeuGc3Gal glycosidic torsion angles (Φ,Ψ = −14°, −33°) are close to literature values for the so-called anticalin conformation (32) and in accord with the conclusions drawn for the P3 antibody, which also binds NeuGc-GM3 (38), albeit with a somewhat broader specificity range. No interactions are observed for the glucose moiety when added to the disaccharide using minimum energy Galβ4Glc glycosidic torsion angles (Φ,Ψ = 54°, −22°).

In the crystal structure, a tightly bound water molecule (B = 34.6 Å²) is positioned midway between Arg H98 εNH and the Tyr H100D hydroxyl group, forming hydrogen bonds to both residues. When superimposed on the modeled complex with NeuGc3Galβ4Glcβ, this water molecule falls into place between Arg H98 εNH and the carboxyl O1A of the sialic acid residue. Only slight adaptations of the Arg H98 and Tyr H100D side chains would be necessary to accommodate the water molecule in the model of the antigen-complex. However, while it is possible that the water molecule remains bound to the cavity even when NeuGc-GM3 is bound to 14F7, it seems more likely that this is not the case, for the following reason. The sialic acid carboxylate is negatively charged and would be stabilized more strongly by a direct rather than by an indirect interaction with a positively charged arginine residue. Salt bridge formation between Arg H98 and the carboxyl group of NeuGc, as predicted by our docking model, is therefore the most likely molecular interaction for electrostatic reasons.

Summarizing, we would like to point out that in the 14F7 NeuGc-GM3 model, ligand contacts occur solely with heavy chain amino acids. Not a single residue of VL is involved in any interactions with the carbohydrate ligand. Trp L94, which is closest to the binding pocket, is still at a distance of 8 Å from the nearest atom of the sialic acid residue. The aim of our investigation was to reveal how the 14F7 monoclonal antibody succeeds in discriminating between the N-glycolyl group of NeuGc and the N-acetyl group of NeuAc present in most other gangliosides (Fig. 1), a difference corresponding to the mere addition of one oxygen atom to the N-acetyl moiety in the context of a trisaccharide. Our results indicate that it is the hydrophilic environment around the hydroxymethyl group of NeuGc that explains the fine discrimination, with Asp H52, Trp H33, and Tyr H50 playing pivotal roles for recognition.

Validation of the Computer Model—While the computer

| Amino acid | Sugar                      | Distance |
|------------|----------------------------|----------|
| Ser H31    | NeuGca3 O9 (glycerol tail) | 2.5 Å    |
| chain CO   |                            |          |
| Trp H33    | NeuGca3 O11 (glycolyl moiety) | 2.9 Å  |
| indole NH  |                            |          |
| Tyr H50    | NeuGca3 O11 (glycolyl moiety) | 4.1 Å  |
| εOH        | NeuGca3 HO7 (glycerol tail) | 3.1 Å    |
| Asp H52    | NeuGca3 HO7 (glycerol tail) | 2.9 Å    |
| εO        | NeuGca3 O8 (glycerol tail)  | 2.9 Å    |
| Thr H54    | NeuGca3 O10 (glycolyl moiety) | 2.9 Å  |
| γOH        | NeuGca3 HO7 (glycerol tail) | 2.9 Å    |
| Arg H98    | NeuGca3 O8 (glycerol tail)  | 2.9 Å    |
| εNH        | NeuGca3 O1A (carboxyl group) | 3.0 Å  |
| Arg H100A  | NeuGca3 O2 (glycosidic oxygen) | 3.1 Å |
| εNH        | NeuGca3 O1A (carboxyl group) | 3.0 Å  |
| Gly H100B  | NeuGca3 O1A (carboxyl group) | 3.0 Å  |
| main chain | NeuGca3 HO8 (glycerol tail) | 2.8 Å    |
| NH         |                            |          |

FIG. 2. Binding of the two Asp H52 point mutants of 14F7 to NeuGc-GM3, as measured by ELISA. For these two mutants (Asp H52 → Val and Asp H52 → Ile), no binding at all was detected. The positive (control) clone is the VL mutant 3Fm, having the original VH domain of 14F7.
docking model presented here certainly cannot be taken as conclusive proof of the structure of the antibody-carbohydrate complex in the same way as would be possible for a crystal structure, several biochemical results support the major features of the structural model. First, we observe that in the 14F7 NeuGc-GM3 model, ligand contacts occur solely with heavy chain amino acids. In this context, it is enlightening to study the results from a chain-shuffling scFv phage display library prepared for other purposes. In this library, the original 14F7 heavy chain variable region was combined with a diverse set of light chain variable regions (VL) from both mice and humans (isotopes kappa and lambda). More than one hundred clones producing phage-displayed antibody fragments able to specifically recognize the antigen were isolated and five of them were selected for sequencing. The VL complementarity determining regions of these clones (CDR L1-L3) differ substantially, both when compared with each other and also compared with the CDR (14F7) (substitutions indicated in Fig. 2A). Despite these differences, the five VL-shuffled variants, however, retained the same binding properties displayed by 14F7, confirming that the VL region is not crucial for recognition of NeuGc-GM3.

Further, according to our computer model, residues critically involved in binding the N-glycolyl group of GM3 are Asp H52, Trp H33, and Tyr H50. Taking Asp H52 as a test case for NeuGc-GM3 recognition, we made two point mutants of this residue, replacing it by either Ile or Val. As predicted, both mutants turned out to no longer exhibit binding to NeuGc-GM3 (Fig. 3).

Analysis of Reported Binding Data for Other Glycolipids, Based on the Theoretical Model—A range of gangliosides having N-glycolyl- or N-acetyl-substituted sialic acids (or a combination thereof) have recently been tested for binding to the 14F7 antibody, resulting in a strong recognition solely of NeuGc-GM3 (5). As shown above, the absolute requirement for an N-glycolyl group rules out binding of any ganglioside carrying an N-acetyl group, whether the sialic acid is bound terminally or at a branching point, such as NeuAc-GM3, NeuAc-GM2, NeuAc-GM1a, etc. Even NeuGc-GM2 tested negative, which contrasts the positive binding of the P3 antibody (38). The negative binding of NeuGc-GM2 to 14F7 can readily be explained from the modeled complex by steric interference between the terminal GalNAcβ residue and arginines H98 and H100A.

However, more surprising was the non-binding of compounds carrying a terminal NeuAc residue such as sialylneolactotetraosylceramide (NeuGc3Galβ4GlcNAcβ3Galβ4Glcβ1Cer) and NeuGc-NeuGc-GD3 (NeuGcααNeuGc3Galβ4Glcβ1Cer). In the case of sialyl-neolactotetraosylceramide, the presence of an additional N-acetyllactosamine segment relative to NeuGc-GM3 results in an ~90° rotation of the sialic acid when superimposing the lactosylceramide part of these two structures. Furthermore, different Glcβ1Cer linkage conformations are most likely at hand, suggesting that an even more disadvantageous binding epitope presentation will occur as a result. Due to the requirement of an unsubstituted glycerol tail on the sialic acid, it must be concluded that the terminal NeuGc of NeuGc-NeuGc-GD3 would have to be involved if binding to 14F7 were to occur. However, superimposing the structures
generated for each of the twelve possible low energy conformations of the NeuGcααNeuGc linkage (38) reveals that the internally located sugar residues will in all cases clash with the protein in different ways, thus accounting for the non-binding of NeuGc-NeuGc-GD3 to 14F7.

Comparison of 14F7 with Other Carbohydrate-binding Antibodies—In most Fab-antigen complexes, the antigen binding site is positioned in a groove at the interface between the light and heavy chain variable domains, and residues from both VL and VH contribute to ligand binding. This is in contrast to our findings, where direct contacts to NeuGc-GM3 are only observed with heavy chain amino acid residues. For Fab 14F7, there is no groove at the VL/VH interface. Instead, the long CDR H3 loop protrudes from this position and divides the potential binding site into two subsites. It therefore comes as no surprise that a superposition of 14F7 with other Fab structures, which were crystalized in complex with their respective carbohydrate antigens (see Table II), revealed that the saccharide ligands in those complexes are in a large majority located in a position overlapping with the long CDR H3 loop of 14F7. A noteworthy exception is the structure of the 2G12 Fab fragment (PDB 10M3, PDB 10P3, PDB 10P5) (17), a broadly neutralizing antibody against HIV-1 gp120. Interestingly, a superposition with this Fab fragment reveals that its bound antigen Manα1–2Man binds at a site in 2G12 that roughly corresponds to the VH subsite of 14F7. However, it should also be noted that the 2G12 antibody has an unusual architecture due to VH domain swapping, involving the tilting of the VH domain, which makes it difficult to compare the two antibodies in detail.

CONCLUSIONS

In the present investigation, we have determined the crystal structure of Fab 14F7, a unique anti-tumor antibody specific for N-glycolyl GM3. Taking the x-ray structure as a starting point, computer-docking studies were performed with the saccharide moiety of the antigen, NeuGc3Galβ4Glcβ. The resulting molecular model of the carbohydrate complex explains the exquisite binding properties of 14F7 and is consistent with glycosphingolipid binding data for a broad set of 14F7 mutants. This model further lays the foundation for future genetic engineering of the antibody to either improve its performance or explore new cellular targets.

Acknowledgments—We thank Dr. Adriana Carr for supplying large quantities of 14F7 for crystallization, Prof. Karl-Anders Karlsson for supporting the work of E. M. at the University of Göteborg, and Georgy Katona and the staff at EL111 at the MAX-lab II synchrotron in Lund for help with data collection.

REFERENCES

1. Malvykh, Y. N., Schauer, R., and Shaw, L. (2001) Biochimie (Paris) 83, 623–634
2. Mukling, J., Steiner, H., Peter-Katalinic, J., Marx, U., Bethke, V., Neuman, V., and Lehmann, J. (1994) J. Biochem. 116, 64–73
3. Marquina, G., Waki, H., Fernández, L. E., Ken, K., Carr, A., Valiente, O., Pérez, R., and Ando, S. (1996) Cancer Res. 56, 5165–5171
4. Varki, A. (2001) Biochimie (Paris) 83, 615–622
5. Carr, A., Mulet, A., Mazorra, Z., Vázquez, A. M., Alfonso, M., mesa, C., Rengifo, E., Pérez, R., and Fernández, L. E. (2000) Hybridomics 19, 241–247
6. Carr, A., mesa, C., del Carmen Arango, M., Vázquez, A. M., and Fernández, L. E. (2002) Hybrid. Hybridomics 21, 463–468
7. Portozzalain, J. (2000) Clin. Rev. Allergy Immunol. 19, 73–78
8. Cygler, M., Rose, D. R., and Bundle, D. R. (1991) Science 253, 442–445
9. Cygler, M., Wu, W., Zdanov, A., Bundle, D. R., and Rose, D. R. (1993) Biochem. Soc. Trans. 21, 437–441
10. Rose, D. R., Przybylska, M., To, R. J., Kayden, C. S., Oomen, R. P., Ver Borg, E., Young, N. M., and Bundle, D. R. (1993) Prot. Sci. 2, 1106–1113
11. Bundle, D. R., Baumann, H., Brisson, J.-R., Gagné, S. M., Zdanov, A., and Cygler, M. (1994) Biochemistry 33, 5183–5192
12. Zdanov, A., Li, Y., Bundle, D. R., Deng, S.-J., MacKenzie, C. R., Narang, S. A., Young, N. M., and Cygler, M. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 6423–6427
13. Jeffrey, P. D., Bajorath, J., Chang, C. Y. Y., Yelton, D., Hellstroem, I., Hellstrom, K. E., and Sherif, S. (1995) Nat. Struct. Biol. 2, 466–471
14. Pichia, S. L., Murali, R., and Burnett, R. M. (1997) J. Struct. Biol. 119, 6–16
15. Villeneuve, S., Souchon, H., Riottot, M.-M., Mazie, J.-C., Lei, P.-S., Glaudemans, C. P. J., Kovalcůná, P., Fournier, J.-M., and Alzari, P. M. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 8433–8438

16. Vyas, N. K., Vyas, M. N., Chervenak, M. C., Johnson, M. A., Pinto, B. M., Bundle, D. R., and Quiocho, F. A. (2002) Biochemistry 41, 13575–13586

17. Calarese, D. A., Scanlan, C. N., Zwick, M. B., Deechongkit, S., Mimura, Y., Kunert, R., Zhu, P., Wormald, M. R., Stanfield, R. L., Roux, K. H., Kelly, J. W., Rudd, P. M., Deew, R. A., Katunger, H., Burton, D. R., and Wilson, I. A. (2003) Science 300, 2065–2071

18. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) Nucleic Acids Res. 28, 235–242

19. Kabsch, W. (1988) J. Appl. Crystallogr. 21, 67–71

20. Kabsch, W. (1988) J. Appl. Crystallogr. 21, 916–924

21. Kabsch, W. (1990) J. Appl. Crystallogr. 23, 795–800

22. Navaza, J. (1994) Acta Crystallogr. Series A 50, 157–163

23. Adams, P. D., Pannu, N. S., Read, R. J., and Brünger, A. T. (1997) Proc. Natl. Sci. U. S. A. 94, 5018–5023

24. Brünger, A. T. (1992) Nature 355, 472–475

25. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., and Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D 54, 905–921

26. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119

27. Ramakrishnan, C., and Ramachandran, G. N. (1965) Biophys. J. 5, 909–933

28. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291

29. Johnson, G., and Wu, T. T. (2001) Nucleic Acids Res. 29, 205–206

30. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) J. Comp. Chem. 4, 187–217

31. Imberty, A., Delage, M.-M., Bourne, Y., Cambillau, C., and Pérez, S. (1991) Glycoconj. J. 8, 456–483

32. Siebert, H.-C., Reuter, G., Schauer, R., von der Lieth, C.-W., and Dabrowski, J. (1992) Biochemistry 31, 6962–6971

33. Nyholm, P.-G., and Pascher, I. (1990) Biochemistry 32, 1225–1234

34. Kuntz, I. (1992) Science 257, 1078–1082

35. Guschwew, D. A., and Kanz, I. D. (1996) J. Comput. Aided Mol. Design 10, 123–128

36. Moreno, E., and León, K. (2002) Proteins 47, 1–13

37. Marks, J. D., Hogenboom, H. R., Bonertz, T. P., McCafferty, J., Griffiths, A. D., and Winter, G. (1991) J. Mol. Biol. 222, 581–597

38. Moreno, E., Lanne, B., Vazquez, A. M., Kawashima, I., Tai, T., Fernández, L. E., Karlson, K.-A., Angstrom, J., and Perez, R. (1998) Glycoconjug. J. 15, 695–705

39. Engh, R. A., and Huber, R. (1991) Acta Crystallogr. Sect. A 47, 392–400

40. Matthews, B. W. (1968) J. Mol. Biol. 33, 491–497