THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 12, Issue of March 25, pp. 11313–11319, 2005

Printed in U.S.A.

Crystal Structure and Binding Properties of the *Serratia marcescens* Chitin-binding Protein CBP21*[S]

Received for publication, June 25, 2004, and in revised form, November 30, 2004
Published, JBC Papers in Press, December 8, 2004, DOI 10.1074/jbc.M407175200

Gustav Vaaje-Kolstad‡, Douglas R. Houston§¶, Anna H. K. Riemen§, Vincent G. H. Eijsink‡, and Daan M. F. van Aalten§‖

From the ‡Department of Chemistry, Biotechnology, and Food Science, Postbox 5003, Agricultural University of Norway, N-1432 Ås, Norway and §Division of Biological Chemistry and Molecular Microbiology, School of Life Sciences, University of Dundee, Dundee DD1 5EH, United Kingdom

Chitin proteins are commonly found in bacteria that utilize chitin as a source of energy. CBP21 is a chitin-binding protein from *Serratia marcescens*, a Gram-negative soil bacterium capable of efficient chitin degradation. When grown on chitin, *S. marcescens* secretes large amounts of CBP21, along with chitin-degrading enzymes. In an attempt to understand the molecular mechanism of CBP21 action, we have determined its crystal structure at 1.55 Å resolution. This is the first structure to be solved of a family 33 carbohydrate-binding module. The structure reveals a "budded" fibronectin type III fold consisting of two β-sheets, arranged as a β-sheet sandwich, with a 65-residue "bud" consisting of three short helices, located between β-strands 1 and 2. Remarkably, conserved aromatic residues that have been suggested previously to play a role in chitin binding were mainly found in the interior of the protein, seemingly incapable of interacting with chitin, whereas the structure revealed a surface patch of highly conserved, mainly hydrophilic residues. The roles of six of these conserved surface-exposed residues (Tyr-54, Glu-55, Glu-60, His-114, Asp-182, and Asn-185) were probed by site-directed mutagenesis and subsequent binding studies. All single point mutations lowered the affinity of CBP21 for β-chitin, as shown by 3–8-fold increases in the apparent binding constant. Thus, binding of CBP21 to chitin seems to be mediated primarily by conserved, solvent-exposed, polar side chains.

Chitin is a linear insoluble polymer of β(1–4)-linked *N*-acetylglucosamine, which is a common constituent of fungal cell walls, shells of crustaceans, and exoskeletons of insects. In nature, two major types of chitin occur that are characterized by an anti-parallel (α-chitin) or a parallel (β-chitin) arrangement of the polymer chains. Chitin binding is an important event in mechanisms varying from plant responses to chitin-containing pathogens to chitin degradation by microorganisms. Genes putatively encoding chitin-binding proteins or proteins containing chitin binding domains have been found in many organisms, ranging from bacteria to man with high diversity. In the classification of carbohydrate active enzymes ((1) afmb.cnrs-mrs.fr/CAZY/), proteins and protein domains with chitin binding properties are classified into several families of carbohydrate-binding modules (families 1, 2, 12, 14, 18, 19, and 33) based on amino acid similarity. The majority of these modules are domains of larger enzymes, where they are thought to increase substrate affinity and enhance catalytic efficiency (2). Crystal structures of complete multidomain chitinases (3, 4), as well as several site-directed mutagenesis studies (5, 6), have shown that aromatic residues on the surfaces of these chitin binding domains are important for substrate affinity.

In addition to appearing as discrete domains in enzymes, chitin-binding modules also exist as independent, noncatalytic, chitin-binding proteins (CBPs).[1] Such noncatalytic CBPs are found in families 14, 18, and 33, where families 14 and 18 constitute small anti-fungal proteins that share a structurally similar chitin-binding motif (7). Family 33 CBPs are mainly found in bacteria and viruses and have so far not been structurally characterized at high resolution, although a low resolution solution x-ray scattering structure has been reported (8). The ability to actually bind chitin has been established for only six CBPs in this family (9–14). Bacterial family 33 CBPs are expressed and secreted during chitin degradation and have been hypothesized to invade the chitin matrix and dissolve individual polymers, making them more accessible to degradation by chitinases (13). Some CBPs in this family also appear as chitin binding domains in putative chitinases, but none of these multidomain enzymes have been characterized.

The Gram-negative soil bacterium *Serratia marcescens* efficiently degrades chitin (15). When the bacterium is grown in the presence of chitin as a carbon source, CBP21 (197 amino acids, including a 27-residue leader peptide) is one of the major proteins produced and secreted, along with three chitinases and an *N*-acetylhexosaminidase (12, 16). The gene for CBP21 is located 1.5 kb downstream of one of the chitinase genes (*chiB*), but the two genes are not in the same operon (12, 17). Several detailed structural and functional studies have recently described the roles of the chitinolytic enzymes (18–20); however, not much is known about CBP21. A study by Suzuki *et al.* (12) has revealed that CBP21 has a specific preference for binding β-chitin and that affinity is strongly pH-dependent.

* This work was supported in part by Norwegian Research Council Grant 140497/420. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 2BEM and 2BEN) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1 and S2.

¶ Supported by BBSRC CASE studentship (Cyclacel).

‖ Supported by a Wellcome Trust Senior Research Fellowship and the EMBO Young Investigator Programme. To whom correspondence should be addressed. Fax: 44-1382-345764; E-mail: dava@davapc1.bioch.dundee.ac.uk.

[1] The abbreviations used are: CBPs, chitin-binding proteins; CHES, 2-(cyclohexylamino)ethanesulfonic acid; r.m.s.d., root mean square deviation; CAPS, 3-(cyclohexylamino)propanesulfonic acid; FnIII, fibronectin type III.

To obtain more insight into the properties of CBP21, we have determined its crystal structure at 1.55 Å resolution. The structure is the first of a family 33 carbohydrate-binding module to be solved and reveals a compact fibronectin III-type fold plus a 65-residue loop/helical pseudo-domain. Most surprisingly, only one of several conserved aromatic residues in CBPs was found to be located on the protein surface. In order to identify functionally important residues, the result of a sequence alignment of bacterial CBPs was combined with the structural data to reveal conserved, surface-exposed residues. The contribution of six of these residues to chitin binding was assessed by studying the effect of mutating them to alanine. Because CPB21 crystallized as a dimer, several experiments were done to detect dimeric forms in solution, and several amino acids in/near the putative dimer interface were mutated, and the effects on chitin binding were examined.

MATERIALS AND METHODS

Gene Cloning and Sequencing—The gene encoding CBP21 was subcloned from plasmid pLES3, which contains a 9.2-kb fragment from the genome of *S. marcescens* BJL200 (21), and cloned into the pRSETB expression vector (Invitrogen). The gene sequence was determined using the Applied Biosystems BigDye Terminator version 3.1 kit and an automated ABI Prism sequencing apparatus. The sequencing reaction was set up according to the supplier's instructions. The sequence has been submitted to the GenBank™ (accession number AY665558).

CBP21 mutants Y54A, E55A, E60A, H114A, Y147A, A152R, Q161A, N163R, D182A, and N185A were made using the Quickchange mutagenesis kit (Stratagene, La Jolla, CA), according to the instructions provided by the supplier. All mutants had retained considerable affinity for chitin under the chromatographic conditions used for purifying the wild type protein and could thus be purified by using the standard procedure described below.

Protein Expression and Purification—The pRSETB vector containing the *cbp21* gene was transformed into BL21(DE3) Star cells (Invitrogen), and the protein was expressed by growing the cells in Luria-Bertani broth containing 50 μg/ml ampicillin for 16 h at 37 °C. Cells were fractionated using the cold osmotic shock method described by Manoil and Beckwith (22), and the periplasmic fraction was used as starting sample for further work. CBP21 was purified using chitin affinity chromatography. 25 ml of periplasmic extract was adjusted to 1 M $NH_4(SO_3)_2$ and 50 mM Tris-HCl, pH 8.0 (running buffer), before being applied on a 1.5 × 10-cm (18 ml) chitin bead column (New England Biolabs, Beverly, MA) equilibrated with running buffer. After application of the sample, the column was washed with 2 volumes of running buffer. CBP21 was eluted from the column by changing the running buffer to 20 mM acetic acid, pH 3.6. The flow rate was 2.5 ml/min throughout the experiment. Fractions containing pure CBP21 (assessed by SDS-PAGE) were collected and concentrated using Centricon Plus-20 ultrafiltration tubes (Millipore, Billerica, MA) to concentrations in the 20–150 mg/ml range. The concentrated protein was dialyzed in a Slide-A-Lyzer (Pierce) cassette against 500 volumes of 50 mM Tris-HCl, pH 8.0, for 24 h at 4 °C with three buffer changes. Protein samples were stored at 4 °C. Protein concentrations were determined using the Bio-Rad Protein Microassay according to the instructions from the supplier (Bio-Rad).

Protein Crystallization—The protein sample was diluted to 20 mg/ml and used in sitting drop vapor diffusion crystallization experiments. Crystals for wild type CBP21 were obtained in several conditions of which 35% (w/v) PEG 3000 and 0.5 M CHES buffer, pH 9.5 (condition *a*), and 1.26 M ammonium sulfate and 0.1 M HEPES, pH 7.5 (condition *b*), resulted in the best diffracting crystals. A heavy atom derivative was obtained by soaking a crystal obtained from crystallization condition *a* for 2 days in a 5 mM solution of sodium perrhenate. The Y54A mutant (at 17.5 mg/ml) was crystallized using a similar approach but from 20% (w/v) PEG-8000, 100 mM CAPS, 200 mM NaCl, and cryoprotected with 20% (v/v) ethylene glycol.

Structure Determination—Data sets for native and derivative crystals were collected on beamline ID23-1 at the European Synchrotron Radiation Facility in Grenoble, France. To remove excess sodium perrhenate, the derivative crystal was soaked in 40% (w/v) PEG 3000 and 0.5 M CHES buffer, pH 9.5. Crystals were flash frozen at 100 K in a nitrogen cryostream. A data set from a native crystal and a complete single anomalous dispersion data set from the rhenium derivative crystal were collected at 1.55 and 1.75 Å, respectively. Data were processed with the HKL suite (23) (Table I). From the native data set, a Matthews coefficient of 2.19 $Å^3$/dalton was obtained assuming two molecules in the asymmetric unit and a space group of $P3_221$. The single heavy atom site was found and refined with SOLVE (24), yielding phases to 1.85 Å resolution with an overall figure of merit of 0.29 (see Table I). Combination of the SAD phases with the 1.55 Å native amplitudes and solvent flattening using DM (25) resulted in an interpretable map. WarpNtrace (26) was used to build an initial model of 137 residues out of a possible total of 350. LSQKAB (27) was used on this partial model to find the noncrystallographic 2-fold axis relating the two monomers in the asymmetric unit. 2-Fold averaging was then applied in another round of solvent flattening in DM, resulting in a further improved electron density map, which allowed WarpNtrace to build a total of 202 residues. Residues that were successfully built into one monomer were fitted to the other monomer by using LSQKAB resulting in two chains of 162 residues and a total of 324 residues out of 350. Initial attempts at refinement, however, gave unfavorable statistics (*e.g.* R = 0.27, R_{free} = 0.31).

A second crystal form was identified, growing from 1.26 M ammonium sulfate and 0.1 M HEPES, pH 7.5, cryoprotected in 20% (v/v) ethylene glycol. A data set collected on a Rigaku rotating anode source was used for the refinement of the final native structure (Table I). Unlike the initial crystals, this had a space group of $P2_12_12_1$ and contained three molecules per asymmetric unit. The structure was solved by molecular replacement using the partially refined structure of the first crystal form as a search model against 8–4 Å data, which, after rigid body refinement in CNS to 3.5 Å, gave an R-factor of 0.33. Simulated annealing, followed by iterations of refinement, using CNS (28) and model building with O (29), resulted in a final model as described in Table I.

The CBP21 Y54A mutant crystallized in the same space group as wild type CBP21 ($P3_221$) and data were collected on a rotating anode (Table I). Refinement, using the same procedures as describe above, was initiated using the native structure as a starting model and resulted in a final model as described in Table I.

Circular Dichroism—CD spectra were recorded with a Jasco J-810 spectropolarimeter (Jasco International Co., Ltd., Tokyo Japan) calibrated with ammonium *d*-camphor-10-sulfonate (Icatayama Chemicals, Tokyo Japan). Measurements were performed at 23 °C, using a quartz cuvette (Starna, Essex, UK) with a path length of 0.1 cm. All measurements were performed with a protein concentration of 0.10 mg/ml in 50 mM sodium phosphate buffer, pH 7.0. Samples were scanned five times at 50 nm/min with a bandwidth of 1 nm and a response time of 0.5 s. The data were averaged, and the spectrum of a sample-free control sample was subtracted. All measurements were conducted at least twice.

Binding Assays—Binding studies were performed using α-chitin from shrimp shells (Hov Bio, Tromsø, Norway), β-chitin from squid bone (<80 mesh-sized particles, France Chitin, Marseille, France), or crystalline cellulose (Avicel; Sigma) as substrates. The (insoluble) substrates were suspended in double distilled H_2O to yield 20 mg/ml stock suspensions that were continuously stirred with a magnetic stirrer in order to obtain consistent amounts when pipetting. In initial binding experiments, reaction mixtures (500 μl) contained 1 mg/ml substrate and 100 μg/ml protein in 50 mM sodium phosphate buffer, pH 7.0. The reaction mixture was rotated vertically at 60 rpm at room temperature. At time points (5, 15, 45, 90, 180, 240, and 480 min), the substrate was spun down in a microcentrifuge for 3 min at 13,000 rpm, and reduction of protein in the supernatant was determined by measuring absorption at 280 nm (A_{280}, Eppendorf Biophotometer, Eppendorf, Hamburg). Binding studies with α-chitin and cellulose were performed with CBP21 wild type only, as these substrates showed little affinity for CBP21.

The CBP21 variants showing reduced binding to β-chitin were analyzed further by using a second binding assay meant to estimate equilibrium binding constants (K_d) and binding capacities (B_{max}). For the binding assays, a CBP21 solution with known protein concentration (Bradford method; see above) was diluted to various concentrations in 50 mM sodium phosphate buffer, pH 7.0. Before adding chitin (from a concentrated stock suspension in 50 mM sodium phosphate buffer, pH 7.0), the A_{280} of the prepared solutions was measured, thus creating individual standard curves for each CBP21 variant. Subsequently, chitin was added, bringing the reaction volume to 1 ml, the β-chitin concentration to 0.5 mg/ml, and the CBP21 concentration to 10, 20, 50, 100, 150, 200, or 300 μg/ml CBP21. After addition of β-chitin, the solutions were mixed by vertical rotation (60 rpm) at room temperature for 16 h. Subsequently, the sample tubes were spun for 3 min at 13,000 rpm in a microcentrifuge to pellet the chitin, and the A_{280} values of the supernatants were measured. Apparent extinction coefficients calcu-

TABLE I
Crystallographic data and refinement statistics
Values in parentheses are reflections from the highest resolution shell. All data were included in the refinement.

	Native1	Native2	Rhenium	Y54A
Space group	$P2_12_12_1$	$P3_221$	$P3_221$	$P3_221$
Unit cell parameters (Å)	$a = 51.93$, $b = 58.05$, $c = 159.83$	$a = b = 84.20$, $c = 81.36$	$a = b = 84.10$, $c = 81.28$	$a = b = 84.45$, $c = 82.17$
Wavelength (Å)	0.89	1.54	1.17624	1.54
Resolution (Å)	25.0–1.55 (1.61–1.55)	15.0–1.55 (1.61–1.55)	15.0–1.75 (1.81–1.75)	30.0–1.80 (1.86–1.80)
R_{merge}	0.038 (0.525)	0.044 (0.180)	0.057 (0.460)	0.069 (0.472)
Completeness (%)	97.9 (83.4)	99.5 (98.9)	100.0 (100.0)	99.4 (98.3)
Redundancy	5.3 (3.4)	3.1 (2.9)	7.6 (7.2)	4.1 (4.2)
$I/\sigma I$	42.2 (6.9)	25.5 (2.8)	40.6 (5.1)	10.4 (2.6)
R_{cryst}	15.6			22.7
R_{free}	17.5			26.3
No. protein atoms	3987			2644
No. water molecules	679			238
R.m.s.d. from ideal geometry				
Bonds (Å)	0.019			0.006
Angles (°)	1.9			1.5
Main chain B (Å^2)	1.5			1.6
$B_{protein}$ (Å^2)	14.1			21.3

lated from the respective A_{280} standard curves were subsequently used to convert A_{280} values to protein concentrations. If a value lower than 20 μg/ml was estimated by measuring A_{280}, the protein concentration was re-determined using the Bio-Rad protein microassay. All assays were performed in triplicate and with blanks (buffer + 0.5 mg/ml β-chitin) and controls to correct for aspecific binding of the protein (buffer + CBP21 variants). The equilibrium dissociation constants, K_d (μM), and substrate binding capacities, B_{max} (μmol/g), were determined by fitting the binding isotherms to the one-site binding equation where P stands for protein: $[P_{bound}] = B_{max}\,[P_{free}]/K_d + [P_{free}]$, by nonlinear regression using the GraphPad Prism software (GraphPad Software Inc., San Diego, CA).

RESULTS AND DISCUSSION

Overall Structure of CBP21—The *cbp21* gene was cloned, overexpressed in *E. coli,* and purified to homogeneity by chitin affinity chromatography. The protein was crystallized, and the structure was solved by a rhenium single-wavelength anomalous dispersion experiment (Table I). The native structure of CBP21 was refined to 1.55 Å resolution with R-factors converging to $R = 0.156$, $R_{free} = 0.175$ (Table I). The asymmetric unit contains three molecules with similar conformation and secondary structure (average pairwise r.m.s.d. on C-α atoms is 0.3 Å). Ramachandran plots created by PROCHECK (30) showed that all residues had dihedral angles within allowed regions, apart from one residue in the A monomer that was positioned in the generously allowed area. The B monomer was used for structure analysis and is shown in the figures.

The structure consists of a three-stranded and a four-stranded β-sheet that form a β-sandwich (Fig. 1). There is a bud-like 65-residue pseudo-domain between β-strands one and two, which consists of loops, one short α-helix, and two short 3_{10} helices (Figs. 1 and 2). The four cysteine residues of CBP21 appear as two disulfide bridges, one in the loop/helical region (Cys-41 and Cys-49) and one joining β-strands 4 and 5 (Cys-145 and Cys-162) (Fig. 1*A*). The core of CBP21 consists of several conserved aromatic amino acids, such as Trp-94, Trp-108, Trp-119, Tyr-121, Trp-128, Trp-178, Phe-187, and Tyr-188 (Figs. 1*A* and 2). The structure represents the first high resolution data for the CBM-33 family, although a low resolution solution scattering envelope for a *Streptomyces* chitin-binding protein (CHB1) has been reported previously (8). The structure of this protein, which is 46% identical at the sequence level with CBP21 (Fig. 2), was described as an elongated globule with a distinct bud-like extension. This is in agreement with the CBP21 structure, where the elongated globule would correspond with the β-sandwich, and the bud-like extension with the 65-residue pseudo-domain between β1 and β2.

Similarity to Other Structures—A search in the DALI data base (31) revealed that the CBP21 fold shares structural similarity with a number of proteins (140 matches with Z-score greater than 2.0). The most similar structures all possess a fibronectin type III (FnIII) domain, which is a common fold appearing in a variety of eukaryotic proteins. In bacteria, FnIII domains have so far only been found among carbohydrate active enzymes (32). Most strikingly, the closest structural match is with the FnIII-like chitin binding domain of chitinase A (ChiA) from *S. marcescens* (3), with an r.m.s.d. of 2.5 Å on 85 structurally matched C-α atoms (Fig. 1*B*). The topology of the β-sandwiches of CBP21 and ChiA-FnIII is similar, but ChiA-FnIII lacks the bud-like extension of CBP21 (Fig. 1*B*). Indeed, it appears that this budded variant of the FnIII fold is not found in other FnIII-containing structures identified in the DALI search. Sequence alignments suggest that the bud is conserved in the CBM-33 family and carries several completely conserved residues (Fig. 2). Most surprisingly, the exposed aromatic amino acids that are known to be responsible for chitin binding in the ChiA-FnIII domain (5) are not present in CBP21, and a structure-based sequence alignment does not reveal any sequence similarity between ChiA-FnIII and CBP21.

The structures of chitinases, chitin binding domains, cellulases, and cellulose binding domains, show that the common carbohydrate binding architecture consists of a surface, groove, cleft, or tunnel lined with aromatic residues that are necessary for binding (3–5, 33–36). The CPB21 monomer shows neither a groove/tunnel nor stretches of solvent-exposed aromatic residues. Searches with the cavity finding option in WHAT IF (37) did not reveal cavities that could potentially play a role in ligand binding.

Analysis of the Crystallographic Dimer—Both crystal forms described in Table I contain more than one independent CBP21 molecule in the asymmetric unit. In both cases, and in a third crystal form in $I2_12_12_1$ (data not shown), two CBP21 molecules are packed together through (pseudo)2-fold symmetry (Fig. 1*C*). The β4-strands of the two molecules make four hydrogen bonds, forming a continuous 8-stranded anti-parallel β-sheet (Fig. 1*C*). Hydrophobic interactions are established through packing of the two 145–162 disulfides and Phe-147, burying 229 Å^2 of surface on each monomer. Through this dimerization two identical clefts on both sides of the dimer are formed (Fig. 1*C*). Preliminary docking studies with short chito-oligosaccha-

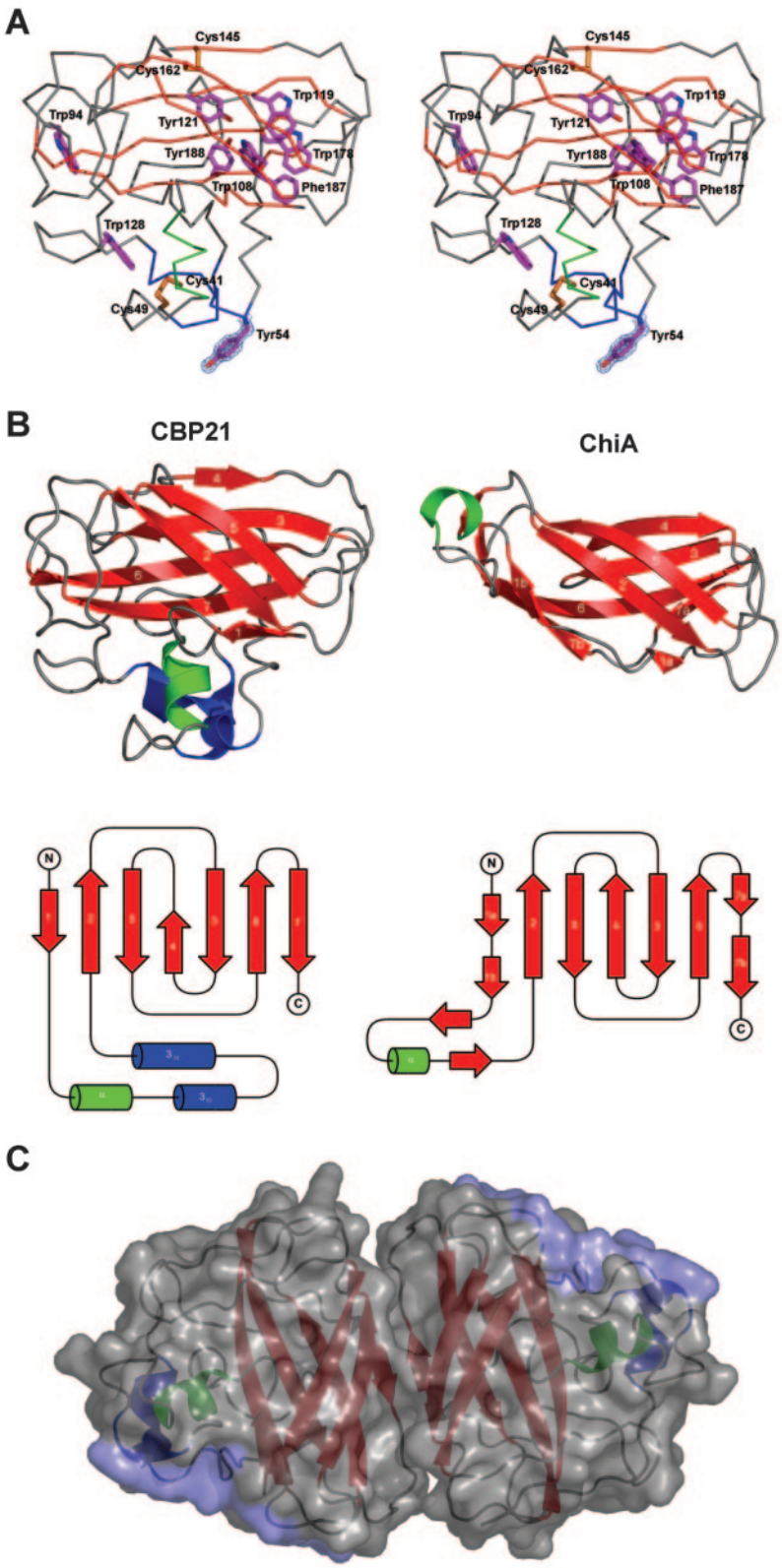

FIG. 1. *A*, α-carbon trace of CBP21 shown in stereo. Conserved aromatic amino acids are shown as *sticks*, with carbon, oxygen, and nitrogen atoms colored *magenta*, *red*, and *blue*, respectively. Disulfide bridges are shown as *orange sticks*. The electron density ($2|F_o| - |F_c|$, ϕ_{calc}) surrounding Tyr-54 is contoured at 2.0σ. *B*, the structure of CBP21 and the FnIII-like chitin binding domain of ChiA (3) shown as schematic and schematic illustrations of the topology. The FnIII-like fold of CBP21 consists of two β-sheets (consisting of β-strands 1, 2, and 5 and β-strands 3, 4, 6, and 7, respectively). The loop/helical domain consists of one α-helix and two 3_{10} helices. *C*, the dimer form of CBP21 as observed in the crystal. The CBP21 molecules are shown in a ribbon representation surrounded by a semi-transparent van der Waals surface. The van der Waals surface of the conserved polar surface patch (see Fig. 4 and text) is colored *light blue*. All panels have β-strands colored *red*, α-helices *green*, and 3_{10} helices *blue*. *A* and *C* were made with PyMol (www.pymol.org). *B* was made with PyMol and TopDraw (47).

rides using AutoDock (38) suggested these clefts could be wide enough to support chito-oligosaccharide binding (data not shown). However, extensive attempts were made to verify this predicted binding mode by co-crystallization and soaking studies with $GlcNAc_{1-8}$ chito-oligosaccharides, without any success. To investigate further the potential relevance of the dimer observed in the crystal, CBP21 mutations directed at preventing dimer formation or/and substrate binding in the dimer cleft were constructed. Two residues responsible for key dimer interactions (Tyr-147 and Gln-161) were mutated to alanines, and two residues on the cleft surface (Ala-152 and Asn-163) were mutated to Arg in order to position bulky side chains in the cleft. Fig. 3 shows that all four mutants bound chitin with equal affinity, suggesting that neither the individual residues nor the cleft formed by dimerization plays a role in the interaction with chitin. It is also worth noting that these residues

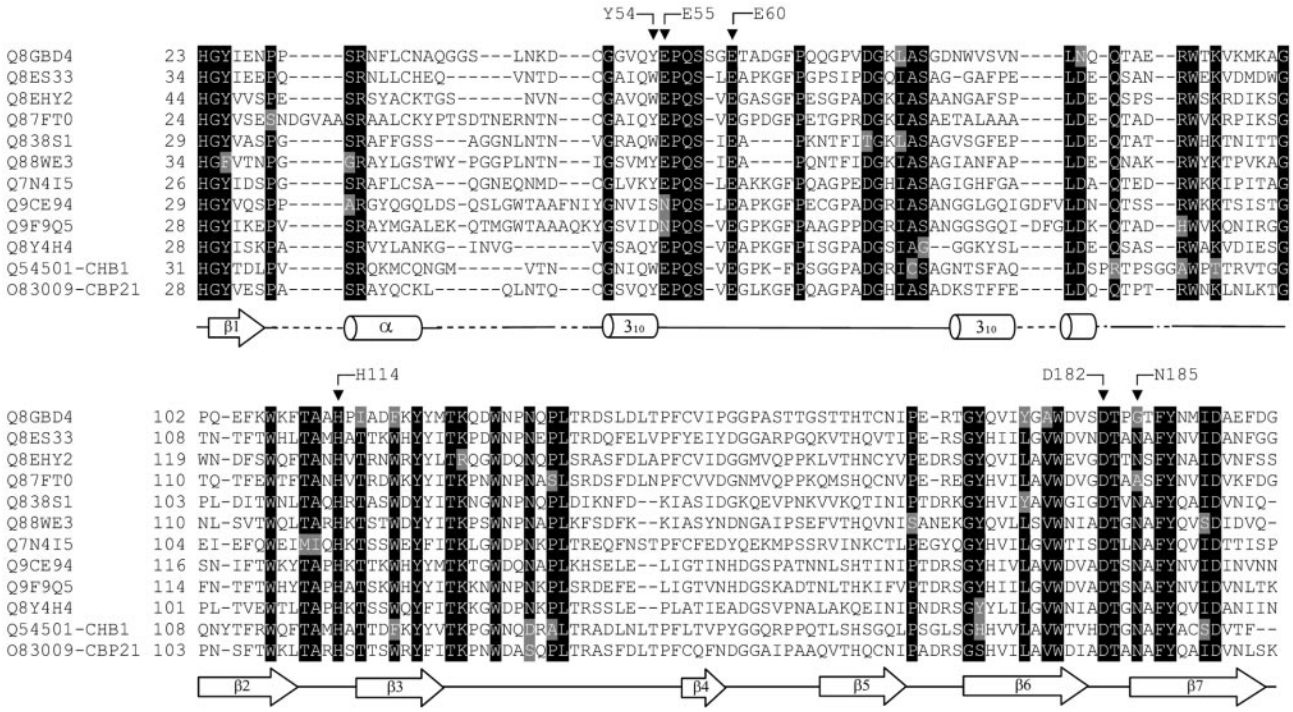

FIG. 2. **Multiple alignment of bacterial CBPs (Q8GBD4, *Yersinia enterocolitica*; Q8ES33, *Oceanobacillus iheyensis*; Q8EHY2, *Shewanella oneidensis*; Q87FT0, *Vibrio parahemeolyticus*; Q838S1, *Enterococcus feacalis*; Q88WE3, *Lactobacillus plantarum*; Q7N4I5, *Photorhabdus luminescens* (subsp. *laumondii*); Q9CE94, *Lactococcus lactis* (subsp. *lactis*); Q9F9Q5*, *Bacillus amyloliquefaciens;* Q8Y4H4, *Listeria monocytogenes*; Q54501*, *Streptomyces olivaceoviridis* (CHB1); O83009*, *Serratia marcescens* (CBP21)).** CBPs marked with *asterisks* in the sentence above have been shown to bind chitin; see Refs. 10–12, respectively. The secondary structure elements of CBP21 are shown and labeled. *Arrows* with residue numbers indicate residues mutated in this study. The multiple alignment was created with ClustalX (48) and edited with T-COFFEE (49).

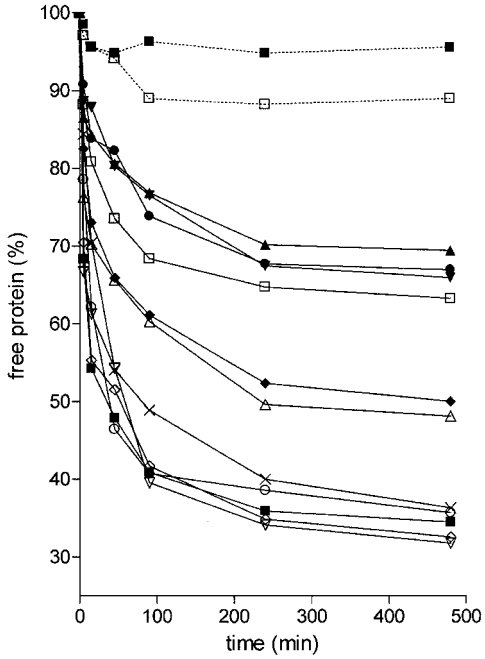

FIG. 3. **Adsorption of CBP21 wild type to cellulose (*filled squares* on a *dotted line*) and α-chitin (*open squares* on a *dotted line*) and adsorption of CBP21 wild type (*filled squares*), Y54A (*filled triangles*), E55A (*filled turned triangles*), E60A (*filled diamonds*), H114A (*filled circles*), Y147A (*open turned triangles*), A152R (*open diamonds*), Q161A (*open circles*), N163R (*tilted cross*), D182A (*open squares*), and N185A (*open triangles*) to β-chitin from squid.**

are not well conserved in the CBM-33 family (Fig. 2). Calibrated gel filtration runs with CBP21 only gave peaks corresponding to a monomeric protein, whereas the position of the CBP21 band in native gels did not change upon addition of $GlcNAc_7$ or glycol-chitin to the gel (see Supplemental Material). Finally, solution x-ray scattering studies on the *Streptomyces* chitin-binding protein CHB1 showed that this CBM-33 family member is a monomer in solution (8). Although we cannot exclude the possibility of CBP21 dimerizing in the presence of long chitin polymers, similar to what was recently observed for a CBM-29 family member (39), there is no evidence to support this in the case of CBP21. It should be noted that carbohydrate-binding proteins and domains with the typical hydrophobic grooves/surfaces (such as CBM-29 (39)) bind single carbohydrate chains/oligosaccharides, whereas CBP21 only binds to insoluble chitin (12), that is to an ordered array of multiple chitin chains.

Binding Studies of Wild Type and Mutant CBP21—By sequence similarity, CBP21 is classified as a family 33 CBM in the CaZy data base (1) and as a chitin binding domain in the Pfam data base (Chitin_bind_3, Pfam entry PF03067). Both data base entries harbor a multitude of viral and bacterial proteins. Of the 80 CBP sequences present in the Pfam data base (February, 2004), 19 other bacterial CBPs cluster with CBP21. On the basis of these sequences a multiple alignment was created in order to identify conserved residues. The alignment (Fig. 2) revealed 40 highly conserved residues (>90% conserved) and a cluster of conserved, mainly hydrophilic residues at the surface of CBP21 (Fig. 4*A*). To investigate the function of these conserved surface residues, a subset was mutated individually to alanine (Tyr-54, Glu-55, Glu-66, His-114, Asp-182, and Asn-185). It should be noted that Tyr-54 is the only conserved surface-located aromatic residue in CBP21 (Fig. 1*A*). A mutational study of a homologous CBP from *Streptomyces olivaceoviridis* had previously indicated that an aromatic residue at this position is important for chitin binding (40).

All CBP21 variants were purified to homogeneity by chitin

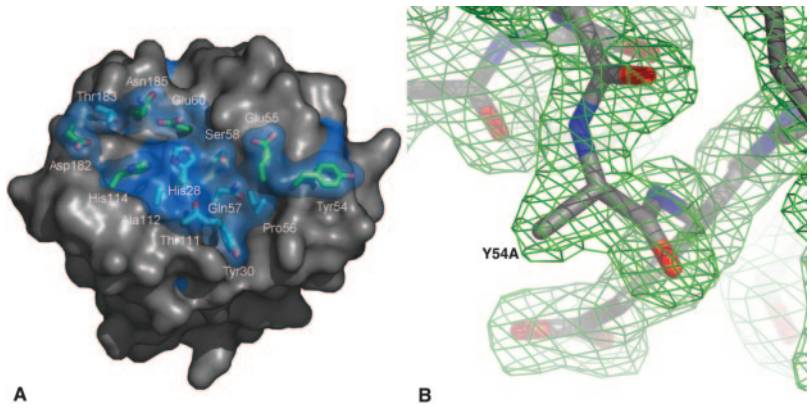

FIG. 4. *A*, surface representation of CBP21 illustrating the patch of conserved residues. Residues that are more than 90% conserved in the alignment shown in Fig. 2 are colored *blue* (there are no such blue patches on the backside of the protein). The side chains of residues that were mutated in this study are shown as *sticks* with carbon, oxygen, and nitrogen atoms colored *green*, *red*, and *blue*, respectively. The side chains of nonmutated residues in the conserved surface patch are shown as *sticks* with carbon, oxygen, and nitrogen atoms colored *light blue*, *red*, and *blue*, respectively. *B*, structure of the CBP21 Y54A mutant. The backbone is shown in a *sticks* representation, with $(2|F_o| - |F_c|, \phi_{calc})$ electron density contoured at 2.0σ. The site of the mutation is labeled.

affinity chromatography, and the obtained pure proteins were used for development of the chitin binding assay and determination of dissociation constants. The protein yields (pure protein) varied from 50 to 100 mg/liter culture. Circular dichroism spectra were recorded to verify the structural integrity of the CBP21 mutants. All CD spectra were essentially identical to that of the wild type (see Supplemental Material). In order to find the incubation time needed for the binding assay, the time course of binding was monitored by measuring the amount of unbound protein for each mutant at 1, 4, and 16 h of incubation. All CBP21 variants had established equilibrium after 16 h of incubation (results not shown; see also Fig. 3). Table II shows the equilibrium dissociation constants (K_d) and binding capacities (B_{max}) of the CBP21 variants, which were obtained from binding isotherms shown in Fig. 5. Although some binding characteristics of wild type CBP21 have been studied previously (12), K_d and B_{max} values had not yet been determined. For crystalline β-chitin, CBP21 shows a K_d of 1.4 μM and a B_{max} of 5.9 μmol/g. These values are in the same range as dissociation constants and binding capacities reported for other carbohydrate-binding proteins and isolated carbohydrate binding domains (41–44). The binding curves of the mutants show decreases in chitin affinity (increase in K_d) for all mutants (Fig. 5 and Table II). The largest increases in K_d were observed for Y54A (7.9-fold) and E60A (5.9-fold). The crystal structure of the Y54A mutant was solved and found to be essentially identical to the wild type structure (r.m.s.d. on C-α atoms is 0.4 Å), except for the lack of the tyrosine side chain (Table I and Fig. 4*B*). Thus, the effect of the Y54A mutation is solely due to the removal of the tyrosine ring and not to conformational changes in CBP21. Together, the mutagenesis data show that residues in the conserved surface area displayed in Fig. 4*A* are important for the ability of CBP21 to bind chitin. It is likely that the binding event is governed by several residues, which may explain why none of the mutations abolished chitin affinity. Only two of the mutants showed a significant reduction in binding capacity (E55A and H114A with a 2.8- and 2.0-fold decrease, respectively). This may mean that these mutants have lost their affinity for parts of the chitin surface.

Concluding Remarks—We have solved the structure of CBP21, the first structure of a family 33 CBM. The structure reveals a budded FnIII fold with no significant sequence conservation with other FnIII folds, but with structural similarity to the ChiA-FnIII domain. In contrast to most other CBMs and ChiA-FnIII in particular, CBP21 does not have a cluster of aromatic amino acids on the surface used for carbohydrate binding. Although an intriguing crystallographic dimer is observed, there is no evidence to support a physiological role for dimerization. The structural and sequence analyses and the mutagenesis studies suggest that binding of CBP21 to chitin could be primarily governed by polar interactions, involving residues in a polar surface patch remote from the observed dimer interface (Figs. 1*C* and Fig. 4*A*). The data also show that the one conserved aromatic residue that is fully surface exposed in CBP21 (Tyr-54) does contribute to chitin affinity. Mutation of the corresponding residue in CHB1 from *S. olivaceoviridis* (Trp-57) to leucine led to a 20-fold increase in K_d (40). Most interestingly, mutation of Trp-57 to Tyr (the residue present in CBP21) also led to a considerable decrease in chitin affinity ((40) equilibrium dissociation constant not determined). Because the crystal structures of CHB1 and its mutants are not known, it is not possible to evaluate the structural consequences of the Trp to Tyr mutation, nor is it at present understood how Tyr-54 or Trp-57 interact with chitin. Perhaps residue 54/57 is not only important for affinity but also for the differences between CBP21 and CHB1 in terms of their preferences for binding β-chitin and α-chitin, respectively.

TABLE II
Binding of CBP21 wild type and mutants to β-chitin

CBP21	K_d	B_{max} (μmol of CBP21/g)
	μM	
Wild type	1.4 ± 0.4	5.9 ± 0.4
Y54A	11.0 ± 3.1	4.0 ± 0.6
E55A	5.1 ± 1.4	2.1 ± 0.2
E60A	8.2 ± 2.6	7.9 ± 1.4
H114A	6.9 ± 1.7	3.0 ± 0.4
D182A	7.4 ± 1.8	5.3 ± 0.6
N185A	6.6 ± 2.0	7.3 ± 1.1

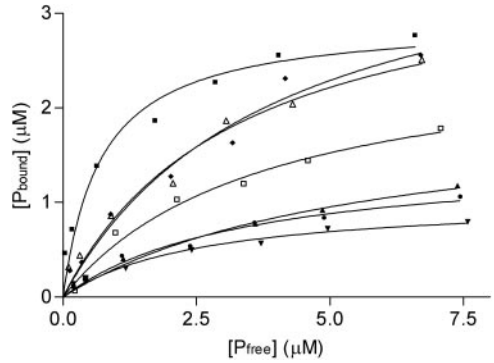

FIG. 5. **Plots of binding data for CBP21 wild type (*filled squares*), Y54A (*filled triangles, bottom down*), E55A (*filled triangles, bottom up*), E60A (*filled diamonds*), H114A (*filled circles*), D182A (*open squares*) and N185A (*open triangles*).** Each point represents the average of values obtained in three independent experiments. P_{bound} corresponds to specifically bound protein (μM), P_{free} corresponds to nonbound protein (μM). All data sets were fitted to the equation for one site binding (see "Experimental Procedures") by nonlinear regression.

Before the structure of the CBM-33 family member was known, the five largely conserved tryptophans characteristic of this family were believed to form a hydrophobic chitin binding surface (analogous to the majority of other carbohydrate-binding modules). However, the structure shows that instead, these tryptophans are part of a tightly packed hydrophobic core (Fig. 1*A*). In the study by Zeltins and Schrempf (40), all conserved tryptophans of CHB1 were mutated to leucines in order to probe their importance for chitin affinity. All mutants resulted in folded protein that also showed reduced affinity to chitin (~9-fold at most). These results are surprising when viewed in light of the CBP21 structure. Mutations of these buried and tightly packed tryptophans would be expected to destabilize the CBM-33 core as a whole, and there is no straightforward explanation for the specific effect of these mutations on chitin affinity.

It is conceivable that CBP21 uses specific hydrogen bonds, in addition to the solvent-exposed Tyr-54, to bind to multiple chitin chains that are at least partially ordered. Specific hydrogen bond-based interactions may explain why certain CBPs preferably bind to α-chitin (11, 13), whereas others, such as CBP21, preferably bind to β-chitin. The main difference between the surfaces of these chitin forms concerns the spatial ordering of groups that may engage in hydrogen bonds with a CBP. In fact, chitin is semicrystalline (mostly crystalline with amorphous regions) (45), and chitin preparations are not likely to represent a source of identical binding sites. It is conceivable that different CBPs have different affinities for the various forms and that mutations can have effects on affinity for certain types of chitin only. Mutational effects on $B_{max,}$ which were observed here and in the study of CHB1 (40), may be explained by assuming that certain residues are of particular importance for binding to certain (sub)types of chitin or certain parts of a chitin surface. Differences in substrate preferences/recognition have also been observed for cellulose-binding modules, where some have been shown to bind specific faces of the crystalline cellulose crystal lattice (36) or to distinguish between crystalline and amorphous variants (35, 46).

Acknowledgments—We thank the European Synchrotron Radiation Facility (Grenoble) for the time at beamline ID23-1 and Dimitrios Mantzilas (Department of Molecular Biosciences, University of Oslo) for assistance with the CD experiments. We thank Stein Ivar Aspmo for help with the gel filtration experiments.

REFERENCES

1. Henrissat, B., and Davies, G. (1997) *Curr. Opin. Struct. Biol.* **7,** 637–644
2. Hashimoto, M., Ikegami, T., Seino, S., Ohuchi, N., Fukada, H., Sugiyama, J., Shirakawa, M., and Watanabe, T. (2000) *J. Bacteriol.* **182,** 3045–3054
3. Perrakis, A., Tews, I., Dauter, Z., Oppenheim, A. B., Chet, I., Wilson, K. S., and Vorgias, C. E. (1994) *Structure* (*Lond.*) **2,** 1169–1180
4. van Aalten, D. M. F., Synstad, B., Brurberg, M. B., Hough, E., Riise, B. W., Eijsink, V. G. H., and Wierenga, R. K. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 5842–5847
5. Uchiyama, T., Katouno, F., Nikaidou, N., Nonaka, T., Sugiyama, J., and Watanabe, T. (2001) *J. Biol. Chem.* **276,** 41343–41349
6. Chang, M. C., Lai, P. L., and Wu, M. L. (2004) *FEMS Microbiol. Lett.* **232,** 61–66
7. Suetake, T., Tsuda, S., Kawabata, S., Miura, K., Iwanaga, S., Hikichi, K., Nitta, K., and Kawano, K. (2000) *J. Biol. Chem.* **275,** 17929–17932
8. Svergun, D. I., Becirevic, A., Schrempf, H., Koch, M. H. J., and Gruber, G. (2000) *Biochemistry* **39,** 10677–10683
9. Folders, J., Tommassen, J., van Loon, L. C., and Bitter, W. (2000) *J. Bacteriol.* **182,** 1257–1263
10. Chu, H. H., Hoang, V., Hofemeister, J., and Schrempf, H. (2001) *Microbiology* (*UK*) **147,** 1793–1803
11. Schnellmann, J., Zeltins, A., Blaak, H., and Schrempf, H. (1994) *Mol. Microbiol.* **13,** 807–819
12. Suzuki, K., Suzuki, M., Taiyoji, M., Nikaidou, N., and Watanabe, T. (1998) *Biosci. Biotechnol. Biochem.* **62,** 128–135
13. Kolbe, S., Fischer, S., Becirevic, A., Hinz, P., and Schrempf, H. (1998) *Microbiology* (*UK*) **144,** 1291–1297
14. Techkarnjanaruk, S., and Goodman, A. E. (1999) *Microbiology* **145,** 925–934
15. Monreal, J., and Reese, E. T. (1969) *Can. J. Microbiol.* **15,** 689–696
16. Watanabe, T., Kimura, K., Sumiya, T., Nikaidou, M., Suzuki, K., Suzuki, M., Taiyoji, M., Ferrer, S., and Regue, M. (1997) *J. Bacteriol.* **179,** 7111–7117
17. Suzuki, K., Uchiyama, T., Suzuki, M., Nikaidou, N., Regue, M., and Watanabe, T. (2001) *Biosci. Biotechnol. Biochem.* **65,** 338–347
18. van Aalten, D. M. F., Komander, D., Synstad, B., Gåseidnes, S., Peter, M. G., and Eijsink, V. G. H. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 8979–8984
19. Suzuki, K., Sugawara, N., Suzuki, M., Uchiyama, T., Katouno, F., Nikaidou, N., and Watanabe, T. (2002) *Biosci. Biotechnol. Biochem.* **66,** 1075–1083
20. Synstad, B., Gaseidnes, S., van Aalten, D. M. F., Vriend, G., Nielsen, J. E., and Eijsink, V. G. H. (2004) *Eur. J. Biochem.* **271,** 253–262
21. Brurberg, M. B., Eijsink, V. G. H., Haandrikman, A. J., Venema, G., and Nes, I. F. (1995) *Microbiology* **141,** 123–131
22. Manoil, C., and Beckwith, J. (1986) *Science* **233,** 1403–1408
23. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
24. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 849–861
25. Cowtan, K. (1994) *Jt. CCP4 and ESF-EACBM Newsl. on Protein Crystallogr.* **31,** 34–38
26. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463
27. Collaborative Computational Project CCP4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763
28. Brunger, A. T., Adams, P. D., Clore, G. M., Gros, P., 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 Biol. Crystallogr.* **54**, 905–921
29. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47**, 110–119
30. Laskowski, R. A., McArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
31. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233,** 123–138
32. Jee, J. G., Ikegami, T., Hashimoto, M., Kawabata, T., Ikeguchi, M., Watanabe, T., and Shirakawa, M. (2002) *J. Biol. Chem.* **277,** 1388–1397
33. Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., and Steitz, T. A. (1996) *EMBO J.* **15,** 5739–5751
34. Bayer, E. A., Chanzy, H., Lamed, R., and Shoham, Y. (1998) *Curr. Opin. Struct. Biol.* **8,** 548–557
35. Carrard, G., Koivula, A., Soderlund, H., and Beguin, P. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 10342–10347
36. Lehtio, J., Sugiyama, J., Gustavsson, M., Fransson, L., Linder, M., and Teeri, T. T. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 484–489
37. Vriend, G. (1990) *J. Mol. Graphics* **8,** 52–56
38. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) *J. Comput. Chem.* **19,** 1639–1662
39. Flint, J., Nurizzo, D., Harding, S. E., Longman, E., Davies, G. J., Gilbert, H. J., and Bolam, D. N. (2004) *J. Mol. Biol.* **337,** 417–426
40. Zeltins, A., and Schrempf, H. (1997) *Eur. J. Biochem.* **246,** 557–564
41. Li, Z. F., Li, C. B., Yang, K., Wang, L. H., Yin, C., Gong, Y. X., and Pang, Y. (2003) *Virus Res.* **96,** 113–122
42. Goldstein, M. A., Takagi, M., Hashida, S., Shoseyov, O., Doi, R. H., and Segel, I. H. (1993) *J. Bacteriol.* **175,** 5762–5768
43. Sheweita, S. A., IchiIshi, A., Park, J. S., Liu, C. C., Malburg, L. M., and Doi, R. H. (1996) *Gene* (*Amst.*) **182,** 163–167
44. Gal, L., Pages, S., Gaudin, C., Belaich, A., ReverbelLeroy, C., Tardif, C., and Belaich, J. P. (1997) *Appl. Environ. Microbiol.* **63,** 903–909
45. Blackwell, J. (1988) *Methods Enzymol.* **161,** 435–442
46. McLean, B. W., Boraston, A. B., Brouwer, D., Sanaie, N., Fyfe, C. A., Warren, R. A. J., Kilburn, D. G., and Haynes, C. A. (2002) *J. Biol. Chem.* **277,** 50245–50254
47. Bond, C. S. (2003) *Bioinformatics* **19,** 311–312
48. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
49. Notredame, C., Higgins, D. G., and Heringa, J. (2000) *J. Mol. Biol.* **302,** 205–217.