Crystal structure of an anti-carbohydrate antibody directed against *Vibrio cholerae* O1 in complex with antigen: Molecular basis for serotype specificity

S. Villeneuve*, H. Souchon†, M.-M. Riottot‡, J.-C. Mazié§, P.-s. Lei¶, C. P. J. Glaudemans¶, P. Kováč§, J.-M. Fournier*, and P. M. Alzari†

†Unité de Biochimie Structurale (Centre National de la Recherche Scientifique, Unité de Recherche Associée 2185), *Unité du Choléra et des Vibrions, ‡Unité d’Immunologie Structurale, and §Laboratoire d’Ingénierie des Anticorps, Institut Pasteur, 25 rue du Dr. Roux, 75724 Paris, France; and ¶Laboratory of Medicinal Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0815.

Communicated by John B. Robbins, National Institutes of Health, Bethesda, MD, January 19, 2000 (received for review November 5, 1999)

The crystal structure of the murine Fab 5-20-4 from a protective anti-cholera Ab specific for the lipopolysaccharide Ag of the Ogawa serotype has been determined in its unliganded form and in complex with synthetic fragments of the Ogawa O-specific polysaccharide (O-SP). The upstream terminal O-SP monosaccharide is shown to be the primary antigenic determinant. Additional perosamine residues protrude outwards from the Ab surface and contribute only marginally to the binding affinity and specificity. A complementary water-excluding hydrophobic interface and five Ab–Ag hydrogen bonds are crucial for carbohydrate recognition. The structure reported here explains the serotype specificity of anti-Ogawa Abs and provides a rational basis toward the development of a synthetic carbohydrate-based anti-cholera vaccine.

Cholera remains a major public health problem. Although the improvement of hygiene is an efficient mechanism for the long-term control of the disease, that improvement is still a distant goal in many developing countries where an efficient, simple vaccine could be of great help. *Vibrio cholerae* strains can be divided into different serogroups based on their lipopolysaccharide (LPS) structures. The serogroup O1, responsible for most cholera outbreaks, includes two major serotypes, Ogawa and Inaba (1). The two types differ only by a single 2-O-methyl group that is present in the upstream (“nonreducing”) terminal perosamine unit of the Ogawa O-specific polysaccharide (O-SP) and that is absent in the terminal perosamine residue of the Inaba O-SP (Fig. 1a) (2, 3). Previous vaccine trials in Asia have shown that killed-cell Inaba vaccines protect against both serotypes, but Ogawa vaccines protect against only the homologous serotype (1). Similar results observed with serotype-specific mAbs tested for protective activity in neonatal mice (4–6) led to the conclusion that the Ogawa serotype contains a specific antigenic determinant, whereas the Inaba serotype contains a different antigenic determinant that crossreacts with the Ogawa serotype. The Ogawa serotype-specific epitope is probably associated with the upstream terminal perosamine residue of the O-SP, because Inaba strains are *rfbT* mutants of wild-type Ogawa strains in which methylation of the terminal perosamine is inactivated (2, 3, 7). Also, binding studies of anti-Ogawa Abs IgG1 S-20-6 and IgG1 S-20-4 with synthetic methyl α-glycosides of fragments, up to the hexasaccharide, of the Ogawa O-SP, as well as with analogs of the terminal monosaccharide, revealed that the terminal residue accounts for approximately 90% of the maximal binding energy (8). This hypothesis raises the question of how such a small antigenic determinant—a single methyl group at the 2-OH position of the terminal perosamine—can dictate a highly specific immune response.

Protective responses against microbial pathogens are frequently based on anti-carbohydrate Abs produced against polysaccharides located on their cell surface (9). However, the extensive immunochemical data available on anti-carbohydrate Abs (10–13) contrasts with the paucity of structural information on these types of systems, because to date only two carbohydrate-specific Abs have been studied in complex with ligands, namely the mAb Se155-4 specific for a *Salmonella* serotype-B O-Ag (14, 15) and the antitumor Ab BR96 (16). To gain further insight into the structural basis of carbohydrate recognition and to understand *V. cholerae* serotype specificity, we have undertaken crystallographic studies of protective anti-cholera Abs in complex with synthetic analogs of the O-Ag. We report here the crystal structure of Fab S-20-4 in its unliganded form and in complex with the synthetic mono- and disaccharide fragments mimicking the terminal sugar residues of Ogawa-LPS (see Fig. 1).

### Materials and Methods

#### Crystallization and Data Collection.

Monoclonal mouse Ab S-20-4 (IgG1, κ) was purified from the ascitic fluid by ion-exchange chromatography. The molecule was eluted at 20 mM phosphate from a DEAE Trysacryl M column (BioSepra, Paris) equilibrated with 10 mM phosphate buffer at pH 8. The Fab fragment was produced by papain digestion with an enzyme to substrate molar ratio of 1:30 in 0.1 M phosphate buffer (pH 7.2), 2.5 mM EDTA, 10 mM L-cysteine at 37°C. The Fab fragment was separated by ion-exchange chromatography, further purified by gel filtration on a GF 250 column (DuPont), and concentrated to 10 mg/ml in 20 mM Tris·HCl (pH 7.2). Crystals were obtained by the hanging-drop vapor diffusion method at 18°C using 15%–18% PEG 6000 and 50 mM sodium acetate at pH 5.5. Larger crystals were obtained by macroseeding. The synthesis of methyl α-glycosides of the mono-(1) and disaccharide (2) of the O-Ag of *V. cholerae* O1 serotype Ogawa (shown in Fig. 1b) has been described elsewhere (17, 18). Crystals of the Fab-Ag complexes were grown by recrystallization with a protein to ligand molar ratio of 1:10 under similar conditions as for the unliganded Fab molecule.

The crystals of the unliganded Fab belong to space group P2₁, with cell dimensions \( a = 45.55 \, \text{Å}, b = 113.36 \, \text{Å}, c = 46.23 \, \text{Å}, \) and \( β = 100.7°, \) and contain one protein molecule in the asymmetric unit. Crystals of the Fab-monosaccharide \( (a = 45.57 \, \text{Å}, b = 113.1 \, \text{Å}) \) were grown in a similar manner.

### Abbreviations

- **CDR**: complementarity-determining regions
- **LPS**: lipopolysaccharide
- **O-SP**: O-specific polysaccharide
- **rfbT**: gene encoding a protein involved in the biosynthesis of the terminal perosamine residue of the O-SP

### Data deposition

The atomic coordinates have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes 1F4W, 1F4X, and 1F4Y).

This work was supported by the French Ministry of Health, the Pasteur Institute, the European Community (TMR network CE5-4071), and the French National Agency for Research on AIDS and Viral Hepatitis (ACIPA 95.011). We thank Dr. D. Chevalier for the generous gift of the Fab 5-20-4 antibody. We also thank Dr. J. Gouy for his helpful comments on the manuscript.
Using the program XPLOR (24), the resulting model (excluding the third hypervariable loop of the heavy chain) refined to an \( R \) factor of 23.5% \( (R_{	ext{free}} = 36.5\%) \) at 2.5 \( \AA \) resolution. Subsequent refinement was carried out by using the program REFMAC from the CCP4 suite of programs (25), alternated with manual model building by using the program O (26). Unless indicated otherwise, all other calculations were performed with programs from the CCP4 package. During the early stages of refinement, solvent-exposed side chains and loops that were not visible in density were omitted from the model. The final model (Table 1) includes all 426 residues of the protein and 51 solvent molecules clearly defined in density. Some amino acid residues with solvent-exposed side chains not visible in density were modeled as alanine. These include Arg L24, Arg L184, Glu L187, Arg L188, His L203, Glu H5, Arg H44, Arg H56, Lys H64, Arg H105, Lys H115, Glu H204, and Lys H218 (Kabat and Wu numbering scheme is used here for the Ab molecule).

The structures of the Fab-monosaccharide and Fab-disaccharide complexes were determined by molecular replacement techniques by using the refined structure of the unliganded Fab S-20-4. A difference Fourier synthesis at this stage clearly showed the bound ligand in both complexes. The structures of the complexes were refined independently as described for the unliganded Fab. The lower resolution model of the disaccharide-bound Fab molecule includes all atomic positions, even those from side-chain residues not visible in density. The latter were positioned assuming the most common rotamer (as defined by the program O; ref. 26) consistent with local steric restraints. The parameters of the final refined models are given in Table 1.

Results and Discussion

Crystallographic Analysis. The three-dimensional structures of the unliganded and carbohydrate-bound forms of Fab S-20-4 were determined by molecular replacement methods. Details of crystallographic parameters, data collection, and crystallographic refinement are given in Table 1. The electron density for the bound monosaccharide and disaccharide ligands is shown in Fig. 2. The stereochemical quality of the models is comparable to that of other structures refined at a similar resolution. With a single exception, all residues of the unliganded Fab and the Fab-monosaccharide complex display main-chain dihedral angles that fall within allowed regions of the Ramachandran plot as defined by the program PROCHECK (27). The only exception is Thr L51 in the L2 hypervariable loop, for which a similar unfavorable conformation has been observed in other crystallographic studies of Fab fragments.

The refined structures of the Fab fragment in the three crystal forms are similar. After least squares superposition of the entire Fab molecules, the rms deviations for all main-chain atoms are 0.32 \( \AA \) between the unliganded and the monosaccharide-bound proteins, 0.52 \( \AA \) between the unliganded and the disaccharide-bound proteins, and 0.47 \( \AA \) between the Fab fragments in the two complexes. The largest rms deviations observed for the disaccharide-bound model are a consequence of different crystal contacts. However, the comparison of the unliganded and monosaccharide-bound Fab molecules (Fig. 3) reveals that the Ab combining site undergoes small conformational changes on ligand binding. For the variable region, the largest shifts are observed for the first and third hypervariable loops of the light chain (Fig. 3); these loops are slightly displaced toward the center of the paratope in the Ag-bound form of the Ab.

Ab-Carbohydrate Interactions. The crystal structures of the Ab–Ag complexes (Fig. 4 a and b) show that the terminal perosamine residue of the serotype O1 Ogawa O-SP is the primary antigenic determinant. Ab S-20-4 binds this monosaccharide in an ex-
tended cavity formed by two adjacent pockets (Fig. 4c); the perosamine sugar ring faces a central hydrophobic pocket formed essentially by residues from the L3 and H3 CDR loops, whereas the tetronic acid moiety of the Ag lies within a lateral shallow pocket primarily defined by the H1 and H3 CDR loops. The total surface area occluded from the bulk solvent on formation of the Fab-monosaccharide complex is 420 Å² (230 Å² from the Ab paratope and 190 Å² from the ligand). Carbohydrate binding affinity is the result of two basic contributions arising from protein–carbohydrate hydrogen bonding interactions and the burial of nonpolar groups at the interface.

The monosaccharide is stabilized by six direct hydrogen bonding interactions with the CDR H1 and H3 loops (Table 2). Four of these interactions involve bifurcated hydrogen bonds formed by two hydroxyl groups of the ligand OH-3 from the perosamine residue and OH-2 from the tetronic acid moiety, with amino acid residues Asp H33 and His H95 at the bottom of the binding cavity (Fig. 4d). An additional hydrogen bond is formed by the carbonyl group of the tetronic acid moiety with the NH group of the main chain at position H98. The primary hydroxyl group OH-4 of the tetronic acid, at the extremity of the molecule, is also involved in hydrogen bonding interactions with the CDR H1 loop in one of the two crystal structures. However, this region of the Ag displays high temperature factors and weak electron density (Fig. 2), suggesting that this interaction is not critical for binding.

The hydrophobic character of the central Ab pocket, which binds the 2-O-methyl group of the Ag (Fig. 4), plays a critical role in carbohydrate recognition. Indeed, the S-20-4 Ab shows no detectable binding to the synthetic terminal mono- and oligosaccharides of the Inaba O-SP lacking the 2-O-methyl group (8). Also, the binding of IgG1 S-20-3 (same clone as S-20-4) to Ogawa LPS could not be inhibited by the Inaba O-SP (6). The central pocket is primarily defined by residues Tyr L32, Trp L91, and Trp L96 from the light chain and residues His H95 and Ala H98 from the heavy chain.
chain (Fig. 4). On binding, the aromatic side chains of Tyr L32 and Trp L91 move slightly toward the terminal perosamine, thus accounting for the structural differences between the complexed and uncomplexed Fab molecules (Fig. 3), which is consistent with previous observations (8, 13) that the tryptophanyl residue L91 is correlated with ligand-induced fluorescence change in Abs. However, few direct contacts are observed between protein residues and the 2-\textit{O}-methyl group of the ligand, and the hydrophobic pocket is large enough to bind the 2-\textit{O}-ethyl derivative 9 (i.e., containing an additional methylene group at the 2-\textit{O} position). Despite this lack of complementarity, no water molecules are trapped within the Ab–Ag interface.

The methylated OH group at the anomeric C-1 carbon of the bound monosaccharide (which is linked to downstream perosamine units in the O-SP) extends away from the Ab binding site. As a consequence, the second perosamine residue in the Fab–disaccharide complex is positioned at the exterior of the Ab binding cavity (see Fig. 5) and makes fewer contacts with the Ab paratope, including a single intermolecular hydrogen bonding interaction between the sugar OH-3 hydroxyl group and the side chain nitrogen atom of Trp L91 (Table 2, Fig. 4d). Despite this lack of complementarity, no water molecules are trapped within the Ag–Ab interface.

The methylated OH group at the anomic C-1 carbon of the

Table 2. Carbohydrate-Ab hydrogen bonding interactions

<table>
<thead>
<tr>
<th>Sugar atom</th>
<th>Residue</th>
<th>Atom</th>
<th>Distance, Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upstream terminal perosamine</td>
<td>O3</td>
<td>Asp H33</td>
<td>O62</td>
</tr>
<tr>
<td></td>
<td>O3′</td>
<td>His H95</td>
<td>N61</td>
</tr>
<tr>
<td></td>
<td>O3″</td>
<td>Ala H98</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>Asp H33</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>O2′</td>
<td>His H95</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>O4″</td>
<td>Thr H31</td>
<td>O</td>
</tr>
<tr>
<td>Second perosamine residue</td>
<td>O3</td>
<td>Trp L91</td>
<td>Nε1</td>
</tr>
</tbody>
</table>

*This interaction is observed only in the structure of the Fab-monosaccharide complex.

Structural Basis of Specificity. The three-dimensional structures of the Ab–carbohydrate complexes provide a rational framework for understanding the fine specificity of mAb S-20-4 (and related Abs such as IgG1 A-20-6) as revealed by using synthetic analogs of the Ogawa O-SP Ag (Tables 3 and 4). The overall position and the glycosidic dihedral angles observed for the Fab-bound disaccharide indicate that additional sugar residues of the O-SP Ag beyond the two terminal residues would make no contacts with the Ab paratope (see Fig. 5), suggesting that Ab recognition would not affect the regular extended conformation of the O-SP Ag in solution (28). These observations correlate well with solution measurements of the reactivities of the Ab for different polysaccharide fragments of O-SP (8), showing that the degree of polymerization of the ligand (between 2 and 6 perosamine residues, compounds 2-6 in Table 3) has no significant effect on the Ab affinity.
The hydroxyl groups OH-3 and OH-2′ of the upstream terminal monosaccharide and its tetronic acid residue, respectively, are essential for binding, because they are engaged in four of the six protein–carbohydrate hydrogen bonding interactions. In agreement with that, mAb S-20-4 showed no detectable binding with the corresponding deoxy monosaccharides derivatives (compounds 11 and 14, Table 4). However, the Ab binds to the D-isomer 13 of the tetronic acid moiety and to the derivative 16, in which the hydroxyl and the methylene groups are transposed. These interactions suggest that the tetronic acid substituent of the terminal residue is flexible, thus allowing the ligand to restore, at least partially, its critical hydrogen bonding interactions with the protein. This hypothesis is consistent with the weak electron density observed for the primary hydroxyl group OH-4′ at the distal end of the tetronic acid moiety in the Fab–disaccharide structure (Fig. 2b), and can also explain why mAb S-20-4 is able to recognize with comparable affinity the sugar derivative 15 lacking this hydroxyl group (Table 4).

The Ogawa and Inaba polysaccharides differ in the presence of methyl group or hydrogen at the O-2 position of their respective upstream terminal sugar residues (Fig. 1). In the crystal structure of the Ogawa Ag–Ab complexes, this methyl group is buried within a central hydrophobic pocket essentially defined by aromatic and aliphatic side chains (Figs. 4c and 5). The important contribution of this interaction to the total binding energy is emphasized by both the lack of Ab recognition for the nonmethylated sugar residue 7 with a free OH-2 group (which mimics the terminal Inaba monosaccharide) and the 2-deoxy derivative 8 lacking both the methyl and hydroxyl groups at C-2 (Table 4). These substitutions do not affect the residual structure of the Ag or interfere with existing Fab–monosaccharide contacts. Therefore, the hydrophobic effect (dehydration of nonpolar surfaces) must contribute significantly to the total binding energy. The exclusion of water molecules from the interaction is probably not due to a high steric Ag–Ab complementarity because the Ab is also able (8) to bind the more bulky 2-O-ethyl derivative 9, a substitution that introduces an additional methylene group within the interface. It therefore appears that the 2-O-methyl group of the O-SP terminal residue is the major determinant of Ogawa serotype specificity, and that recognition of the equivalent epitope on Inaba strains (having a free hydroxyl group at C-2) would require an altogether different Ab combining site.

**Immunological Implications.** Several years ago, Kabat and co-workers (10, 29) proposed that Abs recognizing linear polysaccharide chains through their terminal or nonterminal sugar residues could display either cavity- or groove-shaped binding sites. This hypothesis was later confirmed by solution measurements of the binding of anti-1,6-dextran Igs (30–32). Although sharing some of these features, the only two other anti-carbohydrate Abs for which the structure has been determined in complex with Ag, Se155 (14) and BR96 (16), are both specific for a determinant consisting of multiple saccharides of nonlinear polysaccharides, the *Salmonella* O-Ag and the Lewis Y Ag, respectively.

The work reported here shows unequivocal crystallographic evidence for a cavity type binding Ab that is specific for a single mono saccharide: antigenic terminus (because the terminal perosaminyl residue accounts for 90% of the maximal binding energy). Indeed, mAb S-20-4 binds Ag with a relatively high affinity (Table 3). In general, anti-carbohydrate Abs (and other sugar-binding proteins such as lectins) are known to display low affinities for simple monosaccharides (12, 33), and their determinants often require multiple carbohydrate residues to reach the degree of affinity needed for specific recognition. Thus, antitumor Ab BR96 binds the Lewis Y tetrasaccharide with a binding affinity of 2 × 10^9 M^-1 (16), similar to that of mAb Se155 for the trisaccharide methyl glycoside (34). In contrast, S-20-4 binds the *monosaccharide* ligand 1 with a higher affinity (3.9 × 10^9 M^-1), even though the contact surface area is smaller than that for either of the two foregoing complexes.

The major antigenic determinant of the Ogawa serotype is expressed by the upstream terminal monosaccharide of O-SP. The LPS of the Inaba serotype possesses internal structural units common to both Ogawa and Inaba serotypes (35, 36), and it has been proposed that both the core and the O-SP (see Fig. 1a) may be involved in this common epitope (37). The crystal structure of the cavity type mAb S-20-4 reported herein confirms the serotype specificity of this and related anti-Ogawa Abs by showing the pivotal contribution by so small a structural fragment in the antigenic determinant as a methyl group. It also suggests requirements that should be taken into consideration in

**Table 3. Binding constants K_a (M^-1), determined by fluorescence titration, and deduced free energy of association -∆G° (kcal/mol, in parentheses) for Igs specific for LPS of *V. cholerae* O1 serotype Ogawa, with synthetic methyl α-glycosides of fragments of Ogawa O-SP (8)**

<table>
<thead>
<tr>
<th>IgG</th>
<th>Synthetic fragments of the Ogawa O-SP</th>
<th>Monomer (1)</th>
<th>Dimer (2)</th>
<th>Trimer (3)</th>
<th>Tetramer (4)</th>
<th>Pentamer (5)</th>
<th>Hexamer (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-20-4</td>
<td>3.9 × 10^5 (31.9)</td>
<td>1.2 × 10^6 (34.7)</td>
<td>8.5 × 10^6 (33.8)</td>
<td>1.6 × 10^7 (35.4)</td>
<td>1.1 × 10^8 (34.5)</td>
<td>1.3 × 10^9 (34.9)</td>
<td></td>
</tr>
<tr>
<td>A-20-6</td>
<td>1.9 × 10^5 (30.1)</td>
<td>5.8 × 10^6 (32.9)</td>
<td>7.7 × 10^7 (33.6)</td>
<td>1.5 × 10^8 (35.2)</td>
<td>9.7 × 10^9 (34.2)</td>
<td>1.0 × 10^10 (34.2)</td>
<td></td>
</tr>
</tbody>
</table>
the design of a carbohydrate-based cholera vaccine. The chemical structure of the LPS molecule suggests three distinct epitopic regions as putative targets for Ab recognition: the sugar core (close to the lipid moiety), the middle region of the O-SP, and the upstream terminal perosamine residue (Fig. 1a). As shown by this and previous work (8, 37), Abs specific for the terminal perosamine could selectively protect against the Ogawa serotype but would fail to recognize the Inaba serotype. Therefore, protective Abs against both serotypes should bind to the inner part of the O-SP and/or to sugar residues defining the core of the LPS molecule. However, the proximity of the LPS core to the cell membrane might sterically jeopardize its accessibility by Abs, whereas the O-SP displays a linear extended conformation in solution (28), which renders it a priori more accessible for binding (and probably explains why this region is involved in both Inaba and Ogawa epitopes; see ref. 37). Synthetic compounds mimicking the inner perosamine residues of the O-SP Ag could therefore be expected to elicit Abs specific for both the Ogawa and Inaba serotypes. Such Abs with a groove-shaped binding site were indeed observed for the Brucella A cell wall polysaccharide Ag (38), an O-SP-related molecule that is also composed of a linear chain of α(1,2)-linked perosamine units (39).

Table 4. Binding constants \( K_a \) (M\(^{-1}\)) determined by fluorescence titration, and deduced free energy of association \( -\Delta G^\circ \) (kcal/mol, in parentheses) for IgGs specific for LPS of \( V. \) cholerae O1 serotype Ogawa, with synthetic analogs of the methyl α-glycoside of the terminal residue, and the presumed antigenic determinant of the Ogawa O-SP

<table>
<thead>
<tr>
<th>IgG</th>
<th>Synthetic Analogs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7</td>
</tr>
<tr>
<td>S-20-4</td>
<td>0</td>
</tr>
<tr>
<td>A-20-6</td>
<td>4.3 x 10(^{-3}) (20.7)</td>
</tr>
</tbody>
</table>