



# How Honey Bee Vitellogenin Holds Lipid Cargo: A Role for the C-Terminal

Vilde Leipart<sup>1\*</sup>, Øyvind Halskau<sup>2</sup> and Gro V. Amdam<sup>1,3</sup>

<sup>1</sup>Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway, <sup>2</sup>Department of Biological Sciences, University of Bergen, Bergen, Norway, <sup>3</sup>School of Life Sciences, Arizona State University, Tempe, AZ, United States

Vitellogenin (Vg) is a phylogenetically broad glycolipophosphoprotein. A major function of this protein is holding lipid cargo for storage and transportation. Vg has been extensively studied in honey bees (*Apis mellifera*) due to additional functions in social traits. Using AlphaFold and EM contour mapping, we recently described the protein structure of honey bee Vg. The full-length protein structure reveals a large hydrophobic lipid binding site and a well-defined fold at the C-terminal region. Now, we outline a shielding mechanism that allows the C-terminal region of Vg to cover a large hydrophobic area exposed in the all-atom model. We propose that this C-terminal movement influences lipid molecules' uptake, transport, and delivery. The mechanism requires elasticity in the Vg lipid core as described for homologous proteins in the large lipid transfer protein (LLTP) superfamily to which Vg belongs. Honey bee Vg has, additionally, several structural arrangements that we interpret as beneficial for the functional flexibility of the C-terminal region. The mechanism proposed here may be relevant for the Vg molecules of many species.

**Keywords:** honey bee vitellogenin, alphafold, lipid binding, oxidative stress, C-terminal region

## OPEN ACCESS

### Edited by:

Roberta Spadaccini,  
University of Sannio, Italy

### Reviewed by:

Kannan Sankar,  
Novartis Institutes for BioMedical  
Research, United States  
Steven Cook,  
United States Department of  
Agriculture (USDA), United States

### \*Correspondence:

Vilde Leipart  
vilde.leipart@nmbu.no

### Specialty section:

This article was submitted to  
Structural Biology,  
a section of the journal  
Frontiers in Molecular Biosciences

**Received:** 29 January 2022

**Accepted:** 19 May 2022

**Published:** 09 June 2022

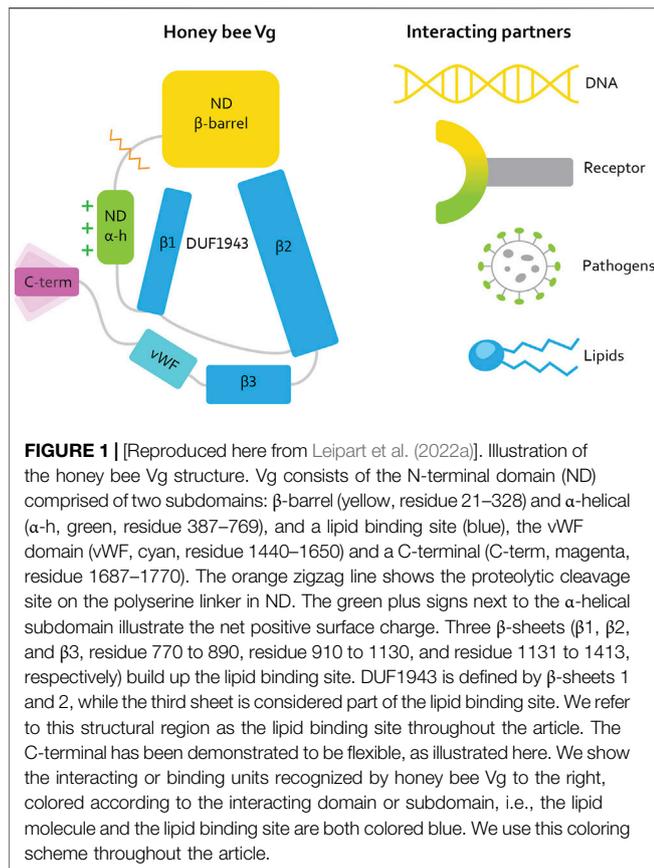
### Citation:

Leipart V, Halskau Ø and Amdam GV  
(2022) How Honey Bee Vitellogenin  
Holds Lipid Cargo: A Role for the C-  
Terminal.  
Front. Mol. Biosci. 9:865194.  
doi: 10.3389/fmolb.2022.865194

## INTRODUCTION

Vitellogenin (Vg) is the most ancient yolk precursor protein in animals (Smolenaars et al., 2007). It is well-known for transporting lipids and other nutrients to developing embryos but is recognized for additional roles in innate immunity and complex behavior (Amdam et al., 2006; Salmela et al., 2015; Sun and Zhang, 2015). As is true of all large lipid transfer protein (LLTP) superfamily members, Vg contains a hydrophobic lipid binding domain that defines a cavity structure. Superfamily members have similar structural landscapes in their binding cavities despite variations in amino acid sequence (Babin et al., 1999; Shelness and Ledford, 2005). Specifically, LLTPs across taxa have amphipathic  $\alpha$ -helical repeats surrounding several amphipathic  $\beta$ -sheets (Van der Horst et al., 2002; Wang et al., 2006). The  $\beta$ -sheets provide a hydrophobic lining to the interior cavity. The connective loops and the flexibility of the  $\alpha$ -helices and  $\beta$ -sheets provide elasticity to expand or compress the cavity during uptake or delivery of the lipid cargo. Beyond the  $\alpha$ -helical and  $\beta$ -sheet structures and features, the characteristics and functions of LLTPs differ (Smolenaars et al., 2007). In terms of the Vg proteins, they all contain a well-conserved N-terminal domain, composed of a  $\beta$ -barrel and  $\alpha$ -helical subdomain (Li et al., 2003; Roth et al., 2013). The remaining domains are otherwise variable but usually include one or several domains of unknown function (DUF) and may include a von Willebrand factor (vWF) domain.

For most animals, the detailed structural composition of Vg remains undescribed. The majority of insight into its structure is derived through template-based modeling from the only experimentally-solved Vg structure, lamprey (*Ichthyomyzon unicuspis*) (Thompson and



Banaszak, 2002). The determination of this crystal structure was a central contribution to understanding Vg proteins, but its partial sequence coverage and distant homology toward many other Vg molecules limit its usefulness for *in silico* structural predictions. Also, few template options for isolated Vg domains or subdomains exist. The massive size and the complex domain organization of Vg proteins, which includes their sizeable hydrophobic (lipid binding) cavity and extensive post-translational modifications (Tufail and Takeda, 2008), have probably contributed to the relative lack of detailed structural insights.

In parallel with these challenges, intriguing data have emerged on the non-reproductive roles of Vg (Corona et al., 2013; Kohlmeier et al., 2019). To date, these have been most studied in honey bees (*Apis mellifera*) (Pan et al., 1969), in which the protein influences social behavior, oxidative stress resilience, and cell-based and trans-generational immunity, in addition to its traditional role in yolk formation (Salmela et al., 2015; Havukainen et al., 2013; Seehuus et al., 2006; Amdam et al., 2004). Recent progress made possible by DeepMind's AlphaFold, a neural network for structure prediction (Jumper et al., 2021), allowed us to generate a full-length structure prediction of honey bee Vg with high confidence (Leipart et al., 2022). This structure prediction reveals the N-terminal domain folding around the lipid binding cavity, as expected for a Vg protein. Surprisingly, four structural units build up the cavity

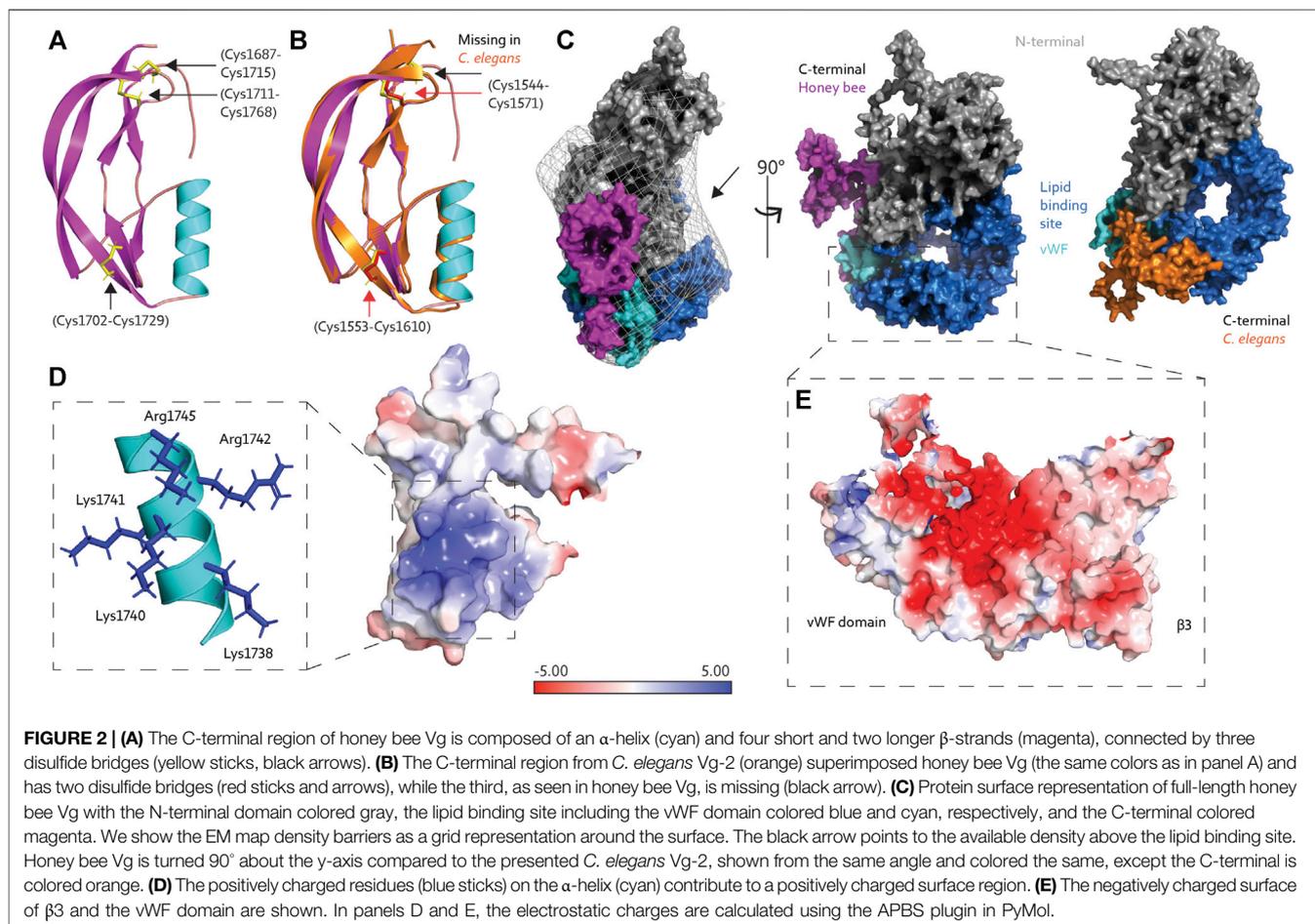
[Figure 1, reproduced here from Leipart et al. (2022a)]. Two  $\beta$ -sheets ( $\beta$ 1 and  $\beta$ 2, also referred to as C- and A-sheet, respectively) comprise the so-called DUF1943 domain. A third  $\beta$ -sheet ( $\beta$ 3) and the vWF domain follow the DUF1943, completing the circular or funnel-like shape of the lipid cavity. The domains and subdomains are interconnected. For example, the longer  $\beta$ 2 sheet extends toward the N-terminal domain, and the  $\alpha$ -helical subdomain covers and scaffolds the DUF1943 structural elements, which reduce the lipid cavity's exposure to the solvent. The C-terminal constitutes a small structural fold connected to the vWF domain through a presumably flexible linker. The folded C-terminal region does not appear to be in direct contact with the lipid binding site but instead appears at the flank of the large Vg structure [(Leipart et al., 2022) for more details].

Our previous study fitted the AlphaFold prediction into a low-resolution EM map (Leipart et al., 2022). However, the C-terminal position was not compatible with EM density barriers. Therefore, we proposed an alternative position of the C-terminal above the lipid binding site, as the fitting revealed available space at the opening of the lipid cavity (marked in Figure 2C). But what is the C-terminal region of Vg possibly doing there? In the current article, we assess this structural organization's feasibility and possible functional relevance.

## C-TERMINAL FLEXIBILITY

To our knowledge, the tertiary structure of the C-terminal region of Vg proteins has not been solved. The structural fold of the C-terminal region is composed of four short  $\beta$ -strands, an  $\alpha$ -helix, and two longer  $\beta$ -strands. Three disulfide bridges connect the short and longer  $\beta$ -strands (Figure 2A). The AlphaFold database (<https://alphafold.ebi.ac.uk/>), a collaboration between DeepMind and EMBL's European Bioinformatics Institute, contains a continuously growing number of predicted proteomes (Varadi et al., 2021), including that of *Caenorhabditis elegans*, which has six Vg-encoding sequences (*vit* gene one–6) (Perez and Lehner, 2019). Vg-2 are most abundantly expressed and best-studied gene (Grant and Hirsh, 1999; Goszczynski et al., 2016). Superimposing the C-terminal region (amino acid 1530–1613) in *C. elegans* Vg-2 with our prediction of the C-terminal in honey bee Vg (amino acid 1688–1770) shows an almost identical fold [RMSD = 1.035 (Schrodinger, 2015), Figure 2B], although Vg-2 has only two disulfide bridges. These predictions have low confidence for the loop connecting the C-terminal to the vWF domain indicating a region of low conservation. No secondary structure folds are predicted for either animal (Jumper et al., 2021; Mariani et al., 2013), suggesting a flexible loop in the region linking the C-terminal to the entrance of the lipid binding cavity. The position of the C-terminal region differs between the bee and worm predictions in reference to the lipid binding site (Figure 2C). However, AlphaFold states that predicting positions for extended linkers or isolated structural elements may be less reliable due to the frequent lack of inter-residue contacts (Jumper et al., 2021).

The vWF and C-terminal regions are often described as single C-terminal domains in Vg proteins. Our prediction of



honey bee Vg, in contrast, describes two separate and distinct structural folds. The vWF domain is packed tightly in the lipid binding site, while the C-terminal is a separate solvent-exposed region (Figures 2C,E). We proposed a possible zinc-coordination site that resides between the two adjacent disulfide bridges in the C-terminal region in manuscript, Leipart et al. (2022b). Similar coordination sites of four cysteine residues are often found in redox switches (Ilbert et al., 2006; Pace and Weerapana, 2014) that can cause conformational changes: During oxidative stress, zinc is released, resulting in oxidative folding and creation of disulfide bridges (Ruddock and Klappa, 1999). A similar mechanism could be relevant for folding at the C-terminal region of honey bee Vg.

Taken together, structural analysis suggests that the C-terminal of honey bee Vg can be flexible and take part in conformational changes such as domain repositioning.

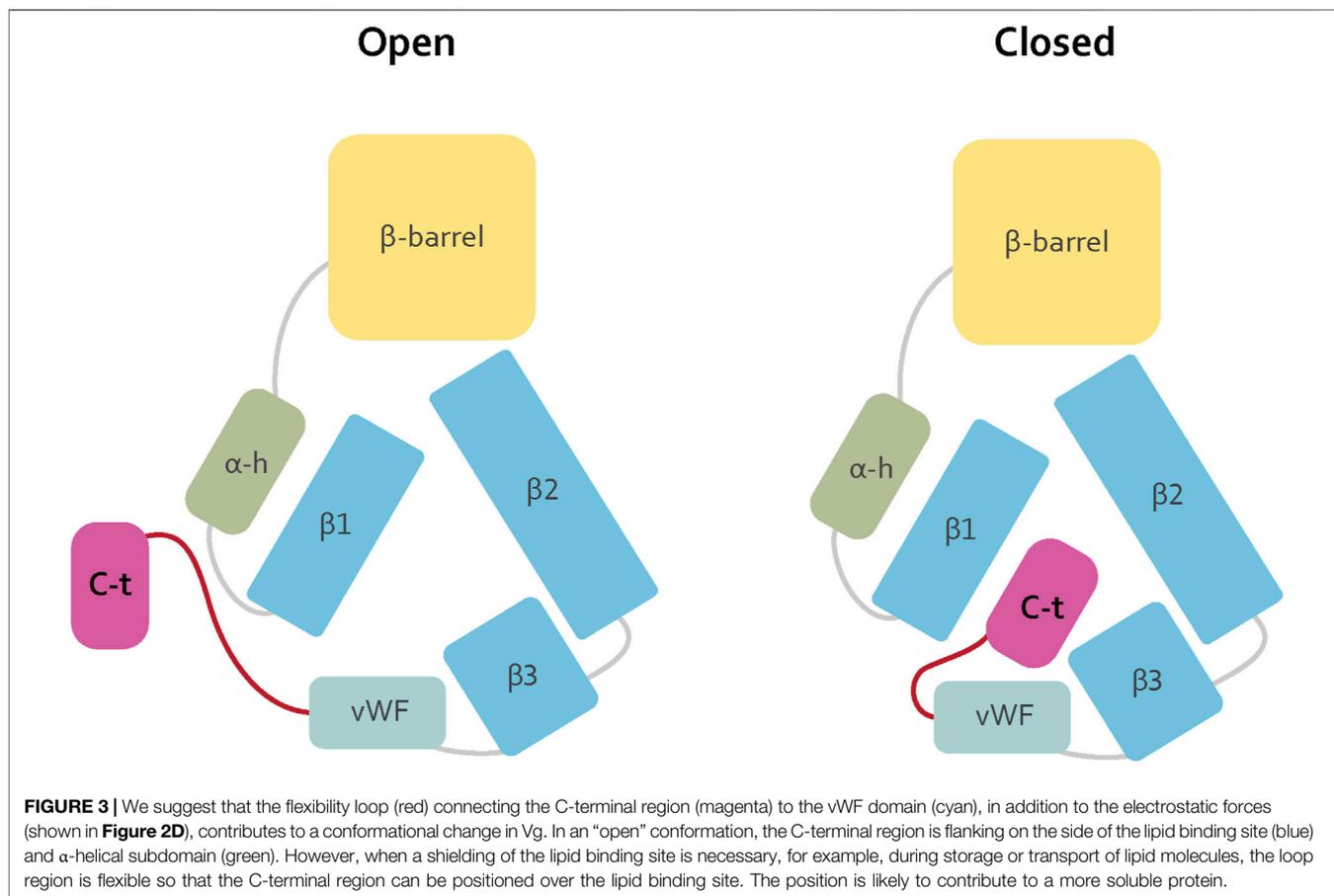
## EXPOSED LIPID BINDING SITE

The proposed C-terminal repositioning is in line with complementary electrostatic forces on the C-terminal region and lipid binding cavity.

Insect Vg proteins have conserved positively charged residues at the C-terminal (Tufail and Takeda, 2008). These reside at the  $\alpha$ -helix, creating a net positive surface charge (Figure 2D). The lipid binding site has a negatively charged center on  $\beta 3$  and the vWF domain (Figure 2E). Additionally, the wide opening of the lipid binding site, exposing the hydrophobic cavity to the solvent, is costly in terms of entropy. Shielding the opening would aid the stability and solubility of Vg, particularly during transport or storage of large lipid cargo. We propose that the C-terminal region provides this shielding. As illustrated in Figure 3, the “closed” position resembles the contour of the EM map (Leipart et al., 2022), while the AlphaFold prediction would represent the “open” (flanking) position. Similar conformational shifts, including an “open” and “closed” state, have previously been reported for LLTPs (Wang et al., 2006; Smolenaars et al., 2007). The precise position of the C-terminal region in our “open” state, however, is uncertain due to reduced inter-residue contacts, as noted above (Figure 2C). A more likely scenario is perhaps a position closer to the Vg structure.

## EXPANSION AND COMPRESSION

LLTPs can bind up to hundreds of lipid molecules (Segrest et al., 2001; Hevonoja et al., 2000). Their packing requires interior stability

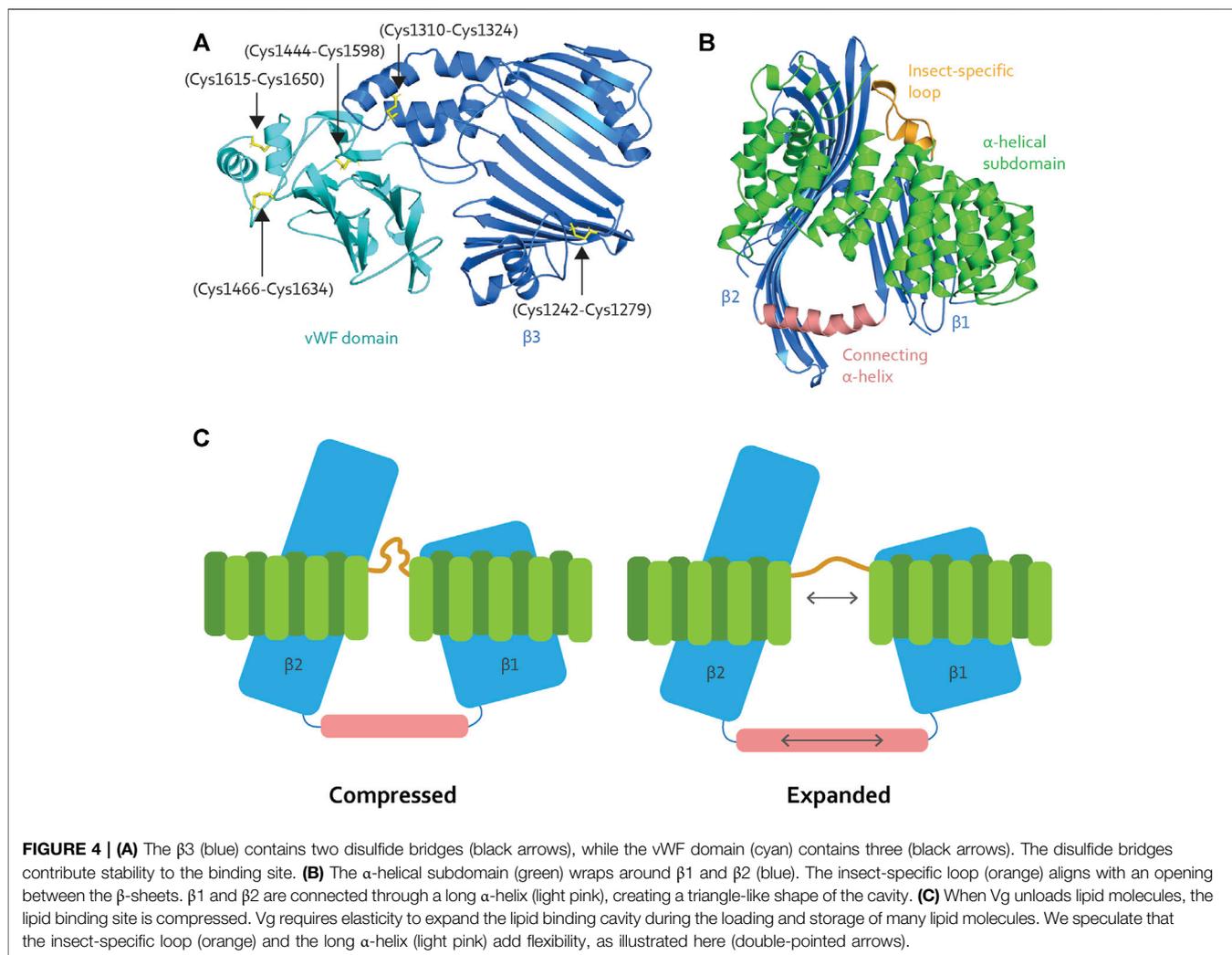


to withstand differences in pressure on the lipid cavity lining and support the elasticity of the lipid core to handle the changing lipid loads. For honey bee Vg, the  $\beta$ -sheet network and the identification of five disulfide bridges distributed between  $\beta$ 3 and the vWF domain may contribute to a stable interior (**Figure 4A**). The lipid binding site of microsomal triglyceride transfer protein (MTP), another LLTP member, has a narrower lipid binding cavity compared to lamprey Vg (Biterova et al., 2019). MTP has a flexible junction to accommodate lipid binding, despite lower lipid binding capacity. We believe that honey bee Vg might require greater flexibility than MTP due to the larger cavity volume.

Interestingly, MTP and lamprey Vg have conserved disulfide bridges in their respective  $\alpha$ -helical subdomains (**Supplementary Figures S1A,B** for a structural representation), which are reported to create stability for the subdomain and the encompassing  $\beta$ 1 and  $\beta$ 2 (Thompson and Banaszak, 2002; Biterova et al., 2019). The disulfide bridges are not conserved in honey bee Vg (**Supplementary Figure S1C** for sequence alignment), suggesting lower fold stability and a greater potential for flexibility. In addition, honey bee Vg has an insect-specific loop region between  $\alpha$ -helices 9 and 10 (residue 582 to 603, **Supplementary Figure S1C**) (Havukainen et al., 2013), which is consistent with an elastic subdomain arrangement. The insect-specific loop aligns with the opening between  $\beta$ 1 and  $\beta$ 2 (**Figure 4B**). The  $\beta$ -sheets interact with the  $\alpha$ -helical subdomain through hydrophobic and electrostatic interactions (**Supplementary**

**Table S1** for a list of the hydrophobic interacting residues and **Supplementary Figures S2, S3** for a structural representation of the hydrophobic and electrostatic interactions, respectively); this is important for maintaining a stable fold. Moreover,  $\beta$ 1 and  $\beta$ 2 are connected through a long  $\alpha$ -helix, creating a triangle-like shape of the lipid binding cavity (**Figure 4B**). The connecting  $\alpha$ -helix is reported as stabilizing for the tertiary structure for other LLTPs (Thompson and Banaszak, 2002; Biterova et al., 2019). In honey bee Vg, the  $\alpha$ -helix is longer (19 amino acids) than in MTP (11 amino acids), lamprey Vg (13 amino acids), and *C. elegans* Vg-2 (11 amino acids followed by a loop region and an additional  $\alpha$ -helix of nine amino acids, folded in parallel with  $\beta$ 2) (**Supplementary Figure S4** for structural comparison). The long and continuous  $\alpha$ -helix in honey bee Vg creates a larger lipid binding triangle. Compared to the other animals, the increased volume for honey bee Vg could perhaps be due to a higher demand for lipids during the maturation of embryos or differences in energy output from various lipid types. For example, the natural environments for honey bees might have a limited lipid availability, and honey bees might compensate by ingesting a higher lipid load.

To summarize: the  $\alpha$ -helical subdomain surrounding the lipid binding cavity in honey bee Vg contains regions that can provide elasticity during expansion and compression (**Figure 4C**). The subdomain and the connecting  $\alpha$ -helix also support an ability to carry very large lipid loads. We note, however, that the expansion of



the hydrophobic core of a lipid binding cavity can result in a less soluble surface (Biterova et al., 2019; Ptak-Kaczor et al., 2021). In this context, we propose that the C-terminal region provides a cover that increases the solubility of Vg, possibly shifting deeper into the cavity in response to increasing loads. Similar shielding has been reported for MTP, which interacts with its  $\beta$ -subunit, protein disulfide isomerase (PDI): PDI binds to the  $\alpha$ -helical subdomain and shields the lipid binding site opening (Biterova et al., 2019). Interestingly, MTP lacks a C-terminal region homologous to that of honey bee Vg. In contrast, honey bee Vg is a single subunit protein that does not pair with a PDI homolog.

## POST-TRANSLATIONAL MODIFICATIONS

Extensive protein modifications, such as ubiquitylation and sumoylation, are not observed for honey bee Vg (Havukainen et al., 2011; Havukainen et al., 2012), but the protein is known to be phosphorylated and glycosylated (Havukainen et al., 2011) (extent and exact positions of these PTMs are unknown). There are well-documented examples of

phosphorylation and glycosylation providing increased solubility (Darewicz et al., 1998; Darewicz et al., 1999; Srikanth et al., 2010; Broyard and Gaucheron, 2015), resistance to disordered elements against proteases (Havukainen et al., 2011; Havukainen et al., 2012; Niu et al., 2015), modulation of the conformational propensities of flexible elements (Fraser et al., 2010; Kurotani and Sakurai, 2015; He et al., 2016), and steric hindrances or complementarity for ligand binding or domain reorganization (Dean and Koshland, 1990; Shipley et al., 1993). Methylation or acetylation of Vg could also conceivably be involved. These modifications are found on lysine and arginine and are associated with epigenetic control but tend to decrease solubility (Ritchie et al., 2015). We acknowledge that PTMs at the folded part of the C-terminal, at its flexible linker, or in the putative binding site could affect conformational propensities of the whole region, regulate or support the correct and timely insertion of the folded element, and protect hydrophobic surfaces from the solvent (Srikanth et al., 2010). However, further

discussion of these possibilities requires more actual and accurate data on Vg PTMs.

## CONCLUDING REMARKS

At this point, we arrive at an explanation for how the lipid binding site of honey bee Vg may be optimized for large lipid cargo. This optimization includes a large lipid binding triangle, an ability to flex and compress to load and unload the lipid cargo, and utilization of the C-terminal to shield the exposed hydrophobic surface. Our model includes predictions about an “open” vs. “closed” protein configuration. This model sets the stage for performing molecular simulation, protein docking, and experimental dynamical studies, in addition to experimental analysis like hydrogen deuterium exchange mass spectrometry (Masson et al., 2019) or cryogenic electron microscopy to test our speculations. We further note that thorough mapping of potential PTM sites using both *in silico* and experimental approaches is required for a complete understanding of the molecular mechanisms. Taken together, these considerations provide a roadmap for future studies of how honey bee Vg holds its lipid cargo and is soluble at high concentrations. We also hope they are inspirational and relevant for research on the Vg molecules of other species.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The datasets analyzed for this study can be found in the PDB at <https://www.rcsb.org/> (PDB-ID: 1LSH and 6I7S) and in AlphaFold database <https://alphafold.ebi.ac.uk/> (UniProt ID: P05690). The AlphaFold prediction of honey bee Vg is supplemented in our recent publication [20].

## AUTHOR CONTRIBUTIONS

VL wrote the paper and made the figures. ØH and GVA contributed to and edited the manuscript.

## REFERENCES

- Amdam, G. V., Simões, Z. L. P., Hagen, A., Norberg, K., Schröder, K., Mikkelsen, Ø., et al. (2004). Hormonal Control of the Yolk Precursor Vitellogenin Regulates Immune Function and Longevity in Honeybees. *Exp. Gerontol.* 39 (5), 767–773. doi:10.1016/j.exger.2004.02.010
- Amdam, G. V., Csondes, A., Fondrk, M. K., and Page, R. E., Jr (2006). Complex Social Behaviour Derived from Maternal Reproductive Traits. *Nature* 439 (7072), 76–78. doi:10.1038/nature04340
- Babin, P. J., Bogerd, J., Kooiman, F. P., Van Marrewijk, W. J. A., and Van der Horst, D. J. (1999). Apolipoprotein II/I, Apolipoprotein B, Vitellogenin, and Microsomal Triglyceride Transfer Protein Genes Are Derived from a Common Ancestor. *J. Mol. Evol.* 49 (1), 150–160. doi:10.1007/pl00006528
- Biterova, E. I., Isupov, M. N., Keegan, R. M., Lebedev, A. A., Sohail, A. A., Liaqat, I., et al. (2019). The Crystal Structure of Human Microsomal Triglyceride Transfer Protein. *Proc. Natl. Acad. Sci. U.S.A.* 116 (35), 17251–17260. doi:10.1073/pnas.1903029116

## FUNDING

The authors acknowledge The Research Council of Norway grant number 262137 for funding toward running costs and positions and BioCat (RCN grant number 249023) for travel grants and conferences support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.865194/full#supplementary-material>

**Supplementary Table S1** | This table provides the 50 hydrophobic interactions between the  $\alpha$ -helical subdomain to  $\beta$ 1 and  $\beta$ 2. The first column numerates the interactions, while the second and third column lists the interacting residues from the  $\alpha$ -helical subdomain and  $\beta$ 1 or  $\beta$ 2, respectively. The hydrophobic interactions were identified using PyMol by selecting the hydrophobic residues for each subdomain with carbon atoms at a 4 Å distance or less.

**Supplementary Figure S1** | (A) Lamprey Vg has a disulfide bridge in the  $\alpha$ -helical subdomain (yellow sticks, black arrow), while (B) MTP has two disulfide bridges (yellow sticks, black arrows) in the same subdomain. (C) The alignment is derived from in the manuscript, Leipart et al. (2022b), and shows two regions in the  $\alpha$ -helical subdomain. The taxa, UniProt ID, and protein type are included on the sequences left. The cysteine residues (yellow) involved in the bridges are not conserved in honey bee Vg and *C. elegans* Vg-2. The insect-specific loop in honey bee Vg is marked with a black line, and the alignment shows gaps in the equivalent region in the other organisms.

**Supplementary Figure S2** | In the  $\alpha$ -helical subdomain (green cartoon), there are 31 residues (magenta spheres) interacting with 25 residues (orange spheres) on  $\beta$ 1 (blue cartoon, right) and  $\beta$ 2 (blue cartoon, left), making 50 hydrophobic interactions. A detailed list of the interactions is included in **Supplementary Table S1**.

**Supplementary Figure S3** | Eight electrostatic interactions connect the  $\alpha$ -helical subdomain (green cartoon) to  $\beta$ 1 (blue cartoon, left) and  $\beta$ 2 (blue cartoon, left). The interactions' residues are shown here and colored according to charge (blue for positive charge and red for negative charge).

**Supplementary Figure S4** | The  $\alpha$ -helix connecting  $\beta$ 1 and  $\beta$ 2 are shown for honey bee Vg, human MTP, lamprey Vg, and *C. elegans* Vg-2. The first and last residue of the  $\alpha$ -helices (red cartoon) are included and illustrate the length differences. The yellow  $\beta$ -strands before and after the  $\alpha$ -helices represented the last and first  $\beta$ -strand of  $\beta$ 1 and  $\beta$ 2, respectively. The dotted green line for MTP and lamprey Vg represents unresolved regions in diffraction data.

- Broyard, C., and Gaucheron, F. (2015). Modifications of Structures and Functions of Caseins: a Scientific and Technological Challenge. *Dairy Sci. Technol.* 95 (6), 831–862. doi:10.1007/s13594-015-0220-y
- Corona, M., Libbrecht, R., Wurm, Y., Riba-Grognuz, O., Studer, R. A., and Keller, L. (2013). Vitellogenin Underwent Subfunctionalization to Acquire Caste and Behavioral Specific Expression in the Harvester Ant *Pogonomyrmex Barbatus*. *PLoS Genet.* 9 (8), e1003730. doi:10.1371/journal.pgen.1003730
- Darewicz, M., Dziuba, J., and Mioduszezwska, H. (1998). Some Physico-Chemical Properties and Structural Changes of Bovine Beta-Casein upon Glycation. *Nahrung* 42 (03-04), 213–214. doi:10.1002/(sici)1521-3803(199808)42:03/04<213::aid-food213>3.0.co;2-i
- Darewicz, M., Dziuba, J., Mioduszezwska, H., and Minkiewicz, P. (1999). Modulation of Physico-Chemical Properties of Bovine B-Casein by Nonenzymatic Glycation Associated with Enzymatic Dephosphorylation. *Acta Aliment.* 28 (4), 339–354. doi:10.1556/aalim.28.1999.4.5
- Dean, A. M., and Koshland, D. E., Jr. (1990). Electrostatic and Steric Contributions to Regulation at the Active Site of Isocitrate Dehydrogenase. *Science* 249 (4972), 1044–1046. doi:10.1126/science.2204110

- Fraser, J. A., Vojtesek, B., and Hupp, T. R. (2010). A Novel P53 Phosphorylation Site within the MDM2 Ubiquitination Signal. *J. Biol. Chem.* 285 (48), 37762–37772. doi:10.1074/jbc.m110.143099
- Goszczynski, B., Captan, V. V., Danielson, A. M., Lancaster, B. R., and McGhee, J. D. (2016). A 44 Bp Intestine-specific Hermaphrodite-specific Enhancer from the *C. elegans* Vit-2 Vitellogenin Gene Is Directly Regulated by ELT-2, MAB-3, FKH-9 and DAF-16 and Indirectly Regulated by the Germline, by Daf-2/insulin Signaling and by the TGF- $\beta$ /Sma/Mab Pathway. *Dev. Biol.* 413 (1), 112–127. doi:10.1016/j.ydbio.2016.02.031
- Grant, B., and Hirsh, D. (1999). Receptor-mediated Endocytosis in the *Caenorhabditis elegans* Oocyte. *MBoC* 10 (12), 4311–4326. doi:10.1091/mbc.10.12.4311
- Havukainen, H., Halskau, Ø., Skjaerven, L., Smedal, B., and Amdam, G. V. (2011). Deconstructing Honeybee Vitellogenin: Novel 40 kDa Fragment Assigned to its N Terminus. *J. Exp. Biol.* 214 (Pt 4), 582–592. doi:10.1242/jeb.048314
- Havukainen, H., Underhaug, J., Wolschin, F., Amdam, G., and Halskau, Ø. (2012). A Vitellogenin Polyserine Cleavage Site: Highly Disordered Conformation Protected from Proteolysis by Phosphorylation. *J. Exp. Biol.* 215 (Pt 11), 1837–1846. doi:10.1242/jeb.065623
- Havukainen, H., Münch, D., Baumann, A., Zhong, S., Halskau, Ø., Krogsgaard, M., et al. (2013). Vitellogenin Recognizes Cell Damage through Membrane Binding and Shields Living Cells from Reactive Oxygen Species. *J. Biol. Chem.* 288 (39), 28369–28381. doi:10.1074/jbc.m113.465021
- He, E., Yan, G., Zhang, J., Wang, J., and Li, W. (2016). Effects of Phosphorylation on the Intrinsic Propensity of Backbone Conformations of Serine/threonine. *J. Biol. Phys.* 42 (2), 247–258. doi:10.1007/s10867-015-9405-0
- Hevonoja, T., Pentikäinen, M. O., Hyvönen, M. T., Kovanen, P. T., and Ala-Korpela, M. (2000). Structure of Low Density Lipoprotein (LDL) Particles: Basis for Understanding Molecular Changes in Modified LDL. *Biochimica Biophysica Acta (BBA) - Mol. Cell Biol. Lipids* 1488 (3), 189–210. doi:10.1016/s1388-1981(00)00123-2
- Ilbert, M., Graf, P. C., and Jakob, U. (2006). Zinc Center as Redox Switch-Nnew Function for an Old Motif. *Antioxid. Redox Signal* 8 (5-6), 835–846. doi:10.1089/ars.2006.8.835
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* 596 (7873), 583–589. doi:10.1038/s41586-021-03819-2
- Kohlmeier, P., Alleman, A. R., Libbrecht, R., Foitzik, S., and Feldmeyer, B. (2019). Gene Expression Is More Strongly Associated with Behavioural Specialization Than with Age or Fertility in Ant Workers. *Mol. Ecol.* 28 (3), 658–670. doi:10.1111/mec.14971
- Kurotani, A., and Sakurai, T. (2015). In Silico Analysis of Correlations between Protein Disorder and Post-Translational Modifications in Algae. *Ijms* 16 (8), 19812–19835. doi:10.3390/ijms160819812
- Leipart, V., Montserrat-Canals, M., Cunha, E. S., Luecke, H., Herrero-Galán, E., Halskau, Ø., et al. (2022). Structure Prediction of Honey Bee Vitellogenin: a Multi-domain Protein Important for Insect Immunity. *FEBS Open Bio* 12 (1), 51–70. doi:10.1002/2211-5463.13316
- Leipart, V., Ludvigsen, J., Kent, M., Sandve, S., To, T.-H., Árnýasi, M., et al. (2022a). Identification of 121 Variants of Honey Bee Vitellogenin Protein Sequence with Structural Differences at Functional Sites.
- Leipart, V., Enger, Ø., Turcu, D. C., Dobrovolska, O., Drablos, F., Ø, H., et al. (2022b). Where Honey Bee Vitellogenin May Bind Zn<sup>2+</sup>-Ions.
- Li, A., Sadasivam, M., and Ding, J. L. (2003). Receptor-Ligand Interaction between Vitellogenin Receptor (VtGR) and Vitellogenin (Vtg), Implications on Low Density Lipoprotein Receptor and Apolipoprotein B/E. *J. Biol. Chem.* 278 (5), 2799–2806. doi:10.1074/jbc.m205067200
- Mariani, V., Biasini, M., Barbato, A., and Schwede, T. (2013). IDDT: a Local Superposition-free Score for Comparing Protein Structures and Models Using Distance Difference Tests. *Bioinforma. Oxf. Engl.* 29 (21), 2722–2728. doi:10.1093/bioinformatics/btt473
- Masson, G. R., Burke, J. E., Ahn, N. G., Anand, G. S., Borchers, C., Brier, S., et al. (2019). Recommendations for Performing, Interpreting and Reporting Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) Experiments. *Nat. Methods* 16 (7), 595–602. doi:10.1038/s41592-019-0459-y
- Niu, C., Luo, H., Shi, P., Huang, H., Wang, Y., Yang, P., et al. (2015). N-glycosylation Improves the Pepsin Resistance of Histidine Acid Phosphatase Phytases by Enhancing Their Stability at Acidic pHs and Reducing Pepsin's Accessibility to its Cleavage Sites. *Appl. Environ. Microbiol.* 82 (4), 1004–1014. doi:10.1128/AEM.02881-15
- Pace, N., and Weerapana, E. (2014). Zinc-binding Cysteines: Diverse Functions and Structural Motifs. *Biomolecules* 4 (2), 419–434. doi:10.3390/biom4020419
- Pan, M. L., Bell, W. J., and Telfer, W. H. (1969). Vitellogenic Blood Protein Synthesis by Insect Fat Body. *Science* 165 (3891), 393–394. doi:10.1126/science.165.3891.393
- Perez, M. F., and Lehner, B. (2019). Vitellogenins - Yolk Gene Function and Regulation in *Caenorhabditis elegans*. *Front. Physiol.* 10 (1067), 1067. doi:10.3389/fphys.2019.01067
- Ptak-Kaczor, M., Banach, M., Stapor, K., Fabian, P., Konieczny, L., and Roterman, I. (2021). Solubility and Aggregation of Selected Proteins Interpreted on the Basis of Hydrophobicity Distribution. *Ijms* 22 (9), 5002. doi:10.3390/ijms22095002
- Ritchie, T. J., Macdonald, S. J. F., and Pickett, S. D. (2015). Insights into the Impact of N- and O-Methylation on Aqueous Solubility and Lipophilicity Using Matched Molecular Pair Analysis. *Med. Chem. Commun.* 6 (10), 1787–1797. doi:10.1039/c5md00309a
- Roth, Z., Weil, S., Aflalo, E. D., Manor, R., Sagi, A., and Khalaila, I. (2013). Identification of Receptor-Interacting Regions of Vitellogenin within Evolutionarily Conserved  $\beta$ -Sheet Structures by Using a Peptide Array. *ChemBioChem* 14 (9), 1116–1122. doi:10.1002/cbic.201300152
- Ruddock, L. W., and Klappa, P. (1999). Oxidative Stress: Protein Folding with a Novel Redox Switch. *Curr. Biol.* 9 (11), R400–R402. doi:10.1016/s0960-9822(99)80253-x
- Salmela, H., Amdam, G. V., and Freitak, D. (2015). Transfer of Immunity from Mother to Offspring Is Mediated via Egg-Yolk Protein Vitellogenin. *PLoS Pathog.* 11 (7). e1005015. doi:10.1371/journal.ppat.1005015
- Schrodinger, L. (2015). *The PyMOL Molecular Graphics System*. Version 1.8.
- Seehuus, S.-C., Norberg, K., Gimsa, U., Krekling, T., and Amdam, G. V. (2006). Reproductive Protein Protects Functionally Sterile Honey Bee Workers from Oxidative Stress. *Proc. Natl. Acad. Sci. U.S.A.* 103 (4), 962–967. doi:10.1073/pnas.0502681103
- Segrest, J. P., Jones, M. K., De Loof, H., and Dashti, N. (2001). Structure of Apolipoprotein B-100 in Low Density Lipoproteins. *J. Lipid Res.* 42 (9), 1346–1367. doi:10.1016/s0022-2275(20)30267-4
- Shelness, G. S., and Ledford, A. S. (2005). Evolution and Mechanism of Apolipoprotein B-Containing Lipoprotein Assembly. *Curr. Opin. Lipidol.* 16 (3), 325–332. doi:10.1097/01.mol.0000169353.12772.eb
- Shipley, J. M., Grubb, J. H., and Sly, W. S. (1993). The Role of Glycosylation and Phosphorylation in the Expression of Active Human Beta-Glucuronidase. *J. Biol. Chem.* 268 (16), 12193–12198. doi:10.1016/s0021-9258(19)03325-8
- Smolenaars, M. M. W., Madsen, O., Rodenburg, K. W., and Van der Horst, D. J. (2007). Molecular Diversity and Evolution of the Large Lipid Transfer Protein Superfamily<sup>2</sup>. *J. Lipid Res.* 48 ((3)), 489–502. doi:10.1194/jlr.R600028-JLR200
- Srikanth, B., Vaidya, M. M., and Kalraiya, R. D. (2010). O-GlcNAcylation Determines the Solubility, Filament Organization, and Stability of Keratins 8 and 18\*. *J. Biol. Chem.* 285 (44), 34062–34071. doi:10.1074/jbc.m109.098996
- Sun, C., and Zhang, S. (2015). Immune-Relevant and Antioxidant Activities of Vitellogenin and Yolk Proteins in Fish. *Nutrients* 7 (10), 8818–8829. doi:10.3390/nu7105432
- Thompson, J. R., and Banaszak, L. J. (2002). Lipid-Protein Interactions in Lipovitellin. *Biochemistry* 41 (30), 9398–9409. doi:10.1021/bi025674w
- Tufail, M., and Takeda, M. (2008). Molecular Characteristics of Insect Vitellogenins. *J. insect physiology* 54 (12), 1447–1458. doi:10.1016/j.jinsphys.2008.08.007
- Van der Horst, D. J., Van Hoof, D., Van Marrewijk, W. J. A., and Rodenburg, K. W. (2002). Alternative Lipid Mobilization: The Insect Shuttle System. *Mol. Cell. Biochem.* 239 (1), 113–119. doi:10.1007/978-1-4419-9270-3\_15

- Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., et al. (2021). AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Res.* 50. D439–D444. doi:10.1093/nar/gkab1061
- Wang, L., Walsh, M. T., and Small, D. M. (2006). Apolipoprotein B Is Conformationally Flexible but Anchored at a Triolein/water Interface: A Possible Model for Lipoprotein Surfaces. *Proc. Natl. Acad. Sci. U.S.A.* 103 (18), 6871–6876. doi:10.1073/pnas.0602213103

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Leipart, Halskau and Amdam. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*