THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 44, Issue of November 2, pp. 41377–41382, 2001

*Printed in U.S.A.*

# Structure of the Induced Antibacterial Protein from Tasar Silkworm, *Antheraea mylitta*

IMPLICATIONS TO MOLECULAR EVOLUTION*

Received for publication, May 22, 2001, and in revised form, August 22, 2001
Published, JBC Papers in Press, August 24, 2001, DOI 10.1074/jbc.M104674200

**Deepti Jain‡§¶, Deepak T. Nair‡§¶, G. Jawahar Swaminathan‡§, E. G. Abraham‖, J. Nagaraju‖**, and Dinakar M. Salunke‡ ‡‡**

*From the ‡National Institute of Immunology, New Delhi 110 067, the ‖Seribiotech Research Laboratory, Bangalore 560 086, and the **Centre for DNA Fingerprinting and Diagnostics, Hyderabad 500 076, India*

**The crystal structure of an antibacterial protein of immune origin (TSWAB), purified from tasar silkworm (*Antheraea mylitta)* larvae after induction by *Escherichia coli* infection, has been determined. This is the first insect lysozyme structure and represents induced lysozymes of innate immunity. The core structure of TSWAB is similar to c-type lysozymes and α-lactalbumins. However, TSWAB shows significant differences with respect to the other two proteins in the exposed loop regions. The catalytic residues in TSWAB are conserved with respect to the chicken lysozyme, indicating a common mechanism of action. However, differences in the noncatalytic residues in the substrate binding groove imply subtle differences in the specificity and the level of activity. Thus, conformational differences between TSWAB and chicken lysozyme exist, whereas functional mechanisms appear to be similar. On the other hand, α-lactalbumins and c-type lysozymes exhibit drastically different functions with conserved molecular conformation. It is evident that a common molecular scaffold is exploited in the three enzymes for apparently different physiological roles. It can be inferred on the basis of the structure-function comparison of these three proteins having common phylogenetic origin that the conformational changes in a protein are minimal during rapid evolution as compared with those in the normal course of evolution.**

The insect immune defense is activated by the invasion of microbes and acts through both cellular and humoral responses. The humoral immune response in insects shares some fundamental characteristics with innate immune response of vertebrates in the sense that it apparently lacks specificity and memory, and nevertheless functions with great efficiency against microbial pathogens. This arm of the insect immune response manifests itself by instantaneous overexpression of an array of potent antibacterial proteins and peptides at the time of bacterial infection. These proteins and peptides provide the first line of defense against pathogens. A large number of these peptides and proteins have been isolated, identified, and characterized biochemically (1–5). They include cecropins (6), defensins (7), attacins (8), proline-rich peptides (9), and lysozymes (10, 11).

As a model for understanding the nature and biology of the innate immune response, insect immunity has come into renewed focus (12, 13). Structural studies on the proteins and peptides of the insect immune system are therefore being increasingly addressed. Elaborate spectroscopic studies have been carried out on cecropins and sapecins (14–16). The NMR structures of the insect immune peptides such as thanatin, drosomycin, and insect defensin have been elucidated (17–19). In addition, the crystal structure of hemolin, a protein belonging to the immunoglobulin superfamily and involved in insect immunity, has been recently determined (20). Here we report the crystal structure of an antibacterial protein from tasar silkworm (*Antheraea mylitta)* (TSWAB). The expression of TSWAB was induced after *Escherichia coli* infection of the larvae, leading to the secretion of large quantities of the protein in the hemolymph. Structural fold of this protein resembles c-type lysozymes and α-lactalbumins, and it belongs to the functional class of chicken lysozymes. The comparison of TSWAB structure with respect to the other c-type lysozymes and α-lactalbumins provides structural implications of rapid *versus* slow molecular evolution.

## EXPERIMENTAL PROCEDURES

*Crystallization and Data Collection*—Purification of the tasar silkworm protein, induced on *E. coli* infection, was carried out by adopting the procedures described previously (21, 22). The N-terminal 32 residues were obtained by protein sequencing using an ABI 270 automated protein sequencer with a 120A analyzer. The crystals of TSWAB were grown in 0.1 M acetate, pH 5.5, with 26.5% polyethylene glycol (8 kDa) as precipitant using hanging drop vapor diffusion method at room temperature. Small wedge-shaped crystals with dimensions of ~0.1 × 0.25 × 0.25 mm grew to their final size in about 2 months. The x-ray diffraction intensity data were recorded using an imaging plate detector system (Marresearch) with the $CuK_\alpha$ radiation from RIGAKU rotating anode x-ray generator (40 kV and 70 mA). The crystals diffracted up to 2.4-Å resolution and were stable in the x-ray beam. X-ray intensity data were processed using DENZO (23). The data completion was 97% at 2.4-Å resolution. The intensity data statistics and the crystal data are given in Table I. Assuming one molecule of the protein in the asymmetric unit, the $V_m$ equals 2.03 Å/Da, which is within the expected range for normal proteins (24). The solvent content of the crystals was calculated to be 39.5%.

*Structure Determination and Refinement*—All molecular replacement calculations were carried out using AMoRe (25). Through N-terminal sequencing, the first 32 amino acids of the sequence were determined. Using the BLAST and ClustalW programs, it was found

* This work was supported in part by a Department of Science and Technology extramural grant (to D. M. S.) and by grants to the National Institute of Immunology from the Department of Biotechnology and to the Seribiotech Research Laboratory from the Central Silk Board of the Government of India. 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.

§ Recipient of a fellowship from the Council of Scientific and Industrial Research (India).

¶ Contributed equally to this work.

‡‡ To whom correspondence should be addressed: Structural Biology Unit, National Inst. of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India. Tel.: 91-11-616-7623 (ext. 234); Fax: 91-11-616-2125; E-mail: dinakar@nii.res.in.

TABLE I
*Data collection and refinement statistics*

| | |
|---|---|
| Space group | $C222_1$ |
| Cell constants (Å) | $a = 47.9$, $b = 51.7$, $c = 94.6$ |
| Solvent content (%) | 39.4 |
| Maximum resolution (Å) | 2.4 |
| Number of reflections | 4807 |
| $R_{merge}$ (%) | 8.3 |
| Completeness (%) | 98.7 |
| Average ($I$)/(Sig$I$) | 23.8 |
| No. of protein atoms | 945 |
| No. of solvent atoms | 56 |
| $R_{crys}$ (%) | 23.8 |
| $R_{free}$ (%) | 27.0 |
| Refinement range (Å) | 100–2.4 |
| r.m.s. deviation bond length (Å) | 0.008 |
| r.m.s. deviation bond angles (°) | 1.70 |
| Average B value of protein atoms (Å$^2$) | 20.4 |
| Average B value of waters atoms (Å$^2$) | 23.3 |

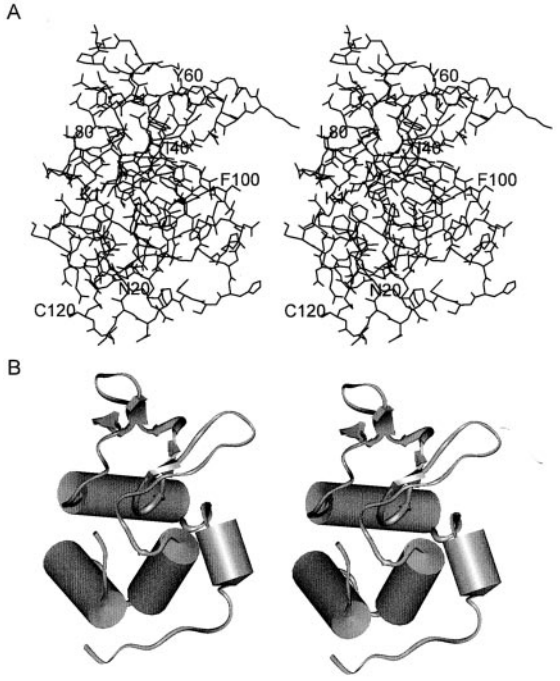


FIG. 1. **Stereo diagram of the tertiary structure of TSWAB.** *A*, atomic structure with every 10th residue numbered. *Single-letter code* is used for amino acids. *B*, Secondary structure rendering representation of the TSWAB highlighting various secondary structural elements.

that there is maximum sequence homology with the N-terminal region of trout (*Onthorhyncus mykiss*) lysozyme (38% sequence identity). The crystal structure of the trout lysozyme (1LMN) was therefore used as a probe model for rotation/translation function calculations between 8- and 4-Å resolution. Other lysozyme models whose N-terminal sequences showed good homology were also used. The trout lysozyme model gave better correlation factor (45%) and $R$-factor (42%) than the other models. The 1LMN model was properly oriented in the unit cell, and the crystal packing was examined to make sure that there are no steric clashes or large voids between symmetry-related molecules. It was subsequently subjected to rigid body refinement in CNS (26) to refine orientation and position of the starting model in the unit cell. The individual atoms were then refined by several cycles of conventional positional refinement with overall B values. Higher resolution data, up to 2.4 Å, were added in a stepwise fashion. Both conventional $R$-factor ($R_{cryst}$) and the free $R$-value ($R_{free}$) (using 10% data) were used to monitor the progress of refinement (27). The electron density map was computed using CNS and was displayed with the help of O (28) on Octane (Silicon Graphics Inc.). This was followed by iterative rebuilding of the model on the basis of $2F_o - F_c$ as well as $F_o - F_c$ maps. Using the electron density maps the sequence of 1LMN was gradually changed to that of TSWAB. Some loops also had to be extensively remodeled. The entire sequence of the TSWAB was obtained only after complete refinement of the structure. Of course, the availability of the homologous

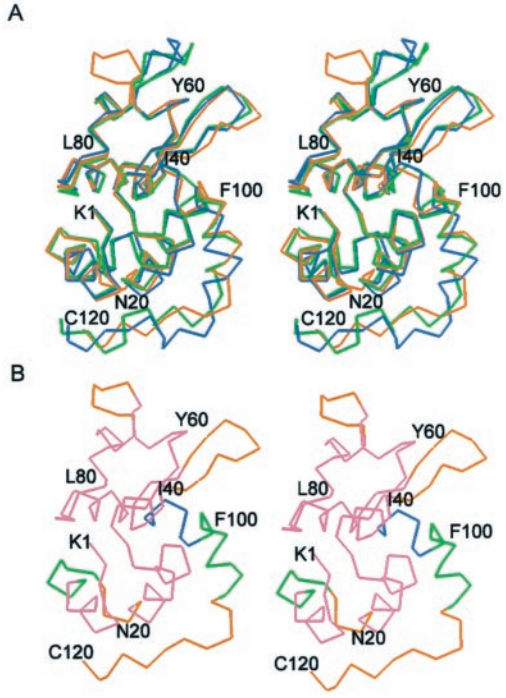


FIG. 2. **Stereo diagram showing the structure of TSWAB and its superimposition with a c-type lysozyme and an α-lactalbumin.** *A*, structural superimposition of TSWAB (*orange*) with chicken lysozyme (Protein Data Bank code 1HEL) (*green*) and human α-lactalbumin (Protein Data Bank code 1HML) (*blue*). The Cα traces of proteins are shown as *thick lines*. *Single-letter code* is used for amino acids. *B*, the Cα trace of TSWAB, colored on the basis of the structural similarity with chicken lysozyme (*green*), human α-lactalbumin (*blue*), unique to TSWAB (*orange*) and the central core region (*magenta*) that is structurally conserved in the three proteins.

sequences from the other lepidopteran insects facilitated incorporation of correct amino acids within the electron density map. Water molecules were included in the model using a contour of $2\sigma$ in $F_o - F_c$ and $1\sigma$ in $2F_o - F_c$ electron density maps and if they were within 3.5 Å from one or more nitrogen or oxygen atoms of the protein or other water molecules.

*Crystal Structure Validation and Analysis*—The deviation from ideal geometry of the refined model was analyzed using XPLOR and PROCHECK (29) in CCP4. The structural alignments were made using the HOMOLOGY module of the InsightII (Molecular Simulations Inc.). The molecular volumes were calculated using GRASP (30). The insect lysozyme sequences were obtained from SwissProt through BLAST server, and their comparative analysis was carried out using ClustalW (31).

## RESULTS

*Overall Structure*—The fully refined structure of TSWAB shows that it is 120 amino acids long. The final model (Fig. 1*A*) includes 961 protein atoms, with 58 atoms corresponding to water molecules. The $\phi$-$\psi$ plot for the backbone torsion angles shows that a total of 83.5% of the non-glycine and non-proline residues have their backbone torsion angles in the most favorable region of the plot, 11.9% in additionally allowed regions, and 1.8% in the generously allowed regions.

TSWAB is a globular protein with approximate dimensions being 69 × 47 × 50 Å. The molecule is bi-lobed, and the structure is predominantly α-helical with two small β-sheets and a number of loops of varying lengths. The helical regions lie in one lobe of the protein, which extend from Arg$^5$ to Gln$^{15}$ (α1), Arg$^{23}$ to Glu$^{32}$ (α2), Ile$^{84}$ to Thr$^{98}$ (α3), and Trp$^{106}$ to His$^{109}$ (α4). The two short β-sheets lie in the other lobe. The first β-sheet is antiparallel, composed of strands from Asp$^{50}$ to Gly$^{52}$ and Ile$^{56}$ to Asp$^{58}$, and the second is a parallel β-sheet composed of strands from Cys$^{62}$ to Lys$^{64}$ and Val$^{74}$ to Cys$^{76}$. Fig. 1*B*

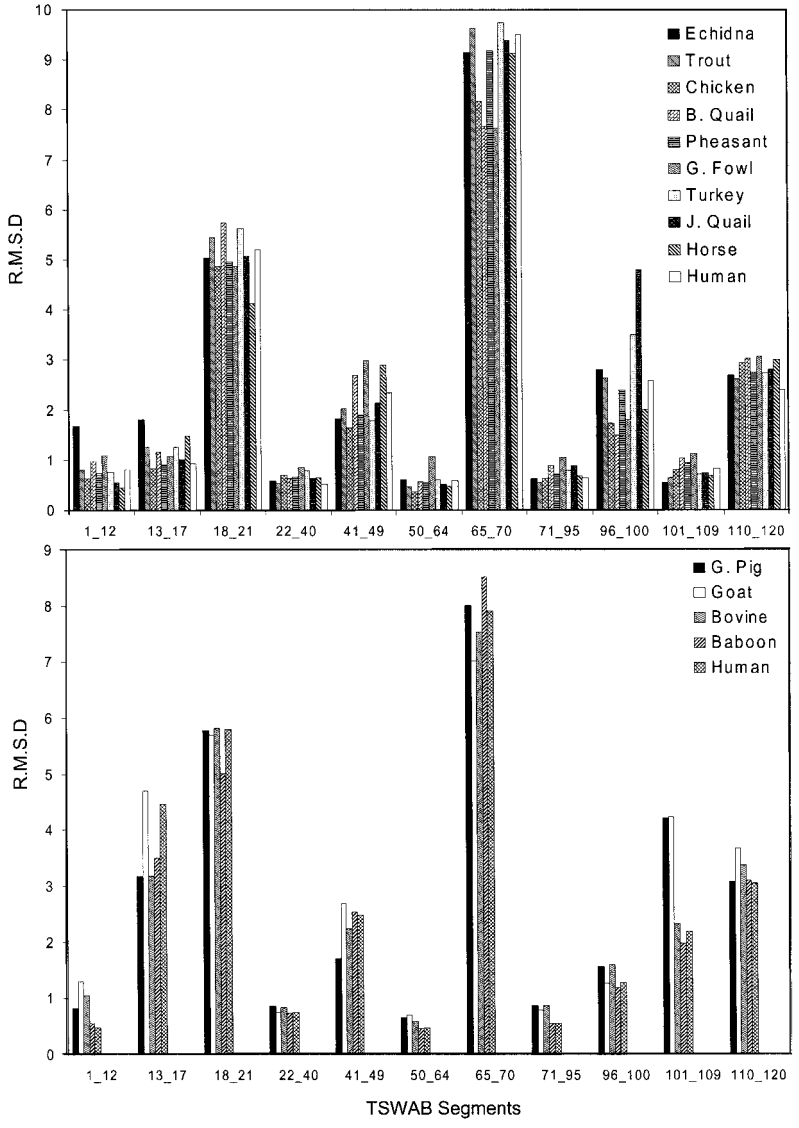


FIG. 3. **Comparison of r.m.s. deviations for different structural segments of TSWAB with c-type lysozymes (*top*) and α-lactalbumins (*bottom*).**

depicts the stereo view of the arrangement of the secondary structural elements of TSWAB. There are four pairs of disulfide bonds present in the structure of TSWAB. The disulfide bonds are present between $Cys^{6}$ and $Cys^{119}$, $Cys^{27}$ and $Cys^{110}$, $Cys^{62}$ and $Cys^{76}$, and $Cys^{72}$ and $Cys^{90}$. All the cysteines present are involved in disulfide bond formation.

*Structural Comparison with Chicken Lysozyme Family*—The N-terminal sequencing of TSWAB was done prior to structure solution, and it gave the first 32 residues of the sequence. Rest of the protein sequence was obtained from the comparative analysis of the sequences of other insect lysozymes aided by the electron density map as the refinement progressed. A BLAST search (32) using the entire sequence revealed significant homology with c-type lysozymes and α-lactalbumins, which have been classified in a common family of structural folds in the SCOP data base (33). The structural alignment of TSWAB with these two proteins (Fig. 2*A*) revealed that it belongs to the lysozyme-like fold, with the core of the molecule being conserved in terms of the number and relative orientations of secondary structural elements. The regions of TSWAB that align well with the chicken lysozyme include residues 1–17, 22–40, 50–64, 71–95, and 101–109, and those that align well with human α-lactalbumin include residues 1–12, 22–40, 50–64, and 71–100. These involve three of the four α-helices and both the β-sheets seen in the TSWAB protein. The disulfide-bonding pattern in TSWAB is consistent with those of the other two proteins. Fig. 2*B* shows the structure of TSWAB colored on the basis of regions that are similar to one or the other of the two proteins to highlight the conserved and variable regions and to define the core of the molecule with respect to chicken lysozyme and human α-lactalbumin.

The above comparison was also extended to all the c-type lysozymes and α-lactalbumins whose structures were available in the protein data bank (34). The core structural features of TSWAB are conserved. However, significant differences were observed in the loop regions, in terms of length, sequence, and conformation, between the three proteins. The structure of TSWAB was divided into 11 segments on the basis of above comparison, and the r.m.s.[1] deviation of each of these segments was compared with the corresponding segments in 10 different c-type lysozymes and 5 different α-lactalbumins. Fig. 3 shows the r.m.s. deviation plotted with respect to each of the 11 segments of TSWAB. As is evident from this figure, five segments of c-type lysozymes show r.m.s. deviation greater than 2 Å with respect to TSWAB. These include residues 18–21, 41–49, 65–70, 96–100, and 110–120. In the case of α-lactalbumins, six segments differ in structure as compared with TSWAB.

[1] The abbreviation used is: r.m.s., root mean square.

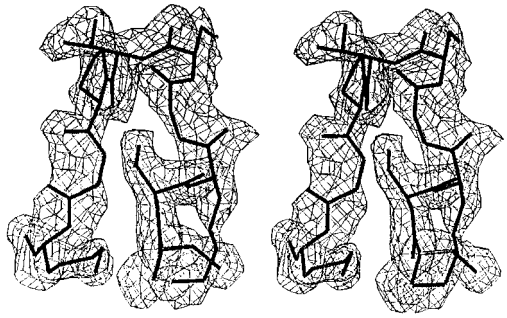

FIG. 4. **Stereo diagram representing section of the $2F_o - F_c$ electron density map for the loop region between 65 and 70 seen at 2.4-Å resolution, contoured at 1$\sigma$.**

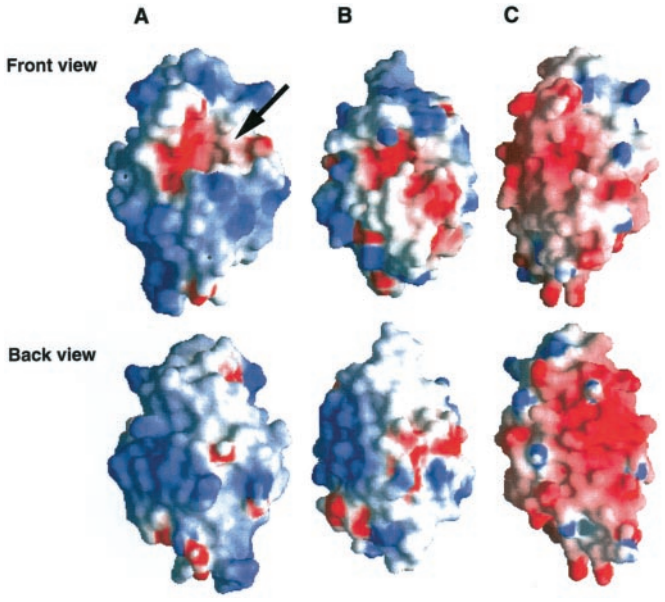


FIG. 5. **Comparison of the surface features and active site geometry of the TSWAB, chicken lysozyme, and human $\alpha$-lactalbumin.** *A*, molecular surface of hen egg white lysozyme; *B*, TSWAB; *C*, human $\alpha$-lactalbumin with front and back view shown. The position of the well defined active site in the case of hen egg white lysozyme has been marked with an *arrow*. The surface is colored according to electrostatic potential, ranging from *blue* (the most positive region) to *red* (the most negative region). Figure was produced using the program GRASP (30).

These comprise residues 13–17, 18–21, 41–49, 65–70, 101–109, and 110–120. The rest of the structure can be defined as the core of the TSWAB molecule, which is largely conserved with respect to the other two proteins. In addition, the plot also indicates that the extent of structural variations of all the c-type lysozymes and $\alpha$-lactalbumins is similar with respect to TSWAB.

The region of TSWAB that is different from lactalbumins but not from c-type lysozymes incorporates residues 13–17. The corresponding segment of lactalbumins includes residues 13–15. This segment of TSWAB forms a part of $\alpha$1, but this helix in lactalbumins is much shorter; as a result, the corresponding segment takes altogether a different conformation. For c-type lysozymes the r.m.s. deviations are below 2 Å for this region.

The loop defined by residues 18–21 of TSWAB with the sequence DENL also shows significant conformational differences with respect to the corresponding loops in all lysozymes and $\alpha$-lactalbumins defined by residues 18–24 and 16–22, respectively. Thus, this loop is shorter in TSWAB by 3 residues, as compared with the other two proteins. This loop connects the $\alpha$1 and $\alpha$2 helices and interacts with $\alpha$3 in the case of $\alpha$-lactalbumins and c-type lysozymes. However, it shows no interactions in the case of TSWAB, probably because of its smaller size.

Another loop of TSWAB that shows significant differences with respect to the other two proteins is the one defined by residues 41–49 incorporating the sequence ANVNKNGSR. This loop joins one of the two strands of the anti-parallel $\beta$-sheet and a loop. The corresponding residues in c-type lysozymes and $\alpha$-lactalbumins are 44–51 and 42–48, respectively. This is the only loop that is slightly longer in TSWAB as compared with the other two proteins.

The conformation of the loop defined by residues 65–70, incorporating the sequence GSTPGK, is particularly striking because it is substantially different with respect to both c-type lysozymes and $\alpha$-lactalbumins. The $2F_o - F_c$ map for this loop in TSWAB is shown in Fig. 4. The loop connects the two $\beta$-strands of the parallel $\beta$-sheet of TSWAB. The analogous loop is 8 residues in length in all lysozymes and $\alpha$-lactalbumins and comprises residues 67–74 and 64–71, respectively. In the case of c-type lysozymes and $\alpha$-lactalbumins, this loop is oriented toward the active site, whereas, in the case of TSWAB, it is oriented away from the active site.

The structural differences are also seen in the loop defined by residues 96–100, of TSWAB integrating the sequence KRTKF. This loop connects $\alpha$3 and a $\beta$-turn. The corresponding residues in lysozymes include 100–105. However, this loop shows r.m.s. deviation of less than 2 Å when compared with $\alpha$-lactalbumins indicating similar conformation. Instead, the segment 101–109 of TSWAB, which incorporates $\alpha$4 region, is closer to c-type lysozymes. The corresponding region in $\alpha$-lactalbumins and c-type lysozymes includes 102–110 and 106–114, respectively. In fact, this region also shows maximum differences in conformation within the various $\alpha$-lactalbumins. In the case of guinea pig and goat proteins, the conformation varies significantly with respect to other members as well as TSWAB. The r.m.s. deviations of c-type lysozymes for the corresponding segment lie below 2 Å.

The C-terminal end region, 110–120, of TSWAB is similar neither to c-type lysozymes (115–129) nor to the $\alpha$-lactalbumins (111–122). This region has altogether different conformations in the three cases. However, within various lysozymes and $\alpha$-lactalbumins, this region has similar conformation except for a few residues at C terminus in the case of $\alpha$-lactalbumins.

*Active Site Geometry and Surface Features*—The active site of TSWAB could be defined by comparison with c-type lysozymes. In fact, this region of TSWAB shows maximum similarity with other lysozymes in terms of the nature of the amino acids, the backbone conformation, and the side chain orientations. Through comparison with chicken lysozyme, the active site could be defined by residues $Asn^{31}$, $Glu^{32}$, $Ala^{34}$, $Val^{43}$, $Asp^{50}$, $Gln^{55}$, $Asn^{57}$, $Tyr^{60}$, $Trp^{61}$, $Ile^{94}$, $Arg^{97}$, $Ala^{102}$, $Trp^{103}$, and $His^{109}$ of TSWAB. The residues $Glu^{32}$ and $Asp^{50}$ of chicken lysozyme have been implicated to have direct role in catalysis (35). The structural alignment illustrates that these residues are conserved in TSWAB as well. Comparison of TSWAB with all other c-type lysozymes suggests that, of these 14 residues defining the active site, 8 are largely conserved. These include residues $Glu^{32}$, $Asp^{50}$, $Gln^{55}$, $Asn^{57}$, $Trp^{61}$, $Ile^{94}$, $Ala^{102}$, and $Trp^{103}$. Major difference in the active site residues in terms of charge between the c-type lysozymes and TSWAB occurs at $Arg^{97}$ of TSWAB, which is Asp in most of the other proteins. The residues in place of $Asn^{31}$, $Ala^{34}$, $Val^{43}$, $Tyr^{60}$, and $His^{109}$ of TSWAB are, respectively, Phe, Asn, Asn, Trp, and Arg in a majority of the c-type lysozymes. A comparison of TSWAB with $\alpha$-lactalbumins, on the other hand, shows that only 5 of the 14 residues defining active site are largely conserved. However, the catalytic residues are not conserved; instead of the residues

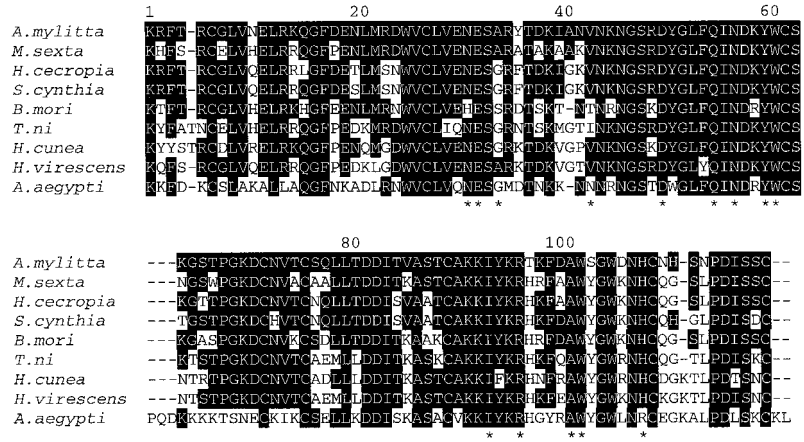


FIG. 6. **Sequence alignment of TSWAB with other insect lysozymes.** Every 20th residue has been labeled in all the three alignments. The residues identical to TSWAB in each case have been highlighted. The residues of TSWAB corresponding to the active site in the case of hen egg white lysozyme have been marked with asterisks.

$Glu^{32}$ and $Asp^{50}$, which have been directly associated with catalysis, $\alpha$-lactalbumins have $Thr^{32}$ and $Glu^{48}$, respectively.

Comparison of the surface electrostatic potentials of the three proteins, TSWAB, chicken lysozyme, and human $\alpha$-lactalbumin, was carried out (Fig. 5). It is clear that the electrostatic potential in the active site region of TSWAB is negative and resembles that of chicken lysozyme. Apart from the active site, the overall surface electrostatic potential in the case of chicken lysozyme is positive. In the case of TSWAB, on the other hand, there are other regions on the surface (in addition to the active site) that show negative electrostatic potential. The charge distribution on the human $\alpha$-lactalbumin surface is distinctly different from the other two proteins, providing it a highly negative electrostatic potential. Additionally, the active site of TSWAB is compact as compared with chicken lysozyme, which is more open and broad, particularly because of the orientation of $Lys^{45}$. The corresponding residue in chicken lysozyme is $Thr^{47}$. The shape of the active site groove of the human $\alpha$-lactalbumin is significantly different compared with the other two proteins.

Fig. 6 shows sequence comparison of TSWAB with the other induced insect lysozymes for which sequences were available. TSWAB shows very high sequence identity (between 70 and 80%) with all of them except *Aedes aegypti*, with which it shows 53% identity. Of the 14 residues defining the active site, 12 are conserved, including the two catalytic residues $Glu^{32}$ and $Asp^{50}$ in all the insect lysozymes of induced origin. The other, conserved active site residues include $Asn^{31}$, $Gln^{55}$, $Asn^{57}$, $Tyr^{60}$, $Trp^{61}$, $Ile^{94}$, $Arg^{97}$, $Ala^{102}$, $Trp^{103}$, and $His^{109}$. Two nonconserved position correspond to $Ala^{34}$, which is Gly in five proteins and Ser in *Bombyx mori*, and $Val^{43}$, which is Val, Thr, Ile, or Asn in other proteins.

## DISCUSSION

The TSWAB is an *E. coli*-induced immune protein expressed as a part of humoral immune response in tasar silkworm from the order Lepidoptera. It is truly a protein among those designed for insect immune defense. The chicken lysozyme, on the other hand, is a constitutive antibacterial protein. However, TSWAB has antibacterial activity comparable with that of the chicken lysozyme against Gram-positive bacteria (21). Therefore, the molecular mechanisms of the action of TSWAB and chicken lysozyme are expected to be similar. On the other hand, $\alpha$-lactalbumins are involved in the regulation of lactose synthase activity. As also observed in many other cases, a common molecular scaffold is used for these three apparently different physiological roles (36, 37). Proteins with similar fold could also be related phylogenetically. $\alpha$-Lactalbumins, c-type lysozymes, and insect lysozymes have been implicated to have a common evolutionary precursor (38).

The physiological roles of TSWAB and chicken lysozyme are different, considering that one is instantaneously induced on bacterial infection and the other acts constitutively. Because the catalytic residues are structurally conserved between the two proteins and their antibacterial activities against Gram-positive bacteria were comparable (21), it is likely that TSWAB employs a similar catalytic mechanism as that of chicken lysozyme to cleave its substrate. Correspondingly, the electrostatic potential surfaces in the active site region are also similar. However, a substantial number of other residues that may be important in stabilizing the substrate in the active site are different from those seen at analogous positions in chicken lysozyme. Therefore, there might be differences in levels of specificity between the two enzymes. On the other hand, the surface electrostatic potential in the regions corresponding to the active site defined in chicken lysozyme is remarkably different in human $\alpha$-lactalbumin. Thus, the correlation of physiological activities of these three proteins sharing common fold can be seen in their surface properties as well, which show obvious gradation in the negative electrostatic potential (Fig. 5).

The TSWAB is a lysozyme recruited for immune function. In the case of lepidopteran lysozymes (of which TSWAB is an example), the expression is induced only on infection. Considering the high sequence homology, the structures of other insect lysozymes of induced origin will also be very similar to that of TSWAB. As seen in Fig. 6, the homology of other lysozymes with *A. aegypti* is relatively less, perhaps because of the fact that the others belong to same order, Lepidoptera, as opposed to *A. aegypti*, which belongs to Diptera, although it is also of induced immune origin. On the other hand, there are other insect lysozymes, *e.g.* lysozyme of fruit fly, the expression of which is constitutive. The structure of TSWAB may not represent those members of insect lysozyme family that have different physiological origins and have much lower (40%) sequence identity. Among the residues defining the active site, four are conserved between insect lysozymes of induced origin, which are not conserved with respect to c-lysozymes (Fig. 6). Perhaps these residues may relate to the specific character of the proteins of this group. Measurement of the molecular volumes showed that TSWAB (10570 $Å^3$) is smaller and more compact than the other c-type lysozymes (11,800–12,200 $Å^3$). The primary reason for this is that the loops are shorter in the case of TSWAB. This compactness might impart greater thermal stability to insect lysozymes, as has been seen for other proteins (39). Greater stability would also ensure that the protein is physiologically present for a longer duration, which would be important, considering that TSWAB expression is induced and not constitutive.

Comparison of the TSWAB structure with those of other

c-type lysozymes and α-lactalbumins led to interesting observations concerning molecular evolution that may be of general significance. It has been suggested that α-lactalbumins have evolved from mammalian lysozyme after divergence of the avian and mammalian lineages and have undergone a rapid evolution, leading to acquisition of a new function and loss of lysozyme activity (38). In contrast, the insect lysozymes diverged earliest from the lineage that led to vertebrate lysozyme and α-lactalbumins. Thus, although the differences in sequences of human α-lactalbumin and TSWAB with respect to chicken lysozyme are comparable, the differences in the corresponding structures are not. Unlike human α-lactalbumin, the structure of TSWAB is significantly different from chicken lysozyme. This is an interesting observation considering the fact that chicken lysozyme shares mechanism of action with TSWAB and not with α-lactalbumin. Apparently, during evolution, TSWAB has undergone structural changes in those parts of the protein that were not involved in the function. Thus, the catalytic residues were selectively conserved and the function was retained. The conservation of function with significant differences in conformation between TSWAB and chicken lysozyme, on one hand, and drastically different function with conserved molecular conformation between α-lactalbumins and c-type lysozymes, on the other, provide an interesting contrast. It appears that structure changes much more slowly during rapid evolution as observed for α-lactalbumins and c-type lysozymes as against that during normal course of evolution as observed between c-type lysozymes and the insect lysozyme.

The crystal structure of TSWAB, an antibacterial protein, described here is the first structure belonging to the insect lysozyme family. It represents the structural class corresponding to the induced lysozymes of innate immune origin and provides insights into the functional aspects of this class. Despite the fact that the mechanism of action of TSWAB is likely to be similar to that of chicken lysozyme, it is expressed only on bacterial infection and has significant structural differences as compared to chicken lysozyme. The insect lysozymes diversified from the lineage of chicken lysozyme while retaining the common function, the lactalbumins evolved rapidly resulting in new function. However, structurally chicken lysozyme and α-lactalbumins are closer to each other compared with insect lysozyme, implying that structure is resistant to changes during rapid evolution. Thus, rapid evolution leads to the acquisition of new function by modification of certain critical residues without drastically changing the structure.

## REFERENCES


1. Mackintosh, J. A., Gooley, A. A., Karuso, P. H., Beattie, A. J., Jardine, D. R., and Veal, D. A. (1998) *Dev. Comp. Immunol.* **22,** 387–399
2. Axen, A., Carlsson, A., Engstrom, A., and Bennich, H. (1997) *Eur. J. Biochem.* **247,** 614–619
3. Chung, K. T., and Ourth, D. D. (2000) *Eur. J. Biochem.* **267,** 677–683
4. Phipps, D. J., Chadwick, J. S., and Aston, W. P. (1994) *Dev. Comp. Immunol.* **18,** 13–23
5. Okada, M., and Natori, S. (1985) *J. Biol. Chem.* **260,** 7174–7177
6. Akuffo, H., Hultmark, D., Engstom, A., Frohlich, D., and Kimbrell, D. (1998) *Int. J. Mol. Med.* **1,** 77–82
7. Fujiwara, S., Imai, J., Fujiwara, M. Yaeshima, T., Kawashima, T., and Kobayashi, K. (1990) *J. Biol. Chem.* **265,** 11333–11337
8. Carlsson., A., Engstrom, P., Palva, E. T., and Bennich, H. (1991) *Infect. Immun.* **59,** 3040–3045
9. Casteels, P. (1990) *Res. Immunol.* **141,** 940–942
10. Powning, R. F., and Davidson, W. J. (1976) *Comp Biochem. Physiol.* **55,** 221–228
11. Jolles, J., Schoentgen, F., Croizier, G., Croizier, L., and Jolles, P. (1979) *J. Mol. Evol.* **14,** 267–271
12. Medzhitov, R., and Janeway, C. (2000) *N. Engl. J. Med.* **343,** 338–344
13. Ulevitch, R. J. (2000) *Immunol. Res.* **21,** 49–54
14. Maget-Dana, R., Bonmatin, J. M., Hetru, C., Ptak, M., and Maurizot, J. C. (1995) *Biochimie* (*Paris*) **77,** 240–244
15. Lee, Y. T., Kim, D. H., Suh, J. Y., Chung, J. H., Lee, B. L., Lee, Y., and Choi, B. S. (1999) *Biochem. Mol. Biol. Int.* **47,** 369–376
16. Lee, K. H., Hong, S. Y., and Oh, J. E. (1998) *FEBS Lett.* **439,** 41–45
17. Cornet, B., Bonmatin, J. M., Hetru, C., Hoffmann, J. A., Ptak, M., Vovelle, F. (1995) *Structure* **3,** 435–448
18. Mandard, N., Sodano, P., Labbe, H., Bonmatin, J. M., Bulet, P., Hetru, C., Ptak, M., Vovelle, F. (1998) *Eur. J. Biochem.* **256,** 404–410
19. Landon, C., Sodano, P., Hetru, C., Hoffmann, J., Ptak, M. (1997) *Protein Sci.* **6,** 1878–1884
20. Su, X. D., Gastinel, L. N., Vaughn, D. E., Faye, I., Poon, P., Bjorkman, P. J. (1998) *Science* **281,** 991–995
21. Nagaraju, J., Abraham, E. G., Sethuraman, B. N., and Datta, R. K. (1992) in *Wild Silkmoth '92* (Akai, H., Kato, Y., Kiuchi, M., and Inouchi, J., eds) pp. 22–30, International Society for Wild Silkmoths, Tokyo, Japan
22. Abraham, E. G., Nagaraju, J., Salunke, D., Gupta, H. M., and Datta, R. K. (1995) *J. Invert. Pathol.* **65,** 17–24
23. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
24. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497
25. Navaza, J. (1994) *Acta Crystallogr. Sect.* A **50,** 157–163
26. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., 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
27. Brünger, A. T. (1992) *Nature* **355,** 472–474
28. Jones, T. A., and Kjeldgaard, M. (1994) *O: The Manual*, Uppsala University, Uppsala, Sweden
29. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
30. Nicholls, A., Bharadway, R., and Honig, B. (1993) *Biophys. J.* **64,** 166–170
31. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
32. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410
33. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) *J. Mol. Biol.* **247,** 536–540
34. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., Bourne, P. E. (2000) *Nucleic Acids Res.* **28,** 235–242
35. Chipman, D., and Sharon, N. (1969) *Science* **165,** 454–465
36. Copley, R. R., and Bork, P., (2000) *J. Mol. Biol.* **303,** 627–640
37. Wang, C, Xi J, Begley, T. P., and Nicholson, L. K. (2001) *Nat. Struct. Biol.* **8,** 47–51
38. Katsutoshi, N., and Shintaro, S. (1989) *Eur. J. Biochem.* **182,** 111–118
39. Thompson, M. J and Eisenberg, D. *(1999) J. Mol. Biol.* **290,** 595–604


PROTEIN STRUCTURE AND FOLDING:

# Structure of the Induced Antibacterial Protein from Tasar Silkworm, *Antheraea mylitta* : IMPLICATIONS TO MOLECULAR EVOLUTION

Deepti Jain, Deepak T. Nair, G. Jawahar Swaminathan, E. G. Abraham, J. Nagaraju and Dinakar M. Salunke

*J. Biol. Chem. 2001, 276:41377-41382.*
*doi: 10.1074/jbc.M104674200 originally published online August 24, 2001*