Cell. Mol. Life Sci. (2009) 66:2585–2598
DOI 10.1007/s00018-009-0063-x

Cellular and Molecular Life Sciences

RESEARCH ARTICLE

Crystal structures of g-type lysozyme from Atlantic cod shed new light on substrate binding and the catalytic mechanism

Ronny Helland · Renate L. Larsen · Solrun Finstad · Peter Kyomuhendo · Atle N. Larsen

Received: 31 March 2009 / Revised: 15 May 2009 / Accepted: 3 June 2009 / Published online: 20 June 2009

Abstract Crystal structures of Atlantic cod lysozyme have been solved with and without ligand bound in the active site to 1.7 and 1.9 Å resolution, respectively. The structures reveal the presence of NAG in the substrate binding sites at both sides of the catalytic Glu73, hence allowing the first crystallographic description of the goose-type (g-type) lysozyme E–G binding sites. In addition, two aspartic acid residues suggested to participate in catalysis (Asp101 and Asp90) were mutated to alanine. Muramidase activity data for two single mutants and one double mutant demonstrates that both residues are involved in catalysis, but Asp101 is the more critical of the two. The structures and activity data suggest that a water molecule is the nucleophile completing the catalytic reaction, and the roles of the aspartic acids are to ensure proper positioning of the catalytic water.

Keywords Atlantic cod · Fish lysozyme · Crystal structure · Active site residues · Substrate binding sites · Surface potential · Muramidase activity

R. Helland (✉) · R. L. Larsen · S. Finstad · P. Kyomuhendo
Department of Chemistry, The Norwegian Structural Biology Centre, University of Tromsø, 9037 Tromsø, Norway
e-mail: Ronny.Helland@uit.no

P. Kyomuhendo
Nofima Marine, P.O. Box 6122, 9291 Tromsø, Norway

A. N. Larsen (✉)
Department of Marine Biotechnology,
Norwegian College of Fishery Science, University of Tromsø,
9037 Tromsø, Norway
e-mail: Atle.Larsen@uit.no

Introduction

Innate immunity represents the first line of defense against invading pathogens and is believed to respond efficiently in fish [1]. One of the innate defense molecules important in the protection against bacterial invaders is lysozyme (muramidase, EC 3.2.1.17), and its role as a defense molecule in fish has recently been reviewed [2]. Lysozyme catalyzes hydrolysis of most bacterial cell walls by cleaving the β-1,4 glycosidic bond between *N*-acetyl-D-glucosamine (NAG) and *N*-acetylmuramic acid (NAM) in the peptidoglycan layer. Two different types of lysozymes are found in vertebrates, i.e., chicken- and goose-type (c-type and g-type, respectively), and although both cleave the glycosidic bond between the NAG and NAM molecules, they share low sequence similarity, differ in molecular weight, and possess different enzymatic properties (see, e.g., [3–5]). In spite of this, the three-dimensional structures of these lysozymes share similarities in the active site where parts of two helices form one side of the binding pocket and a loop comprising three β-strands forms the other [6, 7]. Both enzymes have been proposed to possess binding sites for six NAG and/or NAM molecules, where c-type and g-type lysozymes accommodate sugar molecules in sites named A–F and B–G, respectively [8, 9]. Cleavage of the glycosidic bond occurs between the substrate molecules occupying the D and E sites. So far, substrate binding to the prime side of the binding site has only been structurally described for c-type lysozyme (sites E–F) [10], although the putative E–G sites in g-type lysozyme have also been suggested [8, 11]. The g-type lysozyme was initially identified as an antibacterial enzyme located in egg whites of Embden goose [12, 13], but is now known to exist in both mammals and fish [3]. The first fish g-type lysozyme was reported in Japanese

flounder in 2001 [14] and has since then also been reported from several other fish species [3, 15–19] including Atlantic cod [20]. Recombinant fish g-type lysozymes have been produced and their lytic activity against bacterial fish pathogens has been proven [14, 18, 19]. The g-type lysozyme from Atlantic cod is shown to possess muramidase activity using *Micrococcus luteus* cells as substrate at an unusually low pH [20], but detailed structural knowledge explaining this pH activity profile is lacking. So far, the only crystal structures reported of g-type lysozymes are from the avian species Embden goose and Australian black swan which adopt a well-characterized $\alpha + \beta$ fold consisting of seven α-helices and three β-strands [6, 7, 21, 22]. It is well known that the residues involved in the catalysis of the c-type lysozymes are Glu35 and Asp52. The crucial role of the Glu35 counterpart, Glu73, in catalysis of the g-type lysozymes is well established [7, 21, 23] and is conserved within the g-type lysozyme group. However, structural analysis of the g-type lysozymes has not provided evidence for a counterpart to Asp52 [7, 21]. Several candidates for the second residue have been suggested, including Asp86 and Asp97, but none of these have been experimentally verified.

Here, we present two structures of g-type lysozyme from Atlantic cod with and without NAG molecules bound in the active site, providing the first crystallographic description of a ligand bound to the prime side of the binding site. A detailed comparison with the structures of two avian g-type lysozymes is also performed. Our structural analysis together with experimental muramidase activity data on native and mutant cod g-type lysozymes shed new light on the roles of Asp90 and Asp101 (Asp86 and Asp97, respectively, goose numbering) in catalysis. We also present possible structural explanations for the muramidase activity of cod lysozyme at low pH.

Materials and methods

Crystallization, data collection, structure determination and analysis

Recombinant expression and purification of cod g-type lysozyme (gLYS) is described by Larsen et al. [20]. Crystals of gLYS suitable for X-ray diffraction studies were obtained by the hanging drop vapor diffusion method from 48 to 53% ammonium sulphate, 5 mM $CoCl_2$ and 0.1 M BisTris pH 6.5–7.0 at room temperature. Protein concentration was 10 mg ml^{-1} in 25 mM Hepes pH 7.0 and 150 mM NaCl. The structure of ligand-bound gLYS was obtained by soaking lysozyme crystals overnight in reservoir solution containing a $(NAG)_3$:gLYS ratio (molar) of 6:1. X-ray data were collected at BL14.1 at Bessy. Data were processed in XDS [24] and SCALA and TRUNCATE of the CCP4 program suite [25]. The ligand-free structure was solved by molecular replacement in MOLREP [26] using PDB entry 154L [7] as search model. Automatic tracing of the polypeptide chains was carried out with ARP/wARP [27]. Subsequent improvement of the model was made by manual refitting of amino acids using O [28] based on sigmaA-weighted 2mFo-DFc and mFo-DFc electron density maps. The four molecules in the asymmetric unit were allowed to be refined independently using Refmac5 [29] of the CCP4 suite. The structure of ligand-bound gLYS was obtained by rigid body refinement using the ligand-free structure as starting point and subsequent positional refinement.

Programs for structural comparison and analysis included the CCP4 suite and the DaliLite server (http://www.ebi.ac.uk/DaliLite/) [30]. Main chain atoms of the four molecules in the asymmetric unit of the ligand-bound gLYS were superimposed on each other using LSQKAB of the CCP4 suite (Table 2). In addition, all atoms of the individual molecules in the ligand-bound structure were compared to the respective molecules in the ligand-free structure (MolA to MolA, MolB to MolB, and so forth; values in bold in Table 2). Because of the high similarity of the structures, only the ligand-bound structure will be further described. Significant differences are, however, mentioned.

Modeling of NAG into the D site of gLYS was performed by appending the coordinates of the NAG molecule occupying the D site in the rotated NAG bound GEWL-structure (PDB 154L) to gLYS MolA. The modeled MolA–NAG complex was subjected to 5,000 steps of energy minimization using the MacroModel (version 9.5) option of the Schrödinger software (Schrödinger, LLC, New York, NY, 2008, http://www.schrodinger.com).

Calculation of electrostatic surface potential was done using Delphi [31] with atomic partial charges from the AMBER molecular simulation program [32]. Estimation of intra-molecular hydrogen bonds was carried out using HBPLUS [33] with a 3.4-Å distance criteria and default donor–acceptor angles of 90° Salt bridges, or ionic interactions, are defined as interactions less than 4 Å between side chains of opposite charge.

Illustrations were prepared in PyMOL (DeLano Scientific; http://pymol.sourceforge.net/) with the exceptions of the structural alignment generated at the ESPript server (http://espript.ibcp.fr/ESPript/ESPript/) [34], and the schematic 2D representation of the NAG binding sites generated using LIGPLOT [35].

Generation of cod lysozyme (gLYS) mutants

Site-directed mutagenesis was performed using QuickChange II Site-directed mutagenesis kit (Stratagene)

according to the manufacturer instructions. Primers D90A_F (5′-CCCTGGTTGGGGGGCCAATTATAATGG CTTTGG-3′) and D90A_R (5′-CCAAAGCCATTATA ATTGGCCCCCCAACCAGGG-3′) were used with the "native" expression construct (pET151/D-*gLYS*, as described [20]) to create the mutant D90A, while primer D101A_F (5′-GCTTTGGACTGATGCAGGTTGCTAAG AGATACCACGAACC-3′) and D101A_R (5′-GGTTCGT GGTATCTCTTAGCAACCTGCATCAGTCCAAAGC-3′) were used with either the "native" pET151/D-*gLYS* or pET151/D-*gLYS* D90A construct, to generate the single mutant D101A and the double mutant D90A/D101A, respectively. The expression vectors containing the mutants were transformed into XL-1 Blue (Stratagene) or Top10 (Invitrogen) super-competent cells. Plasmids were isolated using QIAprep Spin Miniprep kit (Qiagen) according to the manufacturer's instructions, and the presence of correct mutations was verified by DNA sequencing. The recombinant expression and purification of gLYS mutants were performed as previously described [20], the main difference being that smaller volumes were used during the expression trials (50 ml pre-culture were used to inoculate 1 l of LB broth).

Determination of lysozyme activity

Lysozyme (muramidase) activity was measured using 1 mg ml^{-1} of lyophilised *M. luteus* cells (Sigma) as substrate (dissolved in 0.1 M NaAc, pH 4.8). Activity was measured by a linear decrease in absorbance at 450 nm using the SpectraMAX M2 microplate reader (Molecular Devices), where one unit of enzyme activity is defined as a decrease in absorbance units of 0.001 min^{-1}. Two independent activity experiments using freshly made substrate solution, and a total of eight parallels were performed. Protein concentration were determined using Bio-Rad Protein Assay according to the microtiter plate protocol as described by the manufacturer (Bio-Rad) using bovine serum albumin (BSA) as standard and in triplicate for each standard and sample. Muramidase activity data (% specific activity) of gLYS D90A, D101A, and D90A/D101A mutants are relative to that of native gLYS (set to 100%).

Results and discussion

Crystallization, data collection and refinement

Thin plate-like diamond or hexagonal shaped crystals of cod lysozyme (gLYS) grew in about a week from ammonium sulphate at pH 6.5–7. Crystals of about $0.2 \times 0.1 \times 0.05$ mm^3 diffracted to 1.9- and 1.7-Å resolution, ligand-free and ligand-bound structure, respectively, and belonged to the monoclinic space group *C*2 with cell parameters of $110 \times 75 \times 78$ $Å^3$, and $\beta = 93°$. The crystal structure of ligand-free gLYS was solved by molecular replacement using a poly-ala model of lysozyme from goose egg white (GEWL) as search model (PDB entry 154L) [7]. Four molecules were identified in the asymmetric unit (MolA, MolB, MolC, and MolD; Fig. 1). ARP/wARP was able to auto-build 475 of 756 residues, and the main chain of 73 additional residues was traced. Manual model building and refinement increased the number of residues which could be fitted into electron density to 728 (MolA: 185; MolB: 183; MolC: 174; MolD: 186). The final structure is refined to crystallographic *R*-factors of 19.1% (R_{work}) and R_{free} of 25.1%, and data collection and refinement statistics are listed in Table 1. The structure of ligand-bound gLYS was obtained by soaking lysozyme crystals overnight in reservoir solution containing a $(NAG)_3$:gLYS ratio (molar) of 6:1. The ligand-free structure was used as starting point for refinement. The structure of ligand-bound gLYS is refined to crystallographic *R*-factors of 20.3% (R_{work}) and 25.3% (R_{free}) (Table 1). Coordinates have been deposited with the Protein Data Bank (accession codes 3GXK and 3GXR for ligand-free and ligand-bound gLYS, respectively).

Overall fold

The majority of residues are well defined in electron density, but a number of side chain and main chain atoms are significantly disordered displaying weak and/or missing electron density in both structures. Particularly, MolC is poorly defined, where residues 79–91 are invisible in the electron density maps, probably due to flexibility in the structures and unfavorable crystal packing interactions. Main chain atoms of molecules A, B, C, and D in the ligand-bound gLYS were superimposed on each other with root mean square displacement (RMSD) values from about 0.4 to 0.8 Å (Table 2). The corresponding RMSD values for the ligand-free gLYS were similar. Most regions with large RMSD variations are involved in close packing interactions being different in the four molecules in the asymmetric unit. Despite requiring extensive manual intervention in model building and possessing a sequence identity of about 55%, main chain atoms of the four molecules in the final models of gLYS superimpose surprisingly well on avian lysozymes (Fig. 2) with RMSD values in the range 0.76–1.13 Å (Table 2) for residues 8–185 (goose lysozyme numbering). This is only slightly higher than the maximum RMSD value for the four gLYS molecules in the asymmetric unit. As shown in Figs. 1 and 2, the gLYS structure adopts the well-characterised $\alpha + \beta$ fold observed for the avian lysozymes, where the conserved catalytic Glu73 is located at the C-terminal rim of

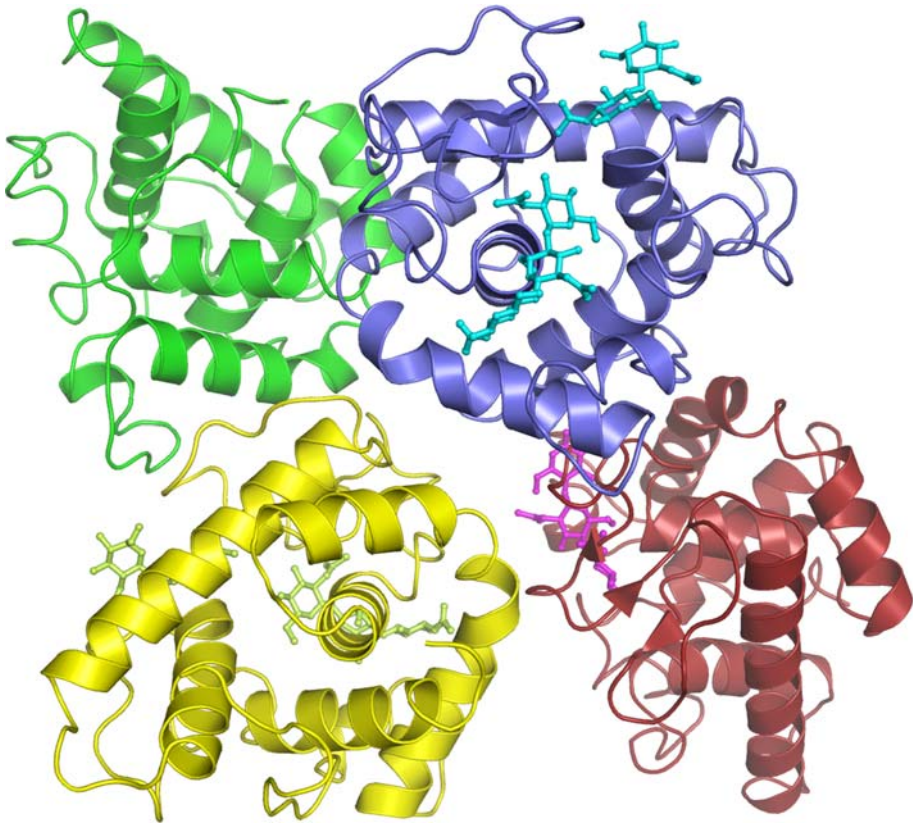

Fig. 1 Illustration of the four molecules of Atlantic cod lysozyme in the crystallographic asymmetric unit. Molecule A is coloured *blue*, molecule B is *red*, molecule C is *green* and molecule D is *yellow*. NAG molecules identified in the crystal structure are illustrated as ball-and-stick models

α-helix 4. The only significant difference in the fold of the fish and the avian lysozymes is in the region 82–92 (cod numbering, corresponding to 82–88 in the avian lysozymes). The start of this region includes a unique insert of four residues (^{82}NTTP^{85}) in gLYS. More interestingly, at the end of this region, Asp90 (Asp86 in goose and swan) suggested to be involved in catalysis, occupies a conformation completely different from the avian lysozymes, and Asp90 in gLYS is, in contrast to the avian lysozymes, folded inwards towards the binding cleft. The difference in the position of Asp90/86 in gLYS and GEWL is not obvious, but may be explained by formation of favorable electrostatic interactions between GEWL Asp86 and Arg155 of a symmetry-related molecule (3.8 and 4.2 Å, PDB 154L and 153L, respectively).

Substrate binding sites

gLYS crystals soaked with $(NAG)_3$ displayed clear electron density for two ligands in the binding cleft of each of the A and D molecules (Figs. 3 and 4), occupying the B–C and E–G sites The electron density was weaker in MolD than MolA, and the ligands were therefore refined with reduced occupancy of 0.5 in MolD. The binding of $(NAG)_3$ in the prime side of the catalytic Glu73 allows the first crystallographic description of the E–G binding sites in g-type lysozymes (Fig. 3). Weak difference electron density was also observed in the E–G sites of MolB, and $(NAG)_3$ was refined with an occupancy of 0.3. NAG cannot bind in the B and C sites in MolB because of steric hindrance due to packing interactions. In MolC, the substrate binding sites are not defined because of the lack of clear electron density for residues 79–91.

The conformation of the binding site of g-type lysozymes has previously been suggested not to change significantly upon binding of substrate [7], and the same appears to be valid for ligand-bound gLYS (Table 2, diagonal). Binding of NAG to gLYS results in an average *xyz* displacement of the main-chain atoms of Asp90, Asn91, and Tyr92 of 0.8, 1.3, and 2.3 Å, respectively, in MolA, and 0.3, 0.3, and 0.2 Å, respectively, in MolD. The lower RMSD values in MolD must be seen in relation to the lower occupancy of the ligands. It should also be noted that MolA is the monomer deviating most from the other monomers (Table 2), particularly at residues 90–92, and the difference is mostly attributed to differences in the crystal packing environments. The side chain of Asp90 moves closer to the ligand in both MolA and MolD (*xyz* displacement of 0.99 and 0.77 Å, respectively), and the same is observed for residue His105. Smaller displacements are seen for Tyr104, Arg103, and Asp101.

Table 1 Data collection and refinement statistics

	Ligand-free gLYS	Ligand-bound gLYS
Data collection statistics		
Diffraction limit	1.9	1.7
Space group	*C*2	*C*2
Unit cell parameters		
a-axis (Å)	110.12	110.19
b-axis (Å)	75.39	75.52
c-axis (Å)	78.37	78.29
β-angle (°)	92.92	93.25
Total no. of reflections	227,168 (30,435)	260,923 (32,950)
No. of unique reflections	50,202 (7,068)	69,961 (9,754)
Completeness (%)	99.5 (96.9)	99.3 (95.6)
I/σ(I)	9.4 (3.3)	7.8 (2.1)
Mean I/σ(I)	14.8 (5.5)	11.1 (2.9)
R_{merge} (%)	6.0 (22.1)	6.2 (35.3)
Multiplicity	4.5 (4.31)	3.7 (3.4)
Wilson B (Å^2)	18.99	22.34
Outer shell resolution (Å)	2.0–1.9	1.79–1.70
Refinement statistics		
R_{work} (%)	19.16	20.34
R_{free} (%)	25.10	25.25
Average *B* factors (Å^2)	18.76	21.49
No. protein atoms	5,735	5,731
No. non-protein atoms (incl. partial occ.)		
NAG (13 molecules)	0	186
Solvent	255	284
RMS deviations		
Bond lengths (Å)	0.019	0.020
Bond angles (°)	1.649	1.842
% Residues in regions of the Ramachandran plot		
Most favored	91.8	91.7
Additionally allowed	8.0	8.0
Generously allowed	0.2	0.3
Disallowed	0	0

Outer shell values are given in parentheses

Substrate binding sites B–D

Residues believed to form the substrate binding sites B–D, including the essential catalytic Glu73, are well conserved in the cod and avian lysozymes with exceptions of Asn91, Tyr92, Tyr104, Arg157, and Arg167 (cod numbering), being Arg, Gly, Ser, Asn, and Gly, respectively, in goose and black swan lysozymes (Fig. 2a). The larger size of the side chains at these positions in gLYS, along with the conformation of residues Asp90, Arg103, Tyr104 (Ser100 in avian lysozymes), Arg157 (Asn153 in avian lysozymes), and Arg167 (Gly163 in avian lysozymes) results in a less accessible B–D binding site in the fish enzyme (Fig. 4). Particularly, the volume of the D site is significantly reduced because of the conformation of the 89–92 loop.

The NAG molecules occupying the B–D region of the lysozyme binding cleft displayed clear electron density for only two of the three NAG molecules (Fig. 3a). There is no NAG molecule occupying the D site in gLYS, and from the present data it is not clear whether the third NAG molecule of $(NAG)_3$ is disordered or if it is cleaved off. The two NAG molecules, occupying the B and C sites, superimpose well on the conformation of NAG in the B and C sites of the goose structure (PDB 154L). Since NAG is observed in the D site in the crystal structure of GEWL, a NAG molecule would be expected to fit into the D site of gLYS as well. However, in order to soak NAG into the D site of gLYS, Asp90 would have to be slightly displaced to avoid steric clashes, but movement of the 89–92 loop region in order to also accommodate ligand in the D site is restricted in gLYS because the neighboring Tyr92 (Gly88 in the avian lysozymes) cannot easily be moved without forming collisions with residues 84–85 and symmetry-related molecules. Attempts to co-crystallize gLYS with $(NAG)_3$, possibly causing NAG to bind in the D site before the crystal was formed, were not successful.

Binding of NAG in site B of MolA in gLYS is characterised by few enzyme-ligand interactions (Fig. 3b), similar to what is observed in GEWL. NAG binds

Table 2 Root mean square (RMS) displacement values for superimposition of g-type lysozymes

	gLYS MolA	gLYS MolB	gLYS MolC	gLYS MolD	GEWL
gLYS MolA	**1.10**	0.78	0.75	0.77	1.13
gLYS MolB		**0.29**	0.37	0.39	0.86
gLYS MolC			**0.39**	0.47	0.75
gLYS MolD				**0.39**	0.76

Main chain atoms of molecules A–D are superimposed on each other for ligand-bound gLYS and on ligand-bound GEWL (PDB 154L). Values in bold (along the diagonal) are RMS displacement values when all atoms of NAG-bound gLYS molecules are superimposed on the corresponding ligand-free gLYS

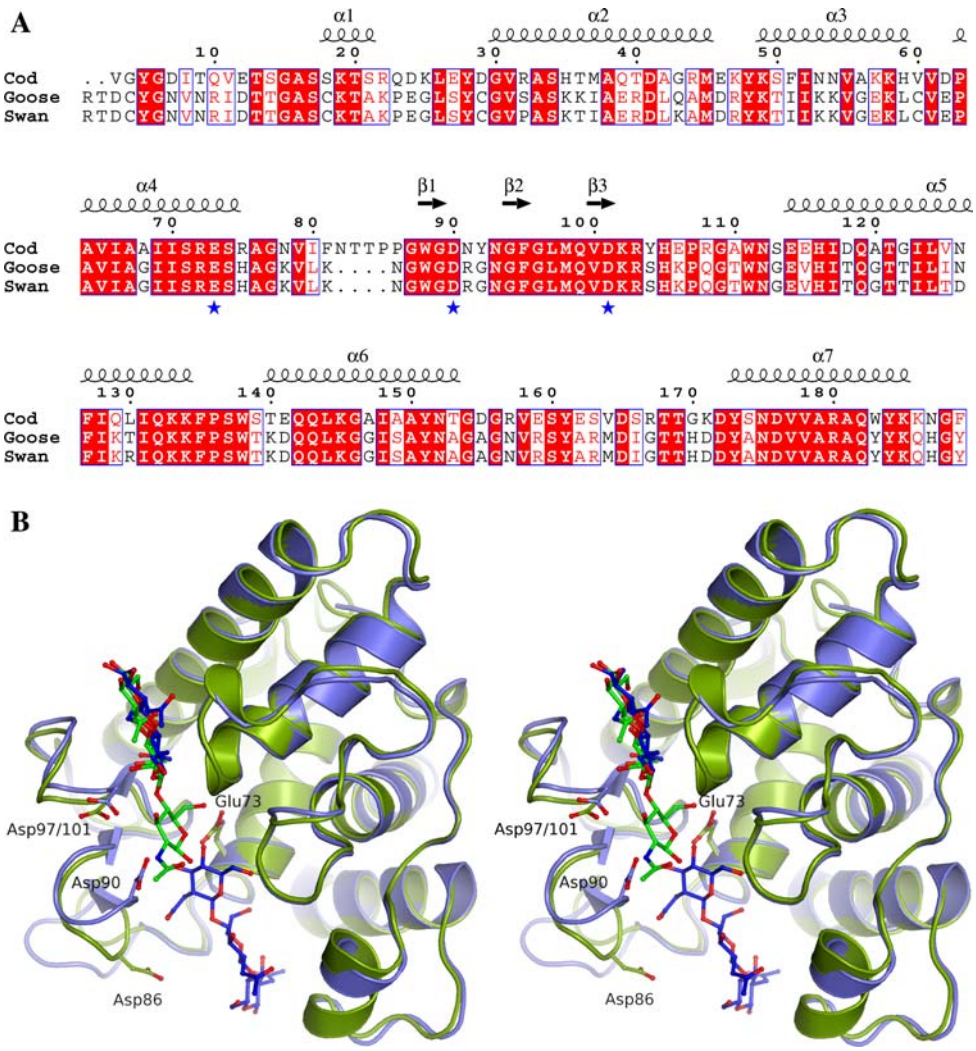

Fig. 2 **a** Structural alignment of cod, goose and black swan lysozymes. The residues proposed to be important in catalysis are marked with a *blue asterisk*. **b** Stereo-plot illustrating superimposition of cod (*blue*) and goose lysozymes (PDB 154L; *green*). NAG bound to cod and to goose lysozymes and residues Glu73, Asp90 (Asp86 in goose) and Asp101 (Asp97 in goose) are illustrated as ball-and-stick models

significantly stronger in site C compared to site B both in gLYS and GEWL as defined by the number of hydrogen bonds, but gLYS apparently possess both a higher number of, and shorter, hydrogen bonds than GEWL, possibly resulting in stronger binding (Fig. 3b). The residue at position 104 is Tyr in gLYS (Ser100 in GEWL), and the bulkier size of the Tyr side chain significantly reduces the volume of site C, and to some extent site B. The conformation of the side chain will also reduce the rotational freedom of Asp101 (Asp97 in GEWL), one of the residues suggested to be involved in catalysis. Tyr104 also forces the $NAG_C{}^{O6}$ atom to occupy a different conformation in gLYS than in GEWL.

Substrate binding sites E–G

The residues forming the E–G binding sites (prime binding sites) in gLYS consists of Gln23, Asp24, Arg72, Glu73, Arg75, Asn78, Val79, Ile80, Gly89, Asp90, Gln99, Asn152, Thr169, and Tyr173 (Figs. 2a, 3b). The residues forming hydrogen bonds to NAG in the E (Arg72, Glu73, Asp90, Gln99, and Thr169) and F (Gln23 and Arg72) sites are conserved in gLYS and GEWL, with the exception of residue 23, which is Pro in GEWL. None of the residues are conserved in the G site (Glu24, His75, and Lys78 in GEWL). The most significant difference is at residue 90 (Asp86 in GEWL) in the E site, where the 89–92 loop (gLYS numbering) have different orientations in the two enzymes (Figs. 2b and 5). NAG_E is the ligand molecule forming most hydrogen bonds to gLYS, while NAG_F and NAG_G are characterized with very few interactions to the enzyme (Fig. 3b). The carboxyl group of the catalytic Glu73 is located only 2.5 Å from $NAG_E{}^{O4}$, the oxygen atom that would form the β-1,4-glycosidic bond between substrate residues occupying the D and E sites. This is only slightly shorter than what was obtained by the molecular dynamics simulations by Hirakawa et al. [11]. It interesting to note that Gln23 in the F site is folded towards the binding cleft in MolA of the ligand-free complex, but it is rotated almost 180° in order to accommodate the NAG

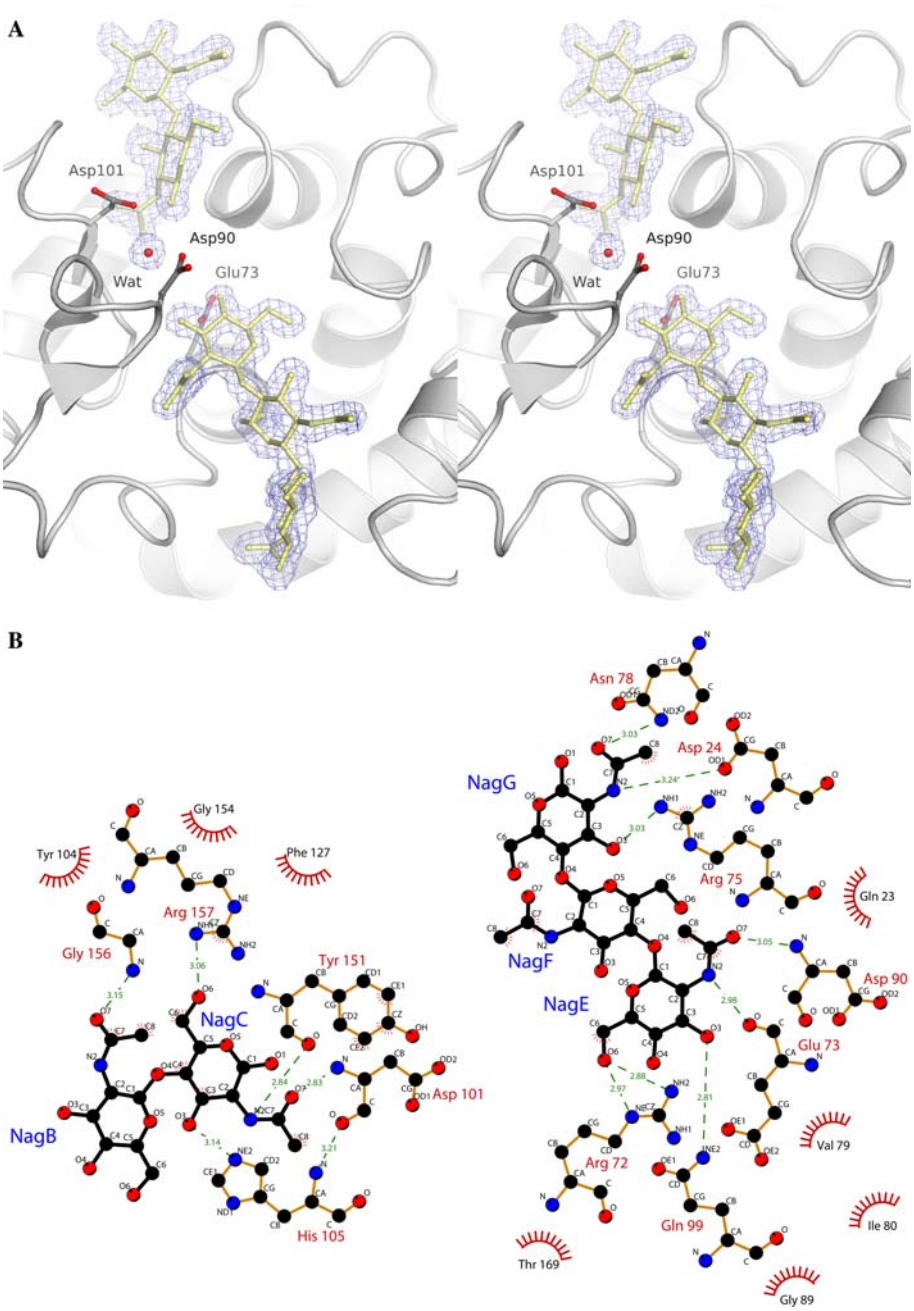

Fig. 3 Substrate binding site in cod lysozyme. **a** Stereo-plot displaying the electron density of NAG oligomers in the cod B, C and E–G sites of MolA. The electron density is contoured in *blue* at 1σ. **b** Schematic 2D illustration of residues interacting with NAG molecules in the B–C sites (*left*) and E–G sites (*right*) of cod lysozyme (MolA)

molecule. The electron density is, however, slightly weaker in the ligand-bound gLYS, suggesting flexibility of the residue. The stability of the F and G sites is partly ensured by salt bridges between residues 24 and 75 in both gLYS and GEWL. Although the bridges differ in composition, with Asp24 and Arg75 in gLYS and Glu24 and His75 in GEWL, the ionic interaction between the residues are in the same order of about 2.7 Å. Arg72 (F site) is also involved in an eight-residue salt-bridge network in the gLYS structure (Fig. 7; discussed in more detail later). Thus, some of the residues involved in forming salt-bridge networks seem to serve a dual role both in maintaining structural integrity as well as provide favorable interactions with substrate bound in the F and G sites.

Differences in amino acids at positions 23, 24, 75, 78, and 80 (Fig. 2) and the conformation of side chains results in a slightly more open E–G site in gLYS than in GEWL (Fig. 4). The difference is most pronounced for residues 23 and 24 (Gln and Asp in gLYS and Pro and Glu in GEWL) of the F and G sites. GEWL has previously reported to hydrolyze NAG oligomers at a much lower rate compared to hen egg-white lysozyme (HEWL, c-type lysozyme) [36]. As seen in Fig. 4, the E–G sites in gLYS and GEWL appears to be less accessible than the E–F sites in human

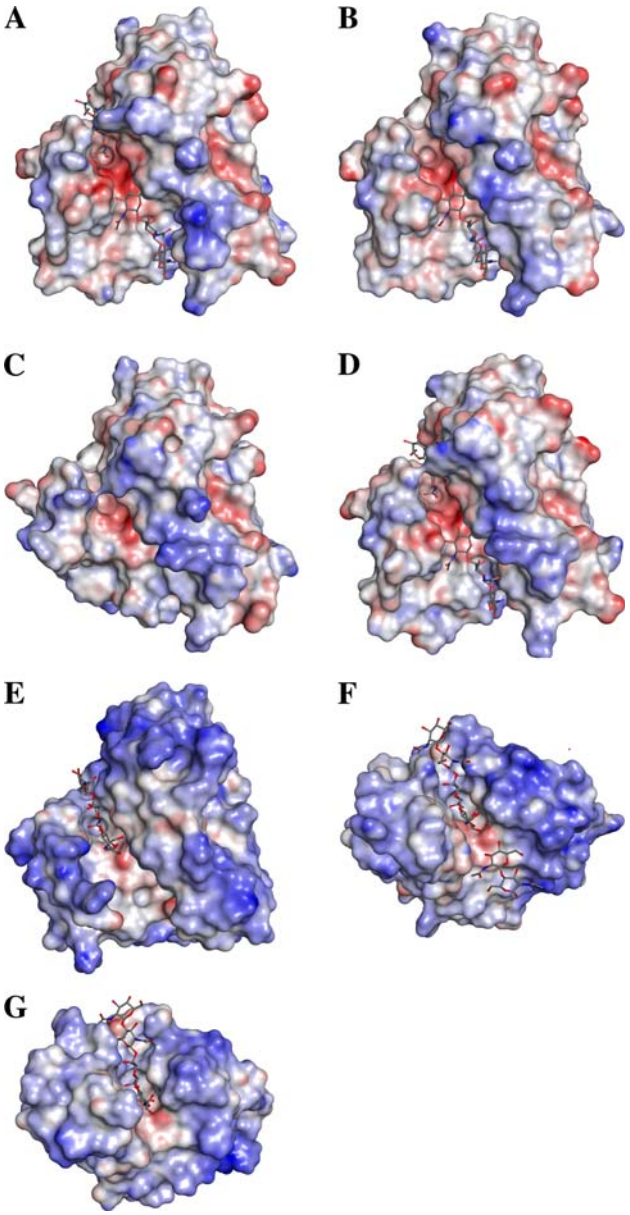

Fig. 4 Surfaces, including electrostatic potential, of cod lysozyme molecules (**a**–**d**), goose (**e**), human (**f**) and hen egg white (**g**) lysozymes (PDB 154L, 1LZS and 1LZC for the three latter, respectively). Residues 79–91 of gLYS MolC (**c**) are omitted from the structure due to unobserved electron density. The electrostatic surface potentials are contoured from −10 (*red*) to 10 (*blue*) kT/q. NAG observed in the active sites of the of lysozyme crystal structures are illustrated as ball-and-stick models

c-type lysozyme (PDB 1LZS [10]) and HEWL (PDB 1LZC [37]), possibly explaining the differences in the hydrolytic efficiency.

Role of Asp90 and Asp101 in the catalysis

Since the structure determination of g-type lysozyme more than two decades ago, the role of Glu73 in catalysis has been established [7, 21, 23], but scientists have still been intrigued by the apparent lack of the second catalytic residue in g-type lysozyme corresponding to Asp52 in c-type HEWL [7, 21]. Although Grutter et al. [6] suggested Asp86 to be the second catalytic residue, this has later been disputed due to the long distance between Asp86 and Glu73 in GEWL, and other possible residues have been suggested (Asn89, Gln95, and Asp97; see, e.g., [7]). It has also been suggested that catalysis by c- and g-type lysozymes proceeds through different mechanisms [38, 39]. Initial analysis of the crystal structure of ligand-free gLYS indicated that Asp90 (corresponding to Asp86 in goose and black swan lysozymes) may indeed function as the long sought after catalytic Asp residue. Although located in an apparent flexible loop region between β1 and β2, as demonstrated by somewhat weak electron density for the side chains and the absence in MolC, three of the four gLYS molecules in both structures show that the residues 89–91 have a conformation different from the avian lysozymes. The location of the Asp90 carboxyl group corresponds surprisingly well with the carboxyl group of the catalytic Asp52 in HEWL (Fig. 5b, PDB 1LZC [37]). The hypothesis was further strengthened by the fact that Asp86 in lysozyme from Atlantic salmon has the same conformation as Asp90 in cod (Helland and Kyomuhendo, unpublished data in preparation).

To further explore the potential role of Asp90 in catalysis, a NAG molecule was modeled into the gLYS D site by superimposing the GEWL (PDB 154L) on MolA of gLYS (Fig. 5). The initial model positions the carboxyl oxygen atom of Asp90 side chain only 2.3 Å from the C1 atom of the NAG_D, hence too close for favorable interactions. Energy minimization of gLYS with $(NAG)_3$ in both the B–D sites and the E–G sites, but with the glycosidic D–E bond cleaved for 5,000 steps using the MacroModel routine of the Schrodinger software, increased the distance to 2.6 Å shifting the position of NAG_D slightly towards the other side of the binding pocket. However, energy minimization resulted in a concomitant displacement of Glu73 towards NAG_E and a slight shift in the position of NAG_E. The results of the energy minimization should therefore be interpreted with care, and could indicate that slight adjustments of large regions in the binding site would have to take place in order to allow $(NAG)_3$ to be soaked, or bound, into in the D site of gLYS. Still, the model demonstrates that Asp90 is close enough to NAG_D to participate in catalysis. The only polar atoms within 5 Å of the C1 atom of the NAG molecule modeled into the D site are those of the side chains of Asp90, Arg103, Asn152, and Thr169. Asp101 is the only acidic residue, in addition to Asp90 and Glu73, within a radius of 8 Å from NAG in the D site and is suggested as the other possible "catalytic candidate" in gLYS. As observed in corresponding residue

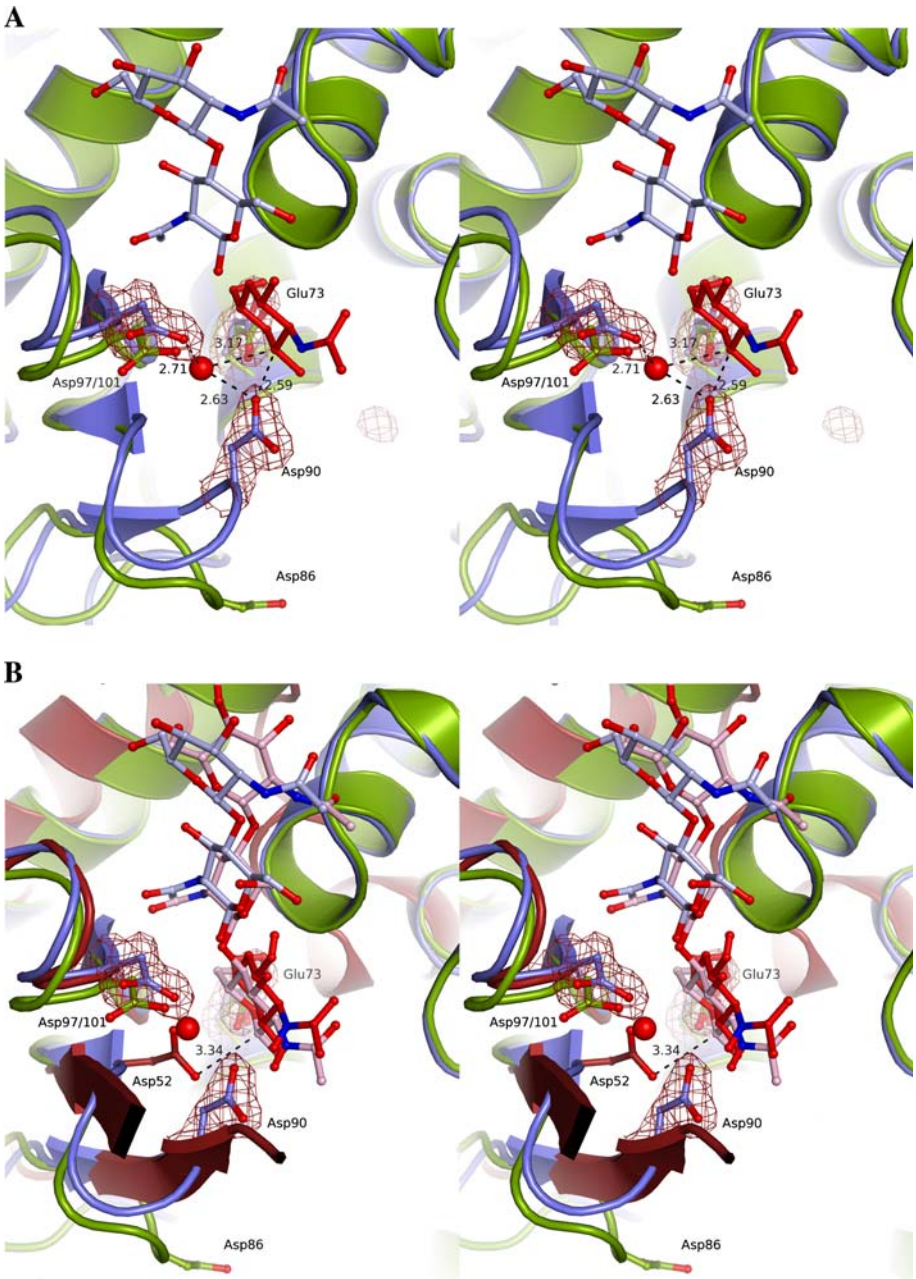

Fig. 5 Stereo-plot illustrating residues involved in catalysis in lysozymes. The active sites of cod, goose, and c-type hen egg white lysozymes (HEWL) are superimposed on each other and are represented as *blue*, *green* and *red* cartoons, respectively. NAG molecules and catalytic residues are displayed as ball-and-stick representations and have the same but paler color as the protein to which they belong. The catalytic water in cod lysozyme is illustrated as a *red sphere*. The electron density of cod residues Glu73, Asp90, and Asp101 (MolA) is contoured in *red* at 1σ. **a** Distances between the catalytic water in cod lysozyme and the catalytic aspartic acids are marked. The putative distance from Asp90 to the C1 atom of a NAG (*red* ball and stick model) molecule modeled into the cod site D is also marked. **b** Asp90 in cod is close to the position of the catalytic Asp52 in HEWL (PDB 1LZC). The putative distances from Asp90 in cod lysozyme and real distance from Asp52 in HEWL to the C1 atom of NAG_D ligand is indicated

in the avian structures (Asp97 in 1LSP/154L), this side chain is extensively hydrogen bonded to the nearby Arg103. Asp101 is also situated too far away to interact directly with the C1 atom of a ligand modeled into the D site. Even if Asp101 in gLYS is modeled into a rotamer closer to NAG_D, the distance would be about 4 Å, the same as obtained in the minimized structure. Asp101 was therefore not expected to act as a second catalytic residue.

To test the hypothesis, we mutated the residues in question to alanine generating two single mutants, D90A and D101A, as well as a double mutant D90A/D101A. The muramidase activity of mutated lysozymes towards *M. luteus* was measured and compared to the activity of the native gLYS (Table 3). Somewhat surprising, our results clearly indicate that Asp101 is most important for muramidase activity of the two candidates. The D101A mutation caused a decrease in the activity towards *M. luteus* by over 100-fold, whereas mutating Asp90 to Ala only decreased the activity by 11-fold. The double mutant caused the highest loss of activity with a decrease of about 300-fold compared to the native variant, rendering the enzyme almost completely inactive.

In a molecular dynamics study, Hirakawa et al. [11] recently found indications for the existence of a water molecule positioned between the $Asp97^{O\delta2}$ atom and the C1 carbon of the NAG molecule located in site D and,

hence, proposed that Asp97 (Asp101 cod numbering) was indeed the second catalytic residue in GEWL. In the gLYS ligand-bound structure (MolA), we clearly found electron density corresponding to a water molecule (HOH 79) located between Asp101$^{O\delta 2}$ (2.71 Å) and the C1 atom of the modeled NAG in site D (3.6 Å) (Figs. 3 and 5). The water molecule is also only 2.6 Å from Asp90 (MolA), and the possibility that the water molecule between Asp101 and Asp90 may act as the nucleophile in catalysis is therefore not unlikely. Hence, we propose that the catalytic mechanism is similar to what is suggested for, e.g., T4 phage lysozyme [38], where the nucleophile is a water molecule positioned between Asp20 and Thr26. Our hypothesis is that both aspartic acids are required in order to ensure efficient catalysis, but Asp101 is the most critical. This is verified by the mutational studies, but it can also be rationalized by the fact that the Asp101 residue is most extensively attached to the rest of the molecule by hydrogen bonds, hence making a preformed binding site for the water molecule. The role of both Asp101 and Asp90 would be to ensure that the water molecule is in a proper location to perform the nucleophilic attack during catalysis (Figs. 3, 5 and 6). The water molecule is not well defined in electron density in the gLYS molecules with no or reduced occupancy of NAG (molecules B–D). However, it is present in MolD of the ligand-free structure (3.12 Å to Asp101 and 3.00 Å to Asp90). This could indicate that binding of substrate and the nucleophile occurs in a concerted action. Mutating Asp90 to Ala will disrupt the hydrogen bond to the nucleophilic water, resulting in a less stable positioning of the attacking water molecule, thereby causing a decrease in hydrolytic activity. It would, however, not be crucial since the hydrolytic water molecule could still be positioned at the nucleophile site through hydrogen bonds to Asp101. The Asp101 mutation would, on the other hand, disturb the binding site to a greater extent and will remove several H-bonds to both main-chain and side-chain atoms of Arg103 possibly causing a more flexible loop region affecting the B–C sites. This may prevent proper/efficient substrate and/or nucleophile binding. Upon binding of NAG in gLYS, the position of Asp101 is almost identical to the unbound state, indicating that the favorable interactions with Arg103 are important for the integrity of the loop comprising the B–C sites. The weaker electron density and higher b-factors of Asp90 suggest higher flexibility of this residue than Asp101. Asp90, therefore, does not possess the same capacity as Asp101 for keeping the hydrolytic water in a correct position for the nucleophilic attack. Our mutational data suggests that Asp101 and Asp90 are involved in the positioning of the catalytic water and the stabilization of the transition state. However, they are not directly involved in the catalytic reaction, e.g., by forming a covalent bond as suggested for the HEWL mechanism [39], since lytic activity is also observed for the double mutant, even though the catalytic rate is extremely low. We attribute this weak activity to the accidental presence of water at the correct position for completing the catalysis.

Table 3 Muramidase activity of cod g-type lysozyme (gLYS) D90A, D101A and D90A/D101A mutants relative to native gLYS (set to 100%)

Name	Relative activity (%)	Fold decrease
Native	100	–
D90A	8.60	11
D101A	0.89	112
D90A/D101A	0.32	312

GEWL has previously been suggested to be an inverting enzyme [11, 38], and the nucleophile completing the hydrolytic reaction was suggested to be a water molecule. Furthermore, Kuroki et al. [38] stated that an inverting glycosidase is typically characterized by a pair of carboxylates about 7–9.5 Å apart. Our observations regarding gLYS support the previous suggestions of g-type lysozymes being inverting enzymes, where the distance from Asp101 to Glu73 is about 9 Å and the water molecule positioned between Asp101 and Asp90 completes the hydrolytic reaction.

Hirakawa et al. [11] proposed that there would be too little space in the prime side binding cleft to accommodate an acceptor molecule and therefore excluded a transglycosylation reaction by GEWL. Kuroki et al. [38] showed that the T4 lysozyme could shift the mechanism from being an inverting enzyme (wild type) to become a retaining enzyme by mutating Thr26 to His. This mutation, in addition to changing the hydrolyzing mechanism, probably leads to a covalent adduct formed between the His side chain and C1-carbon of the substrate molecule actually favoring a transglycosylation reaction. They reasoned that a preference for transglycosylation over hydrolysis could only be explained if His-26 formed a covalent adduct with the C1 atom. Transglycosylation is not expected to occur in gLYS since the E–G sites in gLYS and GEWL are comparable, hence too small to accommodate an acceptor molecule, and the hydrolytic mechanism in gLYS, as suggested in Fig. 6, does not include a covalent intermediate.

Stabilizing structural features

Enzymatic activity at low temperatures has frequently been suggested to be correlated with a more flexible structure at the expense of the stability of the enzyme (for review, see, e.g., [40, 41]). Stability is frequently related to hydrogen

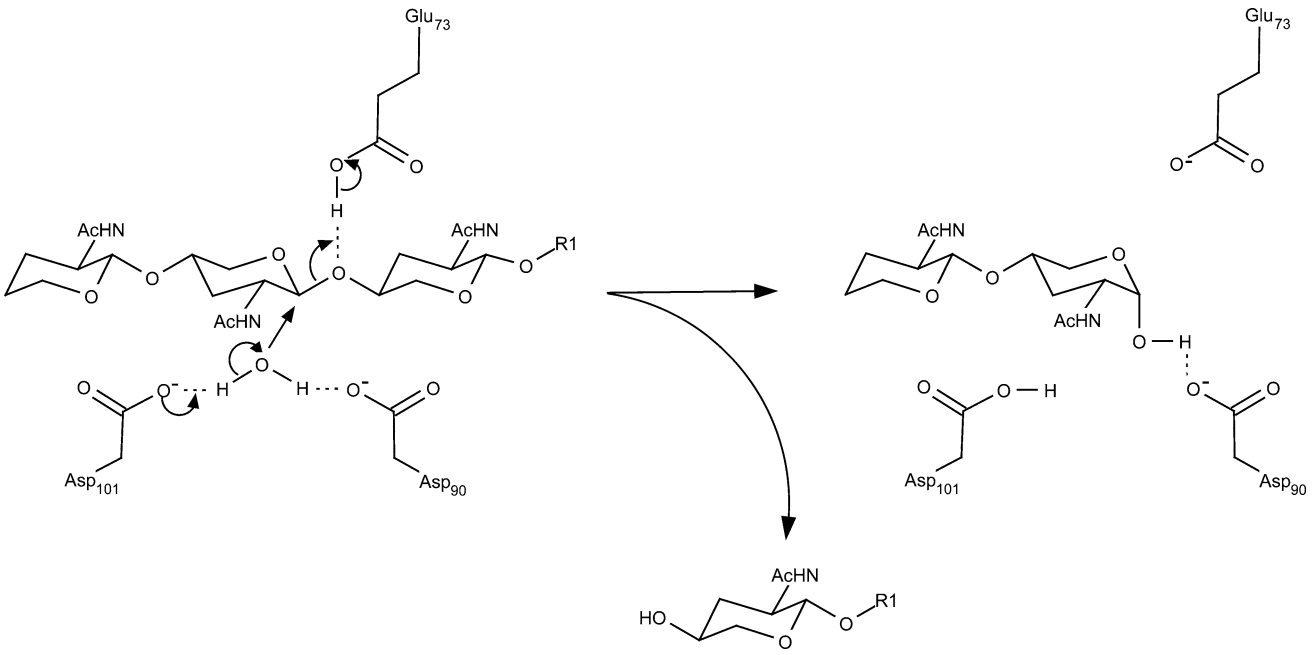

Fig. 6 Schematic illustration of the proposed catalytic mechanism of g-type lysozyme. Substrate [here $(NAG)_3$] and the catalytic water molecule binds in a concerted action. Glu73 initiates the catalytic reaction by donating a hydrogen atom to the 1–4 glycosidic bond. Asp90 and Asp101 position the catalytic water for nucleophilic attack and stabilize the transition state

bonds, ionic interactions and disulphide bonds. The number of hydrogen bonds in gLYS and avian lysozymes does not differ significantly (data not shown), but differences in the salt bridge pattern exist. Both gLYS and avian lysozymes have two four-residue salt-bridge networks (Fig. 7); Asp172–Arg72–Asp176–Arg180 and **Asp24–Arg75**–Asp41–**Arg44** in gLYS and Asp168–Arg72–Asp172–Arg176 and **Glu24–His75**–Asp41–**Lys78** in avian lysozymes (differences in residues in gLYS and avian lysozymes are marked in bold), probably being involved in the stabilization of the prime side of the substrate binding sites. In gLYS, the two networks are connected by an additional ionic interaction between Asp41 and Arg180, hence forming one eight-residue network. Both gLYS and avian lysozymes have one three-residue network each, but the residues involved are different; Lys58–Asp118–His59 and Lys54–Glu110–Lys58, respectively. In addition, GEWL has seven other ionic interactions (l54L/153L), black swan lysozyme has five (1LSP/1GBS), while gLYS has only three to four.

A structural feature common for the avian g-type lysozymes are conserved disulfide bridges located in the N-terminal region (Cys4–Cys60 and Cys18–Cys29). These have recently been proven to be important for structural stability [42]. This structural feature is missing in the gLYS structure (as well as in most other reported fish g-type lysozyme sequences). In general, the very N-terminal region in gLYS (in all four lysozyme molecules found in the asymmetric unit) possess a high flexibility as shown by high structural temperature factors and absence of well-defined electron density. The lack of the disulfide bridges clearly adds to this flexibility as no other stabilizing features such as salt bridges are present to replace the absence of the disulfide bonds. The regions stabilized by the second disulfide bridge (Cys18–Cys29) in the avian lysozymes are to some extent replaced by a hydrogen bond network formed by $Tyr28^{O\eta}$–$Thr36^{O\gamma 1}$ and $Tyr28^{O}$–$Ser21^{O\gamma}$ in cod. The corresponding residues in GEWL, Tyr28, Lys36, and Ala21 are unable to form a similar network. Kawamura et al. [42] showed that mutating the cysteine residues to serine did not affect protein folding nor was it critical for enzyme activity. Since Atlantic cod lives at temperatures close to the freezing point, a valid assumption would be that the lesser number of salt bridges and the lack of the disulfide bonds may create a more flexible structure necessary for the enzyme to perform its function at low temperatures. However, the lack of disulfide bridges is a common trait of most fish g-type lysozymes, even for more temperate fish species. An intracellular role of fish g-type lysozymes have been proposed in general due to a lack of both disulfide bonds and secretion signal peptides [3]. Since then, g-type lysozymes containing possible secretion signal peptides has been detected in both salmon and cod, in addition to zebrafish [3, 15, 20], suggesting that fish possess g-type lysozymes with an extracellular function as well.

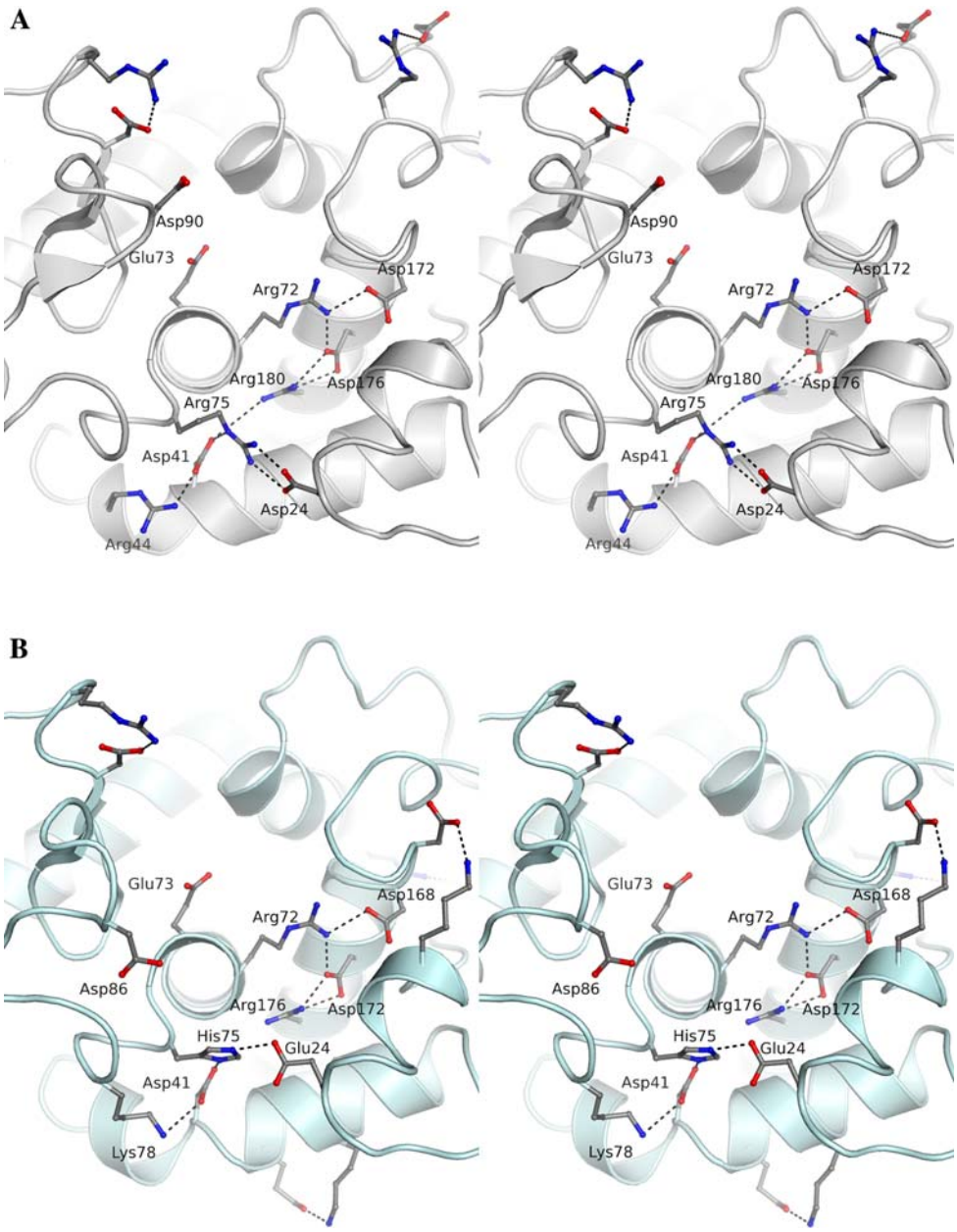

Fig. 7 Salt-bridge network in cod and goose lysozymes. **a** Two-four-residue salt-bridge networks in cod lysozyme Asp172–Arg72–Asp176–Arg180 and Asp24–Arg75–Asp41–Arg44 is connected by the ionic interaction between Asp41 and Arg180 forming an eight-residue network. **b** Two-four-residue salt-bridge networks Asp168–Arg72–Asp172–Arg176 and Glu24–His75–Asp41–Lys78 in goose lysozyme

Molecular surface

The electrostatic surface shows that gLYS have a less cationic character compared to GEWL (Fig. 4a–e), and especially the substrate binding cleft in gLYS have a more anionic character. It has been shown that gLYS possess optimum muramidase activity at or below pH 4.8 with a sharp decrease in activity above pH 4.8 [20], while that of GEWL have an optimum around pH 5.5 [43]. A higher net positive surface charge has previously been attributed to a higher pH optimum of activity in both c-type and g-type lysozymes [43, 44]. The substrate employed in all these studies, *M. luteus* cells, is a polymeric substrate of high molecular weight as well as being negatively charged. The less cationic surface and more electronegative binding cleft of gLYS probably reflects the differences in the observed pH activity profiles of cod and avian g-type lysozymes.

Acknowledgments The present study was supported by the Research Council of Norway (NRC), project number 158952 and the National Functional Genomics Program (FUGE). Provision of beamtime at the Macromolecular Crystallography Beamlines (BL14) of Berliner Elektronenspeicherring (BESSY) and the Swiss-Norwegian Beamlines (SNBL) at European Synchrotron Radiation Facility (ESRF) is gratefully acknowledged.

References

1. Ellis AE (2001) Innate host defense mechanisms of fish against viruses and bacteria. Dev Comp Immunol 25:827–839
2. Saurabh S, Sahoo PK (2008) Lysozyme: an important defence molecule of fish innate immune system. Aquac Res 39:223–239
3. Irwin DM, Gong Z (2003) Molecular evolution of vertebrate goose-type lysozyme genes. J Mol Evol 56:234–242
4. Prager EM, Jollès P (1996) Animal lysozymes c and g: an overview. In: Jollès P (ed) Lysozymes: model enzymes in biochemistry and biology. Birkhäuser, Basel, pp 9–31
5. Jolles P, Jolles J (1984) What's new in lysozyme research? Always a model system, today as yesterday. Mol Cell Biochem 63:165–189
6. Grutter MG, Weaver LH, Matthews BW (1983) Goose lysozyme structure: an evolutionary link between hen and bacteriophage lysozymes? Nature 303:828–831
7. Weaver LH, Grutter MG, Matthews BW (1995) The refined structures of goose lysozyme and its complex with a bound trisaccharide show that the "goose-type" lysozymes lack a catalytic aspartate residue. J Mol Biol 245:54–68
8. Honda Y, Fukamizo T (1998) Substrate binding subsites of chitinase from barley seeds and lysozyme from goose egg white. Biochim Biophys Acta 1388:53–65
9. Strynadka NC, James MN (1991) Lysozyme revisited: crystallographic evidence for distortion of an *N*-acetylmuramic acid residue bound in site D. J Mol Biol 220:401–424
10. Song H, Inaka K, Maenaka K, Matsushima M (1994) Structural changes of active site cleft and different saccharide binding modes in human lysozyme co-crystallized with hexa-*N*-acetylchitohexaose at pH 4.0. J Mol Biol 244:522–540
11. Hirakawa H, Ochi A, Kawahara Y, Kawamura S, Torikata T, Kuhara S (2008) Catalytic reaction mechanism of goose egg-white lysozyme by molecular modelling of enzyme–substrate complex. J Biochem (Tokyo) 144:753–761
12. Canfield RE, McMurry S (1967) Purification and characterization of a lysozyme from goose egg white. Biochem Biophys Res Commun 26:38–42
13. Dianoux AC, Jolles P (1967) Study of a lysozyme poor in cystine and tryptophan: the lysozyme of goose egg white. Biochim Biophys Acta 133:472–479
14. Hikima J, Minagawa S, Hirono I, Aoki T (2001) Molecular cloning, expression and evolution of the Japanese flounder goose-type lysozyme gene, and the lytic activity of its recombinant protein. Biochim Biophys Acta 1520:35–44
15. Kyomuhendo P, Myrnes B, Nilsen IW (2007) A cold-active salmon goose-type lysozyme with high heat tolerance. Cell Mol Life Sci 64:2841–2847
16. Savan R, Aman A, Sakai M (2003) Molecular cloning of G type lysozyme cDNA in common carp (*Cyprinus carpio* L.). Fish Shellfish Immunol 15:263–268
17. Sun BJ, Wang GL, Xie HX, Gao Q, Nie P (2006) Gene structure of goose-type lysozyme in the mandarin fish *Siniperca chuatsi* with analysis on the lytic activity of its recombinant in *Escherichia coli*. Aquaculture 252:106–113
18. Yin ZX, He JG, Deng WX, Chan SM (2003) Molecular cloning, expression of orange-spotted grouper goose-type lysozyme cDNA, and lytic activity of its recombinant protein. Dis Aquat Organ 55:117–123
19. Zheng W, Tian C, Chen X (2007) Molecular characterization of goose-type lysozyme homologue of large yellow croaker and its involvement in immune response induced by trivalent bacterial vaccine as an acute-phase protein. Immunol Lett 113:107–116
20. Larsen AN, Solstad T, Svineng G, Seppola M, Jorgensen TØ (2009) Molecular characterisation of a goose-type lysozyme gene in Atlantic cod (*Gadus morhua* L.). Fish Shellfish Immunol 26:122–132
21. Karlsen S, Hough E, Rao ZH, Isaacs NW (1996) Structure of a bulgecin-inhibited g-type lysozyme from the egg white of the Australian black swan. A comparison of the binding of bulgecin to three muramidases. Acta Crystallogr D Biol Crystallogr 52:105–114
22. Rao Z, Esnouf R, Isaacs N, Stuart D (1995) A strategy for rapid and effective refinement applied to black swan lysozyme. Acta Crystallogr D Biol Crystallogr 51:331–336
23. Kawamura S, Ohno K, Ohkuma M, Chijiiwa Y, Torikata T (2006) Experimental verification of the crucial roles of Glu73 in the catalytic activity and structural stability of goose type lysozyme. J Biochem (Tokyo) 140:75–85
24. Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J Appl Crystallogr 26:795–800
25. Collaborative Computational Project, N (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50:760–763
26. Vagin A, Teplyakov A (1997) MOLREP: an automated program for molecular replacement. J Appl Crystallogr 30:1022–1025
27. Perrakis A, Morris R, Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. Nat Struct Biol 6:458–463
28. Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr A Found Crystallogr 47:110–119
29. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53:240–255
30. Holm L, Park J (2000) DaliLite workbench for protein structure comparison. Bioinformatics 16:566–567
31. Rocchia W, Alexov E, Honig B (2001) Extending the applicability of the nonlinear Poisson–Boltzmann equation: multiple dielectric constants and multivalent ions. J Phys Chem B 105:6507–6514
32. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. J Comput Chem 26:1668–1688
33. Mcdonald IK, Thornton JM (1994) Satisfying hydrogen-bonding potential in proteins. J Mol Biol 238:777–793
34. Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15:305–308
35. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. Protein Eng 8:127–134
36. Fukamizo T, Torikata T, Nagayama T, Minematsu T, Hayashi K (1983) Enzymatic activity of avian egg-white lysozymes. J Biochem (Tokyo) 94:115–122
37. Maenaka K, Matsushima M, Song H, Sunada F, Watanabe K, Kumagai I (1995) Dissection of protein–carbohydrate interactions in mutant hen egg-white lysozyme complexes and their hydrolytic activity. J Mol Biol 247:281–293
38. Kuroki R, Weaver LH, Matthews BW (1999) Structural basis of the conversion of T4 lysozyme into a transglycosidase by reengineering the active site. Proc Natl Acad Sci USA 96:8949–8954
39. Vocadlo DJ, Davies GJ, Laine R, Withers SG (2001) Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. Nature 412:835–838
40. Siddiqui KS, Cavicchioli R (2006) Cold-adapted enzymes. Annu Rev Biochem 75:403–433

41. Smalas AO, Leiros HK, Os V, Willassen NP (2000) Cold adapted enzymes. Biotechnol Annu Rev 6:1–57
42. Kawamura S, Ohkuma M, Chijiiwa Y, Kohno D, Nakagawa H, Hirakawa H, Kuhara S, Torikata T (2008) Role of disulfide bonds in goose-type lysozyme. FEBS J 275:2818–2830
43. Pooart J, Torikata T, Araki T (2005) Enzymatic properties of rhea lysozyme. Biosci Biotechnol Biochem 69:103–112
44. Muraki M, Morikawa M, Jigami Y, Tanaka H (1988) Engineering of human lysozyme as a polyelectrolyte by the alteration of molecular surface charge. Protein Eng 2:49–54