This is to certify that the
thesis entitled

ANALYSIS OF LIGAND-BINDING DOMAINS OF THE
MOSQUITO VITELLOGENIN RECEPTOR

presented by

LIQUN MAO

has been accepted towards fulfillment
of the requirements for the

M. S. degree in Genetics Graduate Program

Major Professor's Signature

Aug. 13, 2003

Date

ANALYSIS OF LIGAND-BINDING DOMAINS OF THE MOSQUITO VITELLOGENIN RECEPTOR

By

Liqun Mao

A THESIS

Submitted to
Michigan State University
In partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Genetics Graduate Program

2003

ABSTRACT

ANALYSIS OF LIGAND-BINDING DOMAINS OF THE MOSQUITO VITELLOGENIN RECEPTOR

By

Liqun Mao

To determine which of the two clusters (CLI and CLII) of complement-type repeats (CRs) in *A. aegypti* vitellogenin receptor (*Aa*VgR) is responsible for the binding of *Aa*Vg, mini-receptors encoding either CLI or CLII were constructed and expressed in *Drosophila* cells. Saturation-binding assays indicated one binding site on each cluster, with dissociation constants (Kd) of 25.9 nM and 53 nM, respectively. *Aa*VgR showed at least two binding sites with the apparent Kd of 3.2 nM. Thus, both clusters contribute to the high affinity to Vg, probably in a synergistic way. Protein modeling shows that both clusters have strong negative surfaces, and the surface of CLI is more negative than that of CLII. This indicates that the CLI has higher affinity and that the force mediating *Aa*VgR-*Aa*Vg interaction is predominantly electrostatic complementation. The modeled CLIIs of both *A. gambiae* VgR and *Aa*VgR have similar surface charge distribution. Modeled *Aa*VgR EGF-like repeats lack strong negative surfaces. Modeled three *Aa*VgR YWTD β-propellers have several surface histidine residues that may bring significant change to the surface charge, which implicates YWTD propellers as false ligands for the ligand-binding domain in regulating *Aa*Vg release and *Aa*VgR recycling. Lamprey lipovitellin has an omni-positive surface. The modeled *Aa*Vg small subunit has a moderately positive surface, which supports the complementation hypothesis. The modeled *Aa*Vg large subunit has the helical domain, C-sheet, and A-sheet.

ACKNOWLEDGMENTS

I want to thank heartedly my mentor, Dr. Alexander S. Raikhel, for his patient and insightful guidance throughout my graduate research. I thank my present advisor, Dr. Suzanne M. Thiem. I appreciate support from my Guidance Committee, Dr. Leslie Kuhn, and Dr. Rebecca Grumet. I wish to thank director of the Genetics Program, Dr. Helmut Bertrand, for his support during these years. I want to thank Mr. Alan R. Hays for his technical support and Dr. Jianxin Sun for his nice suggestions on experiments. I also appreciate suggestions and help from Dr Guoqiang Sun, Dr. Shenfu Wang, Dr. Jinsong Zhu, Dr. Chao Li, and many other members of Dr. Raikhel's laboratory. Finally, I thank Geoffrey Attardo for editing the abstract.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 3

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

KEY TO ABBREVIATIONS

20E	20-hydroxyacdysone
α_2-M	α_2 macroglobulin
α_2-M*	activated form of α_2 macroglobulin
Aa	*Aedes aegypti*
Af	*Acanthogobius flavimanus*
Ag	*Anopheles gambiae*
APP	amyloid precursor protein
ApoE	apolipoprotein E
ApoER2	apolipoprotein E receptor-2
APS	*Aedes* physical saline
At	*Acipenser transmontanus*
BSA	bovine serum albumin
Bmo	*Bombyx mori*
bp	base pair
CAPS	3-cyclohexylamino-1-propane sulfonic acid
Cc	*Cyprinus carpio*
cDNA	complementary DNA
Ce	*Caenorhabditis elegans*
CHAPS	3-[3-Cholamidopropyl)-dimethylammonio]-1-propanesulfonate
CCP	clathrin-coated pits

CPM	count per minite
CR	complement-type repeat
Dm	*Drosophila melanogaster*
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
Dr	*Danio rerio*
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
EP	electrostatic potential
ER	endoplasmic reticulum
FACE	formaldehyde agarose gel electrophoresis
Fh	*Fundulus heteroclitus*
FH	familial hypercholesterolemia
FN3	fibronectin type 3
Gg	*Gallus gallus*
HDL	high density lipoprotein
HDLp	High-density lipophorin
HRP	horseradish peroxidase
Hs	*Homo sapiens*
IB	incubation buffer
IDL	intermediate-density lipoproteins
IgG	immunoglobulin G
IMAC	immobilized metal affinity chromatography

Iu	*Ichthyomyzon unicuspis*
kb	kilobase pair
K_d	dissociation constant
kDa	kilo-dalton
La	*Larus argentatus*
LDLp	low-density lipophorin
LDLR	low-density lipoprotein receptor
LL/LI	leucine-leucine or leucine-isoleucine
Lp	lipophorin
LpR	lipophorin receptor
LR11	LDLR relative with 11 ligand-binding repeats
LRP	low density-lipoprotein receptor-related protein
LV	lipovitellin
Ma	*Melanogrammus aeglefinus*
MAP	mitogen activated protein
mRNA	messenger RNA
$_m$VgR	VgR minireceptor
NCBI	national center for biotechnology information
NMR	nuclear magnetic resonance
NPxY	asparigine-proline-x-tyrosine (x can be any amino acid residue)
Ol	*Oryzias latipes*
Om	*Oncorhynchus mykiss*
Pa	*Periplaneta Americana*

PAGE	polyacrylamide gel electrophoresis
PAI-1	plasminogen activator inhibitor 1
PBS	phosphate buffered saline
PEG	polyethylene glycol
Pp	*Pimephales promelas*
PTU	phenylthiourea
PVDF	polyvinylidene difluoride
PCR	polymerase chain reaction
RAP	receptor associated protein
Rc	*Riptortus clavatus*
RNA	ribonucleic acid
R.T.	room temperature
S2 cell	*Drosophila melanogaster* Schneider 2 cell
SDS-PAGE	sodium dodecylsulfate-polyacrylamide gel electrophoresis
Sj	*Sillago japonica*
SorLA	sorting protein-related receptor containing LDLR class A repeats
SRP	surface plasmon resonance
TC	tubular compartments
TNT	transcription and translation
uPA	urokinase-type plasminogen activator
VCB	vitellogenic cathepsin-B
VCP	vitellogenic carboxypeptidase
Vg	vitellogenin

VgR	vitellogenin receptor
VHDLp	very high-density lipophorin
VLDLR	very low-density lipoprotein receptor
Vn	vitellin
Xl	*Xenopus laevis*
YPP	yolk protein precursor
YPR	yolk protein receptor
VPS	vacuolar protein sorting
Yxxϕ	tyrosine-x-x-ϕ (ϕ is a residue with a bulky hydrophobic side chain)
YWTD	tyrosine-tryptophan-threonine-aspartic acid

Chapter 1

Literature Review

LDLR FAMILY AND INSECT YOLK PROTEIN RECEPTORS

Introduction

In mosquitoes, the egg maturation is activated by a blood meal that initiates the accumulation of yolk proteins— mainly vitellogenin (Vg)—in the developing oocyte, which increases in size over 300-fold within 36 hours post blood meal (PBM) during vitellogenesis. Coupled with vitellogenesis, mosquitoes transmit numerous pathogens to humans and animals. Every year over one million people die worldwide from mosquito-borne malaria. Mosquito-vectored diseases include parasitic diseases, such as malaria and filarial diseases (such as dog heart worm), and virus diseases, such as dengue, encephalitis, and yellow fever. Therefore, elucidating the fundamental mechanisms of mosquito vitellogenesis is critical in controlling mosquito-borne diseases.

During vitellogenesis of a female mosquito, *Aedes aegypti*, Vg is synthesized in trophocytes of fat bodies (insect metabolic tissue analogous to the vertebrate liver), circulated in the hemolymph, and accumulated via receptor-mediated endocytosis into the developing oocyte. This internalization process is mediated by the Vg receptor (VgR), a member of the low-density lipoprotein receptor (LDLR) family.

LDLR Family

The LDLR family is a group of constitutively recycling cell-surface receptors that recognize and internalize extracellular ligands for degradation by lysosomes or storage as yolk granules that provide essential nutrients for cellular functions. The prototype of this family, LDLR, plays a major role in cholesterol homeostasis. Low-density lipoproteins (LDL) contain apolipoprotein B-100 (apoB-100) and are cholesteryl ester-rich,

triglyceride-poor macromolecules, arising from the intravascular lipolysis of triacylglycerol-rich very-low-density lipoproteins (VLDL) produced in small intestine and liver by lipoprotein lipase. The LDLR is responsible for the uptake of cholesterol-containing lipoprotein into cells. Mutations in the LDLR gene result in the accumulation of LDL cholesterol in the circulation, which leads to familial hypercholesterolemia (FH), a genetics autosomal dominant disorder.

Apart from the basic cargo function, quite a few family members bind cytosolic signaling proteins and scaffold proteins with cytoplasmic tails, and thus have signaling function (Willnow *et. al.* 1999, Howell and Herz 2001). The LDLR-related protein 2 (LRP2, previously named megalin) is also involved in the regulation of biological function of retinoids and steroids by endocytosis (Howell and Herz 2001). Tables 1 through 3 list tissue specificities, functions, mutational phenotypes, and ligand spectrums of selected members of the LDLR family.

Each member of the LDLR family has a long extracellular portion, one transmembrane helix, and a short cytoplasmic tail. The extracellular portion contains at least three kinds of modules: the cysteine-rich complement-type repeats (CR, or called LDLR class A repeat) that are candidates for the ligand-binding site, the cysteine-rich epidermal growth factor (EGF)-like repeats, and the tyrosine-tryptophan-threonine-aspartic acid (YWTD) β-propeller fold. The YWTD β-propeller and juxtaposed EGF-like repeats compose the EGF precursor-homology domain. Some family members also have a serine/threonine-rich, *O*-linked sugar domain in the juxta-transmembrane region, but this sugar domain is not important for the ligand binding. The intracellular cytoplasmic

Table 1. Tissue expression of the LDLR family members.*

Receptors	Liver	Brain	Heart	Intestine	Kidney	Muscle	Adipose	Adrenal	Lung	Placenta	Ovary	Testis
LDLR	+++	+	+	+	+	+		+++	+	+	+	+
VLDLR	-	+	+++		+	+++	+++	+	+	+	++	++
LRP1	+++	+++		+	+			++	+++	+	+	+
LRP2	-	-	-	+	+++	-		-	++	+	+	-
ApoER2	-	+++	-		-				-	++	+	+
VgR/YPR	-	-	-	-	-	-	-	-	-	-	+++	-
LR11	++	+++						+			++	++

* Mostly adopted from Hussain's review (Hussain *et. al.* 1999). For abbreviations, refer Fig. 1.

Table 2. Functions and mutational phenotypes of LDLR family members.

Receptor	Tissue Expression Pattern and Function	Mutation Phenotype
LDLR	• Keep cholesterol homeostasis by regulating plasma cholesterol	• Familial hypercholesterolemia (FH) • Premature coronary artery disease
VLDLR	• Mainly expressed in extrahepatic tissues • Uptake of triglyceride-rich, apoE-containing lipoproteins in tissues with active fatty acid metabolism, role in lipid metabolism. • Mediating oogenesis in premammalian species • Function in Reeln signaling pathway and neuronal development	• Female sterility • Premature atherosclerosis • Neuronal defect • Reduced adipose (fat) tissue mass
LRP1	• Clearance of plasma-activated α_2-M and ApoE-enriched lipoproteins • Essential for early embryonic development • Signal transduction	• Hyperlipidemia • Associated with Alzheimer's disease • Fetal death
LRP2	• Mainly expressed in absorptive epithelial cells, and important role in absorption by intestine, kidney. • Transport across blood-brain barrier. • Vitamin homeostasis • Essential for development of forebrain • Signal transduction • Major target antigen of rat autoimmune disease Heymann nephritis	• Heymann nephritis in rat • Defect in forebrain development
ApoER2	• Predominantly expressed in brain. • Important role in Reeln signaling pathway, critical functions during brain development	• Affect the pathogenesis of Alzheimer's disease.
LRP5, LRP6	• Expressed in most of the adult and fetal tissues. Highest in aorta • Wnt/Wingless signaling pathway • Mitogenic activity	• Developmental defects
LR11 (SorLA)	• Predominantly expressed in brain • Lack of regulation by cholesterol & estrogen • Maybe involved in cell-cell interactions, especially between neurons	
Invertebrate VgR/YPR	• Endocytosis of yolk protein	• Female sterile

Table 3. Ligand spectrums of the mammalian LDLR family members.

Ligands	LDLR	VLDLR	ApoER2	LRP1	LRP2	LR11
Midkine				+		
α_2-M* (activated α_2 macroglobulin)				+	–	
ApoB100	+				+	
Reelin		+	+			
PAI-1 (plasminogen activator inhibitor 1)				+	+	
pro-uPA (urokinase-type plasminogen activator)		+		+	+	
uPA-PAI-1		+		+	+	
IDL (intermediate-density lipoprotein)	+	+				
VLDL (very LDLR)	+	+	+			
β-VLDL	+	+	+	+	+	
ApoE (apolipoprotein E)	+	+	+	+	+	+
RAP (receptor-associated protein)	+	+	+	+	+	+
Total ligands	8	16	5	29	21	2

tail has one or more endocytosis signal(s), and the most frequent one is the NPxY motif. Figure 1 shows structural organization of the LDLR family members.

The number and composition of cluster(s) of CRs are two essential features of LDLR family members. Except for insect yolk protein receptors, all LDLR family members have either one cluster (CLI) or four clusters (CLI, CLII, CLIII, and CLIV) of CRs. Insects have two kinds of highly homologous yolk protein receptors, the vitellogenin receptor (VgR, in the case of most insects, including the mosquito), and the yolk protein receptor (YPR, in higher dipteran insects such as fruitfly and housefly). In the LDLR family, the insect VgRs/YPRs represent a unique two-cluster subfamily with two clusters of CRs (5 CRs for CLI and 6~8 CRs for CLII). Table 4 summarizes numbers of clusters and CRs in each cluster.

Insect Yolk Protein Receptors

The major yolk protein precursor, Vitellogenin (Vg), is a lipophosphoglycoprotein produced in the fat bodies of insects or livers of higher animals, and transported in the circulation to the female gonads. In egg-laying species, the VgR transports Vg into the developing oocyte by endocytosis during vitellogenesis. Thus far, all characterized vertebrate VgRs have a single cluster of 8 CRs, while the insect VgRs/YPRs have two clusters of CRs.

In both fruitfly, Drosophila *melanogaster* (*Dm*), and mosquito, *Aedes aegypti* (*Aa*), the yolk protein receptors (*Aa*VgR and *Dm*YPR) are present only in oocytes, primarily in the cortex (Sappington *et. al.* 1995; Schonbaum *et. al.* 2000). The *Dm*YPR is expressed very early during the development of the oocyte, as is the *Aa*VgR. During the

Fig. 1. Structural organization of members of the low-density lipoprotein receptor family. Abbreviations: ApoER2, apolipoprotein E receptor 2; *Ce*, *Caenorhabditis elegans*; *Dm*, *Drosophila melanogaster*; LDLR, low-density lipoprotein receptor; LL/LI, leucine-leucine or leucine-isoleucine; LRP, LDLR related protein; LRP5/6, LDL receptor related proteins with 5 or 6 ligand-binding repeats; LR11, LDLR relative with 11 ligand-binding repeats, or called SorLA (Sorting protein-related receptor containling LDLR class A repeats); NPxY, asparigine-proline-x-tyrosine (x could be any amino acid residue); NPxY-like (NxxY), asparigine-x-x-tyrosine (first x is an amino acid residue other than proline); VgR, vitellogenin receptor; VLDLR, very low-density lipoprotein receptor; YPR, yolk protein receptor; Yxxf, tyrosine-x-x-f (f is a residue with a bulky hydrophobic side chain); YWTD, tyrosine-tryptophan-threonine-aspartic acid; +/-, included or excluded; 3-8, 3 to 8 complement-type repeats. **Images in this thesis are presented in color.**

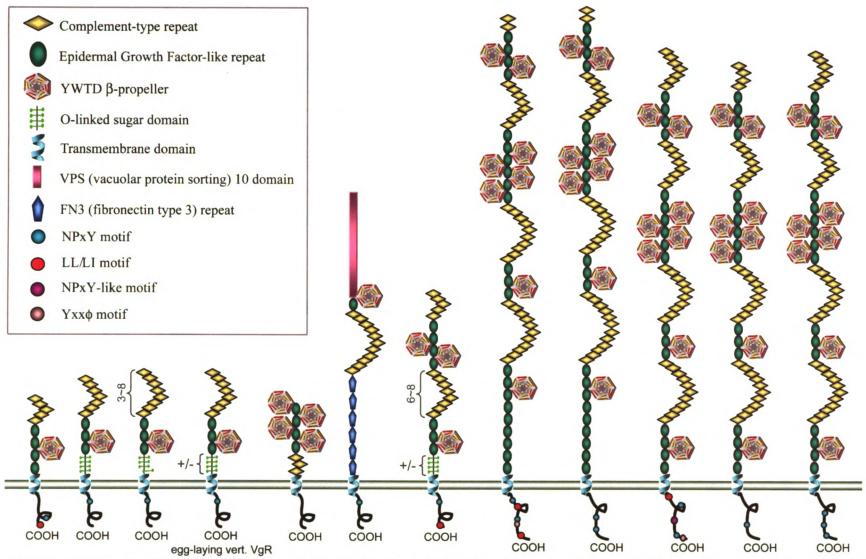

Complement-type repeat
Epidermal Growth Factor-like repeat
YWTD β-propeller
O-linked sugar domain
Transmembrane domain
VPS (vacuolar protein sorting) 10 domain
FN3 (fibronectin type 3) repeat
NPxY motif
LL/LI motif
NPxY-like motif
Yxxϕ motif
3~8
+/-
6~8
+/-
COOH
egg-laying vert. VgR
CeYPR
LDLR
ApoER2
VLDLR
LRP5/6
LP11
insect VgR/YPR
LRP1
LRP1b
LRP2 (Megalin)
DmLRP2
CeLRP2

Table 4. Number of CRs of LDLR family members. Abbreviations: CR, complement-type repeat; CL, cluster.

No. of CLs	LDLR Family Member	No. of CRs in Each CL			
		CL I	CL II	CL II	CL IV
1	LRP5/6	3			
1	nematode YPR	5			
1	LDLR	7			
1	ApoER2	3~8			
1	VLDLR	8			
1	bird/reptile/amphibian/fish/crustacean VgR	8			
1	LR11	11			
2	insect VgR/YPR	5	6~8		
4	LRP1	2	8	10	11
4	LRP1B	2	8	10	12
4	mammalian LRP2	7	8	10	11
4	*Drosophila* LRP2	3	8	10	11
4	nematode LRP2	6	8	10	11

previtellogenic stage of oogenesis, *Dm*YPR is evenly distributed throughout the arrested oocyte. After transition to the vitellogenic stage, *Dm*YPR relocalizes predominantly to the cortex of oocytes (Schonbaum *et. al.* 2000). Such redistribution of yolk protein receptors upon the onset of vitellogenin uptake also occurs in vertebrates. In previtellogenic oocytes of chicken, *Gallus gallus* (*Gg*), *Gg*VgR/VLDLR is detected in vesicular structures of oocytes. During fast yolk protein uptake, *Gg*VgR/VLDLR relocalizes mainly to cortex of the oocyte (Shen *et. al.* 1993). Mosquito *Aa*VgR migrates as a 205 kDa band under a non-reducing condition or a 214 kDa under a reducing condition, a size similar to that of *Dm*YPR. *Aa*VgR is likely a dimer *in vivo*. In the mosquito, *Aa*VgR presents in the pre-vitellogenic ovary as early as the day of eclosion, and its level rises and decreases paralleling the rate of Vg uptake after initiation of vitellogenesis by a blood meal (Sappington *et. al.* 1995).

Oligomerization of Some LDLR Family Members

Oligomeric forms have been observed in several LDLR family members, including LDLR, *Aa*LpR, *Aa*VgR. On a nonreducing western blot, apart from the monomer-formed *Aa*VgR band, a secondary band was observed with native *Aa*VgR preparation (Sappington *et. al.* 1995, Figure 7), and a 390 kDa rather than a 205 kDa band was observed to be the native-formed *Aa*VgR on a native western blot (Sappington *et. al.* 1995, Figure 6*A*). These secondary bands could represent aggregates or polymerized forms of receptors. The ligand blot (or called far-western blot) result showed strong binding of *Aa*Vg to the monomer form of *Aa*VgR, indicating sufficient affinity of monomer to *Aa*Vg (Sappington *et. al.* 1995).

In *Aedes aegypti*, another LDLR family member, the ovary-formed lipophorin receptor (*Aa*LpR, insect equivalent of the mammalian VLDLR) also forms a dimer. The *Aa*LpR protein expressed in a TNT system showed as a 145 kDa band on a reducing SDS-PAGE. When *Aa*LpR purified from the ovary membrane was subjected to a non-reducing SDS-PAGE, a 250 kDa receptor band was detected by its ligand, lipophorin (personal communication from Jianxin Sun), which suggests the dimmer-formed *Aa*LpR under the nonreducing condition.

In addition to *Aa*VgR and *Aa*LpR, the LDLR was also found to have oligomeric forms. In van Driel's experiment, after the bovine LDLR was solubilized from bovine adrenal cortex membranes and immuno-precipitated, most receptors were found to be presenting as dimers via disulfide bonds (van Driel *et. al.* 1987). These bonds were formed only after homogenization, because covalent bonds could not be formed if tissue was homogenized in the presence of sulfhydryl alkylating agents (alkylated cysteines do not form S-S). The authors demonstrated that the cytoplasmic domain of LDLR is responsible for self-association, and stated that native formed bovine and human LDLRs have capacity to self-associate noncovalently and form dimers and higher order structures (van Driel *et. al.* 1987). Because monomeric forms of LDLR from SDS-PAGE gels bind LDL and β-VLDL on ligand blots, it appears that oligomerization is not crucial for ligand binding (van Driel *et. al.* 1987). In another experiment, when an LDLR mini-receptor covering the ligand-binding domain and one following EGF-like repeat was subjected to a non-reducing SDS-PAGE, at least three bands were observed, with the smallest band corresponding to the size of a monomer (Dirlam *et. al.* 1996). This suggested that LDLR exists as both monomers and multimers, formed possibly throught intermolecular

disulfide cross-linking. The fact that on ligand blots all three bands bound human LDL in a calcium-dependent manner supported the notion that oligomerization is not very important for ligand binding.

MOSQUITO YOLK PROTEINS AND INTERNALIZATION OF VITELLOGENIN BY ITS RECEPTOR

Mosquito Vitellogenesis and Yolk Proteins

In the mosquito *A. aegypti*, vitellogenesis process can be divided into the pre-, post-, and vitellogenic phases. The previtellogenic phase can be further split into a 3-day long preparatory developmental stage when the fat body and ovary gain competence to a blood meal, and a developmental arrest stage when the mosquito seeks a blood meal. The vitellogenisis phase is triggered by a blood meal ingestion, when yolk protein genes are activated and massive yolk proteins are produced in fat bodies and deposited into ovaries. In the postvitellogenic phase, production of yolk proteins is shut off, and the mosquito is ready for another blood meal and a subsequent new cycle of vitellogenesis (Raikhel 1987).

In the female mosquito, *A. aegypti*, most yolk protein precursors (YPPs) are synthesized in fat bodies during vitellogenisis. These YPPs include the most abundant Vg, the second most abundant proenzyme, vitellogenic carboxypeptidase (VCP) (Cho *et. al.* 1991), the proenzyme, vitellogenic cathepsin B (VCB) (Cho *et. al.* 1999), and the lipophorin (Lp) (Sun *et. al.* 2000). During pre-and post-vitellogenisis, these four YPPs are expressed at low levels in female mosquitoes. Following a blood meal that initiates commencement of vitellogenesis, the expressions of YPPs increase rapidly and reach

their peak levels around 24 hours PBM, then decline gradually. This expression pattern is also similar to those of the *Aa*VgR (Sappington *et. al.* 1995) and *Aa*LpR (Cheon *et. al.* 2001).

AaVCP: *A. aegypti* VCP is a serine carboxypeptidase that is synthesized as an inactive latent prenzyme 53 kDa in size by fat bodies of female mosquitoes during vitellogenesis, and secreted to the hemolymph and internalized into the develop oocytes. At the onset of embryonic development, *Aa*VCP is activated and its size is reduced to 48 kDa (Cho *et. al.* 1991). This active form of VCP is maximally present in the middle of embryonic development and disappears by the end (Cho *et. al.* 1991). Noticing that some of the serine carboxypeptidases are implicated in proteolytic activation of a number of enzymes or biologically active molecules, it was suspected that *Aa*VCP activates hydrolytic enzymes involved in the degradation of yolk proteins in developing embryos, or directly degrades yolk proteins (Cho *et. al.* 1991). On a preliminary ligand blot from Cho's laboratory (personal communication from Wenlong Cho), radioactively labeled *Aa*VCP bound an ovarian membrane protein band larger than the 214 kDa *Aa*VgR monomer, which suggested a putative new receptor on the ovary membrane that recognizes *Aa*VCP.

AaVCB: *A. aegypti* VCB was given its name for the sake of its high similarity to mammalian cathepsin B (Cho *et. al.* 1999). The secreted *Aa*VCB in the hemolymph is a large proenzyme, likely a hexamer that is consisted of 44 kDa subunits. After internalization into oocytes, the size of *Aa*VCB decreases to 42 kDa. In the mature yolk body, *Aa*VCB is located in the matrix surrounding *Aa*Vn. At the onset of embryogenesis,

*Aa*VCB is further processed to a 33 kDa active form that is possibly involved in the embryonic degradation of Vg (Cho *et. al.* 1999).

***AaLp and AaLpR*:** The oocyte development in insects involves the accumulation of large amounts of lipids mostly originated extra-ovarianly and delivered by Lp in the insect hemolymph. Lp is basically composed of a 230~250 kDa apoLipophorin-1 (apoLp-1) and a 70~85 kDa apoLp-II. Based on the lipid amount it carries, Lp could be classified as high-density lipophorin (HDLp) or low-density lipophorin (LDLp). In the insect body, Lp mainly functions as a reusable lipid shuttle transporting lipids among lipid loading sites in the midgut, storage sites, and metabolism sites.

The Lp-mediated lipid delivery into the developing oocytes could take place in two ways. The major vehicle is LDLp. In response to a stimulatory signal, HDLp associated with two apoLp-IIIs loads its major lipid component, diacylglycerol, from fat bodies and converts to LDLp. LDLp is associated with apoLp-III in proportional to the amount of diacylglycerol incorporated (Soulages and Wells 1994). Upon reaching delivery sites, LDLp unloads its lipid, releases apoLp-III subunits, and converts back to HDLp. Internalization by cells is not needed for this lipid loading and unloading event. There are also some lipids that are delivered by HDLp, and internalized by the developing oocytes via receptor-mediated endocytosis. The internalized HDLp unloads lipid and apoLp-III, and is converted to very-high-density Lipophorin (VHDLp).

The Lps from mosquitoes, *A. aegypti*, *Anopheles albimanus* (Wiedemann), and *Culex quinquefasciatus* (Say) are all HDLps. These mosquito Lps have triacylglycerol as their major neutral lipid component, in contrast to diacylglycerol in other insect species. *Aa*Lp has a hydrated density of 1.113 g/ml and contains 49% of lipid and 3.2% of

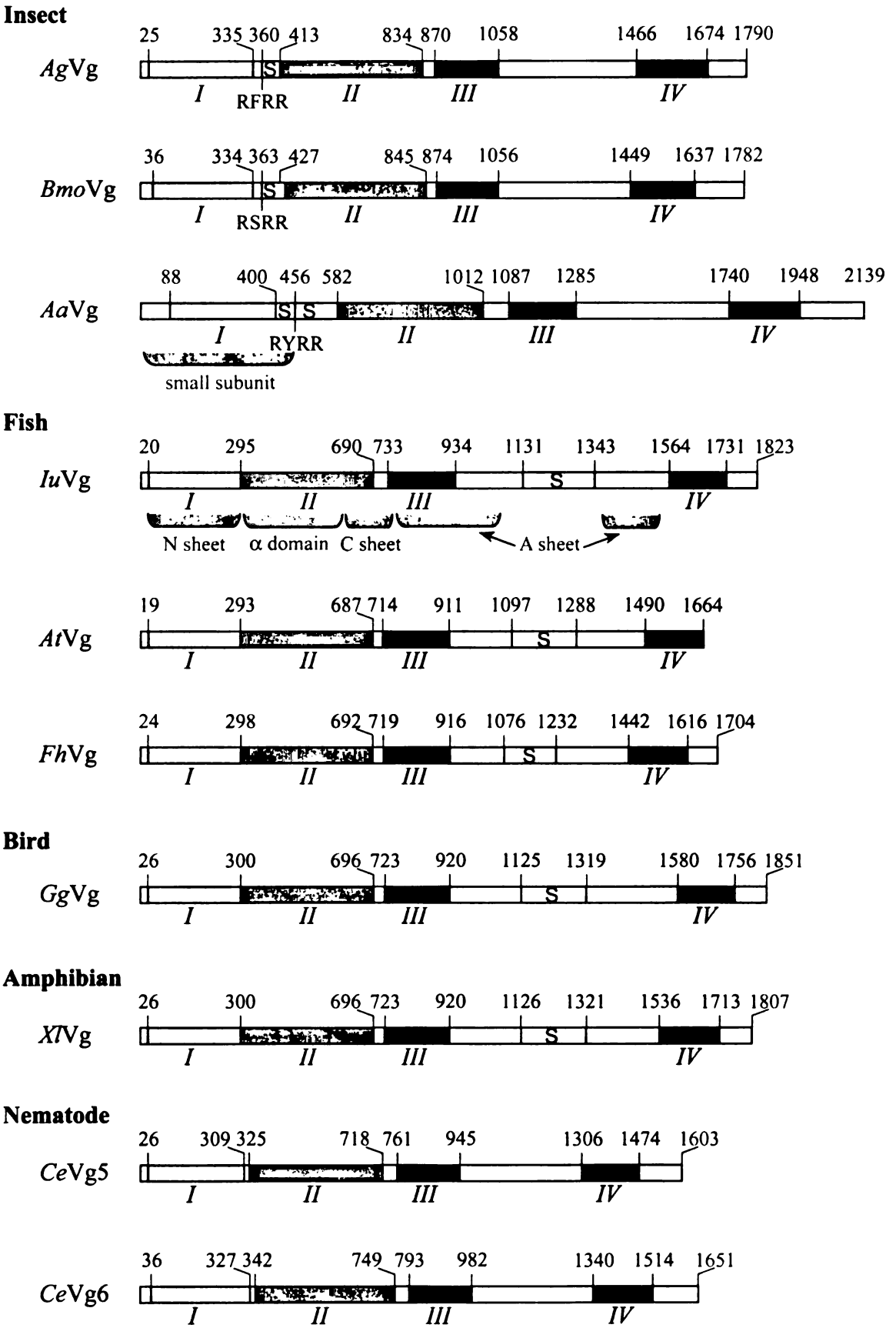
carbohydrate (Ford and Van Heusden 1994). The native formed *Aa*Lp is 480 kDa in size, and contains one 238 kDa apoLp-1 and one 73 kDa apoLp-II subunits that are both glycosylated (Ford and Van Heusden 1994).

The ovary-formed *Aa*LpR is a 129 kDa long membrane protein that is present only in female germ line cells and expressed early during oocyte development (Cheon *et. al.* 2001). *Aa*LpR shares highest similarity with the locust LpR and high homology with numerous VLDLRs. The ovarian *Aa*LpR internalizes *Aa*Lp into the developing oocyte (Cheon *et. al.* 2001). Data from Alexander Raikhel's laboratory suggested that both *Aa*VCP and *Aa*VCB bind *Aa*Lp, which binds the *Aa*LpR. In another word, the *Aa*VgR binds only *Aa*Vg but not other three minor yolk proteins on the oocyte cell surface.

*Aa*Vg Synthesis in Fat Bodies and Internalization by Ovaries

In egg-laying animals, Vg is a very-high-density lipoprotein secreted by the liver of vertebrates, the intestine of nematode, and the fat body of insect. Vg plays a fundamental role during vitellogenesis (Wallace 1985; Byrne *et. al.* 1989). Insect Vg is a high molecular weight oligometric phosphoglycolipoprotein with 7-15% lipids, mainly phospholipids and diacylglycerol (Raikhel and Dhadialla 1992). In most Insects, Vg precursor is cleaved in fat bodies into the large and small subunits that oligomerize to form a monomer Vg, while in vertebrate, Vg is synthesized in the liver and cleaved into several subunits only after internalization by the oocyte. Figure 2 shows the conservation of Vg sequences among invertebrate and vertebrate animals.

Fig. 2. Conservation of vitellogenin sequences among egg laying animals. The vitellogenin (Vg) sequences from insects, nematode, and vertebrates share five homologous regions (shaded regions) and first four were numbered. The region I in the *Iu*Vg forms the N sheet, which is corresponding to the small subunit of *Aa*Vg. Numeral, amino acid position; S, polyserine region; RXRR, cleavage signal in insect Vgs; *Aa*, *Aedes aegypti* (yellow fever mosquito); *Ag*, *Anopheles gambiae* (African malaria mosquito); *At*, *Acipenser transmontanus* (white sturgeon); *Bmo*, *Bombyx mori* (domestic silkworm); *Ce*, *Caenorhabditis elegans* (nematode); *Fh*, *Fundulus heteroclitus* (mummichog), *Gg*, *Gallus gallus* (chicken); *Iu*, *Ichthyomyzon unicuspis* (lamprey); *Xl*, *Xenopus laevis* (African clawed frog).


Insect
AgVg
25 335 360 413 834 870 1058 1466 1674 1790
S
I RFRR II III IV
BmoVg
36 334 363 427 845 874 1056 1449 1637 1782
S
I RSRR II III IV
AaVg
88 400 456 582 1012 1087 1285 1740 1948 2139
S S
I RYRR II III IV
small subunit
Fish
IuVg
20 295 690 733 934 1131 1343 1564 1731 1823
S
I II III IV
N sheet α domain C sheet A sheet
AtVg
19 293 687 714 911 1097 1288 1490 1664
S
I II III IV
FhVg
24 298 692 719 916 1076 1232 1442 1616 1704
S
I II III IV
Bird
GgVg
26 300 696 723 920 1125 1319 1580 1756 1851
S
I II III IV
Amphibian
XlVg
26 300 696 723 920 1126 1321 1536 1713 1807
S
I II III IV
Nematode
CeVg5
26 309 325 718 761 945 1306 1474 1603
I II III IV
CeVg6
36 327 342 749 793 982 1340 1514 1651
I II III IV

In *A. aegypti*, a blood meal activates the translation of the 224 kDa long pre-pro-Vg in the rough endoplasmic reticulum (ER). This pre-pro-Vg is cotranslationally glycosylated and posttranslationally phosphorylated in the rough ER to produce a 250 kDa long pro-Vg (Dhadialla and Raikhel 1990), which is rapidly cleaved into the 190 and 62 kDa long subunits by a Vg convertase (Chen and Raikhel 1996). This Vg convertase is a member of a subtilisin-like proprotein endoprotease family that recognizes the motif, (R/K)x(R/K)R or RxxR, with a juxtaposed β turn for optimal recognition (Brakch *et. al.* 1993), and cleaves immediately after this motif (Barr 1991; Rouille *et. al.* 1995). Both Vg fragments then enter the Golgi complex, are sulfated and further glycosylated to form a 200 kDa long large subunit and a 66 kDa long small subunit. In the Golgi complex, two subunits oligomerize to form the 380 kDa long mature *Aa*Vg, which is packaged into condensing vacuoles that develop further into large secretory granules, and is finally released into the hemolymph (Dhadialla and Raikhel 1990). Figure 3 depicts the synthesis and processing of *Aa*Vg in the fat body.

In vertebrates, Vg is synthesized in the liver, secreted into the blood stream and internalized by oocytes. In oocytes, Vg is proteolytically cleaved into several polypeptide chains. The lipid-binding product is renamed lipovitellin (LV), which is comprised of a heavy chain (LV1) of approximately 120 kDa and a light chain (LV2) of about 30 kDa (Byrne *et. al.* 1989). LV can load varying amounts (up to 16% w/w) of noncovalently bound lipid, 2/3 of which are phospholipids (Ohlendorf et. al. 1977; Norberg and Haux 1985). The polyserine region of vertebrate Vg is also released. This released small peptide is named phosvitin (28-35 kDa) or phosvette (13-19 kDa), because it is serine-

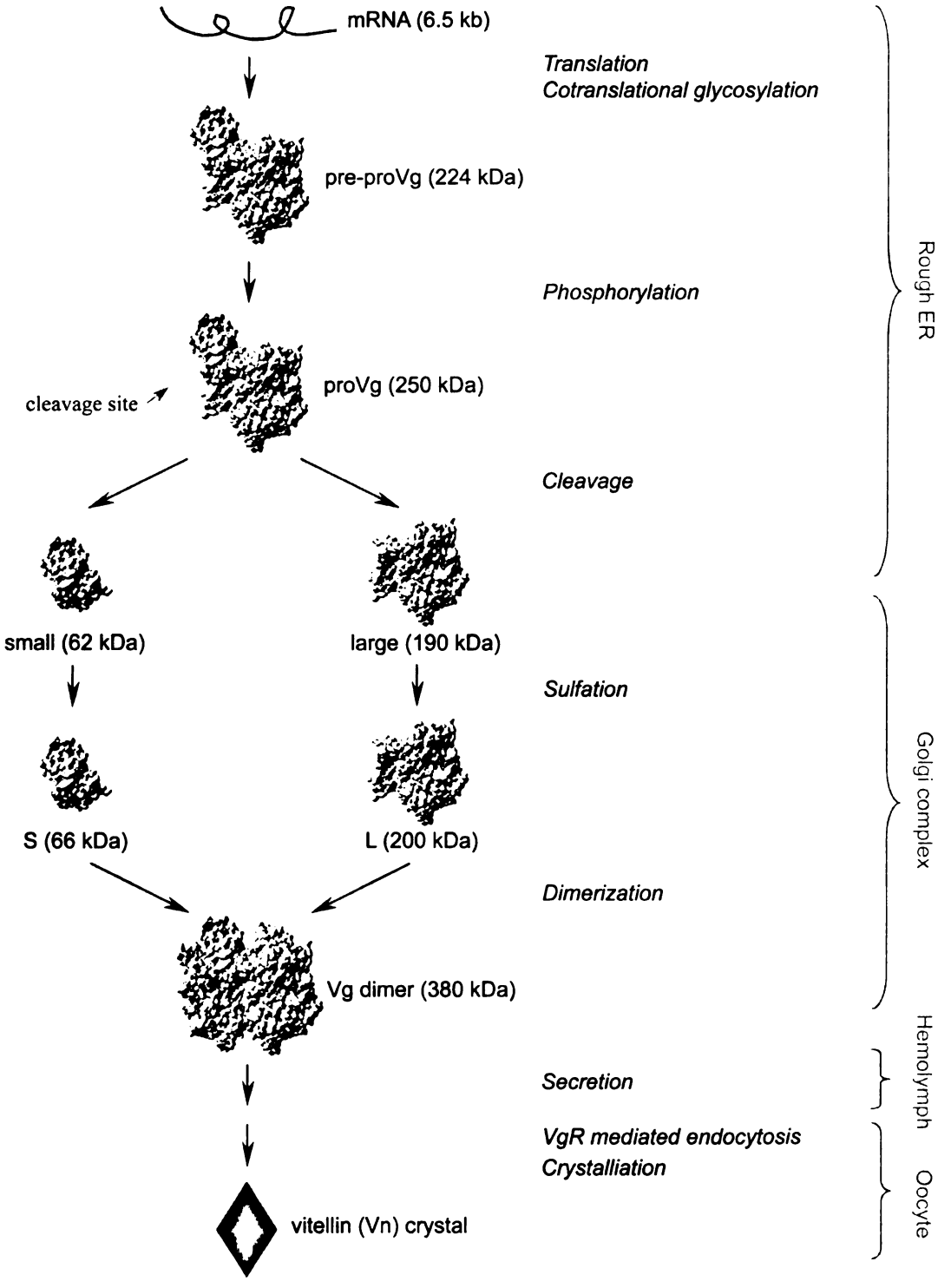


Fig. 3. Synthesis and processing of *Aa*Vg in the fat body. (modified after Dhadialla and Raikhel 1990)

enriched (up to 50%), and most of serine residues are phosphorylated (Byrne *et. al.* 1989).

After *Aa*Vg is internalized into the oocyte by the *Aa*VgR, it is discharged from the receptor, stripped of its lipids, and crystallized in the storage form called vitellin (Vn). The *Aa*VgR is then recycled to the oocyte cell surface (Snigirevskaya *et. al.* 1997). The details of this process are shown in Figure 4.

Forces Mediating Receptor-Ligand Interaction

All of the CRs of LDLR family members have a highly conserved acidic Ca^{2+}-binding motif that contributes to the negatively charged surface of the C terminal moiety. Negative charges have long been proposed to be primarily responsible for the recognition of ligands by receptors. Comparisons among 3-D structures of CR5 and CR6 of LDLR and CRII-6 (the sixth module in the CLII of CRs) of LRP showed a region of negatively charged surface electrostatic potential (EP) surrounding the coordinated Ca^{2+} ion (North and Blacklow 2000, Figure 5). In contrast to CRs, The two EGF-like repeats of LDLR lack the concentration of negatively charged residues found in CRs.

Among numerous ligands of the LDLR family members, only ApoE and receptor-associated protein (RAP) bind to all mammalian LDLR family members. The several identified potential receptor-binding sites on various ligands share no significant sequence homology except that they are all rich in basic residues (lysine (K) and arginine (R)) (Table 5). More than one decade ago, Roehrkasten and Ferenz reported K/R residues on locust Vg to be important for binding of VgR on the oocyte membrane (Roehrkasten and Ferenz 1992). Suramin is a negatively charged compound called polysulfated

Fig. 4. Internalization of Vg and recycling of VgR in mosquito oocytes. After mosquito Vg is delivered to the oocyte though the hemolymph, Vg enters between follicle cells and through pores of the vitelline envelope, and binds to VgR on the oocyte plasma membrane. Vg/VgR complexes cluster in clathrin-coated pits (CCP), which pinch-off to form coated vesicles. Upon losing the clathrin coat, the coated vesicles are transformed into early endosomes, which fuses with one another to form late endosomes, or transitional yolk bodies. At endosomal pH, Vg is released from VgR and delivered to the mature yolk bodies where it is crystallized and stored until the onset of embryonic development. Released VgRs are recycled to the oocyte surface via tubular compartments (TC). The figure was drawn after Snigirevskaya *et. al.* (Snigirevskaya *et. al.* 1997).

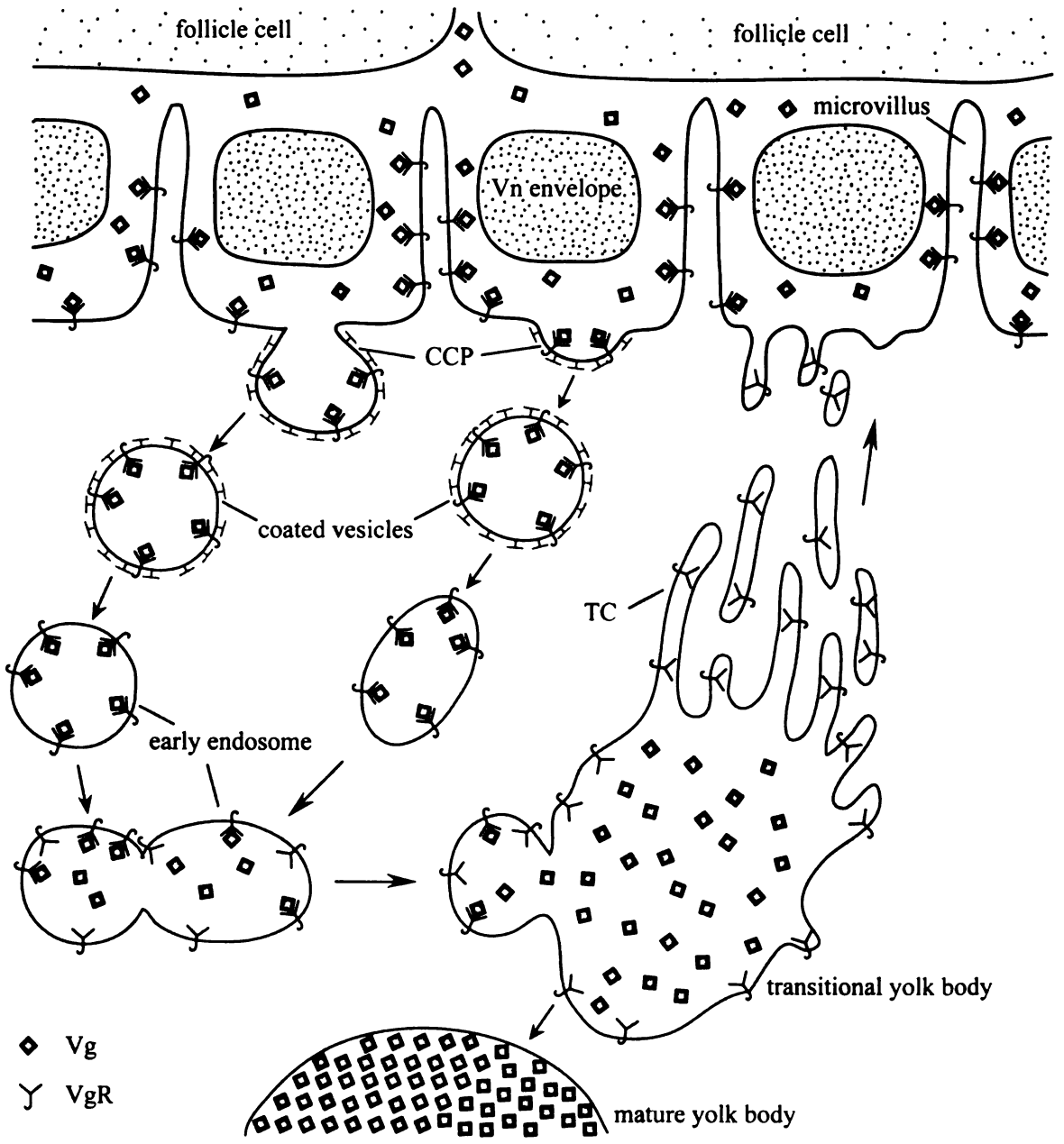

follicle cell
follicle cell
microvillus
Vn envelope
CCP
coated vesicles
TC
early endosome
transitional yolk body
Vg
VgR
mature yolk body

Table 5. Sequences in various ligands critical for receptor binding.

Ligands	Location	Sequence*	References
Activated form of α_2-M (α_2-M*)	1366–1392	**FIPLKPTVKMLERSNHVSRTEVSSNHV**	Nielsen 1996
ApoE	134–150	**RVRLASHLRKLRKRLLR**	Wilson *et. al.* 1991; Zaiou *et. a.* 2000
ApoB	3359–3367	**RLTRKRGLK**	Boren *et. al.* 1998
Aprotinin	38–47	**CRAKRNNFKSA**	Moestrup *et. al.* 1995
Lipoprotein lipase	380–384 403–425	**LKWKS** **KIRVKAGETQKKVIFCSREKVSHL**	Nielsen *et. al.* 1997
plasminogen activator inhibitor 1 (PAI-1)	69–87	**DKGMAPALRHLYKELMGPWN**	Stefansson *et. al.* 1998
Pseudomonas exotoxin A	54–67	**DALKLAIDNALSIT**	Kounnas *et. al.* 1992
vitellogenin	182-189	**HLTKTKDL**	Li *et. al.* 2003

* bald blue letters, basic residues at neutral pH (K, R); bald green letter, acidic residues (D, E).

polycyclic hydrocarbon. Suramin blocks the binding of many ligands (including vertebrate and insect Vgs) to their receptors, and dissociates binding by competition. This implies that ionic bonds play a crucial role in Vg/VgR interaction. Recently, Li *et. al.* claimed a K/R-rich VgR-binding site on tilapia Vg (Li *et. al.* 2003).

INTRODUCTION TO THIS THESIS WORK

Questions Raised on *Aa*VgR and My Hypotheses

It is already accepted that the single cluster of CRs in one-cluster LDLR family members mediates binding of ligands, while for two- and four-cluster family members things are much more complicated. Taking LRP as an example, LRP has 4 clusters of CRs with 2, 8, 10, 11 modules in each cluster. Studies with LRP mini-receptors showed that both the CLII and CLIV bind the same 12 kinds (including α_2-M*) of known ligands (Neels *et. al.* 1999; Willnow *et. al.* 1994). In contrast, the CLIII binds only ApoE and binds RAP weakly (Neels *et. al.* 1999), and the CLI has only low affinity to α_2-M* (Mikhaihenko *et. al.* 2001). In comparison to either the CLI or CLII, a LRP mini-receptor covering both clusters has much higher affinity to α_2-M* (Mikhaihenko *et. al.* 2001), which indicates that the CLI and CLII cooperate to generate a high-affinity binding site for α_2-M*.

From the LRP binding results, we learned that for four-cluster family members such as LPR, different clusters of CRs could have anywhere from nearly the same to very different ligand spectrum(s). But how about the two clusters in insect VgRs/YPRs? The VgRs in egg-laying vertebrate animals (for example, bird, fish, and amphibian) have only

one cluster of eight CRs, and the nematode YPR also has one cluster of merely five CRs. Is it possible that the first cluster of five CRs in the insect VgR/YPR and the single cluster of five CRs in nematode YPR are evolved from a common ancestor, while the second cluster of CRs in the insect VgR/YPR and the single cluster of CRs in some one-cluster LDLR family members are evolved from another common ancestor?

Another question on the ligand-binding domain of *Aa*VgR concerns the function of one additional cluster of CRs in insect VgRs/YPRs. One possible answer is that there are different ligand spectrums for CLI and CLII. While a surprising fact is, unlike chicken VgR/VLDLR that also imports VLDL, riboflavin-binding protein, and α_2-M in addition to Vg into oocytes (Schneider 1996), Vg is the only ligand for known insect (mosquito, fruitfly, and cockroach) VgRs. Another possible explanation is the strengthened binding of Vg through forming either a large interdomain ligand-binding site with higher affinity to Vg or two separate ligand-binding sites, with each site binding one copy of Vg independently.

Until now, no comparison on the ligand-binding properties of CLI and CLII of CRs in insect VgR/YPR has been reported. An interesting question on *Aa*VgR is which one of the two clusters of CRs is predominantly or exclusively responsible for high affinity to its ligand, *Aa*Vg?

The negative surface charges of some LDLR family members have been proposed to be primarily responsible for recognition of ligands. While Fass *et. al.* solved the structure of the CR5 of LDLR, and showed that acidic residues in the conserved acidic motif of CR5 are buried to participate in Ca^{2+} coordination, rather than being exposed to the surface (Fass *et. al.* 1997). Based on nearly neutral surface EP of the LDLR CR5 they

showed (Fass *et. al.* 1997, Figure 4), they suggested a hydrophobic concave face of the LDLR CR5 rather than a negative surface to be responsible for interaction with ligands.

To elucidate the force mediating the *Aa*VgR-*Aa*Vg interaction, a check on the molecular surface EPs of individual modules from both clusters of CRs in *Aa*VgR can provide a good answer. Comparison of surface charges of modules from both clusters can also tell us which cluster has a stronger negatively charged surface and thus higher affinity to *Aa*Vg, given that most modules have predominantly negative surfaces. Modeling the small and large subunits of *Aa*Vg can also allow me to evaluate, from the ligand side, the hypothesis that the *Aa*VgR-*Aa*Vg interaction is mediated mainly by the negative-positive electrostatic attraction.

The YWTD β propeller of LDLR has a negatively charged top face on the cell surface. In the crystallized form, this top face contacts the Ca^{2+}-binding loops of CR4 and CR5 of LDLR at endosomal pH (pH5.3) (Rudenko *et. al.* 2002). Because both CR4 and CR5 are critical for lipoprotein binding, their association with YWTD β propellers (rather than ligands) at endosomal pH proposed a mechanism for lipoprotein release in the endosome by replacing a real ligand of LDLR with a false one, the YWTD β propeller of LDLR (Rudenko *et. al.* 2002). An investigation of the CR4 and CR5 docking face of the LDLR YWTD β propeller found on the docking face two histidine residues that potentially change the surface charge of this face from negative to positive at endosomal pH (Jeon and Blacklow 2003, Figure 4). To evaluate whether the proposed mechanism of LDLR ligand release is applicable to other LDLR members—for example, *Aa*VgR—an investigation must be done of the surface properties of three assumed YWTD β propellers of *Aa*VgR and evidence of histidine residues must be sought of the

surface of assumed YWTD β propellers of *Aa*VgR that can potentially switch the surface charge to positive.

Although CRs of the LDLR family members are primarily responsible for ligand binding, the EGF precursor homology domain of LDLR (including the YWTD β propeller and its flanking EGF-like repeats) were shown to be also important for binding of LDLR to some ligands. Davis *et. al.* reported that an LDLR mini-receptor lacking the EGF precursor homology domain showed markedly reduced affinity to LDL on the cell surface, and complete degradation in the transfected cells after incubation with β-VLDL on the cell surface (Davis *et. al.* 1987). The EGF precursor homology domain was also found to be important for *in vivo* binding of LDLR to LDL on the cell surface, but not for *in vitro* binding (Davis *et. al.* 1987). In another experiment, binding analyses of LDLR mini-receptors showed that the first EGF-like repeat was required for binding of LDL, but not β-VLDL, while the second EGF-like repeat was not required for ligand binding (Esser *et. al.* 1988). To investigate whether any of seven EGF-like repeats of *Aa*VgR can contribute to the assumed electrostatic complementarities between *Aa*VgR and *Aa*Vg, it is necessary to check the surface charge distribution of each EGF-like module in *Aa*VgR.

Recently, a "VgR-binding motif" on the small subunit of blue tilapia Vg was reported by Li *et. al.* (Li *et. al.* 2003). It would be useful to investigate the corresponding regions on lamprey Vg and *Aa*Vg and evaluate whether these corresponding regions could be the VgR-binding motifs or not.

Methodology and Predicted Outcomes

To find evidence to support or oppose my assumption that the CLI in *Aa*VgR and the cluster of five CRs in nematode YPR are evolved from one common ancestor, while the CLII in *Aa*VgR and the single cluster of CRs in some of the one-cluster LDLR family members are evolved from another common ancestor, a multiple-sequence alignment of CRs from investigated species can provide a sound prediction. If the alignment shows that the five CRs of nematode YPR are more homologous to the CLI than the CLII of CRs in insect VgR/YPR, and the single cluster of CRs in some one-cluster members is more homologous to the CLII in insect VgR/YPR, then my assumption is correct. If the alignment shows that the five CRs of nematode VgR are more homologous to the CLII than the CLI in insect VgR/YPR, then the CLI of five CRs in insect VgR/YPR is evolved from a different ancestor.

To test whether one cluster of CRs in *Aa*VgR is predominantly or exclusively responsible for the binding of *Aa*Vg, saturation binding assays need to be done to directly measure the dissociation constant for each cluster. For this purpose, mini-receptors with either one or both clusters of CRs, together with the flanking EGF-precursor homology domain, need to be constructed. To make sure mini-receptors have correct post-translational modification and thus biological activities, mini-receptor genes need to be expressed in an insect cell line. Because the *Drosophila melanogaster* VgR (*Dm*VgR) is structurally very similar to the *Aa*VgR, and because the *Drosophila* S2 cell line is commercially available, it is reasonable to express mini-receptor genes in *Drosophila* cells. Even though the reproduction system of *Drosophila* is somewhat similar to that of the mosquito, it is still possible that the truncated VgR may not be fully processed and modified in the *Drosophila* system. If saturation ligand-binding assays show no affinity

of mini-receptors to *Aa*Vg and there are no mutations in the coding regions of mini-receptors, other insect expression systems (preferably mosquito), with the exception of fruit fly, should to be tried later. If ligand-binding results show high affinity for one cluster and much lower affinity for another and the higher affinity approaches that of the full-length *Aa*VgR, then one cluster is predominantly responsible for *Aa*Vg binding. If the affinities of both clusters are much lower than that of the full-length *Aa*VgR, then both clusters contribute to the high affinity, probably in a synergistic way.

To test whether the negative-positive charge attraction is the main force mediating the *Aa*VgR-*Aa*Vg interaction, evidence from both sides needs to be obtained. On the receptor side, it will be necessary to check to see if most of the CRs of *Aa*VgR have negatively charged surfaces and if some of them have strongly negatively charged surfaces. This can be done by modeling thirteen CRs of *Aa*VgR and calculating surface EP of each module. If the results show that some modules do have negatively charged surfaces, and some modules have strong negative surfaces, then I can compare the surface EPs of the two clusters. Theoretically, the cluster with stronger negative surface has higher affinity to *Aa*Vg. If the results show that most modules have nearly neutral or even positive surfaces, then the force mediating the *Aa*VgR-*Aa*Vg interaction is primarily not negative-positive charge attraction. On the ligand side, a partial or whole *Aa*Vg needs to be modeled, and surface EP of the modeled region needs to be calculated to find expected positively charged surface patches. Because the 3-D structure of the mature form of lamprey Vg has been solved by X-ray crystallography (Anderson *et. al.* 1998; Raag *et. al.* 1988; Thompson and Banaszak 2002), modeling *Aa*Vg becomes possible. If

the surface of mature lamprey Vg has no large positively charged patches, then electrostatic attraction is not the dominant force mediating the VgR-Vg interaction.

To investigate whether EGF-like repeats of *Aa*VgR contribute to the assumed electrostatic complementarities between *Aa*VgR and *Aa*Vg, models for each module of seven EGF-like repeats will be made, and then the surface EP of each modeled module will be calculated. One possible result would be that some of the EGF-like repeats, especially the EGF-like repeats juxtaposed the boundary CRs in each cluster, may have concentrated negatively charged residues. If that is true, then these EGF-like repeats are possibly involved in binding to *Aa*Vg, together with CRs. A second possibility is that there are no concentrated negatively charged surface patches on any EGF-like repeats of *Aa*VgR. It could then be concluded that no EGF-like repeats contribute to negative-positive charge complementarities between *Aa*VgR and *Aa*Vg.

To evaluate whether any of the assumed three YWTD β propellers of *Aa*VgR potentially act as a false ligand of the ligand-binding domain of *Aa*VgR, three assumed YWTD β propellers of *Aa*VgR will be modeled, if possible. If they really have β propeller structures, then calculation of the surface EP of each YWTD domain will be done and a search for surface histidine residues will be performed. If the three YWTD domains do not have six-bladed β propeller structures, then it is impossible for YWTD domains of mosquito VgR to play a critical role in Vg release and VgR recycling. If at least some YWTD domains have six-bladed β propeller structures but there is only one histidine residue on the surface, or if there are several histidine residues on the surface but they are scattered and located not in positive or nearly neutral surface areas, then it would still be somewhat possible that the YWTD β propellers of *Aa*VgR contact a

ligand-binding domain in endosomes, even though the conversion to the false ligand in eondosomes is not switched by histidine residues. If some YWTD domains have six-bladed β propeller structure, there are at least two histidine residues on the surface, and some surface histidine residues are gathered and located in a negative or nearly neutral surface area, then these gathered surface histidine residues potentially could switch the YWTD β propeller into false ligand of the ligand-binding domain of *Aa*VgR at endosomal pH.

To test whether in *Aa*VgR there is a counterpart of the "VgR-binding motif" found in tilapia Vg, a multiple alignment of Vgs will need to be done. An inspection of the corresponding region of lamprey Vg would be very persuasive in evaluating popularity of this "VgR-binding motif" in fishes and insects. If in some insect Vg the corresponding regions lack the feature of a receptor-binding motif, and/or if the corresponding region in lamprey Vg does not make the surface EP of this region more positive than most other surface regions, then this clamed "VgR-binding motif" is at least not applicable to the species investigated. If the opposite is the case, then it is quite possible that the counterpart in *Aa*Vg is a VgR-binding motif.

Chapter 2

Both Clusters of Complement-type Repeats in *Aa*VgR Bind *Aa*Vg

INTRODUCTION

This chapter focuses on two questions. First, a multiple alignment of the CLI and CLII of CRs in insect VgRs/YPRs with the single cluster of CRs in one-cluster LDLR family members will be performed to evaluate whether it is possible that the single cluster of five CRs in nematode YPR and the CLI of five CRs in insect YPRs/VgRs are evolved from a common ancestor. Another question is whether one cluster of CRs in *Aa*VgR is predominantly or exclusively responsible for the binding of *Aa*Vg. If both the CLI and CLII in *Aa*VgR bind *Aa*Vg, then we should try to find out which cluster has higher affinity to AaVg. To answer this question, mini-receptors with either one or both cluster(s) of CRs will be constructed and transfected into a *Drosophila* line, and the expressed mini-receptor proteins will be used in saturation binding assays to measure the dissociation constant for each cluster.

MATERIALS AND METHODS

Animals

Mosquitoes, *Aedes aegpti*, were reared as described (Hays and Raikhel 1990). Vitellogenesis was initiated 3~5 days after eclosion with a blood meal on rats.

Construction of Mini-receptors

Two *Aedes aegpti* vitellogenin mini-receptor (Aa_{m}VgR) genes, which encode the signal peptide, one of the two clusters of CRs, and one following β propeller fold with its three juxtaposed EGF-like repeats, were obtained by Polymerase Chain Reaction (PCR) with PCR SuperMix High Fidelity (Invitrogen). The third $_{m}VgR$ gene encoding the

extracellular portion of *Aa*VgR except for the *O*-linked sugar domain was constructed as a positive control. The plasmid pBlue-VgR2.7 was constructed earlier by inserting a 2.7 kb long *Aa*VgR cDNA fragment from an *A. aegpti* ovary cDNA library to a pBluescript vector (Stratagene) (Sappington *et al.* 1996). The plasmid pG-VgR3.2 was built earlier by inserting a 3.2 kb long *Aa*VgR cDNA fragment from a 3′ RACE of an *A. aegpti* ovary total RNA into the vector pGEM-5 (Promega) (Sappington *et al.* 1996). Identities of all constructs were checked by both restriction analysis and sequencing.

Mini-receptor $_m$VgR1: A 1920 bp-long *AaVgR* cDNA fragment encoding the CLI of CRs and the first YWTD β propeller (YWTD1) juxtaposed by three EGF-like repeats (EGF1, EGF2, and EGF3) was amplified from the template pBlue-VgR2.7. The primer set was composed of the forward primer P1, 5′-AATGAGCTCTCGAGTTTGATGGGCGCGATC-3′ with restriction sites *Sac* I and *Xho* I introduced, and the reverse primer P3, 5′-GAAGGGCCCGCGGCAAGAATGCTTG-3′ with *Apa* I site introduced. The 1897 bp-long *Sac* I-*Apa* I fragment of the amplified DNA was inserted to the same sites on the cloning vector, pGEM5Zf(+) (3003 bp, Promega) to construct the 4831 bp-long plasmid, pGmVgR1. The 1898 bp-long *Xho* I-*Apa* I fragment of pGmVgR1 was inserted to the same sites on the expression vector, pAc5.1/V5-HisA (Invitrogen), in the 5′ end of a 83 bp-long sequence encoding a V5 epitope and a six-residue polyhistidine (6×His) tag. The constructed 7266 bp-long plasmid was named pAmVgR1.

Mini-receptor $_m$VgR2: A 115 bp-long cDNA encoding the signal peptide of *Aa*VgR was amplified from the template, pBlue-VgR2.7, with the forward primer P1, and the reverse primer P2, 5′-TGCCTGCAGGATCCACGCTTCCGAAG-3′, with *Bam*H I

and *Pst* I sites introduced. The 97 bp-long *Sac* I-*Pst* I fragment of the amplified DNA was inserted to the same sites on pGEM5Zf(+) to construct the 3088 bp-long plasmid, pG-signal. A *Bam*H I site was conveniently introduced in the 3′ end of the signal peptide coding sequence. The 2223 bp-long *AaVgR* cDNA fragment encoding the CLII of CRs and YWTD3 juxtaposed by 3 EGF-like repeats (EGF5, EGF6, and EGF7) was amplified from the template, pG-VgR3.2. The primer set was composed of the forward primer P6, 5′- GCGGATCCCTGCGAGTTCAAGTGTACC -3′ with *Bam*H I site introduced, and the reverse primer P7, 5′- GTTGGGCCCTTGCGGACAGATGTCC-3′ with *Apa* I site introduced. The 2206 bp-long *Bam*H I-*Apa* I fragment of the amplified DNA was inserted to the same sites on pG-signal to construct the 5233 bp-long plasmid, pGmVgR2(a). The 2305 bp-long *Xho* I-*Apa* I fragment of pGmVgR2 was inserted to the same sites on pAc5.1/V5-HisA to construct the 7668 bp-long plasmid, pAmVgR2.

***Mini-receptor $_m$VgR1-2*:** A 2254 bp-long *AaVgR2.7* fragment was obtained by PCR with the template, pBlue-VgR2.7, and the primer set, P1/P5. The sequence of the reverse primer, P5, is 5′-CCCGGTGGCGAGTCTGGAG-3′. The 2866 bp-long *Aa*VgR3.2 fragment was obtained by PCR with the template, pG-VgR3.2, and the primer set, P4/P7. The sequence of the forward primer, P4, is 5′-CCCGGTGGCGAGTCTGGAG-3′. The purified *AaVgR2.7* and *AaVgR3.2* PCR fragments were taken as two overlapping templates in the second tier of PCR with the primer set, P1/P7. The obtained 5059 bp-long *Sac* I-*Apa* I fragment encoding the extracellular portion of *Aa*VgR except for the *O*-linked sugar domain was inserted to the same sites on pGEM5Zf(+) to construct the 8008 bp-long plasmid, pG_mVgR1-2(a). A

5080 bp-long *Xho* I-*Apa* I fragment of pG_mVgR1-2 was inserted to same sites on pAc5.1/V5-HisA to construct the 10443 bp-long plasmid, pA_mVgR1-2.

Reversion of Point Mutations on Mini-receptors

Sequencing was performed in the Genomics Technology Support Facility at Michigan State University. Three mini-receptor genes were fully sequenced, and four clones of *VgR3.2* cDNA—pG-VgR3.2, pG-VgR3.2-1a, pG-VgR3.2-2a, and pG-VgR3.2-2b—were partially sequenced. A nonsense mutation (C3046G) was found near the 5′ portion of a sequence coding for the CLII of CRs in plasmids pG_mVgR2, pG_mVgR1-2, pG-VgR3.2, pG-VgR3.2-2a, and pG-VgR3.2-2b. pG-VgR3.2-1a does not have this mutation. In the CLII of CRs in the *$_mVgR1$-2* gene, one more point mutation (A3331G) was detected that replaced the asparagines (N) residue with the aspartic acid (D) residue.

***Reversion of pG_mVgR2 mutants*:** The mVgR2 coding sequence was reamplified from an alternative template, pG-VgR3.2-1a, to avoid the nonsense mutation from the original template, pG-VgR3.2. The amplified fragment was cloned on pG-signal to construct the plasmid, pG_mVgR1-2(b). Sequencing of three clones of pG_mVgR2(b) and the template, pG-VgR3.2-1a, detected in the C terminal region eight common point mutations from pG-VgR3.2-1a that are absent in pG-VgR3.2. To reverse these eight point mutations in the C terminal region of the *$_mVgR2(b)$* gene, a 887 bp-long *Xho* I-*Pst* I fragment of pG_mVgR2(a), which encodes the signal peptide and ¾ of the CLII of CRs, was replaced with its equivalent on pG_mVgR2(b) to construct the reversed plasmid, pG_mVgR2.

***Reversion of pG $_m$VgR1-2 mutants*:** There are respectively 1, 2, and 3 *Eco*R I site(s) on pG $_m$VgR2, pG $_m$VgR1-2 and pG-VgR3.2-1a. A 1122 bp-long fragment from *Eco*R I digestion of pG-VgR3.2-1a was ligated to a 6886 bp-long fragment from *Eco*R I digestion of pG $_m$VgR1-2 to reverse the nonsense mutation. Preferred orientation of the insert was indicated by a 468 bp (rather than 653 bp) fragment from *Eco*R V/*Pst* I double digestion. The reversed construct was named pG $_m$VgR1-2.

***In Vitro* Expression of Mini-receptor Genes in a Coupled Transcription and Translation System**

The mini-receptor gene driven by a phage SP6 RNA polymerase promoter on the pGEM5 vector was expressed in the TNT® SP6 quick-coupled reticulocyte lysate transcription and translation system (Promega), and labeled with [^{35}S] methionine (PerkinElmer) as described in technical manual. The Luciferase SP6 Control DNA (Promega) was used as a positive control. Reactions were subjected to the reducing sodium dodecylsulfate–polyacrylamide gel (SDS-PAGE) and autoradiography.

Maintenance of Insect Cell Line

Drosophila melanogaster Schneider 2 (S2) cell (Invitrogen) culture was initiated from frozen stock and maintained generally as described in the manual. The GIBCO™ *Drosophila* Serium-Free Medium (*Drosophila*-SFM, Invitrogen) with L-glutamine, penicillin, and streptomycin at the final concentrations of 16.4 mM, 10 U/μl, 10 μg/μl, respectively, was used for cell maintenance. Before long storage of cells, three-day-old cells were loosen by taping shock and the cell concentration was determined on a

hemacytometer. Cells were then pelleted and colleted by spinning at 840*g* for 4 min at 4°C. After washing in the Phosphate Buffered Saline (PBS) and respinning, cells were resuspended at a density of 1.1×10^7 cells/ml in the freezing medium (11 ml of SFM, 7 ml of conditioned medium, and 2ml of dimethyl sulfoxide (DMSO, Sigma)). Aliquot of 1ml was dispensed to each 2 ml cryogenic vial, and chilled gradually to -80°C. Frozen vials were then submerged in liquid nitrogen for the long-term storage.

Stable Transfection of Insect Cell Line

The *Drosophila* Expression System (DES, Invitrogen) was applied to the stable transfection of S2 cells. The plasmids used for transfection were, the constitutive expression vector, pAc5.1/V5-HisA, the selection vector, pCoHYGRO, which carries a hygromycin resistance gene, and the positive control vector, pAc5.1/V5-His/lacZ, which expresses β-galactosidase protein.

Stable transfection of S2 cells with Effectene reagent: 10^6 S2 cells were suspended in 5 ml of Schneider's *Drosophila* complete medium (Invitrogen) and seeded in a 25 ml flask or a 60 mm Petri dish. After cultured at 28°C for one day, the flask is expected to be 40-80% confluent on the day of transfection. 1 μg of pAc5.1/V5-HisA (mock transfection), pA_mVgR1, pA_mVgR2, pA_mVgR1-2, or pAc5.1/V5-His/lacZ (positive control), 50 ng of pCoHygro, and the DNA condensation buffer (Qiagen) were mixed to make a total volume of 150 μl. 8 μl of enhancer (Qiagen) was then added and mixed by vortexing for 1 sec. After the mixture was incubated at room temperature (R.T.) for 5 min, 10 μl of Effectene transfection reagent (Qiagen) was added, and the mixture was further incubated for 10-15 min to form a complex. While complex formation was

taking place, cells were washed once in the flask with PBS, and 4 ml of the fresh medium was added to the flask. When incubation was done, the transfection complex was mixed with 1ml of the *Drosophila* complete medium by pipetting up and down twice, and immediately applied to the cell layer in the flask. 24-48 hours after transfection, cells were passed into selective media with penicillin/streptomycin and hygromycin at a concentration of 100-200 μg/ml depending on the cell density.

***Stable transfection of S2 cells with Cellfectin reagent*:** S2 cells were cultured in the *Drosophila*-SFM without antibiotics. 10^6 cells were seeded in each 25 cm^2 flask and cultured for one day. On the day of transfection, 30 μg of plasmid construct, 1.5 μg of pCoHygro, and 2.5 ml of the *Drosophila*-SFM were mixed in a tube, while 120 μl of the CellFectin reagent (Invitrogen) was mixed with 2.5 ml of the SFM in another tube. Solutions from two tubes were mixed gently, and the mixture was incubated for 15-45 min at R.T.. S2 cells were washed once in the flask with the PBS or SFM, and overlaid with the lipid-DNA complex. After 5-24 hours, the transfection mixture was replaced with the SFM, and cells were cultivated for two days before passed into the selective medium.

β-Galactosidase Enzyme Assay

To monitor the expression of β-galactosidase in the S2 cells transfected with positive control vector, pAc5.1/V5-His/lacZ, adherent cell lysate was prepared using the reporter lysis buffer (Promega) as described in the technical bulletin. The β-galactosidase assay was done using the β-Gal Assay Kit (Invitrogen) as described in the manual.

Extraction of poly A^+ mRNA from Transfected S2 Cells and Northern Blot of Mini-receptors

***Extraction of poly A^+ mRNA*:** S2 cells stably transfected with mini-receptor genes were harvested 7, 25, 49, and 72 hours after activation by passage of cells, and were used for extraction of poly A^+ mRNAs with the Oligotex Direct mRNA Kit (QIAGEN) as described in the protocol. Extracted mRNA was then subjected to formaldehyde agarose gel electrophoresis (FAGE).

***Northern blot of mini-receptors*:** The 3-morpholinopropanesulfonic acid (MOPS, Sigma) buffer (Pata and Truve 1998) was used to be the buffer system for the formaldehyde agarose gel (2.5 M formaldehyde and 2mM Ethylenediaminetetraacetic acid (EDTA) in the MOPS buffer), the RNA gel loading buffer (final 3 M formaldehyde, 34% formamide, 4.5% glycerol, and bromophenol blue in the MOPS buffer), and the gel running buffer (0.61 M formaldehyde and 2 mM EDTA in the MOPS buffer). The poly A^+ mRNA extracted from the same mass of cells was mixed with sheared salmon sperm DNA (Invitrogen), denatured at 68°C for 10 min, and subjected to FAGE. After run at 25 volts overnight, mRNA bands were transferred to a positively charged nylon filter (Hybond-N^+, Amersham Biosciences) overnight as described (Pata and Truve 1998). The nylon blot was rinsed briefly in 2× SSC, and crosslinked in a Stratalinker® UV crosslinker (Stratagene) as described in the manual. The RNA marker slot was cut off the blot, and stained with methylene blue as described (Sambrook and Russell 2001). The Cross-linked blot was rinsed briefly in H_2O to remove salts, and prehybridized in the ULTRAhyb™ hybridization buffer (Ambion) overnight at 42°C. The linearized $_m$*VgR1-2* DNA was used as the template to make the DNA probe with the random primed DNA

labeling kit (Roche) as described in the manual. The labeled probe was purified by passage through a spin column packed with Sephadex G-50 (Amersham Biosciences). The purified probe was denatured at 95°C for 10 min and added to the hybridization bottle. The hybridization was carried out at 42°C overnight. After intensive washing as described (Pata and Truve 1998), the blot was subjected to autoradiography.

Purification of Mini-receptors by Immobilized Metal Affinity Chromatography

His-tagged mini-receptor proteins expressed in stably transfected S2 cells and secreted to media were purified with the cobalt-based TALON® resin (CLONTECH) packed in a gravity-flow column. The conditioned cell culture media were harvested by spinning for 5 min at 800*g*, and 20 min at 1200*g* at 4°C. Proteins were precipitated with 18% of polyethylene glycol (PEG) 6000-8000 with rotation at 4°C overnight followed by centrifugation at the highest speed for 40-50 min. Pellet was resuspended in the PBS with 5-10 mM imidazole, the 100× diluted protease inhibitor cocktail for mammalian tissues (PI-tissue, Sigma), and 12.5 mM EDTA, and was dialyzed against the PBS with 5 mM imidazole in a 15-ml Slide-A-Lyzer® dialysis cassette (Pierce). The 10 ml chromatography columns (Bio-Rad) were packed with the PBS equilibrated TALON® metal affinity resin (CLONTECH) and loaded with dialyzed samples. After washing with 5 mM imidazole in the PBS, the bound proteins were eluted with 150 mM imidazole in the PBS. Fractions with protein eluate were detected by a Bradford protein microassay (1.25-25 μg/ml) as described in the technical note (Bio-Rad), pooled, and diafiltrated against the Incubation Buffer (IB, Sappington *et al.* 1995) in an Amicon® Ultra-4

centrifugal filter device (Millipore) as described in the user guide. The protein concentration was determined by a Bradford protein assay with bovine serum albumin (BSA) as a standard.

Isolation of *Aedes* Ovary Membranes and Enrichment of *Aa*VgR

The preparation of mosquito ovary membrane and enrichment of *Aa*VgR was essentially according to previous descriptions (Dhadialla *et al.* 1992; Sappington *et al.* 1995). Purification was done either on ice or at 4°C.

***Isolation of mosquito ovary membranes*:** Ovaries were dissected from female mosquitoes fed with a blood meal 20-24 hours before dissection as described previously (Sappington *et al.* 1997) and stored at -85°C until use. 2000 pairs of frozen ovaries were suspended in 5 ml of the *Aedes* Physiological Saline (APS, Hagedorn *et al.* 1977) with 50× diluted PI-tissue and 12.5 mM EDTA and homogenized in a glass-Teflon homogenizer for 5 bouts. After centrifugation at 500*g* for 10 min, the surface lipid was wiped off, if possible, and the supernatant was saved. The pellet was extracted one more time with 5 ml of the APS (PI+), and two times with 5 ml of the APS/suramin (APS with PI, EDTA, and 5 mM suramin (EMD Biosciences)) as above to release bound VgR from Vn. The supernatants from above steps were combined, and spun at 100,000*g* for 45 min. The supernatant and lipid layer were carefully removed, and the pellet was extracted with 5 ml of the B2 buffer/suramin (B2 buffer with 100× diluted PI-tissue and 5 mM suramin at pH 8.4) to release crystallized and bound Vn. After spinning at 100,000*g* for 45 min, the pellet was extracted one more time with the B2 buffer/suramin, two times with the B2 buffer (PI+) to wash off suramin, one time with the APS (PI+, pH7) to change buffer and

neutralize, and one time with the Tris-Maleate buffer (with 100× diluted PI at pH6) to change buffer and further reduce the pH. The pellet was resuspended in a minimum volume of the Tris/Maleate buffer (PI+) by pipetting and sonicating (Vibra Cell^{TM} VC40, Sonics & Materials) at the lowest setting for 5 sec. The protein concentration of the suspension was determined by a Bradford protein microassay, and adjusted to 5-6 μg/μl by dilution.

***Solubilization of the ovary membrane and enrichment of AaVgR*:** The above membrane preparation was mixed with an equal volume of Tris-Maleate/CHAPS (Tris-Maleate buffer with PI and 80 mM CHAPS (Sigma)) and incubated for 1 hour to solubilize membrane proteins. After spinning at 100,000*g* for 1 hour, the supernatant was recovered and diluted with the Tris-Maleate buffer to reduce the concentration of CHAPS to 15 mM. After spinning at 100,000*g* for 45 min to precipitate partially soluble proteins, the supernatant was recovered and treated with 9% of PEG4600 for 1-2 hours with rotation to selectively precipitate soluble proteins. After another spinning at 100,000*g* for 40 min, the VgR-enriched pellet was washed in 100 μl of the IB by pipetting and sonicating as above, and spun at 100,000*g* for 30 min. The pellet was finally resuspended in the IB and stored at -80°C.

Metabolic Radiolabelling of Vitelloginin in the Fat Body Tissue Culture

Fat body tissues were dissected from female mosquitoes 17-22 hour PBM and cultured *in vitro* generally as described previously (Raikhel *et al.* 1997; Dhadialla *et al.* 1991). In the complete fat body culture medium, 200× diluted PI cocktail for tissue culture media (PI-culture, Sigma), 10^{-5} M β-ecdysone, and 0.16 mM phenylthiourea

(PTU, Sigma) resolved in ethanol were included to diminish protein degradation, reinforce the induction of total protein synthesis, and prevent the oxidation of amino acids, respectively. Vg was metabolically labeled with 0.4 μM [^{35}S]-methionine (PerkinElmer) in the complete culture medium with 8 μM unlabled methionine for 2-3 hours. The conditioned medium was collected and diluted in 1/5 volume of the B2 buffer (Dhadialla *et al.* 1992) with the 6× diluted PI-tissue and 75 mM EDTA. After clarification by spinning at 19,000*g* for 8 min at 4°C, the supernatant was stored at -85°C until use.

Purification of Vitelloginin and Vitellin by Anion Exchange Chromatography

The Vg samples were from the mosquito fat body tissue culture, and the Vn sample was made by resolving crystallized Vn from the homogenized mosquito ovaries with the B2 buffer as described (Koller *et al.* 1989). The raw protein preparation was diafiltrated against the 150 mM anion exchange buffer (with 150 mM salt, Dhadialla *et al.* 1991) in an Amicon® Ultra-15 centrifugal filter device (Millipore) as described in the user guide. The 100× PI-tissue was added back to the sample and the sample was loaded three times to the DEAE-Sepharose Fast Flow (Amersham Biosciences) resin equilibrated with the 150 mM AE buffer and packed in an Econo-Pac® chromatography column (Bio-Rad). After washing the column with 150 mM and 200 mM AE buffers (with 150 mM and 200 mM salt), pure Vg was eluted with the 400 mM AE buffer (with 400 mM salt). The elution peak was determined by transferring 2 μl from each fraction to a 20 ml scintillation vial with 3 ml of the BioCount™ complete counting scintillation

cocktail (Research Products International) and measuring the counts per minute (CPM) with the Beckman Liquid Scintillation System (LS 1801, Beckman). The peak fractions were pooled and diafiltrated against the IB. The final protein concentration was determined and the CPM measured to calculate the labeling strength of Vg.

Silver Staining of Polyacrylamide Gel

After electrophoresis, the polyacrylamide gel was fixed in the fixing solution (50% ethanol, 12% acetic acid, and 0.05% formaldehyde) for at least 90 min. The gel was washed with 50% ethanol several times, and rinsed with pure H_2O for 20 sec only. The gel was transferred to another tray, submerged in 0.02% sodium thiosulfate for exactly 1 min, and washed with pure H_2O three times each for 20 sec. The gel was submerged in the chilled silver solution (0.1% silver nitrate and 0.075% formaldehyde) for 15 min at 4°C and rinsed with pure H_2O two times each for 20 sec at R.T.. The stained gel was transferred to a new tray and incubated in the developing solution (3% Na_2CO_3, 0.05% formaldehyde, and 0.0005% $Na_2S_2O_3$) until the development was done. The reaction was stopped with 5% acetic acid.

Western Blot of Mini-receptors

The purified mini-receptor proteins were subjected to a nonreducing Criterion™ 4-15% Tris-HCl SDS-PAGE gel (Bio-rad) at 4°C according to the instruction manual and blotted at 30 volts overnight to a polyvinylidene difluoride (PVDF) membrane in the transfer buffer (10 mM 3-cyclohexylamino-1-propane sulfonic acid (CAPS, EMD Biosciences) and 10% methanol) at 4°C. After rinsing for several times to renature

proteins, the blot was treated with the blocking solution (0.5% casein-Hammersten (US Biochemical) and 1% BSA in the IB) overnight at 4°C. After intensive washing in the IB with 0.05% Tween20, the blot was incubated with rabbit anti-native-formed *Aa*VgR IgGs (in the blocking solution) for at least 3 hours at 4°C. After intensive washing, the blot was incubated with the horseradish peroxidase (HRP) conjugated goat anti-rabbit IgG (Cappel Research Reagents) for 1-2 hours at 4°C. After intensive washing, the blot was incubated with the SuperSignal® West Pico chemiluminescent substrate (Pierce) for 5 min, and promptly exposed to a Biomax light film (Kodak). After exposion, the blot was stained with the Colloidal Gold Total Protein Stain (Bio-Rad) solution as a reference.

Solid Phase Saturation-Binding Assay of Mini-receptors

The test wells on 96-well high-affinity Reacti-Bind™ EIA plates (Pierce) were coated with the same amount (no more than 400 ng) of protein for the same receptor sample in 100 μl overnight at 4°C. After the sample solution was aspirated, the wells were washed briefly with the IB for two times and blocked with 330 μl of the blocking solution (0.05% Tween-20 and 5% BSA in the IB) first for 1 hour at R.T. and then for several hours at 4°C. The blocking solution was aspirated, and the wells were washed briefly with the blocking solution for 2 times to remove residue receptor proteins. Aliquots of the 100 μl blocking solution with geometrically increasing amounts of [^{35}S]-labeled Vg were loaded successively, and the plates were incubated for 2-5 hours at 4°C. One hundred times of unlabeled Vn was added to the geometrically diluted "hot" Vg solutions as a parallel, nonspecific binding control for each receptor. To calculate the concentration of unbound Vg, aliquots of 60 μl were transferred from wells, mixed with 5

ml of the BioCount™ complete counting scintillation cocktail, and counted with the Liquid Scintillation System (Beckman). After washing promptly with the blocking solution for three times in a well-to-well way, the bound Vg was lifted with 0.1 M NaOH in the diluted RADIACWASH radiodecontamination solution (Biodex Medical Systems) at R.T.. To calculate the concentration of bound Vg, aliquots were removed from wells for liquid scintillation accounting. The radioactive saturation binding data were fitted into 1- or 2-site binding equations to get regression curves in the nonlinear regression analysis with the GraphPad Prism® software. Either 6 or 9 data points were used in the regression analysis, and each data point represents the mean of either 3 or 4 determinations. Nonlinearly regressed nonspecific binding was deduced from the total binding to obtain the specific binding, and was not shown.

RESULTS

The CLII of CRs in insect YPR and the single cluster of CRs in VgRs/YPRs from other egg-laying animals, VLDLR, LDLR, and ApoER2 share homologous modules in different combinations

The potential ligand-binding domains of insect VgR/YPR are composed of a varying number of CRs in different combinations (Figure 5 and 6). Although all modules in the second cluster (CLII) of CRs of *Anopheles gambiae* (malaria mosquito) VgR (*Ag*VgR) are more homologous to their counterparts in *Aa*VgR than in *Dm*VgR, the CLII of *Ag*VgR lacks CRII-5 (fifth CRs in the CLII). There is no possibility of alternative splicing, because on the *A. gambiae* genome the region from the C terminal 2/3 of CRII-4 to the C terminus of the protein is coded by one large exon (strain PEST, genome ID:

Fig. 5. The CLII of CRs in insect VgR/YPR and the single cluster of CRs in egg-laying vertebrate and nematode VgR/YPR share homologous modules in different combinations. CRs from the CLII of insect VgR/YPR and from egg-laying vertebrate and nematode VgR/YPR were aligned and grouped by homology. The five CRs in *Ce*YPR were found less homologous to those from the CLI of insect VgR/YPR (alignment not shown). Each CR was assigned a group number. In the consensus sequences: a, aromatic (F, H, W, Y); d, (D,N); h, hydrophobic (A, F, G, H, I, K, L, M, P, R, T, V, W, Y); l , aliphatic (I,L,V); p, polar (D, E, H, K, R, S, T); t, turnlike (D, E, G, H, K, Q). Disulfide bonds are diagramed as lines on top. --, missing residues; ▼, acidic Ca^{2+}-coordinating residues with side chain carboxylate oxygen in human LDLR (Fass *et. al.* 1997); Δ, position of residues coordinate Ca^{2+} through backbone carbonyl oxygen in human LDLR (E, R, W); *Gg*VgR, *Gallus gallus* (chicken) VgR/VLDLR (Bujo *et. al.* 1994); *Xl*VgR, *Xenopus laevis* (African clawed frog) VgR (Okabayashi *et. al.* 1996); *Ce*YPR, *Caenorhabditis elegans* (nematode) YPR (Brant and Hirsh 1999). DmYPR, *Drosophila melanogaster* (fruit fly) YPR (Schonbaum *et. al.* 1995); *Pa*VgR, *Periplaneta americana* (American cockroach) VgR (Tufail and Takeda 2002, protein ID: BAC02725); *Ag*VgR, *Anopheles gambiae* (malaria mosquito) VgR (Holt *et. al.* 2002). The *Ag*VgR sequence is from conceptual translation (ID: EAA06264) except that the last exon starts at 3682837 instead of 3682852 on the genome sequence AAAB01008846.

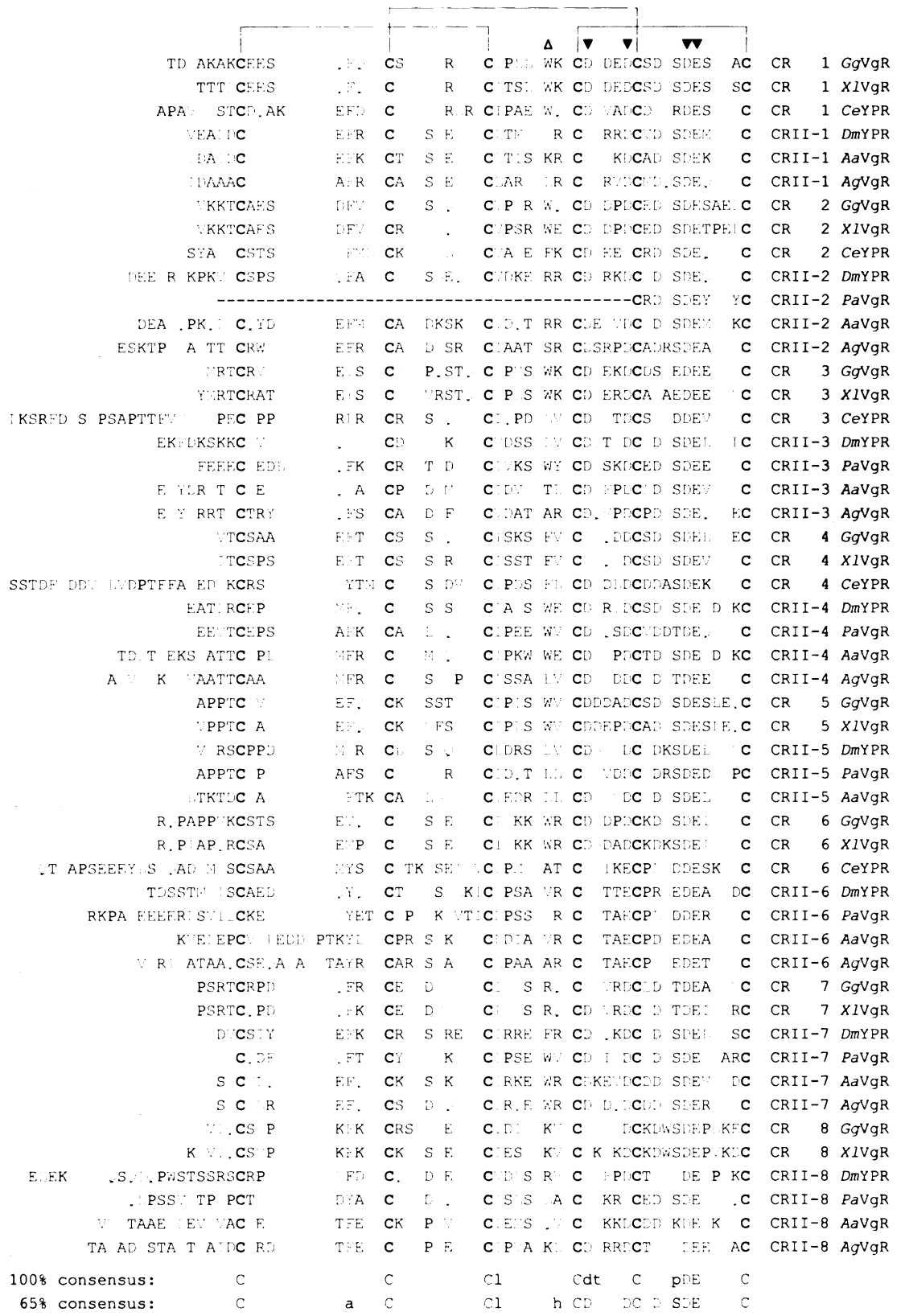
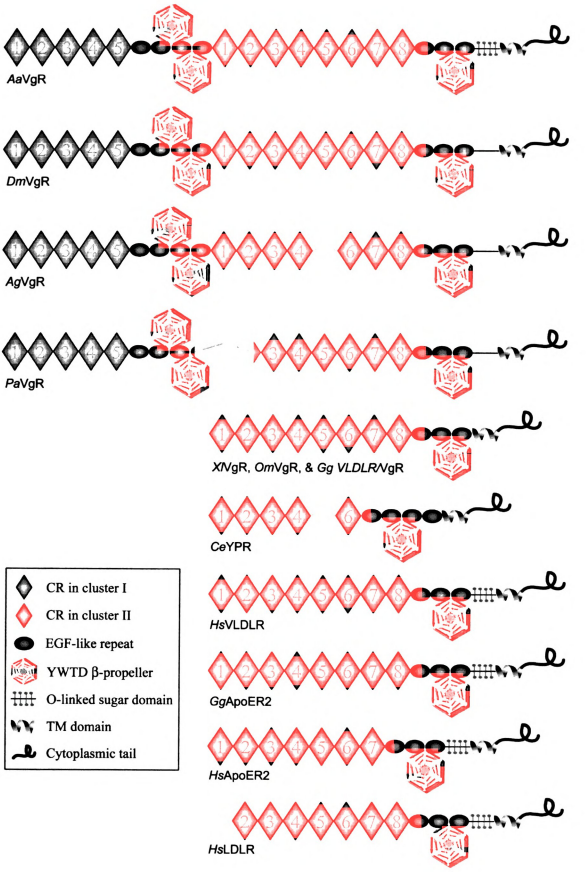

Δ
CR 1 GgVgR
CR 1 XlVgR
CR 1 CeYPR
CRII-1 DmYPR
CRII-1 AaVgR
CRII-1 AgVgR
CR 2 GgVgR
CR 2 XlVgR
CR 2 CeYPR
CRII-2 DmYPR
CRII-2 PaVgR
CRII-2 AaVgR
CRII-2 AgVgR
CR 3 GgVgR
CR 3 XlVgR
CR 3 CeYPR
CRII-3 DmYPR
CRII-3 PaVgR
CRII-3 AaVgR
CRII-3 AgVgR
CR 4 GgVgR
CR 4 XlVgR
CR 4 CeYPR
CRII-4 DmYPR
CRII-4 PaVgR
CRII-4 AaVgR
CRII-4 AgVgR
CR 5 GgVgR
CR 5 XlVgR
CRII-5 DmYPR
CRII-5 PaVgR
CRII-5 AaVgR
CR 6 GgVgR
CR 6 XlVgR
CR 6 CeYPR
CRII-6 DmYPR
CRII-6 PaVgR
CRII-6 AaVgR
CRII-6 AgVgR
CR 7 GgVgR
CR 7 XlVgR
CRII-7 DmYPR
CRII-7 PaVgR
CRII-7 AaVgR
CRII-7 AgVgR
CR 8 GgVgR
CR 8 XlVgR
CRII-8 DmYPR
CRII-8 PaVgR
CRII-8 AaVgR
CRII-8 AgVgR
100% consensus:
65% consensus:

Fig. 6. The Structural organization of the VgR/YPR, VLDLR, LDLR, and ApoER2. CRs from the CLII of insect VgR/YPR and CRs from other VgR/YPR, VLDLR, LDLR, and ApoER2 were assigned into one of eight groups by homology. Abbreviations: ApoER2, apolipoprotein E receptor 2; *Ce*, *Caenorhabditis elegans*; *Dm*, *Drosophila melanogaster*; *Gg*, *Gallus gallus*; *Hs*, *Homo sapiens*; *Om*, *Oncorhynchus mykiss*; *Pa*, *Periplaneta Americana*; *Xl*, *Xenopus laevis*; LDLR, low-density lipoprotein receptor; VgR, vitellogenin receptor; VLDLR, very low-density lipoprotein receptor; YPR, yolk protein receptor; YWTD, tyrosine-tryptophan-threonine-aspartic acid.


AaVgR
DmVgR
AgVgR
PaVgR
XlVgR, OmVgR, & Gg VLDLR/VgR
CeYPR
HsVLDLR
GgApoER2
HsApoER2
HsLDLR
CR in cluster I
CR in cluster II
EGF-like repeat
YWTD β-propeller
O-linked sugar domain
TM domain
Cytoplasmic tail

AAAB01008846.1, join 3682837..3684961). *Periplaneta americana* (American cockroach) VgR (*Pa*VgR, Tufail and Takeda 2002) has at least six intact CRs (CRII-3 to CRII-8) and the C terminal ¼ remnant of CRII-2 in the CLII (Figure 5 and 6). The juxtaposed EGF repeat N terminal to the CL II has only an N terminal remnant. A surprising observation is that the five CRs of *Caenorhabditis elegans* (nematode) YPR (*Ce*YPR, Grant and Hirsh 1999) apparently have higher homology to modules from the CLII than to those from the CLI in insect VgR/YPR (Figure 5 and 6). *Ce*YPR has homologues of CRII-1 to CRII-4 and CRII-6 of the insect VgR/YPR, and no homologues of CRII-5, CRII-7, and CRII-8. A similar observation comes from human LDLR and ApoER2, which lack homologues of the first and eighth CRs of VLDLR (Figure 5 and 6).

Construction, Expression, and Purification of *Aa*VgR Mini-receptors

***Construction of the AaVgR mini-receptors*:** To determine which cluster of CRs in *Aa*VgR is responsible for exclusive or predominant *Aa*Vg binding, two mini-receptor genes ($_m$*VgR1* and $_m$*VgR2*) encoding either the CLI or CLII of CRs and its C terminal adjacent EGF precursor homology domain were constructed by PCR and cloned into cloning vectors and expression vectors (Figure 7 and 8). A third mini-receptor gene ($_m$*VgR1-2*) encoding the extracellular portion (except for the *O*-linked polysaccharide domain) was also constructed as a positive control (Figure 7 and 8). Point mutations found in the constructs were reversed either by a copy-and-paste work, or by amplification with different cDNA clones. The details of construction and reversion of point mutations are described in the Materials and Methods in this chapter.

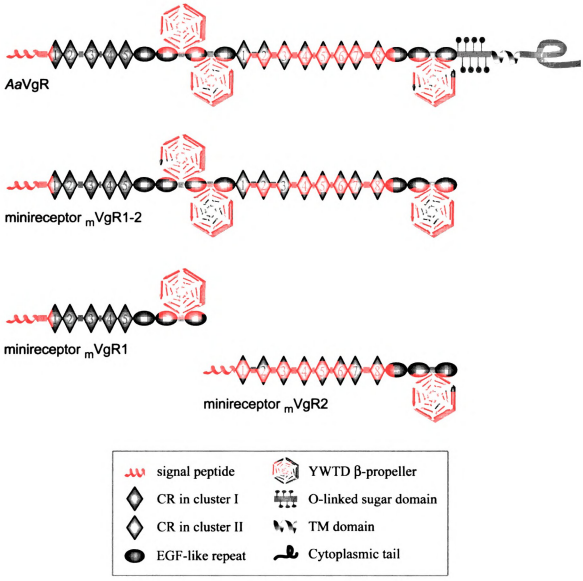


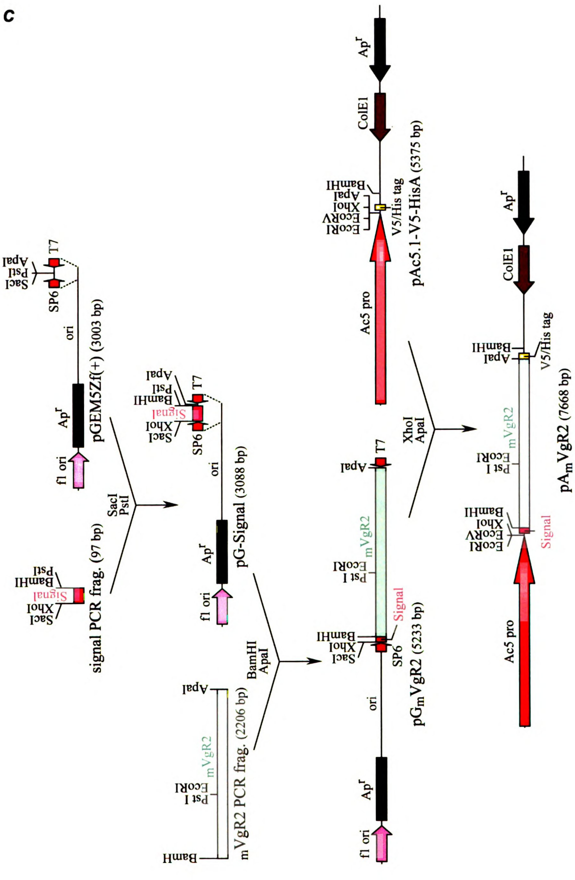
Fig. 7. The schematic representation of *Aa*Vg mini-receptors.

Fig. 8. Construction of *Aa*VgR mini-receptors. Mini-receptors, $_{m}$VgR1, $_{m}$VgR2, and $_{m}$VgR1-2, were constructed as described in the Materials and Methods. For the legend, refer Figure 7. ***A.*** Amplification of the signal PCR fragment, $_{m}$VgR1 PCR fragment, $_{m}$VgR2 PCR fragment, and mVgR1-2 PCR fragment. ***B.*** Construction of the mini-receptor $_{m}$VgR1. pGEM5Zf(+) and pAc5.1-V5-HisA were cloning vector and expression vector respectively. ***C.*** Construction of the mini-receptor $_{m}$VgR2. ***D.*** Construction of the mini-receptor $_{m}$VgR1-2.

A
mVgR1-2 PCR fragment
P1/P7
P7
P1
P1/P5
P4/P7
P6
P4
pEGM5-VgR3.2
P7
P6/P7
P3
P5
P1
P2
pBlue-VgR2.7
P1/P3
BI
mVgR2 PCR fragment
mVgR1 PCR fragment
P1/P2
BI
signal PCR fragment

B

SacI
XhoI
Signal
mVgR1
ApaI
mVgR1 PCR frag. (1897 bp)
SacI
ApaI
SacI
PstI
ApaI
SP6
T7
ori
Ap^r
f1 ori
pGEM5Zf(+) (3003 bp)
SacI
ApaI
XhoI
SP6
Signal
mVgR1
T7
ApaI
ori
Ap^r
f1 ori
pGmVgR1 (4831 bp)
XhoI
ApaI
Ap^r
ColE1
EcoRI
EcoRV
XhoI
ApaI
BamHI
V5/His tag
Ac5 pro
pAc5.1-V5-HisA (5375 bp)
ApaI
BamHI
V5/His tag
ColE1
Ap^r
mVgR1
Signal
EcoRI
EcoRV
XhoI
Ac5 pro
pAmVgR1 (7266 bp)


c
signal PCR frag. (97 bp)
SacI
XhoI
Signal
BamHI
PstI
pGEM5Zf(+) (3003 bp)
f1 ori
Ap^r
ori
SP6
T7
SacI
PstI
ApaI
pG-Signal (3088 bp)
mVgR2 PCR frag. (2206 bp)
BamH
Pst I
EcoRI
mVgR2
ApaI
BamHI
ApaI
pG$_m$VgR2 (5233 bp)
Ac5 pro
V5/His tag
EcoRI
EcoRV
XhoI
ApaI
BamHI
ColE1
pAc5.1-V5-HisA (5375 bp)
XhoI
ApaI
pA$_m$VgR2 (7668 bp)

D

pGEM5Zf(+) (3003 bp)
SacI
PstI
ApaI
SP6
T7
ori
Ap^r
f1 ori
mVgR1-2 PCR frag. (5059 bp)
ApaI
PstI
EcoRI
mVgR2
mVgR1
Signal
XhoI
SacI
pGmVgR1-2 (8008 bp)
T7
mVgR2
mVgR1
Signal
SacI
XhoI
SP6
ori
Ap^r
f1 ori
pAc5.1-V5-HisA (5375 bp)
Ac5 pro
EcoRI
EcoRV
XhoI
ApaI
BamHI
V5/His tag
ColE1
Ap^r
pAmVgR1-2 (10443 bp)
Ac5 pro
Signal
mVgR1
mVgR2
V5/His tag
ColE1
Ap^r
SacI
ApaI
XhoI
ApaI

***Expression of mini-receptors in vitro*:** The constructed mini-receptor genes were expressed in vitro in a coupled reticulocyte lysate Transcription and Translation (TNT®) system. The *$_m$VgR1* gene product gave a major band at anticipated size of around 80 kDa and a minor band with a slightly lower molecular weight (Figure 9, lane 3). The *$_m$VgR2* gene product was a broad band with a molecular weight between 115 and 140 kDa (Figure 9, lane 4), which is higher than the anticipated 92 kDa before modification of the protein. This is probably because of the limited co-translational and post-translational modification (phosphorylation, glycosylation, and sulfation). A major band of around 210 to 230 kDa was shown to be the *$_m$VgR1-2* gene product (Figure 9, lane 5), although the size virtually exceeds the size limit of 176 kDa for the TNT system.

***Expression of mini-receptors in the insect cell line and purification*:** To get mini-receptor proteins properly modified after translation, three mini-receptor genes were stably transfected into the *D. melanogaster* S2 cell line with either Effectene reagent (Qiagen) and Cellfectin reagent (Invitrogen). The Vector plasmid, pAc5.1/V5-HisA, and the expression control plasmid, pAc5.1/V5-His/lacZ, were also transfected into S2 cells as the negative and positive controls, respectively. The efficiencies of two transfection methods were roughly equivalent as indicated by a β-galactosidase enzyme assay (data not shown).

Stably transfected S2 cells were harvested at different time intervals and the poly A^+ mRNAs were extracted from cell lysates. The poly A^+ mRNAs extracted from the same mass of cells were subjected to formaldehyde agarose gel electrophoresis and northern blot with *$_m$VgR1-2* DNA as a probe. Auto-radiography of the blot showed a

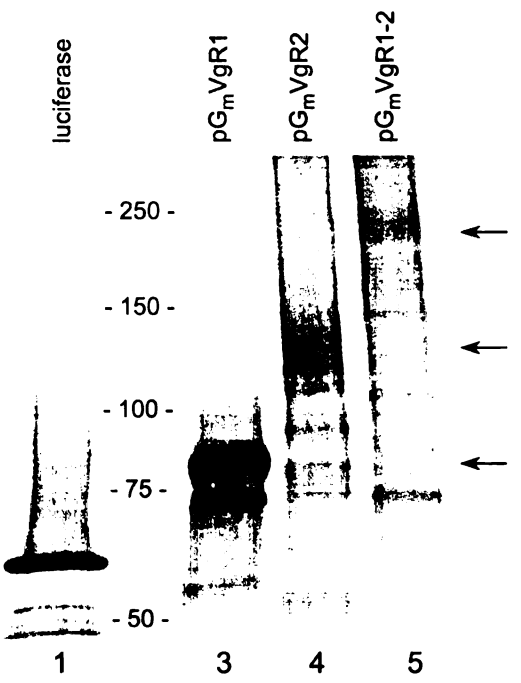


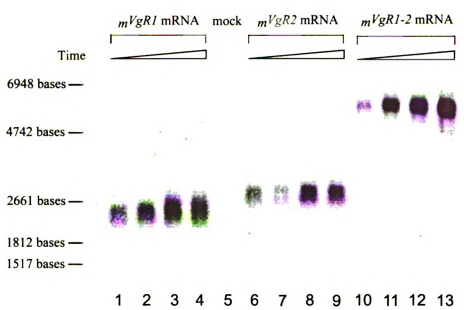
Fig. 9. *In vitro* expression of the VgR mini-receptors in a coupled transcription and translation system. Each mini-receptor gene driven by a SP6 promoter on the pGEM5 vector was expressed in the TNT SP6 coupled reticulocyte lysate transcription and translation system (Promega) in the presence of [^{35}S] methionine. Reactions were subjected to a reducing SDS–polyacrylamide (PAGE) gel and autoradiography. The anticipated sizes of nascent unmodified products are 78-, 92-, and 194-kDa for the $_{m}$VgR1 (lane 3), the $_{m}$VgR2 (lane 4), and the $_{m}$VgR1-2 (lane 5) respectively. The size of the product of the luciferase control gene should be 61-kDa (lane 1).

clear single band for each mini-receptor gene transcript in the correct size (Figure 10*A*: lanes 1-4, 2337 bases for $_{m}$*VgR1* mRNA; lanes 6-9, 2739 bases for $_{m}$*VgR2* mRNA; lanes 10-13, 5541 bases for $_{m}$*VgR1-2* mRNA), which suggested the normal initiation and termination of transcription. Overall, the transcripts of three mini-receptor genes from the same cell mass increased in quantity over a period of three days, which implied the stability of mini-receptor transcripts and/or increased transcription rates after activation of cells by passage. No band was observed from the mock transfection control (Figure 10*A*, lane 5).

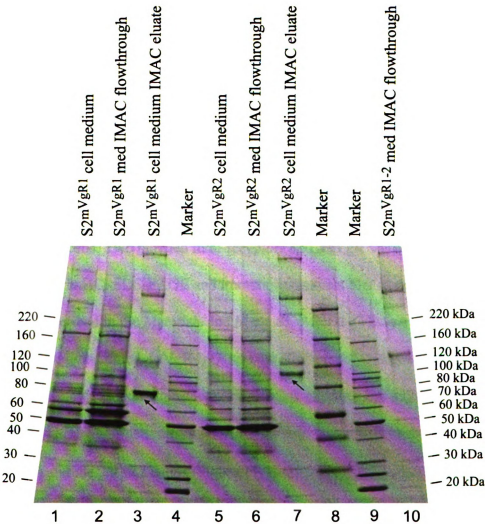
Total proteins in the transfected S2 cell media were precipitated with 18% PEG, and the His-tagged mini-receptor proteins were purified by the cobalt-based immobilized metal affinity chromatography (IMAC). Figure 10*B* shows the purified $_{m}$VgR1 and $_{m}$VgR2 proteins on an SDS-PAGE gel stained with Coomassie blue. The abundances of both $_{m}$VgR1 and $_{m}$VgR2 were very low in the transfected S2 media (Figure 10*B*, compare lane 1 with lane 2, and lane 5 with lane 6), as indicated by virtually the same band spectra of S2 media before and after loaded to the affinity column. The $_{m}$VgR1 eluate gave a major $_{m}$VgR1 band of around 80 kDa, which accounted for nearly 40% of total proteins (Figure 10*B*, lane 3). The $_{m}$VgR2 eluate also gave a major band of around 100 kDa (Figure 10B, lane 7), although the $_{m}$VgR2 preparation was less pure than the $_{m}$VgR1 preparation. The purified His-tagged $_{m}$VgR1-2 protein, which was predicted to be around 200 kDa, was invisible on the gel (Figure 10*B*, lane 10), indicating either the very low efficiency of expression of the around 200 kDa protein in S2 cells or lack of an intact His-tag at the C terminus of $_{m}$VgR1-2, possibly caused by the premature

Fig. 10. Transcription and translation of mini-receptor genes in the *Drosophila* S2 cell line. *Drosophila* S2 cells were transfected with mini-receptor genes. Mock transfection was made with vector plasmid only. ***A***, Stably tranfected S2 cells were harvested 7, 25, 49, and 72 hours after passage, and poly A^+ mRNAs were extracted. Mini-receptor mRNAs from the same mass of cells were subjected to a Northern blot and detected with the $_m$*VgR1-2* probe. The anticipated sizes of transcripts (excluding poly A tail added during polyadenylation) are 2337, 2739, and 5514 bases for $_m$*VgR1* (lane 1-4), $_m$*VgR2* (lane 6-9), and $_m$*VgR1-2* (lane 10-13) respectively. Lane 5 is mock transfection with vector plasmid only. ***B***, His-tagged VgR mini-receptor proteins were purified from S2 cell culture media by TALON immobilized metal affinity chromatography (IMAC). The eluates were subjected to SDS-PAGE, and the gel was stained with Coomassie brilliant blue. Arrows, $_m$VgR1 (lane 3) and $_m$VgR2 (lane 7) protein bands.

A


mVgR1 mRNA
mock
mVgR2 mRNA
mVgR1-2 mRNA
Time
6948 bases
4742 bases
2661 bases
1812 bases
1517 bases
1 2 3 4 5 6 7 8 9 10 11 12 13

B


S2mVgR1 cell medium
S2mVgR1 med IMAC flowthrough
S2mVgR1 cell medium IMAC eluate
Marker
S2mVgR2 cell medium
S2mVgR2 med IMAC flowthrough
S2mVgR2 cell medium IMAC eluate
Marker
Marker
S2mVgR1-2 med IMAC flowthrough
220
160
120
100
80
60
50
40
30
20
220 kDa
160 kDa
120 kDa
100 kDa
80 kDa
70 kDa
60 kDa
50 kDa
40 kDa
30 kDa
20 kDa
1 2 3 4 5 6 7 8 9 10

termination of translation. The 5,154 bp-long coding region of the $_{m}$VgR1-2 gene was not fully sequenced after reversion of a nonsense mutation found in the CLII of CRs. All of the visible bands from $_{m}$VgR1-2 sample eluate had equivalent bands from both $_{m}$VgR1 and $_{m}$VgR2 eluates (Figure 10*B*, compare lane 10 with lane 3 and lane 7). Multi-bands demonstrated insufficiency of one-step purification of mini-receptors. Another reason is the inclusion of 5 mM instead of 10 mM imidazole in both the sample loading buffer and the washing buffer, which increased the nonspecific binding to columns. To make the mini-receptor preparation purer, a gel filtration step could be added.

To substitute $_{m}$VgR1-2 with a new positive control, the full length *Aa*VgR was enriched from *A. aegypti* vitellogenic ovaries 20-24 hours PBM as described in the experimental procedures. The purified $_{m}$VgR1 and $_{m}$VgR2 were subjected to a reducing Tris-HCl SDS-PAGE and blotted to a PVDF membrane. After the blot was incubated with the HRP conjugated anti-V5 antibody (Invitrogen), a chemiluminescence reaction detected anticipated-sized single band for both the $_{m}$VgR1 and $_{m}$VgR2 mini-receptors (data not shown). This indicated that the full length $_{m}$VgR1 and $_{m}$VgR2 have been purified from S2 cells. Two purified mini-receptors and *Aa*VgR were further subjected to a nonreducing Tris-HCl SDS-PAGE and immunofluorescent blot under a nondenaturing condition as described in the Materials and Methods. After the blot was incubated with the rabbit anti-native-formed *Aa*VgR IgG, and the HRP conjugated goat anti-rabbit IgG, protein bands were detected with a chemiluminescent substrate and exposed to a light film (Figure 11). The anti-native-formed VgR antibody bound to both the $_{m}$VgR1 (Figure 11, lanes 1-2) and the $_{m}$VgR2 bands (Figure 11, lanes 5-6) at anticipated sizes. A much weaker band at anticipated size was also detected to be potentially the $_{m}$VgR1-2 protein

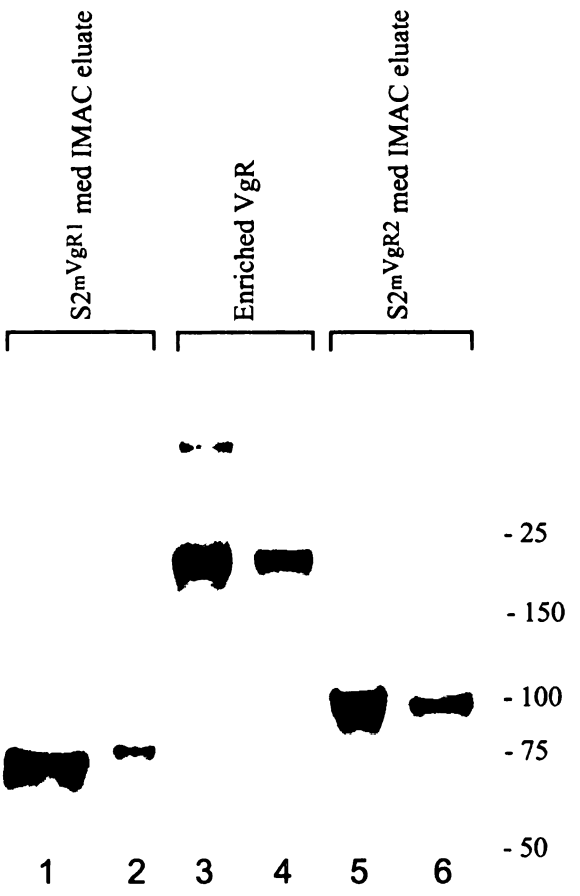


Fig. 11. Nonreducing Western blot of VgR mini-receptors. Enrichment of native VgR from vitellogenic mosquito ovaries and western blot are described in the Materials and Methods. Protein samples were subjected to a nonreducing 4~15 % Tris-HCl SDS-PAGE, and blotted onto a PVDF membrane. The blot was incubated with the primary antibody, rabbit anti-native *Aa*VgR IgG, and the HRP conjugated secondary antibody. Vg receptor bands were detected with a chemiluminescent substrate and exposed to a Biomax light film.

on the blot by anti-VgR IgG, and was visualized on the blot after colloidal gold staining (data not show), which suggested that a trace amount of the His-tagged $_m$VgR1-2 protein has been obtained. For each of the three mini-receptors and *Aa*VgR, there was always one additional, very faint band with a much higher molecular weight in proportion to the size of the major band (Figure 11, lane 3 for clear secondary band, too week to be seen here for the $_m$VgR1 and $_m$VgR2, and not shown for the $_m$VgR1-2).

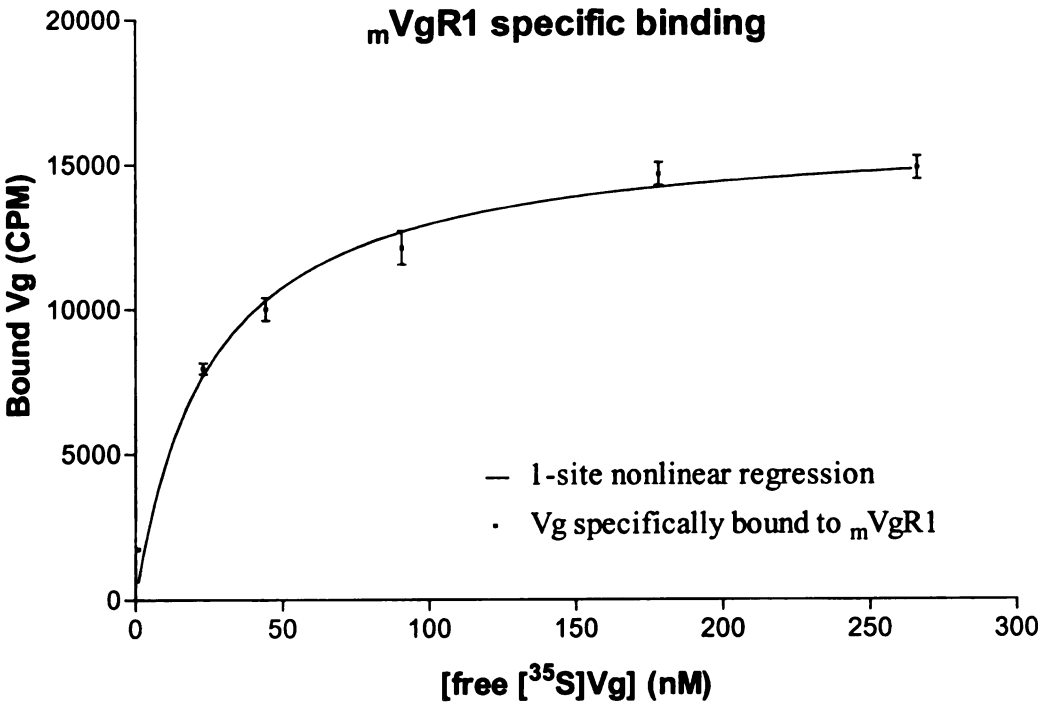
Binding of The *Aa*VgR Mini-receptors to *Aa*Vg

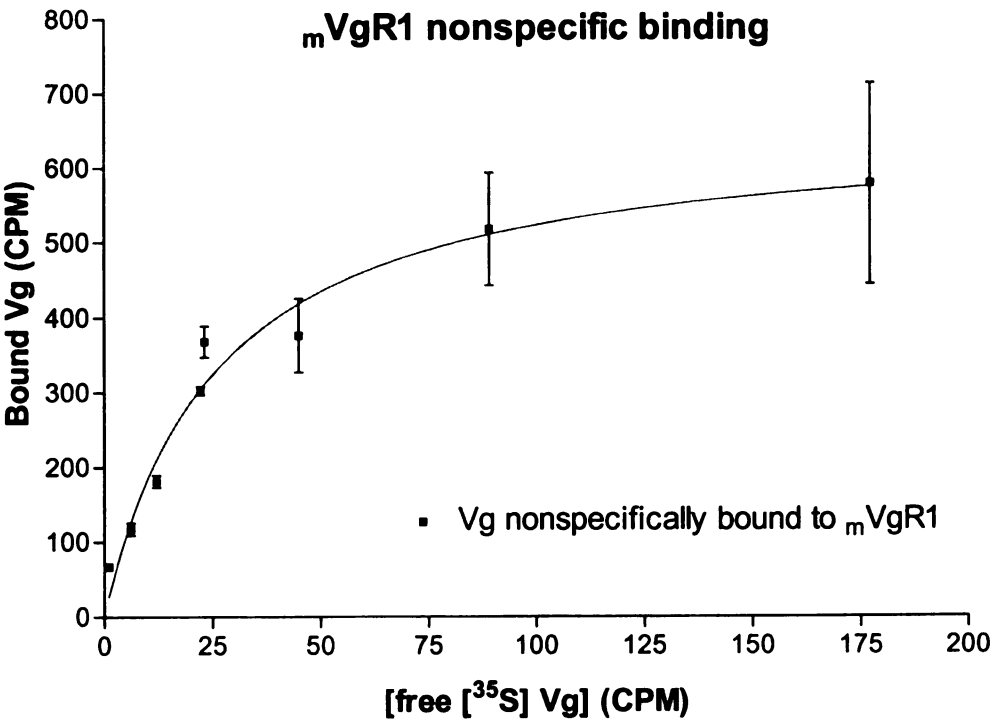
***Both clusters of CRs contribute to the high affinity of the AaVgR to AaVg*:** To prepare for VgR/Vg binding assays, Vg was metabolically radiolabeled with [^{35}S]-methionine in the *Aedes aegypti* fat body tissue culture. The crystallized Vn was resolved from homogenized *Aedes aegypti* ovaries. Both Vg and Vn were purified by anion exchange chromatography, and detected by SDS-PAGE (data not shown).

The 96-well microtiter plate-based, solid phase saturation-binding assays were performed to measure the binding affinities of $_m$VgR1, $_m$VgR2, and *Aa*VgR to *Aa*Vg (Figure 12). One hundred times of unlabeled (cold) Vn was added to geometrically diluted [^{35}S]-Vg solutions as a parallel, nonspecific binding control for each receptor. The radioactive saturation binding data for $_m$VgR1 (Figure 12*A*) and $_m$VgR2 (Figure 12*B*) fit one-site nonlinear regression, and the dissociation constants (K_d) of 25.9 ± 3.3 nM and 53 ± 8.5 nM were calculated for $_m$VgR1 and $_m$VgR2, respectively (Table 6). In contrast to two mini-receptors, the binding data for *Aa*VgR fit two-site nonlinear regression better than one-site (Figure 12*C*), which indicated the presense of at least two binding sites on the full-length *Aa*VgR. The apparent K_d of VgR was calculated to be 3.2

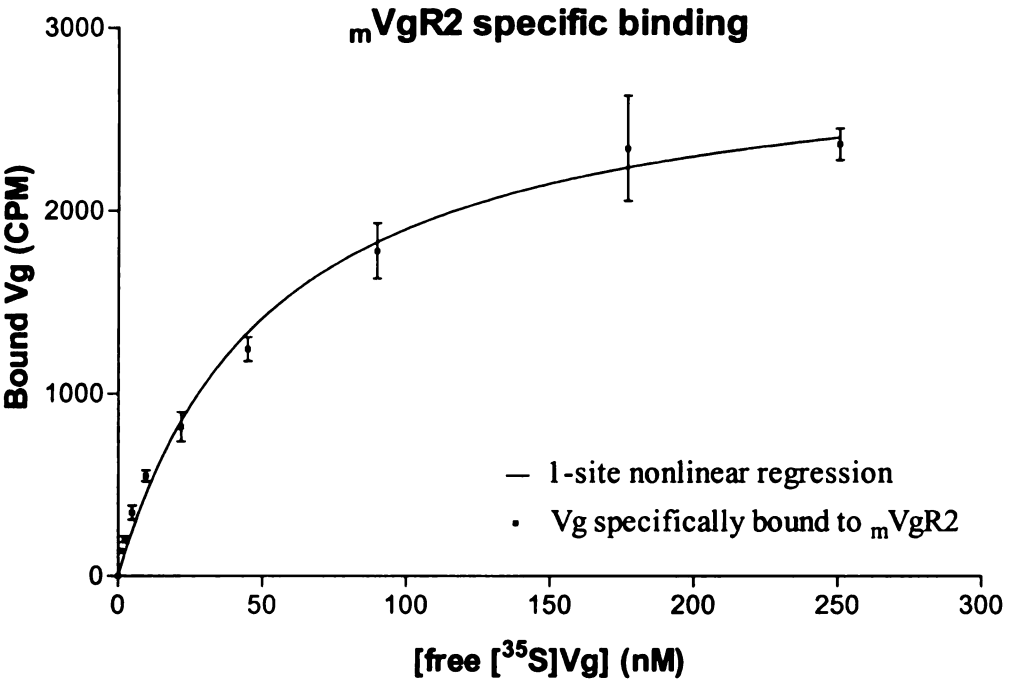
Fig. 12. The Solid phase saturation binding assays of the mini-receptors. 96-well microtiter plates were used in the solid phase saturation binding assay. Test wells were coated with same amount of protein for the same kind of receptor. After blocking test wells, geometrically increasing amount of [^{35}S]-Vg were loaded in order. After incubation, aliquots of incubation solution were accounted for free [^{35}S]-Vg signal. After washing, the bound [^{35}S]-Vg was lifted and aliquots were counted. Nonspecific binding controls were set by involving 100 times of cold Vn in the binding solutions for each receptor. 1- and 2-site nonlinear regression analyses were performed with software Prism (GraphPad), and a primary model was chosen based on the significance of 2-site model over the 1-site one on data fitting. Either six or nine data points were included in the regression analysis, and each data point represents the mean of three or four determinations. Graphs show the specific binding and nonspecific binding data. ***A.*** Nonlinear regression of $_{m}$VgR1 specific and nonspecific binding. ***B.*** Nonlinear regression of $_{m}$VgR2 specific and nonspecific binding. ***C.*** Nonlinear regression of VgR specific and nonspecific binding.

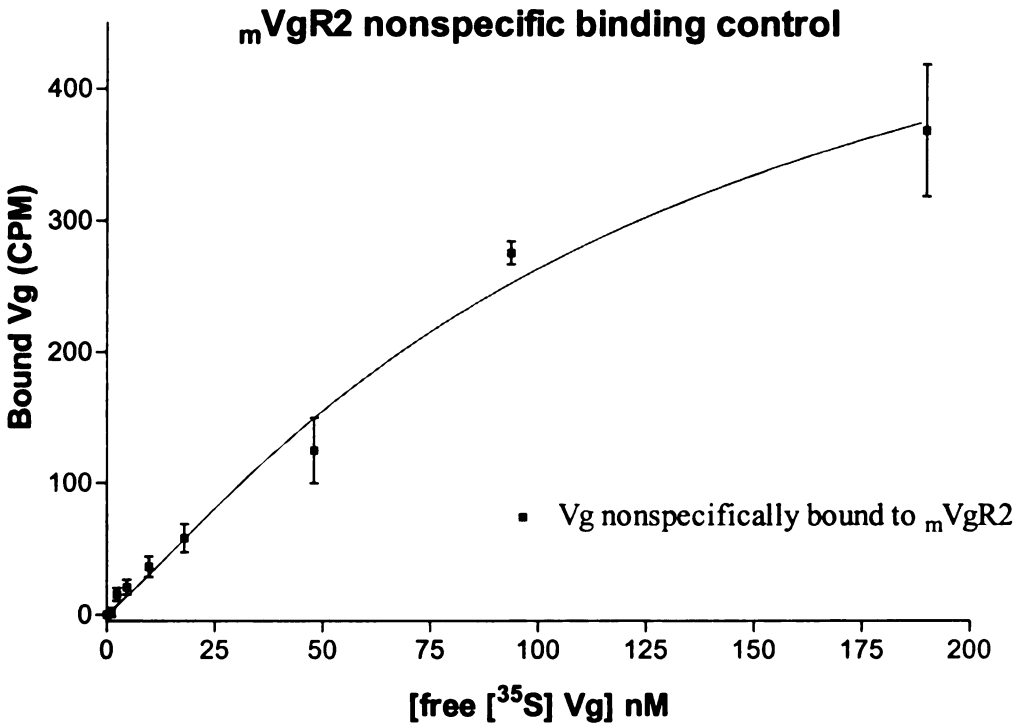
A


mVgR1 specific binding
Bound Vg (CPM)
20000
15000
10000
5000
0
0
50
100
150
200
250
300
[free [35S]Vg] (nM)
— 1-site nonlinear regression
· Vg specifically bound to mVgR1

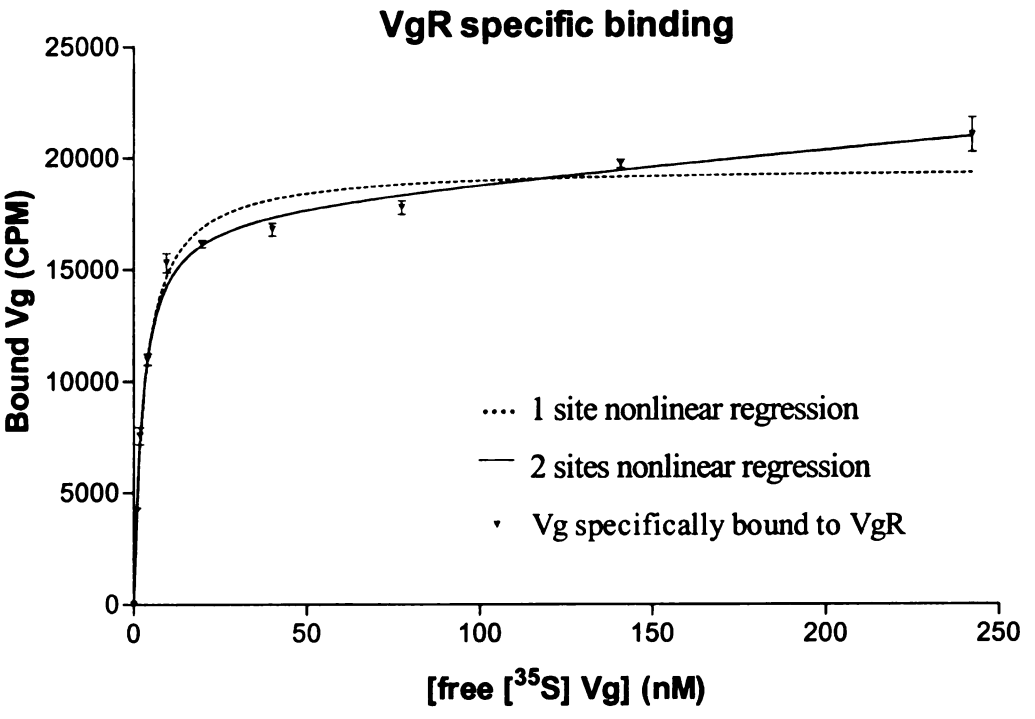

mVgR1 nonspecific binding
Bound Vg (CPM)
800
700
600
500
400
300
200
100
0
0
25
50
75
100
125
150
175
200
[free [35S] Vg] (CPM)
▪ Vg nonspecifically bound to mVgR1

B


mVgR2 specific binding
Bound Vg (CPM)
3000
2000
1000
0
0
50
100
150
200
250
300
[free [35S]Vg] (nM)
— 1-site nonlinear regression
• Vg specifically bound to mVgR2


mVgR2 nonspecific binding control
Bound Vg (CPM)
400
300
200
100
0
0
25
50
75
100
125
150
175
200
[free [35S] Vg] nM
▪ Vg nonspecifically bound to mVgR2

C


VgR specific binding
25000
20000
15000
10000
5000
0
Bound Vg (CPM)
0
50
100
150
200
250
[free [35S] Vg] (nM)
1 site nonlinear regression
2 sites nonlinear regression
Vg specifically bound to VgR

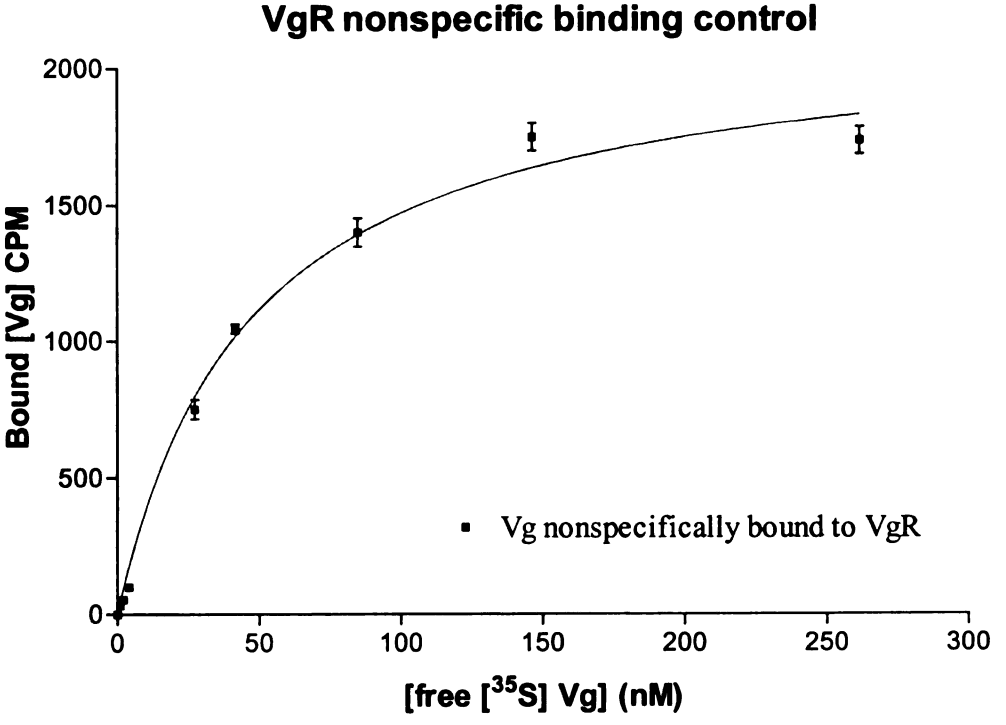

VgR nonspecific binding control
2000
1500
1000
500
0
Bound [Vg] CPM
0
50
100
150
200
250
300
[free [35S] Vg] (nM)
Vg nonspecifically bound to VgR

Table 6. K_d values for saturation binding assays.

Receptor	No of determinations		No of points	$(K_d)_{app}$ or K_d	$K_{d\,(high)}$
	sample	control		*nM*	*nM*
VgR	3	2	9	3.17 ± 0.27	2.50 ± 0.26
$_m$VgR1	3	2	6	25.9 ± 3.3	n/a
$_m$VgR2	4	3	9	53.0 ± 8.5	n/a

± 0.3 nM in one experiment and 3.49 nM on average over three independent experiments (data not shown). The apparent K_d of *Aa*VgR to *Aa*Vg is much higher than those of $_m$VgR1 and $_m$VgR2 and is also much higher than those reported earlier (Sappington *et. al.* 1995; Dhadialla and Raikhel 1991). Based on the two-site binding model, there are at least one very high-affinity ($K_{d\,(high)}$ of 0.93-2.5 nM) Vg-binding site and one very low-affinity site on *Aa*VgR based on the results from three independent binding assays. The high-affinity site was responsible for 70%-90% of specific binding. The calculated K_d of the very-low-affinity Vg-binding site on *Aa*VgR varied significantly in three independent binding assays.

DISCUSSION

In the LDLR family, insect VgRs/YPRs represent a subfamily with two clusters of CRs (Figure 5). Noticing that the cDNA clone of *Pa*VgR encodes a protein missing most of the C terminal 5/6 of the EGF-like module right before the CLII of CRs and the N terminal ¾ of CLII-2 (or even CLII-1, if it has), there is a possibility that *Pa*VgR

transcript was subjected to an alternative splicing event (personal communication from M. Tufail). Quite a few alternative splicing examples have been reported in the LDLR family, which include VLDLR in human (Christie et. al. 1996) and ApoER2 in mouse (Brandes *et. al.* 2001), chicken (Brandes *et. al.* 1997), and human (Brandes *et. al.* 1997). Another possibility is that *Pa*VgR lacks CRII-1 and most of CRII-2. This is also possible because compared with *Aa*VgR, which has eight contact CRs in the CLII, the CLII in *Ag*VgR has only seven contact modules. To summarize, the CLII of CRs in insect VgR/YPR can have a varying amount of CRs, and the maximum number of modules in the CLII is eight, because in both the single cluster of one-cluster members and the CLII of four-cluster members the maximum number of CRs is eight.

A striking observation is that the five CRs of nematode *Ce*YPR have apparently higher homology to modules from the CLII than those from the CLI in insect VgR/YPR (Figure 5 and 6, the multiple alignment of CRs from the CLI is not shown). This implies that the single cluster of CRs in the one-cluster VgR/YPR (including five-CR-bearing *Ce*YPR) and the CLII in the two-cluster insect VgR/YPR evolve from a common ancestor and that the more conserved CLIs of insect VgRs/YPRs evolve from a different common ancestor. One conspicuous fact is that some VgRs/YPRs have fewer CRs than others. For example, *Ce*YPR has homologues of CRII-1,2,3,4 and CRII-6 of insect VgR/YPR, but no homologue of CRII-5, CRII-7, or CRII-8. *Ag*VgR has homologues of most modules of *Aa*VgR except CRII-5. These findings suggest that not all CRs are important for yolk protein binding. In other words, some CRs are unimportant to binding of Vg. The CRII-5 and CRII-8 of VgRs might be such unimportant modules.

The microtiter plate-based solid-phase saturation binding results showed that, both the CLI and CLII in *Aa*VgR have one binding site for *Aa*Vg, while the full length *Aa*VgR has at least two binding sites (Figure 12). The affinity of the full-length *Aa*VgR is much higher than that of either cluster, which suggests a synergistic effect of two clusters in achieving the high affinity to *Aa*Vg. Because the affinities of both clusters are not low, it is apparent that one additional cluster of CRs in *Aa*VgR supplies strengthened binding of *Aa*Vg.

For the LDLR family members with multiple clusters, LRP is one that has been studied extensively to elucidate the roles of modules from different clusters of CRs in the binding of different ligands. LRP has four clusters of CRs with 2, 8, 10, and 11 modules per cluster, respectively. Studies with LRP mini-receptors have shown that both the CLII and the CLIV bind twelve kinds (including α_2-M*) of known ligands (Neels *et. al.* 1999; Willnow *et. al.* 1994) and both clusters have only minor differences in ligand-binding kinetics (Neels et. al. 1999, Table I and II). This suggests a functional redundancy within LRP. In the CLII of LRP, Neels *et. al.* found that CRII-1 to -5 are essential for the binding of ligands (Neels et. al. 1999, Table III). In contrast to the CLII and CLIV, the CLIII only bound ApoE and weakly bound RAP, both of which also bound the CLII and CLIV. Although Neels *et. al.* did not find any ligand for the CLI of LRP, Mikhaihenko *et. al.* reported that not only the CLII (and CLIV) of LRP, but also the CLI have affinity to α_2-M*, albeit lower affinity (Mikhaihenko *et. al.* 2001). In comparison with the CLI and CLII, a LRP mini-receptor covering both the CLI and CLII-1,2,3 has much higher affinity to α_2-M* (Mikhaihenko *et. al.* 2001). This indicates that the CLI and CLII cooperate to generate a high-affinity binding site for α_2-M*. The results from Neels *et. al.*

also confirmed the discovery that the CRII-1,2,3 is important to binding of α_2-M*, because CRII-1,2,3,4,5 bound seven ligands (including α_2-M*), while CRII-1,2 and CRII-3,4,5 did not (Neels *et. al.* 1999).

CONCLUSIONS

First, modules in the CLII of CRs in insect VgRs/YPRs are more homologous than those in the CLI to modules in the single cluster of CRs in VgRs/YPRs from other egg laying animals, VLDLR, LDLR, and ApoER2. The CLII of CRs in insect VgRs/YPRs has from six to eight modules. The five CRs of nematode *Ce*YPR have apparently higher homology to modules from the CLII than those from the CLI in insect VgR/YPR. This implies that the one-cluster VgRs/YPRs and the CLII of insect receptors might evolve from a common ancestor. Second, both the CLI and CLII in *Aa*VgR have one binding site for *Aa*Vg, while the full-length *Aa*VgR has at least two binding sites. The CLI has higher affinity to *Aa*Vg than the CLII, and both clusters have moderate affinity. The affinity of the full-length *Aa*VgR is much higher than either cluster, which suggests a synergistic effect of two clusters in achieving the high affinity to *Aa*Vg. One additional cluster of CRs in *Aa*VgR supplies strengthened binding of *Aa*Vg.

Chapter 3

Protein Modeling of *Aa*VgR and *Aa*Vg: Structural Basis of The *Aa*Vg-*Aa*VgR Interaction

INTRODUCTION

In this chapter, sequence analyses and protein modeling efforts will be performed in order to answer several questions. First, each module in both the CLI and CLII of CRs will be modeled, and the surface charges of modules from both clusters will be compared to see if some CRs in CLI and/or CLII of CRs have negatively charged surfaces and to determine which cluster of CRs has stronger negatively charged surface. Second, both the small and large subunits of *Aa*Vg will be modeled to see if there is a positive surface on *Aa*Vg and to learn the 3-D structure feature of *Aa*Vg. Third, supposed three YWTD β propeller folds in *Aa*VgR will be modeled, and the surface charges of each model will be investigated. The distribution of histidine residues will also be checked to see if there are enough histidine residues on the propeller surface to potentially switch a surface patch from weak negative, nearly neutral, or weak positive to strong positive. This will help to determine the mechanism of *Aa*Vg release and *Aa*VgR recycling. Fourth, all of the EGF-like repeats in *Aa*VgR will be modeled, and their surface charge distribution will be checked to see if any of them can potentially contribute to the assumed electrostatic complementarities between *Aa*VgR and *Aa*Vg. Finally, the claimed VgR-binding motif on the small subunit of blue tilapia Vg will be investigated by multiple-sequence alignment, by checking its counterpart on lamprey Vg, and by modeling tilapia Vg and *Aa*Vg.

MATERIALS AND METHODS

Protein Sequence/Structure Analysis and Protein Modeling

Protein sequence and structure analyses were performed with the National Center for Biotechnology Information (NCBI) services (http://www.ncbi.nlm.nih.gov/), the SeqWeb® package (Version 2.02, Accelrys), and the PredictProtein Service at Columbia University (http://cubic.bioc.columbia.edu/pp/). 3-dimensional (3-D) protein modelings of *Aa*VgR were done on the bioinformatics server at Ben Gurion University (http://www.cs.bgu.ac.il/~bioinbgu/) for protein structure fold recognition (threading) modeling, with the LOOPP program at Cornell University (http://ser-loopp.tc.cornell.edu/loopp_old.html) for sequence/structure alignments, and on the SWISS-MODEL protein modeling server (Schwede *et. al.* 2003) for protein structure homology modeling. The 3-D figures were generated and molecular surface electrostatic potential calculated with the Swiss-PdbViewer (Version 3.7). A target protein sequence was firstly aligned with protein sequences from other species by hand. Based on the multiple-alignment, the target protein sequence was ligned with all potential 3-D templates with known 3-D structures by hand and alignments were submitted to the SWISS-MODEL server. The returned models were evaluated based on three disulfide bond connections, the conservation on secondary structures, the degree of homology to the template sequences, the total energy, the additional secondary structures, and the number and sizes of uncertain regions. Improvement in hand alignment was made until no better model could be obtained with each potential template. For each target sequence, acceptable models predicted with different templates were compared and a best model was chosen for that target sequence.

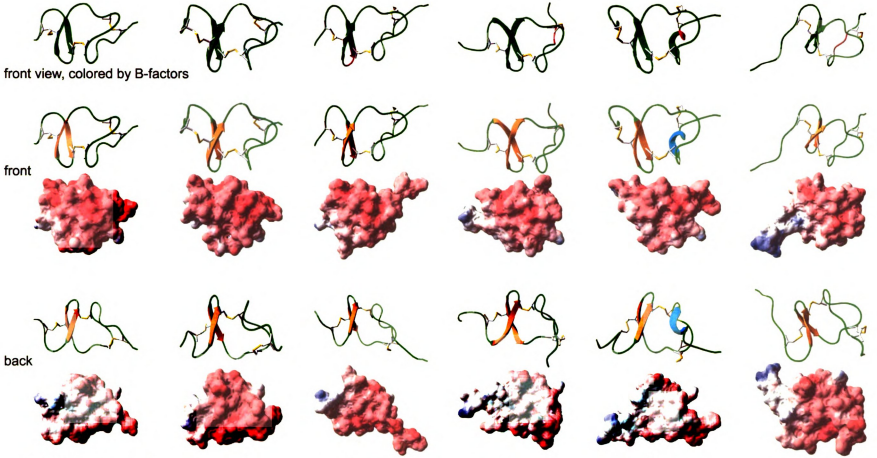
RESULTS

*Aa*VgR Protein Modeling

***The CLI of CRs in AaVgR has stronger negative surface electrostatic potential than the CLII*:** All of the thirteen modules from both clusters of CRs in *Aa*VgR were modeled (Figure 13). With the exception of CRII-1, which missed the C_1-C_3 disulfide bond in the model, the modules were all modeled with three S-S bonds (C_1-C_3, C_2-C_5, and C_4-C_6) correctly linked. All modeled modules have, overall, similar 3-D structures, with two anti-parallel β strands in the β loop region near the N terminus (Figure 13), and most models have negative total molecular energy. For eleven modules that have alternative models, the alternative models gave similar molecular surface electrical potentials (EPs). Comparing surface EPs of modules from two clusters showed that all thirteen modeled modules from both clusters have, overall, negative surface EP. In the CLI, CRI-1 and CRI-5 are the two modules with the highest negatively charged surfaces, and CRI-3 has a very strongly negative surface. In the CLII, CRII-3 has the most strongly negatively charged surface, and CRII-2, CRII-4, and CRII-7 also have strongly negative surfaces. In general, the CLI has more strongly negative EP than the CLII, and the surface charges of both clusters fall within the same scale.

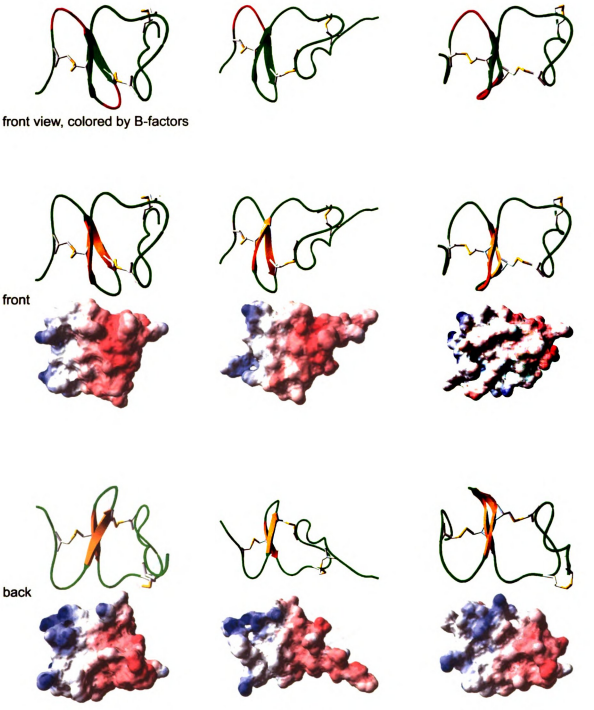
***The surface charge distribution of the A. gambiae VgR CLII is similar to that of the AaVgR CLII*:** All seven modules from the CLII of CRs in *Ag*VgR were modeled (Figure 14). Alternative models were generated for all modules in the *Ag*VgR CLII, and each chosen model had negative total energy. For all seven modules, alternative models for each module gave similar molecular surface EPs. In the CLII of *Ag*VgR, CRII-7 had the most strongly negatively charged surface. Both CRII-3 and CRII-4 had very strongly

Fig. 13. Surface electrostatic potentials (EPs) of thirteen modeled modules from two clusters of CRs in *Aa*VgR. The top row shows ribbon representation of *Aa*VgR CRs colored by model B-factors (model confidence factors, shown as color temperature). The green and red colors represent regions with certainty and uncertainty respectively. The second and fourth rows show ribbon diagrams of modeled *Aa*VgR CRs. β strands are colored in golden and α helices in aqua. Only side chains of cysteine residues are displayed to show three disulfide bonds in yellow. All the modules are oriented with N terminals on the left and C terminals on the right. The third and fifth lines show corresponding molecular surfaces colored with EP (Red: -12, strongly negative; white: neutral; blue: +6 positive). Acceptable models for each module were chosen by alternative hand alignments with all potential templates and by comparison of models made on Swiss-Model modeling server. Graphs were drawn with the Swiss-PdbViewer. Alternative models for CRI-1, CRI-2, CRI-3, CRI-5, CRII-2, CRII-3, CRII-4, CRII-5, CRII-6, CRII-7, and CRII-8 were compared to choose the best model for each module. Surface charges of best models for all *Aa*VgR CRs were compared.

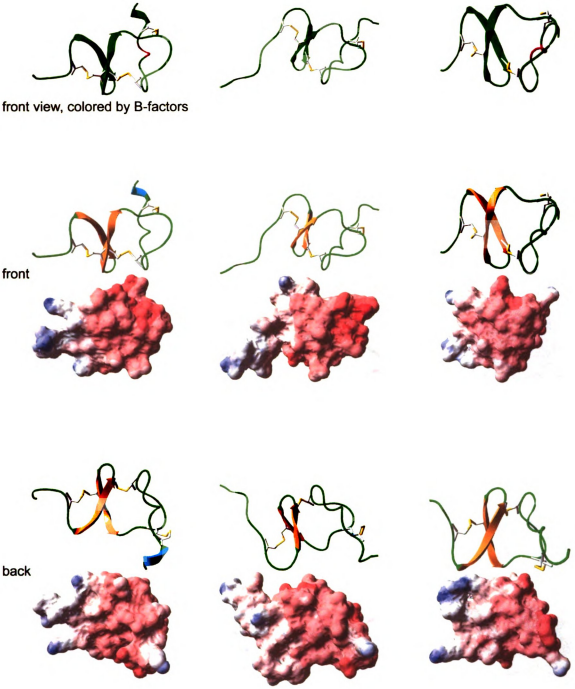


Modeled residue: 7-43	residue 5-44	residue 5-45	residue 4-45	residue 3-45	residue 1-45
E_{tot}: -859 KJ/mol	125 KJ/mol	-406 KJ/mol	-879 KJ/mol	-522 KJ/mol	-372 KJ/mol
Idn/Sim: 41%/70%	40%/73%	38%/69%	34%/68%	36%/67%	36%/60%
1AJJ model	1D2J model	1CR8 model	1J8E model	1D2L model	1F5Y model

*Aa*VgR CRI-1 model comparison



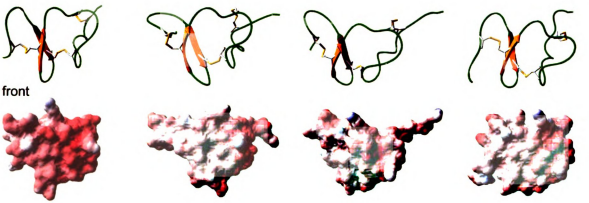
*Aa*VgR CRI-2 model comparison



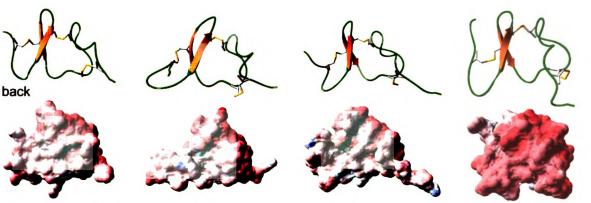
Modeled residue: 4-46	residue 1-46	residue 6-45
E_{tot}: 78 KJ/mol	942 KJ/mol	806 KJ/mol
Idn/Simi: 39%/70%	40%/60%	40%/58%
1J8E model	1F5Y model	1F8Z model

*Aa*VgR CRI-3 model comparison

front view, colored by B-factors



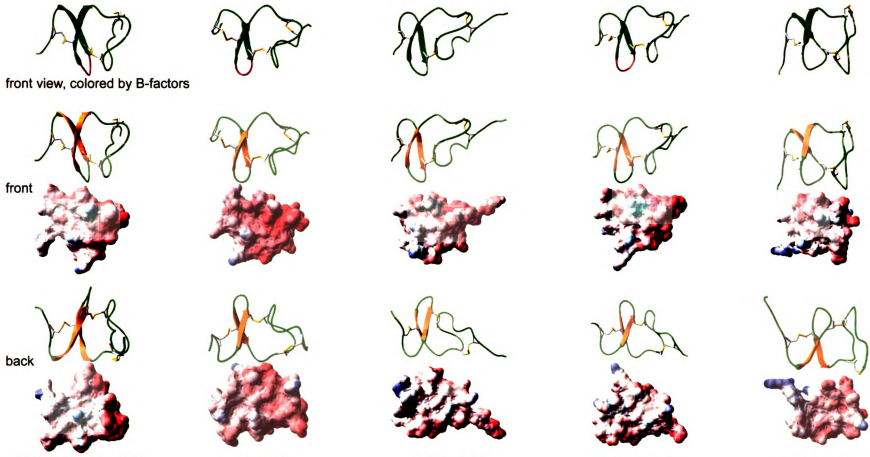
front



back

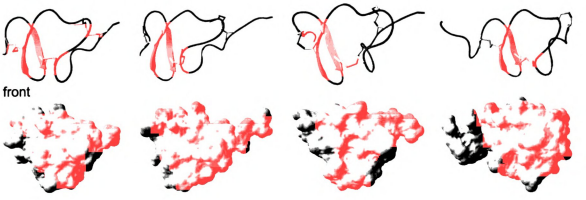
Modeled residue 3-43	residue 1-45	residue 1-45	residue 1-46
E_{tot}: -20 KJ/mol	500 KJ/mol	630 KJ/mol	485 KJ/mol
Idn/Sim: 34%/59%	39%/67%	38%/62%	36%/62%
1AJJ model	1D2L model	1CR8 model	1F5Y model

*Aa*VgR CRI-5 model comparison

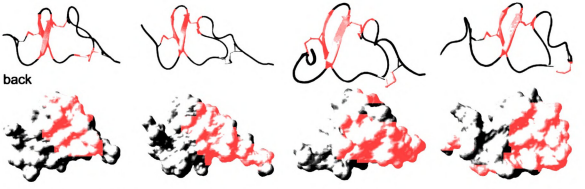


*Aa*VgR CRII-2 model comparison

front view, colored by B-factors



front

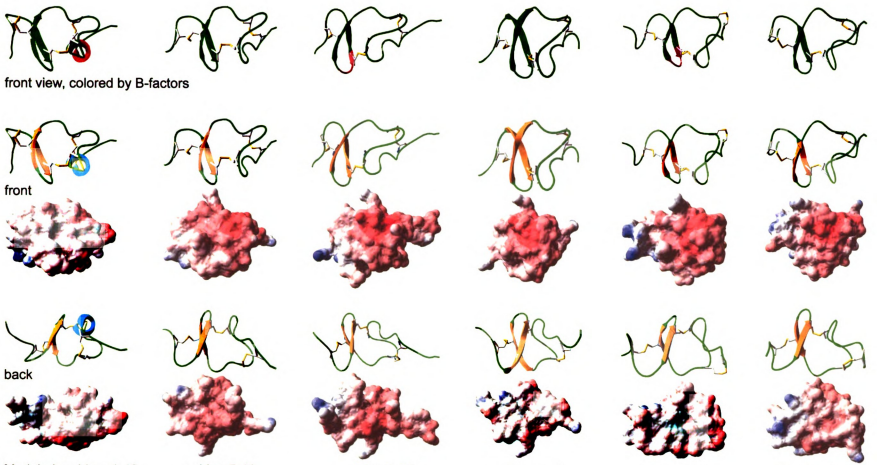


back

Modeled residue: 7-43	residue 5-45	residue 3-45	residue 1-42
E_{tot}: -77 KJ/mol	393 KJ/mol	207 KJ/mol	26 KJ/mol
Idn/Sim: 43%/65%	44%/73%	37%/67%	36%/60%
1AJJ model	1CR8 model	1D2L model	1F5Y model

*Aa*VgR CRII-3 model comparison

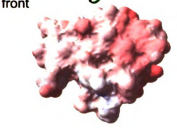
AeVgR CRII-4 model comparison

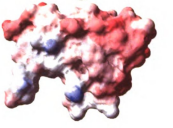


*Aa*VgR CRII-5 model comparison

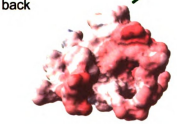
front view, colored by B-factors

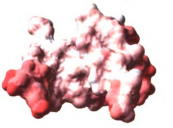
front





back





Modeled residue: 5-50

E_{tot}: 167 KJ/mol

Idn/Sim: 33%/50%

1F5Y model

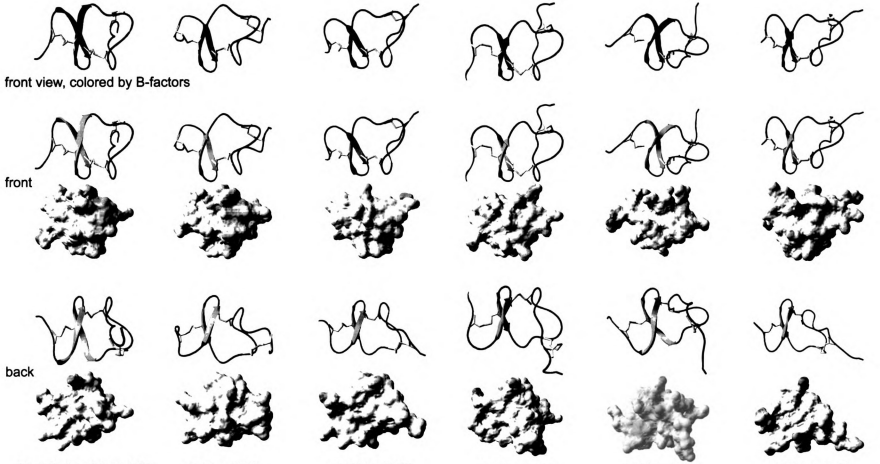
residue 5-50

52 KJ/mol

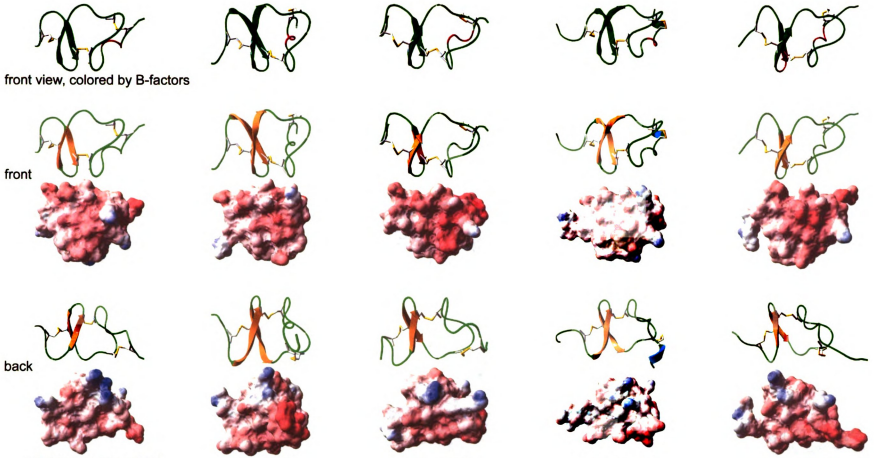
31%/52%

1LDL model

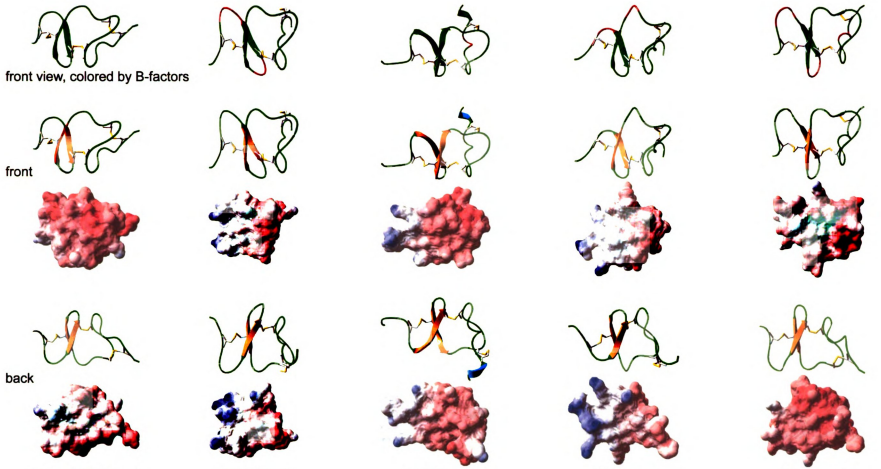
*Aa*VgR CRII-6 model comparison



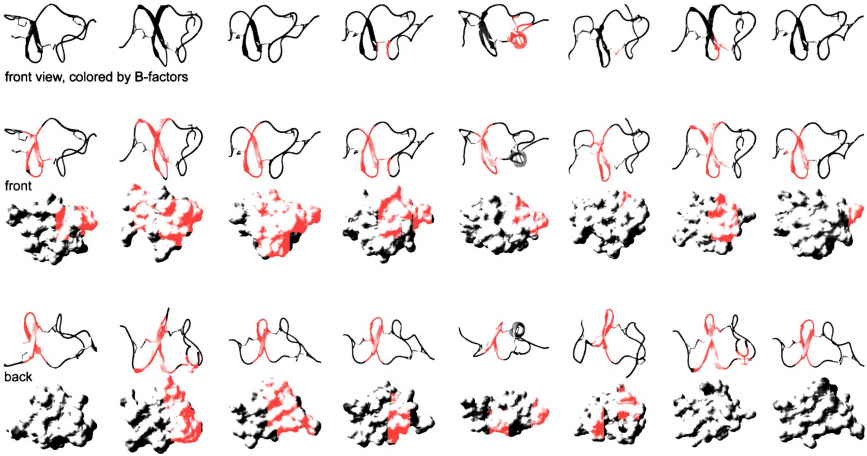
*Aa*VgR CRII-7 model comparison



*Aa*VgR CRII-8 model comparison

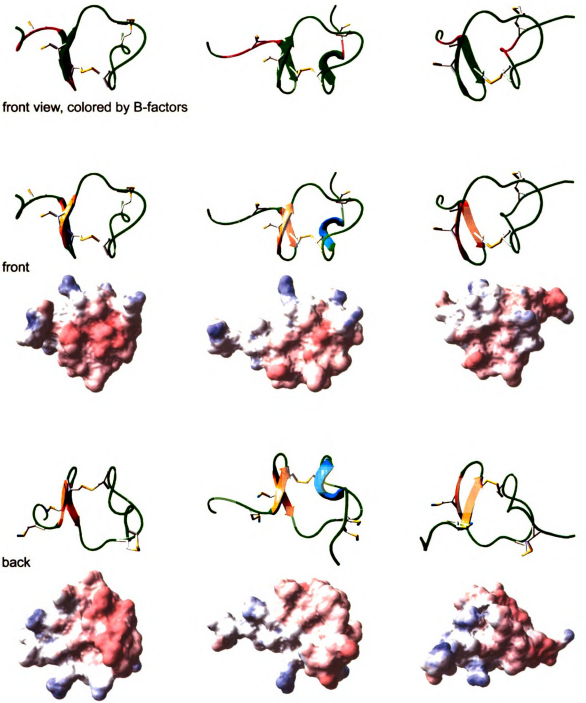


Cluster I of CRs of *Aa*VgR

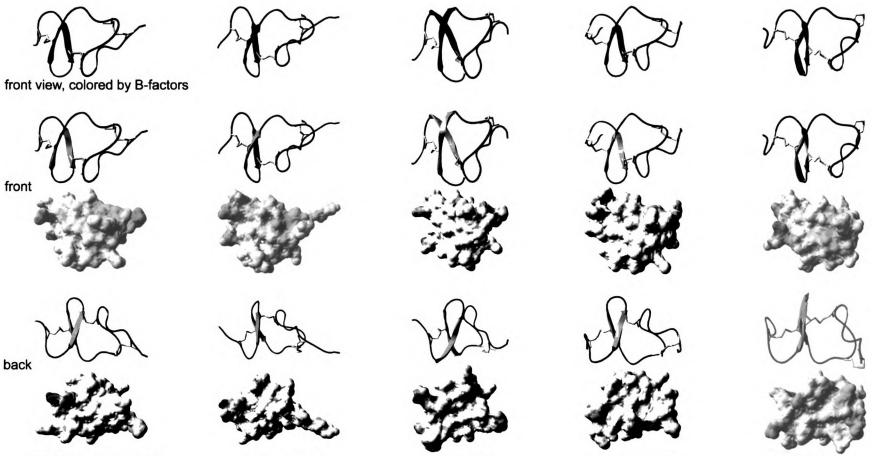


Cluster II of CRs of *Aa*VgR

Fig. 14. The surface EPs of seven modules in the CLII of CRs of the *Anopheles gambiae* VgR. The alternative models for each module in the CLII of CRs of *Anopheles gambiae* VgR (*Ag*VgR) were made and compared in the same way as *Aa*VgR models. Alternative models for *Ag*VgR modules CRII-1, CRII-2, CRII-3, CRII-4, CRII-6, CRII-7, and CRII-8 were compared, and the surface charges of the best models for all the *Aa*VgR CRs were compared. Graphs are arranged in the same way as in Figure 13. The color scale for molecular surfaces is, red: -12, strongly negative; white: 0, neutral; blue: +6, positive.



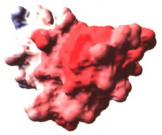
A. gambiae VgR CRII-1 model comparison

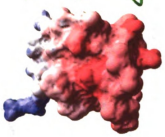


A. gambiae VgR CRII-2 model comparison

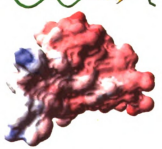
front view, colored by B-factors

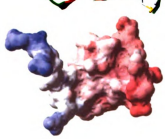
front





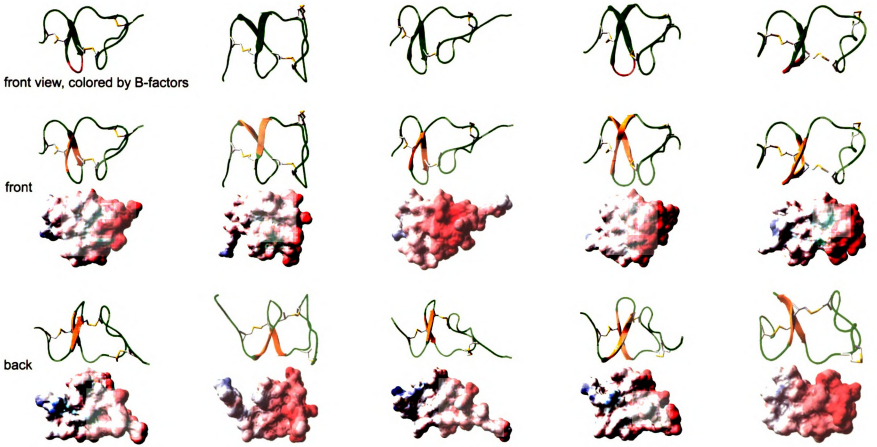
back





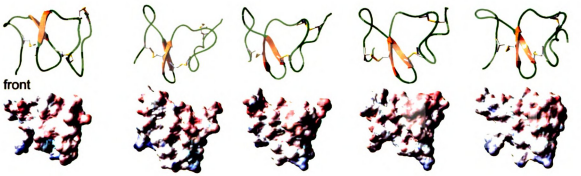
Modeled residue: 7-43	residue 4-42
E_{tot}: -1150 KJ/mol	-550 KJ/mol
Identity/similarity: 46%/73%	39%/61%
1AJJ model	1LDR model

A. gambiae VgR CRII-3 model comparison

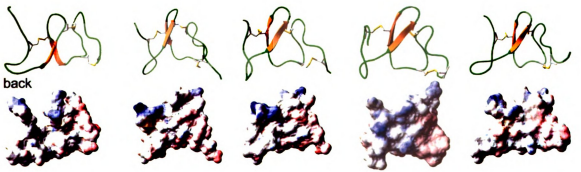


A. gambiae VgR CRII-4 model comparison

front view, colored by B-factors



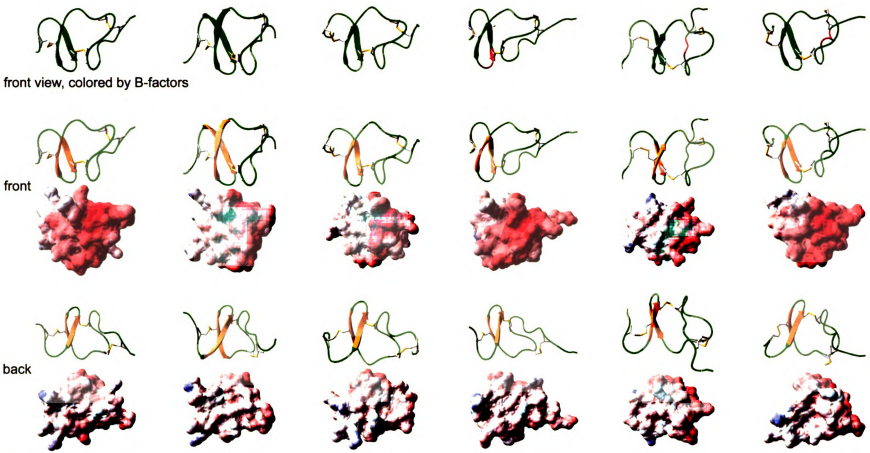
front



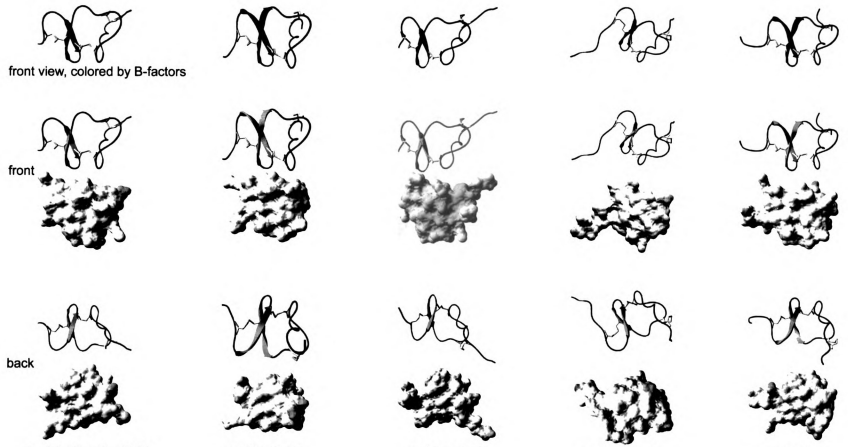
back

residue 4-49	residue 5-51	residue 7-50	residue 5-51	residue 6-51
E_{tot}: -431 KJ/mol	-772 KJ/mol	-489 KJ/mol	-297 KJ/mol	-678 KJ/mol
Idn/sim: 28%/55%	26%/60%	27%/55%	26%/55%	26%/54%
1LDR model	1CR8 model	1AJJ model	1D2J model	1F8Z model

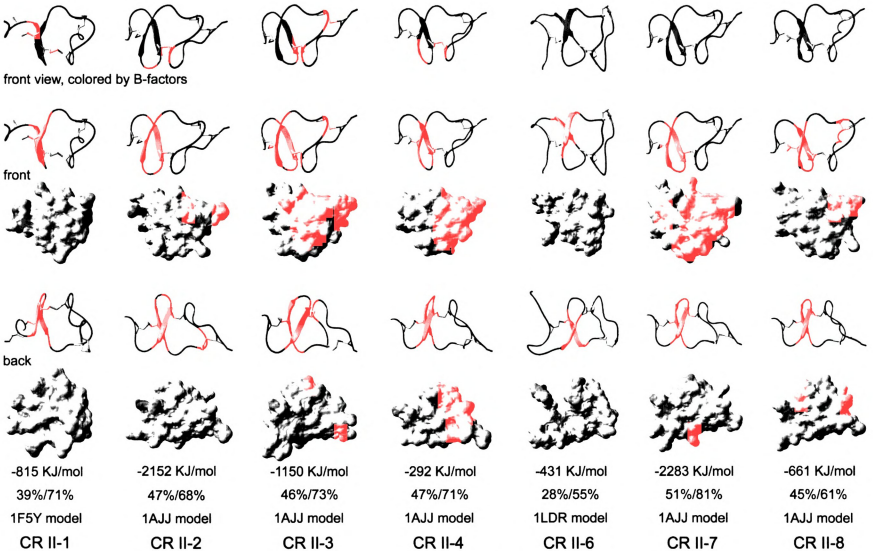
A. gambiae VgR CRII-6 model comparison



A. gambiae VgR CRII-7 model comparison



A. gambiae VgR CRII-8 model comparison



Cluster II of CRs in *A. gambiae* VgR

negative surfaces. Comparison of the surface charge distribution of the *Aa*VgR CLII with that of the *Ag*VgR CLII (Figure 15) showed three common modules with strongly negative surface charges: CRII-3, CRII-4, and CRII-7. One big difference is that in the *Ag*VgR CLII, CLII-7 rather than CLII-3 has the most strongly negative surface EP. Table 7 summarized modules with strongly negative surface charges in *Aa*VgR and *Ag*VgR.

Table 7. Comparison of surface charge distribution of the CLII of CRs in *Aa*VgR with that in *Ag*VgR.

	negative surface charge	
module	*Aa*VgR	*Ag*VgR
CLII-2	strong	nearly neutral
CLII-3	strongest	very strong
CLII-4	strong	very strong
CLII-7	strong	strongest

***The EGF-like repeats of AaVgR very possibly do not contribute to the binding of AaVg*:** The EGF1-2, 3, 4, 5-6, and 7 of *Aa*VgR were modeled with Swiss-Model server (Figure 16). All the EGF-like modules modeled have three correctly linked S-S bonds (C_1-C_3, C_2-C_4, and C_5-C_6) and overall similar 3-D structures, with two β sheets each made of two anti-parallel β strands. Only EGF4 and EGF7 have alternative models, and these alternative models are only approximately 30% identical to their templates. Despite the limited homology, the alternative models showed similar surface EPs for both EGF4 and EGF7. The total energy of each selected model for the seven modules is negative. EGF1 has a very strongly positive surface, and EGF4, EGF5, and EGF 6 have

Fig. 15. Comparison of surface charge distribution of the CLII of CRs in *Aa*VgR with that in *Ag*VgR. The left two columns are the front and back views of surface EPs of best models for modules from the CLII of CRs in *Aa*VgR, and the right two columns show their counterparts in *Ag*VgR. The color scale for molecular surfaces is, red: -12, strongly negative; white: 0, neutral; blue: +6, positive.

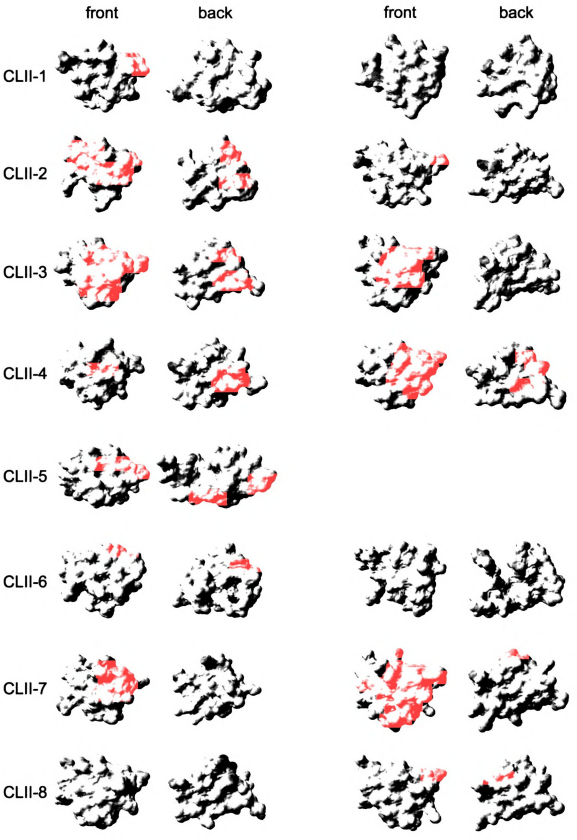
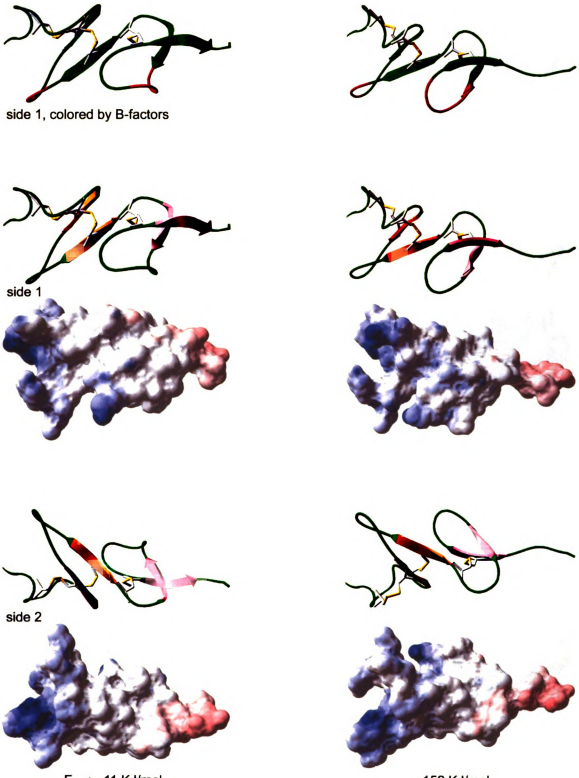
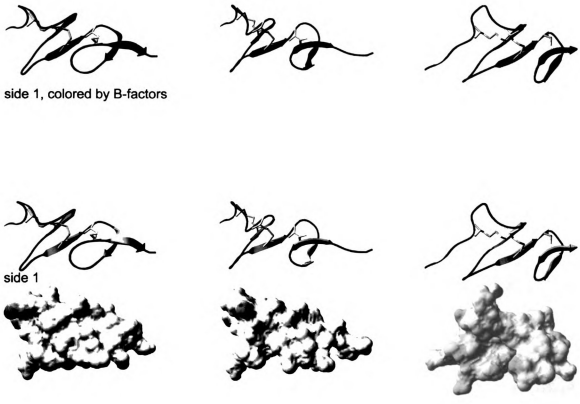

front
back
front
back
CLII-1
CLII-2
CLII-3
CLII-4
CLII-5
CLII-6
CLII-7
CLII-8
A. aegypti VgR
A. gambiae VgR

Fig. 16. The surface EPs of seven EGF repeats in *Aa*VgR. The surface EPs of models for EGF1,2, EGF3, EGF4, EGF5,6, and EGF7 are shown. Alternative models for EGF4 and EGF7 are compared. EGF1,2, EGF3, and EGF5,6 each had only one acceptable model. In each graph, the top row shows the models of *Aa*VgR EGF-like repeats colored by B-factors (model confidence factors, shown as color temperature). The green and red colors represent regions with certainty and uncertainty respectively. The second and fourth rows are ribbon diagrams of the modeled *Aa*VgR EGF-like repeats. The four β strands in each module are colored either in golden or in pink, and α helices are colored in aqua. Only side chains of cysteine residues are displayed to show three disulfide bonds in yellow. Calcium ions are represented by small spheres. All the EGF repeats are oriented with N terminals on left and C terminals on right. The third and fifth rows show corresponding molecular surfaces colored with EP (Red: -12, negative; blue: +6 positive).



*Aa*VgR EGF4 model comparison

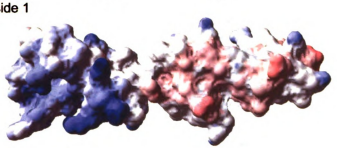


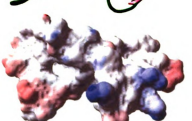
Modeled residue: 6-47	residue: 6-47	residue 4-45
E_{tot}: -354 KJ/mol	-68 KJ/mol	-209 KJ/mol
Idn/Sim: 29%/52%	29%/52%	29%/57%
1HJ7 N' model	1I0U N' model	1HJ7 C' model

*Aa*VgR EGF7 model comparison

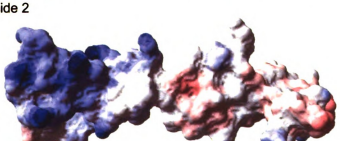
side 1, colored by B-factors

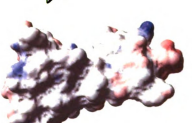
side 1





side 2





E_{tot}: -1427 KJ/mol

Idn/Sim: 34%/63%

EGF1,2 1I0U model

-101 KJ/mol

35%/60%

EGF3 1J8E model


side1 colored by B-factors
side1
side2
E_{tot}: -11 KJ/mol
Idn/Sim: 31%/58%
EGF4 1HJ7 model
-1793 KJ/mol
37%/63%
EGF5,6 1I0U model
-354 KJ/mol
29%/52%
EGF7 1HJ7 model

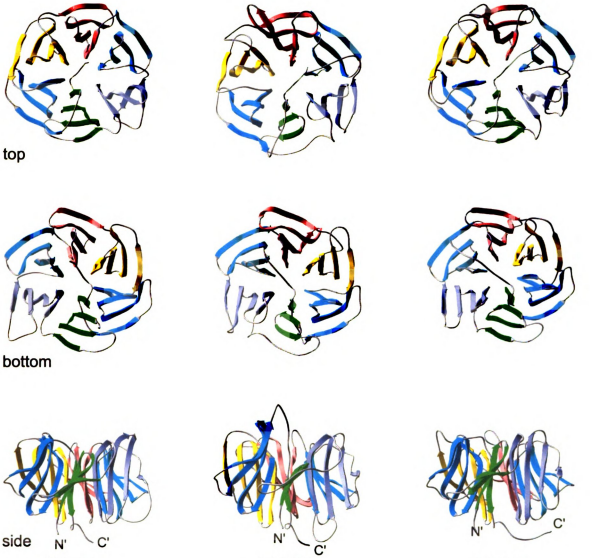
weakly positive surfaces. In comparison, EGF3 has a nearly neutral surface, and both EGF2 and EGF7 have weakly negative surfaces. The positive or weakly negative surface properties of EGF-like modules in *Aa*VgR exclude them from being involved in the binding of *Aa*Vg through primarily electrostatic complementarities.

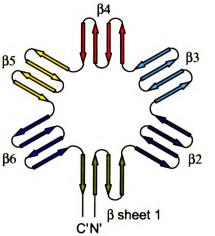
***The AaVgR YWTD β propellers and their role in Vg dissociation and receptor recycling*:** The three YWTD β propellers of *Aa*VgR were modeled with Swiss-Model server (Figure 17 and 18). The topology of the three propellers is the same as that of the six-bladed YWTD β propeller in human LDLR. There are 24 highly compact β strands organized in six anti-parallel β sheets (called blades), with four strands in each sheet (Figure 17).

The YWTD1 has, overall, highly negative surface EP, although the side surface, composed of β1 (green) and β6 (blue), is weakly positive. There are five histidine residues on both the top and bottom faces of YWTD1. Because histidine residues convert from neutral on the cell surface at around pH7 to bearing positive charges in the endosome at pH5, there is a possibility that these histidine residues can create a patch of positive surface area on the top and/or bottom of YWTD1 in the endosome.

In comparison with the YWTD1, the YWTD2 has, overall, a weakly to moderately positive surface, although there is a moderately negative side face composed of β4 (pink), which breaks the continuity of the positive surface. Significantly, there are eight histidine residues on the surface of the YWTD2, and five of them aggregate around the β4 sheet (four on the β4 sheet and one on the last β strand of β3) in a region with moderately negative EP. When receptors transfer from the cell surface to endosomes, the low pH may cause a significant change to the EP of this negatively charged surface patch,

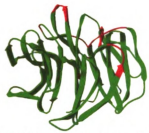
Fig. 17. Ribbon diagram and topology of the *Aa*VgR YWTD β propellers. The top threes rows are the top, bottom, and side views of models of the *Aa*VgR YWTD β-propeller 1,2, and 3. The β strands are shown as arrows. The six β sheets (propeller 'blades') are marked in different colors and numbered and the loop regions are colored in grey. The top view looks over the pseudosymmetry axis, with the N terminus coming out of the page and the C terminus entering the page. The bottom row is the topology diagram of three models. The six β sheets are colored and numbered as above, with centripetal and centrifugal arrows representing ascending and descending β strands, respectively.

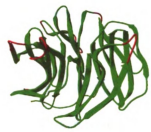




*Aa*VgR YWTD topology

Fig. 18. The surface EPs of the *Aa*VgR YWTD β propellers at pH7. The top row shows the ribbon diagrams of three *Aa*VgR YWTD β-propeller models colored by B-factors. The green and red colors represent regions with certainty and uncertainty, respectively. The second and fourth rows show models with β sheets colored as in Figure 17. The third and fifth rows show the corresponding molecular surfaces colored with EP (Red: -6 kEV, negative; white: 0, neutral; blue: +6 kEV, positive) at pH7. Histidine residues are drawn in red.

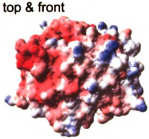
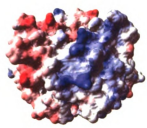
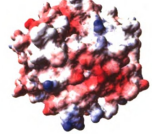




top & front, colored by B-factors

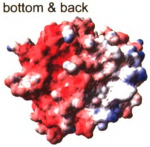
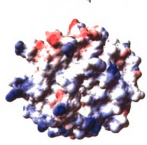
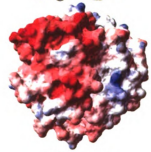


top & front



bottom & back

E_{tot}: -7307 KJ/mol	552 KJ/mol	78 KJ/mol
Idn/high sim: 31.1%/57.2%	18.2%/39.1%	16.8%/50.2%
Similarity: 64.8%	51.6%	63.7%
YWTD-1	YWTD-2	YWTD-3

and may possibly switch most of the surface of YWTD2 to positive, making YWTD2 a strong competitor to *Aa*Vg for the ligand-binding domain of *Aa*VgR.

The YWTD3 has a very strongly negative bottom surface. On the side of β5 (yellow), there is a strongly negative surface patch, and the rest of the surface is nearly neutral. On the surface of YWTD3, there are seven histidine residues: two on β2 (indigo) and five on the top. These five histidine residues may act as pH-sensitive switches that convert the surface charge of the top face from nearly neutral at the cell surface to negative in the endosome.

***The Insect VgR/YPR has one or two kind(s) of endocytosis signal(s)*:** Searches for potential sorting signals in the cytoplasmic tails of insect and nematode VgRs/YPRs found at least one or two LL/LI internalization motif(s) (Figure 19). Yxxϕ motifs and NPxY-like (or NxxY) motifs exist in vertebrate VgRs, while, in contrast, they are not found in invertebrate VgRs/YPRs, possibly because the task of VgR/YPR is simply to transfer yolk proteins to oocyte yolk storage vesicles. Both *Aa*VgR and *Dm*YPR have an FxNPxY-homologous sequence, yet a non-aromatic residue occupies the sixth position in both receptors, and thus it cannot form a tight turn essential to recognition by adaptor proteins. Unlike *Aa*VgR and *Dm*YPR, which both have a nonfunctional NPxY motif; both *Pa*VgR and *Ce*YPR have one functional NPxY motif. The NPxY motif in *Ce*YPR (YGNPMY) theoretically has the highest efficacy of internalization, and the one in *Pa*VgR (FKNPTF) should have a less-than-typical efficiency of endocytosis with no tyrosine residue at either position 1 or position 6.

*Aa*Vg modeling—Evidence of A Positive Surface

***The "receptor-binding motif" on blue tilapia Vg is not a receptor-binding site*:** Recently a "receptor-binding motif", HLTKTKDL (position 182-189), in the small subunit of *Orechromis aureus* (blue tilapia) Vg was reported (Li *et. al.* 2003). Mutation of either H182 or K187 had no effect on VgR binding, and K185A mutant attenuated the binding to approximately 50% (Li *et. al.* 2003, Figure 8), which suggested that the basic

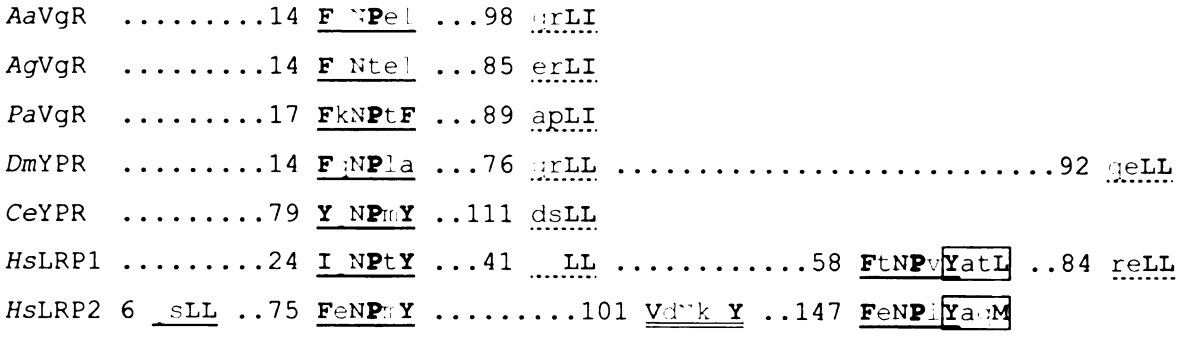


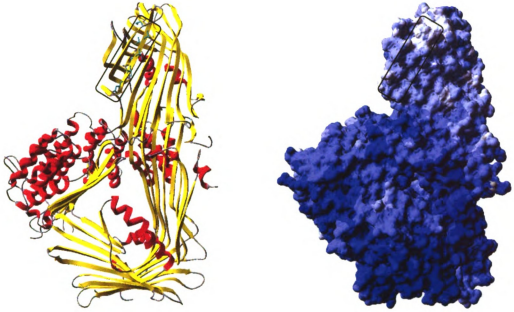
Fig. 19. Potential sorting signals in the cytoplasmic tails of invertebrate VgRs/YPRs, human LPR1, and human LPR2. A number on the left of each underlined sequence indicates the distance from the leftmost residue in the motif to the transmembrane helix of receptor. *Aa*VgR, *Aedes aegypti* VgR; *Ag*VgR, *Anopheles gambiae* VgR; *Ce*YPR, *Caenorhabditis elegans* YPR; *Dm*YPR, *Drosophila melanogaster* YPR; *Hs*LRP1, *Homo sapiens* LRP1; *Hs*LRP2, *Homo sapiens* LRP2; PaVgR, *Periplaneta americana* VgR; single and double underlines, NPxY motif and NPxY-like motif respectively; framed residues, Yxxϕ motif; bald upper case and lower case, conserved and not conserved residues in a motif; dotted line, omitted sequence.

Fig. 20. The claimed "receptor-binding motif" of blue tilapia Vg is not a receptor-binding site. ***A***, The ribbon diagram (left) of the *Ichthyomyzon unicuspis* lipovitellin monomer (*Iu*LV, protein database ID: 1lsh) and the surface EP (right) are shown to highlight the region in *Iu*LV corresponding to the "receptor-binding motif" of tilapia Vg. The left face of *Iu*LV is oriented with the helical domain facing left, and the N-sheet domain and the lipid cavity on the top and bottom, respectively. α helices and β strands are colored in red and yellow, respectively. Side chains of residues on *Iu*LV corresponding to the "receptor-binding motif" of tilapia Vg are drawn in aqua, and boxed. The right figure shows the corresponding molecular surface colored with EP (Red: -0.1, slightly negative; white: 0, neutral; blue: +14, strongly positive). ***B***, Partial sequences of the N-sheets of 40 vitellogenins from 16 insects, 13 fishes, 1 amphibian, and 2 birds were aligned using Pileup of the SeqWeb package (Accelrys) and adjusted by hand. The multiple-alignment shows the counterparts of the claimed "receptor-binding motif" of tilapia Vg and the locations of two gaps in vertebrate Vgs and the first gap in some cockroach Vgs. Gray loops, disulfide bonds; red-colored dotted underline, sequence stretch corresponding to the proposed VgR-binding motif of blue tilapia Vg; single underlines, α helices; double underlines, β strands. Abbreviations: *Ap*, *Antheraea pernyi* (Chinese oak silkmoth); *Ay*, *Antheraea yamamai* (Japanese oak silkmoth); *Bmo*, *Bombyx mori* (domestic silkworm); *Scr*, *Samia cynthia ricini* (Indian eri silkmoth); *Ld*, *Lymantria dispar* (gypsy moth); *Aa*, *Aedes aegypti* (yellow fever mosquito); *Aga*, *Anopheles gambiae* (African malaria mosquito); *Ar*, *Athalia rosae* (coleseed sawfly); *Pn*, *Pimpla nipponica* (parasitoid wasp); *Ps*, *Plautia stal*i (brown-winged green stink bug); *Rc*, *Riptortus clavatus* (bean bug); *Agr*, *Anthonomus grandis* (boll weevil); *Gn*,

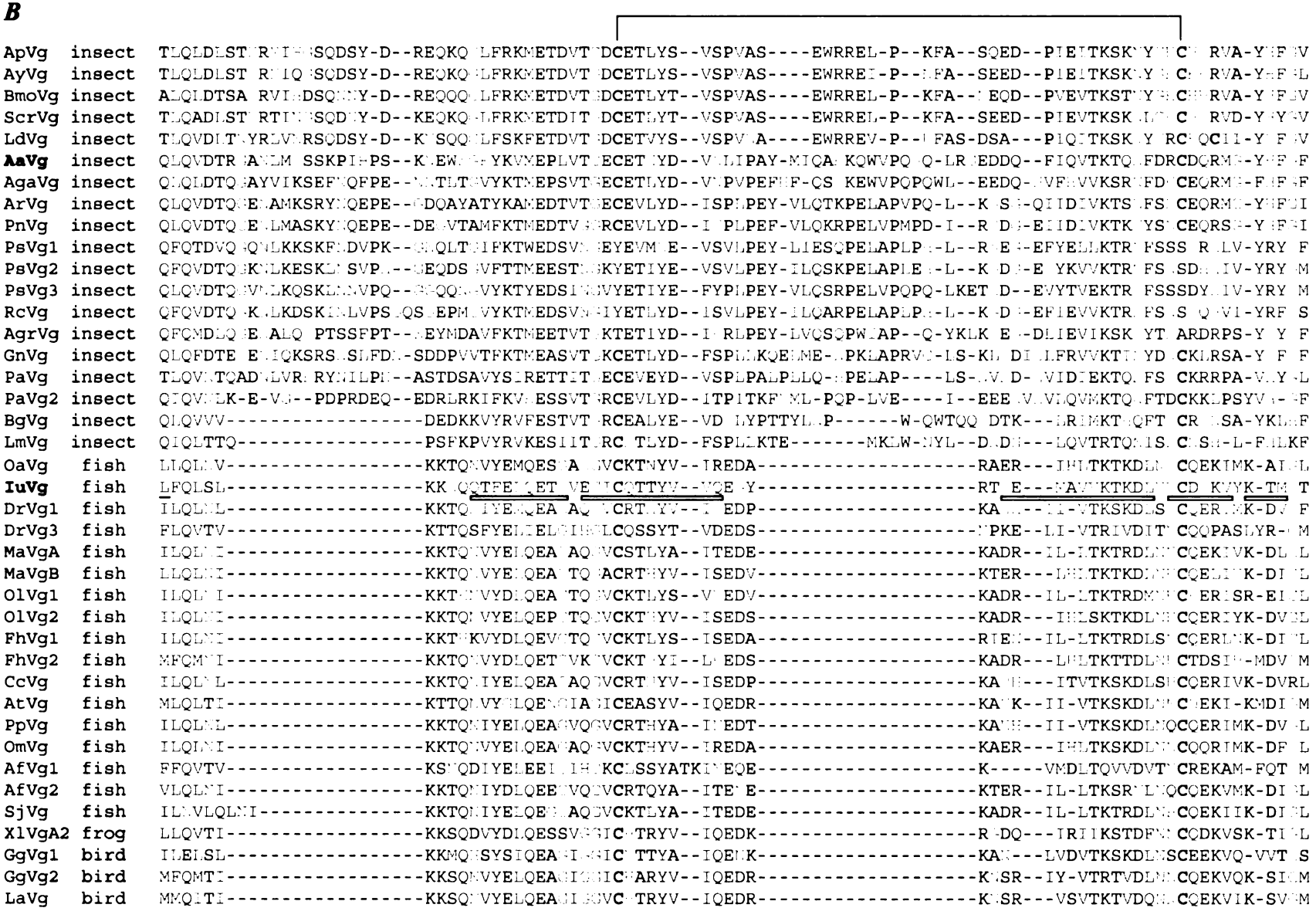
Anopheles gambiae (African malaria mosquito); *Af*, *Athalia rosae* (coleseed sawfly); *Pn*, *Pimpla nipponica* (parasitoid wasp); *Ps*, *Plautia stali* (brown-winged green stink bug); *Rp*, *Riptortus pedestris* (bean bug); *Ags*, *Acyrthosiphon pisum* (pea aphid); *Tc*,

Graptopsaltria nigrofuscata (cicada); *Pa*, *Periplaneta americana* (American cockroach); *Bg*, *Blattella germanica* (German cockroach); *Lm*, *Leucophaea maderae* (Madeira cockroach); *Oa*, *Oreochromis aureus* (blue tilapia); *Iu*, *Ichthyomyzon unicuspis* (silver lamprey); Dr, *Danio rerio* (zebrafish); *Ma*, *Melanogrammus aeglefinus* (haddock); *Ol*, *Oryzias latipes* (Japanese medaca); *Fh*, *Fundulus heteroclitus* (mummichog); *Cc*, *Cyprinus carpio* (common carp); *At*, *Acipenser transmontanus* (white sturgeon); *Pp*, *Pimephales promelas* (fathead minnow); *Om*, *Oncorhynchus mykiss* (rainbow trout); *Af*, *Acanthogobius flavimanus* (yellowfin goby); *Sj*, *Sillago japonica* (ray-finned fish); *Xl*, *Xenopus laevis* (African clawed frog); *Gg*, *Gallus gallus* (chicken); *La*, *Larus argentatus* (herring gull).

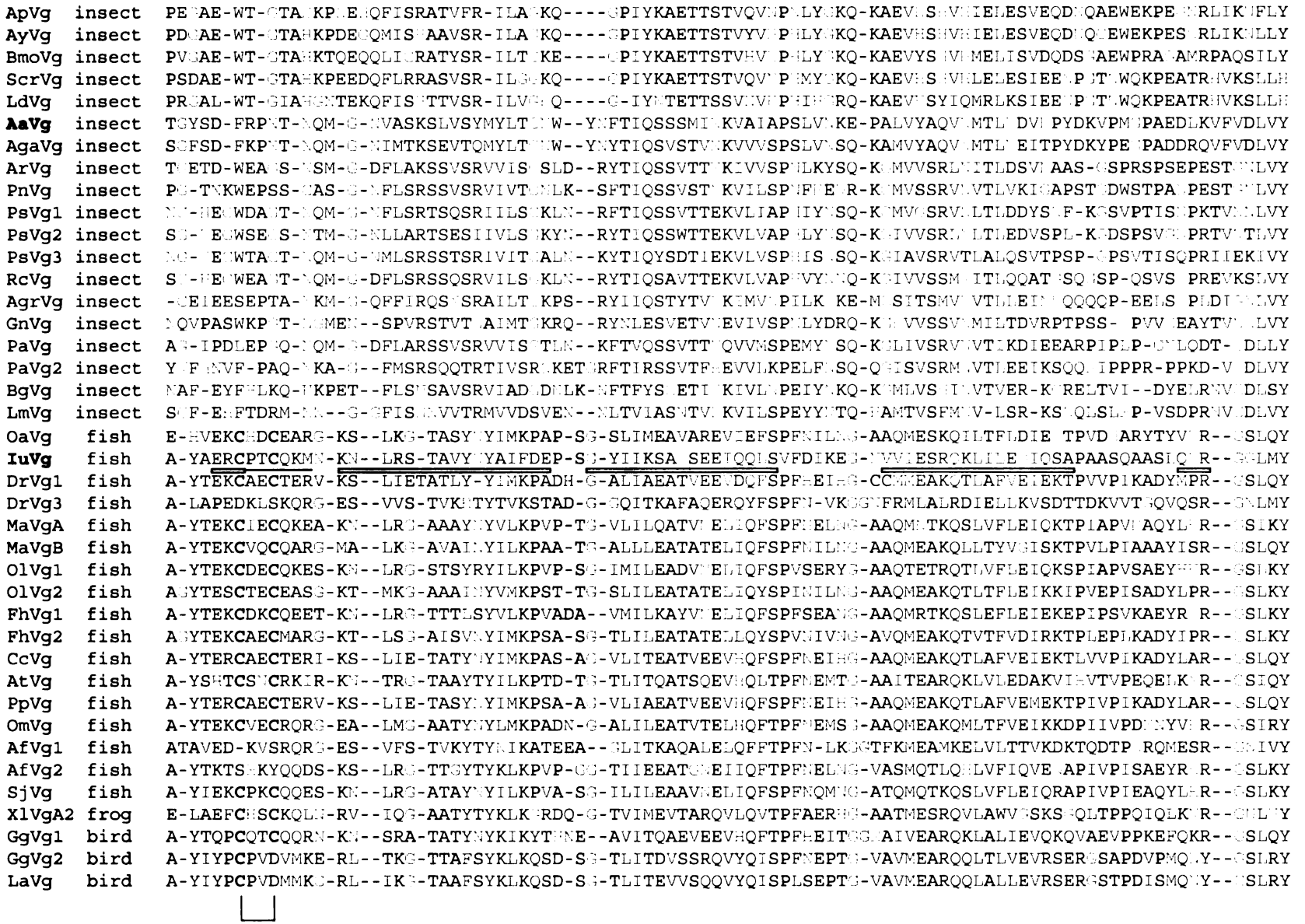
A



B



ApVg	insect	TLQLDLST RVI SQDSY-D--REQKQ LFRKMETDVT DCETLYS--VSPVAS----EWRREL-P--KFA--SQED--PIEITKSK Y C RVA-Y F V
AyVg	insect	TLQLDLST R IQ SQDSY-D--REQKQ LFRKMETDVT DCETLYS--VSPVAS----EWRREI-P--KFA--SEED--PIEITKSK Y C RVA-Y F L
BmoVg	insect	ALQLDTSA RVI DSQ Y-D--REQQQ LFRKMETDVT DCETLYT--VSPVAS----EWRREL-P--KFA--EQD--PVEVTKST Y C RVA-Y F V
ScrVg	insect	TLQADLST RTI SQD Y-D--KEQKQ LFRKMETDIT DCETLYT--VSPVAS----EWRREL-P--KFA--SEED--PIEVTKSK L C RVD-Y Y V
LdVg	insect	TLQVDLT YRLV RSQDSY-D--K SQQ LFSKFETDVT DCETVYS--VSPV A----EWRREV-P--I FAS-DSA---PIQITKSK Y RC QC I-Y F V
AaVg	insect	QLQVDTR A LM SSKPI PS--K EW Y KVMEPLVT ECET YD--V LIPAY-MIQA KQWVPQ Q-LR EDDQ--FIQVTKTQ FDRCDQRM -Y F F
AgaVg	insect	Q QLDTQ AYVIKSEF QFPE-- TLT VYKTMEPSVT ECETLYD--V PVPEF F-QS KEWVPQPQWL--EEDQ VF VVKSR FD CEQRM -F F F
ArVg	insect	QLQVDTQ E AMKSRY QEPE-- DQAYATYKAMEDTVT ECEVLYD--ISP PEY-VLQTKPELAPVPQ-L--K S -QIIDIVKTS FS CEQRM -Y F I
PnVg	insect	QLQVDTQ E LMASKY QEPE--DE VTAMFKTMEDTV RCEVLYD--I PLPEF-VLQKRPELVPMPD-I--R D -EIIDIVKTK YS CEQRS -Y F I
PsVg1	insect	QFQTDVQ Q LKKSKF DVPK-- QLQLT IFKTWEDSV EYEVM E--VSVLPEY-LIESQPELAPLP -L--R E --EFYELLKTR FSSS R LV-YRY F
PsVg2	insect	QFQVDTQ K LKESKL SVP --QEQDS VFTTMEEST KYETIYE--VSVLPEY-ILQSKPELAPLE -L--K D -E YKVVKTR FS SD IV-YRY M
PsVg3	insect	QLQVDTQ V LKQSKL VPQ-- QQQ VYKTYEDSI VYETIYE--FYPLPEY-VLQSRPELVPQPQ-LKET D--EVYTVEKTR FSSSDY IV-YRY M
RcVg	insect	QFQVDTQ K LKDSKI LVPS QS EPM VYKTMEDSVN IYETLYD--ISVLPEY-ILQARPELAPLP -L--K D --EFIEVVKTR FS S Q VI-YRF S
AgrVg	insect	QFQMDLQ E ALQ PTSSFPT-- EYMDAVFKTMEETVT KTETIYD--I RLPEY-LVQSQPW AP--Q-YKLK E --DLIEVIKSK YT ARDRPS-Y Y F
GnVg	insect	QLQFDTE E IQKSRS SLFDI SDDPVVTFKTMEASVT KCETLYD--FSPL KQE ME- PKLAPRV -LS-KL DI LFRVVKTI YD CKLRSA-Y F F
PaVg	insect	TLQV TQAD LVR RY ILPI--ASTDSAVYSIRETTIT ECEVEYD--VSPLPALPLLQ- PELAP----LS- V D- VIDIEKTQ FS CKRRPA-V Y L
PaVg2	insect	QIQV LK-E-V --PDPRDEQ--EDRLRKIFKV ESSVT RCEVLYD--ITPITKF ML-PQP-LVE----I--EEE V VLQVMKTQ FTDCKKLPSYV - F
BgVg	insect	QLQVVV----------DEDKKVYRVFESTVT RCEALYE--VD LYPTTYL P------W-QWTQQ DTK---RIMKT QFT CR SA-YKL F
LmVg	insect	QIQLTTQ----------PSFKPVYRVKESIIT RC TLYD--FSPLLKTE-------MKLW-NYL--D D---LQVTRTQ IS C S -L-F ILKF
OaVg	fish	LLQL V----------KKTQ VYEMQES A VCKTNYV--IREDA----------RAER---IHLTKTKDL CQEKIMK-AI L
IuVg	fish	LFQLSL----------KK QQTEE ET VE IC TTYV--VQE Y----------RT E---NA TKTKDL CD KVYK-TM T
DrVg1	fish	ILQL L----------KKTQ IYE QEA AQ CRT YV--I EDP----------KA --I-VTKSKDLS CQER K-D F
DrVg3	fish	FLQVTV----------KTTQSFYELIELP LCQSSYT--VDEDS----------PKE--LI-VTRIVDIT CQQPASLYR- M
MaVgA	fish	ILQL I----------KKTQ VYEIQEA AQ VCSTLYA--ITEDE----------KADR--IL-ITKTRDL CQEKIVK-DL L
MaVgB	fish	LLQL I----------KKTQ VYE QEA TQ ACRT YV--ISEDV----------KTER---L TKTKDI CQEL K-DI L
OlVg1	fish	ILQL I----------KKT VYDIQEA TQ VCKTLYS--V EDV----------KADR--IL-LTKTRDM C ER SR-EI L
OlVg2	fish	ILQL I----------KKTQ VYELQEP TQ VCKT YV--ISEDS----------KADR---I LSKTKDL CQERIYK-DV L
FhVg1	fish	ILQL I----------KKT KVYDLQEV TQ VCKTLYS--ISEDA----------RIE --IL-LTKTRDLS CQERL K-DI L
FhVg2	fish	MFQM I----------KKTQ VYDLQET VK VCKT YI--L EDS----------KADR---L LTKTTDL CTDSI -MDV M
CcVg	fish	ILQL L----------KKTQ IYELQEA AQ VCRT YV--ISEDP----------KA ---ITVTKSKDLS CQERIVK-DVRL
AtVg	fish	MLQLTI----------KTTQ VY LQE IA ICEASYV--IQEDR----------KA K--II-VTKSKDL C EKI-KMDI M
PpVg	fish	ILQL L----------KKTQ IYELQEA VQ VCRTHYA--I EDT----------KA --II-VTKSKDL QCQERIMK-DV L
OmVg	fish	ILQL I----------KKTQ VYELQEA AQ VCKTHYV--IREDA----------KAER---I LTKSKDL CQQRIMK-DF L
AfVg1	fish	FFQVTV----------KS QDIYELEE I IH KCLSSYATKI EQE----------K-----VMDLTQVVDVT CREKAM-FQT M
AfVg2	fish	VLQL I----------KKTQ IYDLQEE VQ VCRTQYA--ITE E----------KTER--IL-LTKSR L QCQEKVMK-DI L
SjVg	fish	IL VLQL I----------KKTQ IYELQE AQ VCKTLYA--ITEDE----------KADR--IL-LTKTRDL CQEKIIK-DI L
XlVgA2	frog	LLQVTI----------KKSQDVYDLQESSV IC TRYV--IQEDK----------R DQ---IRIIKSTDF CQDKVSK-TI L
GgVg1	bird	ILEISL----------KKMQ SYSIQEA I IC TTYA--IQEIK----------KA ---LVDVTKSKDL SCEEKVQ-VVT S
GgVg2	bird	MFQMTI----------KKSQ VYELQEA I IC ARYV--IQEDR----------K SR--IY-VTRTVDL CQEKVQK-SI M
LaVg	bird	MMQITI----------KKSQ VYELQEA I VC TRYV--IQEDR----------K SR---VSVTKTVDQ CQEKVIK-SV M


ApVg insect
AyVg insect
BmoVg insect
ScrVg insect
LdVg insect
AaVg insect
AgaVg insect
ArVg insect
PnVg insect
PsVg1 insect
PsVg2 insect
PsVg3 insect
RcVg insect
AgrVg insect
GnVg insect
PaVg insect
PaVg2 insect
BgVg insect
LmVg insect
OaVg fish
IuVg fish
DrVg1 fish
DrVg3 fish
MaVgA fish
MaVgB fish
OlVg1 fish
OlVg2 fish
FhVg1 fish
FhVg2 fish
CcVg fish
AtVg fish
PpVg fish
OmVg fish
AfVg1 fish
AfVg2 fish
SjVg fish
XlVgA2 frog
GgVg1 bird
GgVg2 bird
LaVg bird

residues in this motif are not very conserved. A multiple alignment among 40 Vgs from sixteen insects, thirteen fishes, one amphibian, and two birds showed that the two basic residues in this motif are not highly conserved. In some species, only the first basic residue is conserved, and in some others, there is not even a basic residue in the corresponding region. The surface area of *Ichthyomyzon unicuspis* lipovitellin (*Iu*LV) corresponding to the "VgR-binding motif" of *Oa*Vg is on the left side of the N sheet domain, which corresponds to the small subunit of *Aa*Vg (*Aa*Vg-S). On *Iu*LV, this region is not more positive than most other areas on the N sheet, and it is much less positive than the A sheet domain, which corresponds to a portion of the big subunit of insect Vg (Figure 20*A*). Because tilapia Vg and lamprey LV are highly homologous, this motif in the modeled tilapia Vg is also not more positive than most other regions in the N-sheet (data not shown). The most important thing is that in *Iu*LV this motif is located at the monomer interface and is very near a disulfide bond. Just beside highly conserved first basic residue (K175) in this motif, threonine 176 in one monomer interacts with alanine 562 in another (Anderson *et. al.* 1998, in table 2). These are two reasons why this motif is conserved. Because the motif does not have to be positive and is buried in the docking face of two monomers, it is virtually impossible for this motif to be a part of the receptor-binding site.

The lamprey lipovitellin has an omni-positive surface EP: *Iu*LV has an omni-positive surface, and the surface of the N-sheet domain is much less positive than that of the lipid-binding cavity structure, which corresponds to the large subunit of *Aa*Vg (Figure 21). The lamprey vitellogenin has a very strongly positive surface around the lipid-

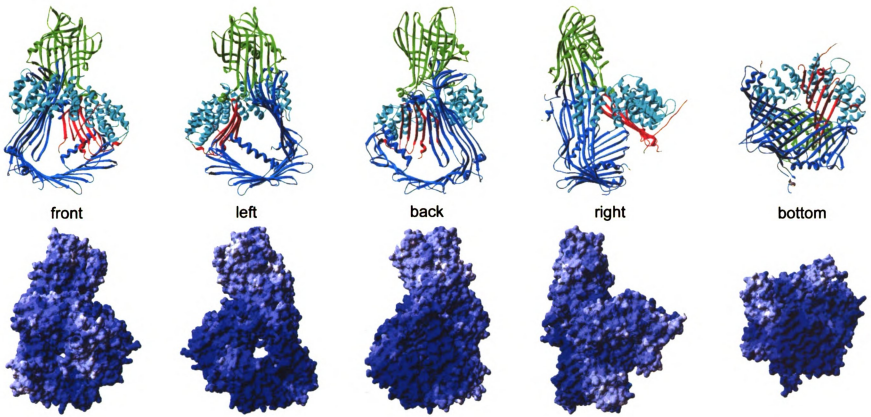


Fig. 21. Lamprey lipovitellin has a strongly positive surface surrounding the lipid-binding cavity. The top row shows ribbon diagrams of *I. unicuspis* (lamprey) lipovitellin (*Iu*LV, protein database ID: 1lsh) monomer. The N-sheet, helical domain, C-sheet, and A-sheet are shown in green, aqua, red, blue, and gray respectively. The bottom row shows the corresponding molecular surfaces colored with EP (Red: -0.1, slightly negative; white: 0, neutral; blue: +10, strongly positive).

binding cavity (bottom row in Figure 21). The strongest positive surface is located on the A-sheet, which corresponds to a part of the large subunit of *Aa*Vg (*Aa*Vg-L).

The AaVg small subunit has a positively charged surface: The *Aa*Vg-S was aligned with the N-sheet domain of *Ichthyomyzon unicuspis* lipovitellin (*Iu*LV) by hand (Figure 22 and Figure 20*B*). *Aa*Vg-S residue 85 to 399 was modeled based on the alignment (Figure 23). The structure of *Aa*Vg-S is very similar to that of the N-sheet domain of *Iu*LV. Because the β11 strand of the C sheet joins the A-sheet and the β12 strand of the C sheet joins the C-sheet in the *Aa*Vg-S regional model, the β12 and β13 strands do not show up. Two significant differences between *Aa*Vg-S and *Iu*LV are two insertion sequences in *Aa*Vg-S. The first gap in the alignment was predicted to be a pair of additional anti-parallel β strands (referred as β-A and β-B). Because these two additional β strands are beside β7 in the model, it is possible and quite reasonable that they join the N-sheet by contacting β7 at the left rim of the N-sheet. The second gap was predicted to be a long loop with an internal helix (α-A) by the Swiss-Model server. The 2-D structure prediction program, PROF, also predicted this region to have a helix with low reliability, while the program PHD made no prediction for this region. A check on the surface charge of the modeled *Aa*Vg-S showed that *Aa*Vg-S has a moderately positive surface and that the front and right sides are more positive (Figure 23).

The AaVg large subunit has the helical domain, C-sheet, and A-sheet: Regions in 40 Vg sequences corresponding to the helical domain, the C-sheet, and a part of the N terminal region of the A-sheet of *Iu*LV were aligned by hand (Figure 24). Residue 601 to 1301 of *Aa*Vg-L was modeled based on the multi-alignment (Figure 25). The structure of

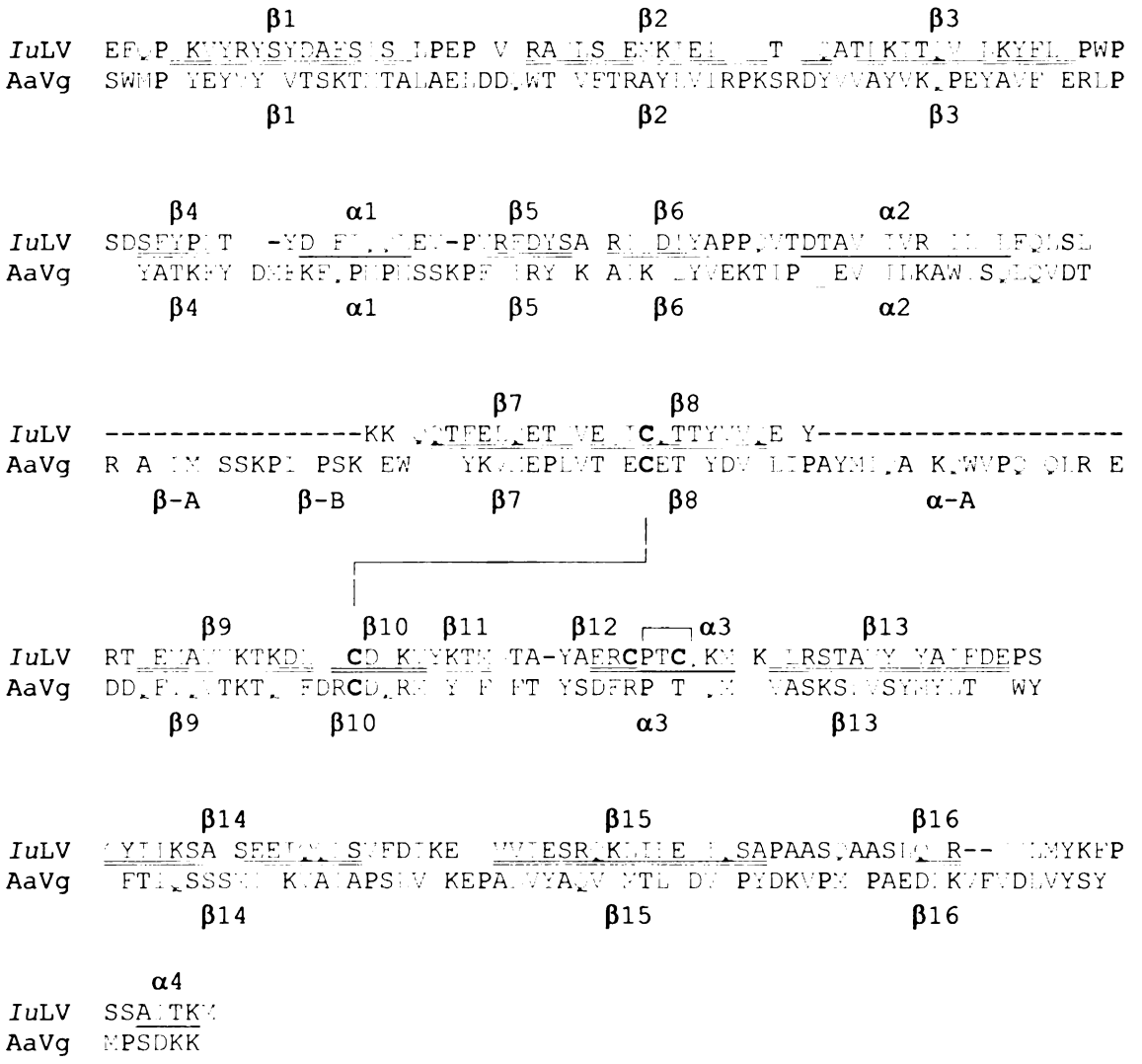


Fig. 22. 2-D structural alignment of the *Aa*Vg small subunit with the N-sheet domain of lamprey lipovitellin. The *Aa*Vg small subunit sequence was aligned with the N-sheet domain of *Ichthyomyzon unicuspis* lipovitellin (*Iu*LV) by hand. Two disulfide bonds in the N-sheet of *Iu*LV are represented as gray loops. Double underlines, β stands; single underlines, α helices; dark underlines, secondary structures of *Iu*LV (protein database file 1lsh); shallow underlines, secondary structures predicted in this model; α, α helix; β, β strand.

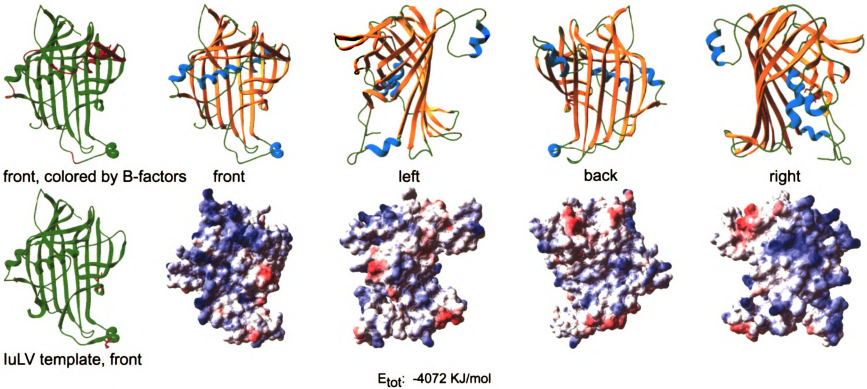
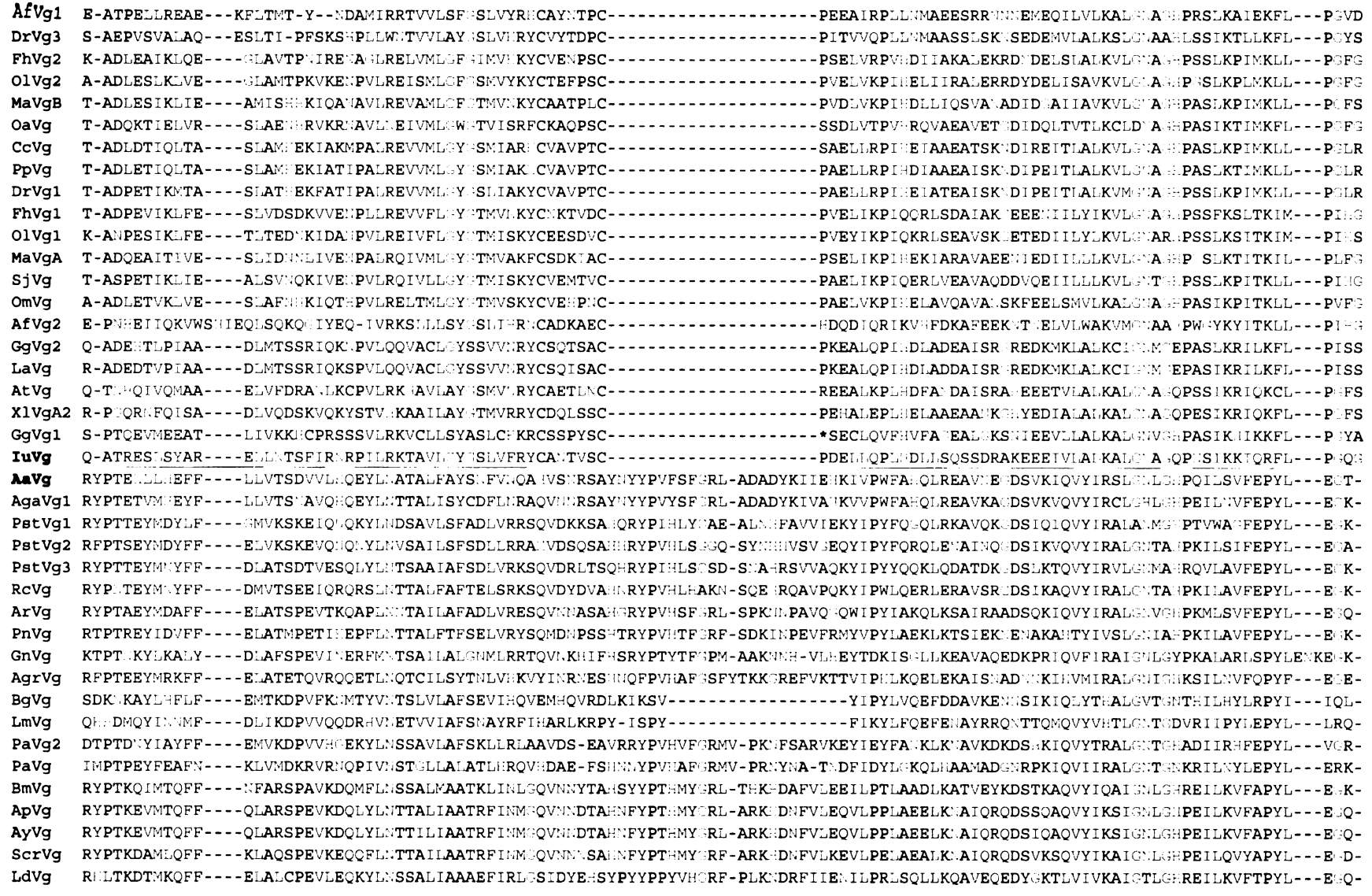
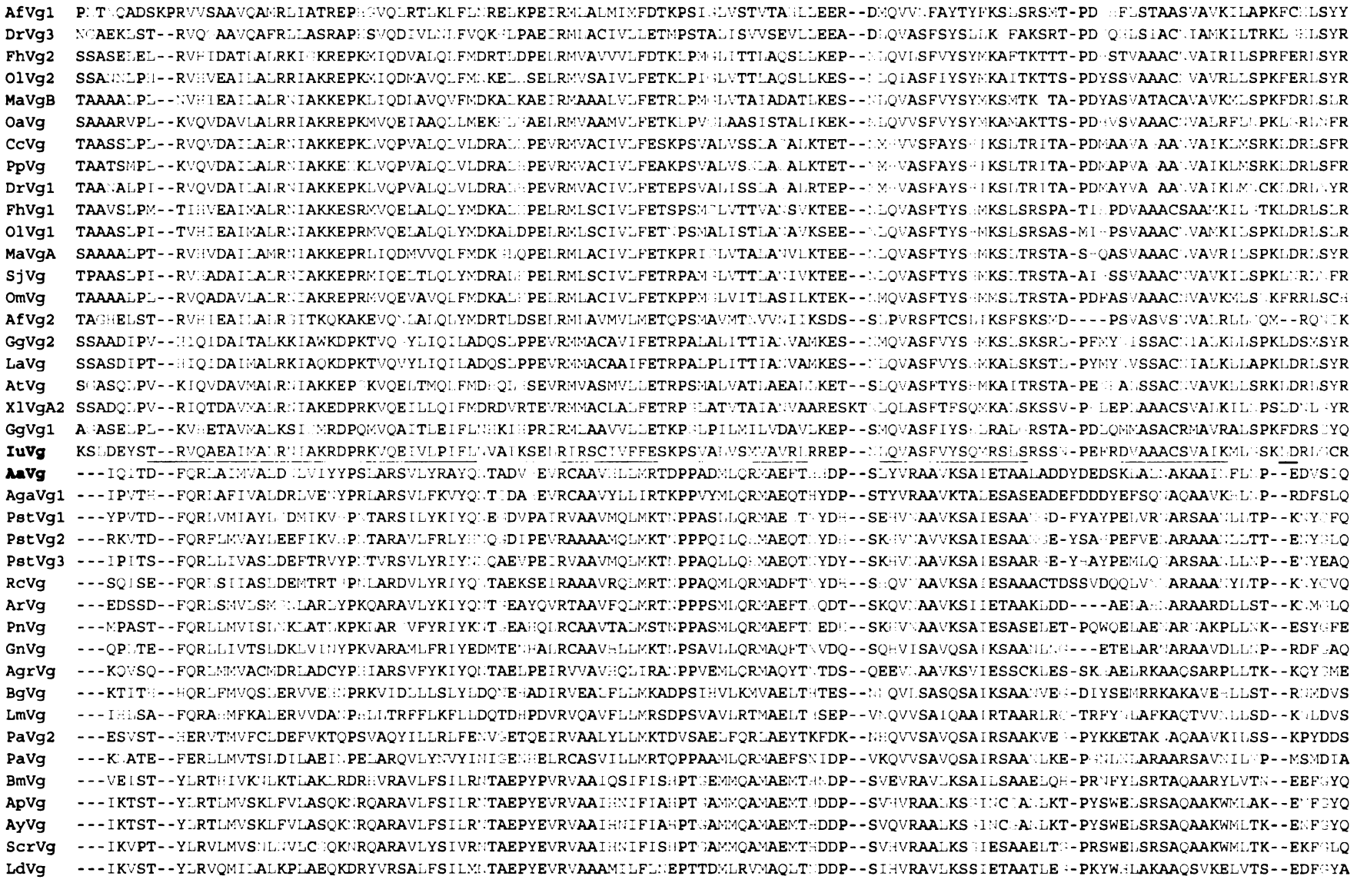


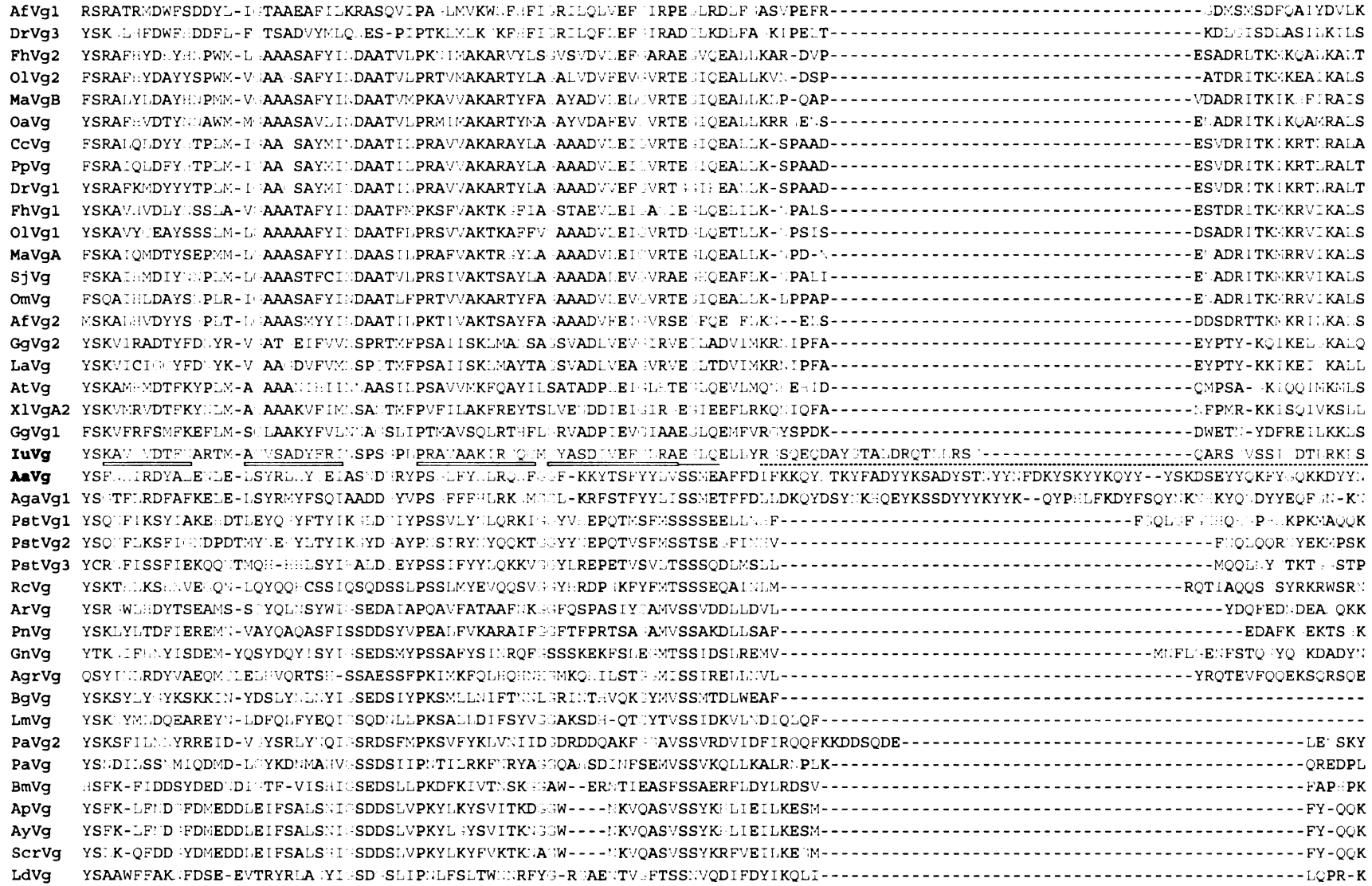
Fig. 23. The *Aa*Vg small subunit has a positively charged surface. The top row is the ribbon representation of *Aa*Vg-S. In the first ribbon diagram colored by B-factors, green and red colors represent regions with certainty and uncertainty, respectively. In the rest ribbon diagrams β strands are colored in golden and α helices in aqua. The side chains of two cysteine residues in a S-S bond are displayed in red. The corresponding molecular surfaces are colored with EP (Red: -6 kEV, negative; blue: +6 kEV positive).

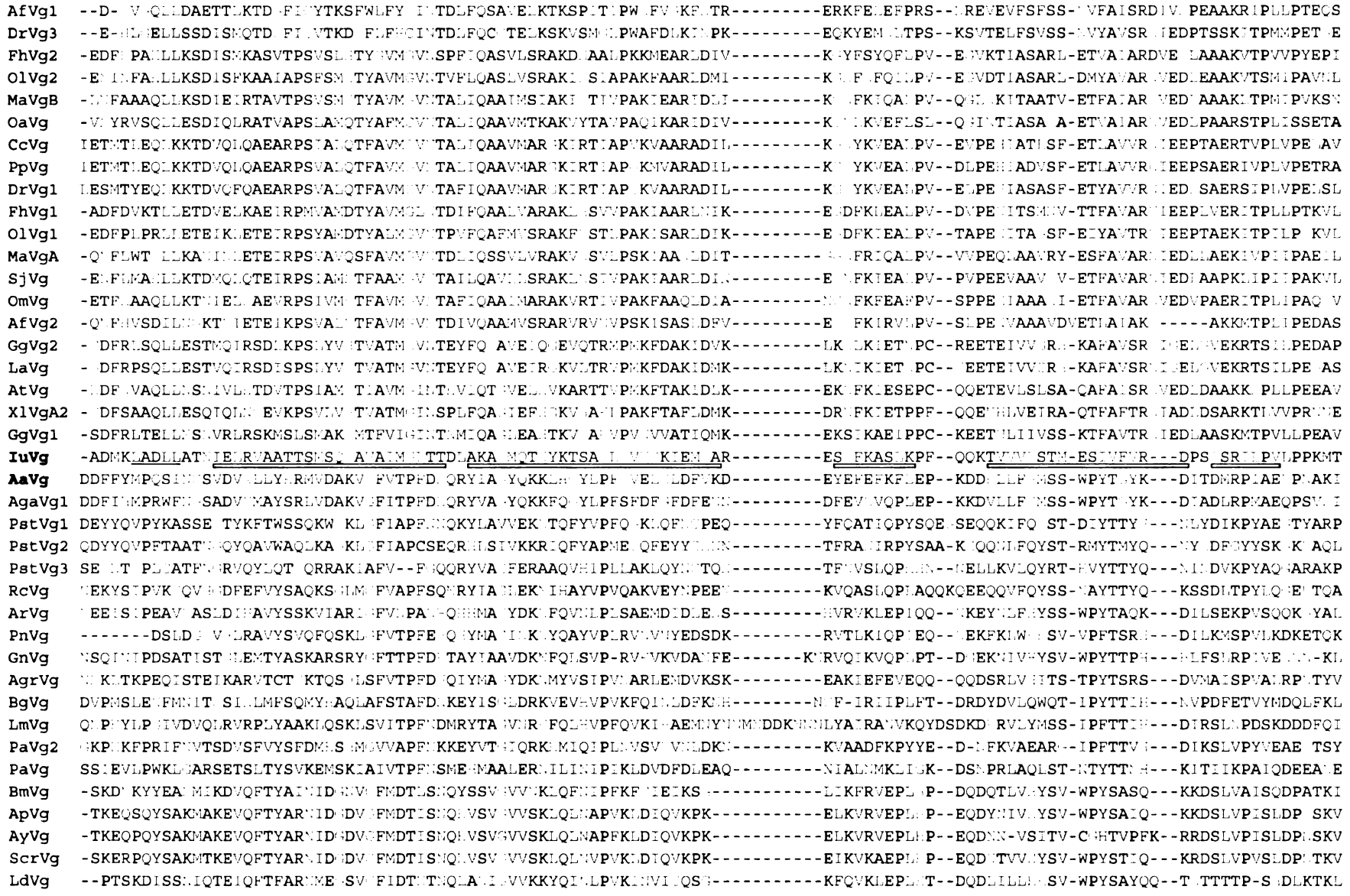
Fig. 24. Sequence alignment of the N terminal region of the *Aa*Vg large subunit with other vitellogenins. Regions in 40 Vg sequences from 16 insects, 13 fishes, 1 amphibian, and 2 birds corresponding to the helical domain, the C-sheet, and a part of the A-sheet of *Iu*LV were aligned by hand. The aligned region (amino acid residue 601 to 1301) of *Aa*Vg was fully modeled. Abbreviations are the same as in Figure 22. Single underlines, α helices; double underlines, β strands; golden-colored underlines, secondary structures the helical domain of *Iu*LV; red-colored underlines, secondary structures the C-sheet of *Iu*LV; blue-colored underlines, secondary structures in the N terminal region of the A-sheet of *Iu*LV; dotted underlines, invisible regions in the X-ray crystal structure of *Iu*LV.


AfVg1 KLDDPVPKATELIKRLVQANTNQ----LDSTTTEDAIKLYQLLRVI----PLEKLEKMWME----------VEHNRNERNWFLQTVVEVNDARILALLERLLREHKLNRLEAITVIVRAFNHL
DrVg3 NLNDPVPKILDLIKRLAQANIYH----VDSETSTEILDLIQLMRVT----TLDNLEHLWKQ----------VSNDEHRRWFLDLVVEVTDERILKFLEARYKA DITANEA QALVVAFNHL
FhVg2 RITNVEAQIVESLKNLVSLNM H----AHEDSPLKFIELIQLLRVA----KYESIEALWSQ----------FKTKIDHRHWLLSSIPAI THVALKFIKEKIVA EVTAAEAAQAIMSSTHLV
OlVg2 RISNAEAQIVETLNKLVTLNM K----AHEDSPLKFIELIQLLRVA----RYESIEALWDQ----------YKARSDYRTWLLSSIPAI NHVALKFIKEKIIS DLTSFEAAEALLSSLHLV
MaVgB RITNAEAQIVEVLNHLVTYNVAK----VHEDAPLKFVELIQLLRVV----SYERVDALWSQ----------YKVKPAYRHWFLNAVPAI THASLKFIKEKFLLDDISVVEAAQALI SIHMV
OaVg RVTHAEAQIVSTLNHLVASNVAK----VHEDAPLKFVELIQVMRVA----RFETIESLWAQ----------FKSRPDHRYWLLNAVPHIRTHAALKFLIEKLLANELSETEAAMALLECLHSV
CcVg KISDAPAQIIEVLKHLVANNVAM----VHEDAPLKFVQLIQLLRVS----TLENIEAIWAQ----------FKDKPAYRRWLLDALPSV TPVIIKFIKEKFLA ELTLPEFIQALVVALQMV
PpVg KISDARAQITEVLKHIVENNVAM----VHDDAPLKFVELIQLLRAA----TLENTEAIWAQ----------FKDKPVYRRWLLDALPAV TPVIVKFIKEKFLA DLTIPEFIQALVVALQMV
DrVg1 KISDAPAQIVEVLKHLVSNNKDM----VHDDAPFKFVQLVQLLRVA----SLEKIEAIWSQ----------FKDKPVYRRWLLDALPAV TPVIIKFIKEKFLA EFTTPEFIQTLVIALQMV
FhVg1 KISDAPAQVAEVLKHLATYNIED----VHENAPLKFLELVQLLRIA----RYEDLEMYWNQ----------YKKMSPHRHWFLDTIPAT TFA LRFIKEKFMAEEITIAEAAQAFITAVHMV
OlVg1 KVTDERAQTEELLNHLVTHNAEK----VNDHAPLKYLELIQFLRLA----RYEDLEVLWSK----------YRNMPSHRFWLLEAIPAI TSAAVRFIKEKFQAEDLSVAEAVRTLVAAVHMV
MaVgA KITNPQAQIVELLNHLVINNMQT----VHEEAPIKFLELIQVLRKS----NIADIDQIWSQ----------FK KPVYRQWILDTLPAI TSTTLRFIREKLMARDISILEAAQALVASIHMV
SjVg KVNNVQAQIVEVLNHLVSHNVER----VHEDAPLKFLEFVQLLRAA----RLEDLEMIWSQ----------YRARPASRQWILDAIPAVETPAALRFIKEKIMVDELTVAEAAQALIASIHMV
OmVg KISNARAQAVKILNHLVTYNTAP----VHEDAPLKFLQFIQLLRMA----SSETINAIWAE----------FKAKPAYRHWILDAVPSI SSVAVRFIKEKFLA DITIFEAAQALVAAVHMV
AfVg2 KMKEVQEQIPEVLNHLVVNNRDK----VHEDAPLKFLELIQLLRAS----DITQLRNLLTT----------HSMSAEKRQWLLDAIPNI TMAALEIVMAEVLKREISVPEATQLLI TLHMV
GgVg2 KTKNPEQRIVETLQHIVLNNQQD----FHDDVSYRFLEVVQLCRIA----NADNLESIWRQ----------VSDKPRYRRWLLSAVSAS TTETLKFLKNRIRNDDLNYIQTLLTVSLTLHLL
LaVg KTKNPEQRIVETLQHIVQNNQQD----FHYDVPYRFLELVQLCRVA----SADTLESIWRQ----------ASDKPRYRRWLLSAVSAA TTEALKFIKSRIRNDDINYLQTLLSVSFALHSM
AtVg KTRSPETKIKEVLQHLVQNNQQQ----VQSDAPSKFLQLTQLLRAC----THENIE IWRQ----------YEKTQLYRRWILDALPAAATPTAFRFITQRIMKRDLTDAEAIQTLVTAMHLV
XlVgA2 KTKSPEAQAVEVLQHLVQDTQQH----IREDAPAKFLQLVQLLRAS----NFENLQALWKQ----------FAQRTQYRRCLLDALPMA TVDCLKFIKQLIHNEELTTQEAAVLITFAMRSA
GgVg1 KIKDVERQIEERLQDLVETTYEQ----LPSDAPAKALKLMHLLRAA----NEENYESVWKQ----------FSSRPAYRRYLLDLLPAAASHRSLRFLRHKMERQELTNWEIAQTVLVALHSS
IuVg K KNLESEIHTVLKHLVENNQLS----VHEDAPAKFLRLTAFLRNV----DA VLQSIWHK----------LRQQKDYRRWILDAVPAMATSEALLFLKRTLASEQLTSAEATQIVASTLSNQ
AaVg RNVDAIRSV KLVEEIASDL-ENPSNLPKANTMSKFNILTRAIRAM YEDIYELAQKYFVSQKERQ-VAQFSDKKFSKRVDAWVTLRDAVAEA TPSAFKLIFDFIKEKKLR YEAATVIASLAQSI
AgaVg1 HNVDASRYAYKLAYEIADEL-QEISQVPKSNTLNKFTILARVLRTMHYQDIYDVCQKLFVSQKEREE-GSNHSESFAKKVDAWNTFRDALAQA TPPAFKVIKELIEEKKLR NEAASVIATLPKTI
PstVg1 SG-HTVDNIVKVARQI KFL-EYPSNIPKDNVLTYYSLLPHLIAAANSKQLEQATEKLYYPEERIQ ---STNKEVYHYYRSWQVYRDAVAQA T PALVTIYSWLKSAKVRENE AQLLAVLPKAA
PstVg2 E EQVVERMVRLARQMALSF-NNPSEMYKQDPLVQYLIMSRTASIASYEQLQEATEKLYYSEEKLSG---G-KPESYQSYRAWQAYRD VARA T PALVAMSSWIKS KVH NEAAHLLSVLPRTA
PstVg3 K-SQVVEKMVNVAKRI ETI-LEPNRIIDENILTSFSILARLVS ANSQQLQQAVDRMYYSETNVQ L S E TSVYQNKQ WVAFRDAIANA T PALVTIRALIDS KLK EEAAEVLATLPSAA
RcVg QKQNVPEQIKSLAKKV QWV-QNP QIPEENRLPLFSMLTRMVQNANSQQLEQVTQSLYHQQSKNQQ----QSDAERQQYQSWAAFRDAVAQA T PALVTIKQWIQSKKVQ EEAAQLLAVLPYTA
ArVg SQVNPVSVAKKLASQI EEL-QKPSNIPKQNTLEKFSHLAALLRTMSAKQIEQAAHDMYAPAKEN------ASKSESAKNAAWKAFRDAVAQA T PAFSVIKNWIQTNKIRNNQAADLIAVLPSTT
PnVg DEIEPITMAKTLASQI SDF-QDPNSITDEQTLEKFTLLVRIIRTMSTQQIAEAERDLYQSNNEIDP----HDESQSVRR TWAAFRDAVAQA TAPALVTITLWIKQKKIR VEAASVV VLAKTA
GnVg NSSKTYDNVVKLANEM EEIVQDASSMS LNTVEKFALLSYII TLNARTIEYI EQLYYPQSQVSN----DNREDSVKHQAWVAYRDAVVDA T PALVALSSWIQNKKVY MEAAKLLSALPDTV
AgrVg KDFDIRQNVENLVTEISDEIKQSEKTISK-HTLDKYTILNTLVRLMDEDDIQFVAEQMY----------SQMKN QQRYTWSIFRDSVAEA TI PALLNIKKWIETKKIQKTEAAQVI TLAQST
BgVg VDVDPV VAVRLSKDIAADLQ EPRV EDRHILPRFTILVRLLKQLKVSQIMEAARKL-YK------LE--NDHPNYMNWDTWRVYRDAVSQA TWSALNSIQQFISSEMVEPKEASHLITVLPAAV
LmVg SDVDPVKAVVQLANDI HDL-IDPDSLPDKDTITKFIIMVRVLRNLQLSEILDIAQQLQVK------LDSQMVRKDSPQWEAWKSFRDAVSQT THAAVHSIIIFLSRRYISQSEAQDLFNVLPAAV
PaVg2 EVDPEKIV-LLARTISSEL-QEPDTMVKKNILSRFSILTNLVRAASFSQLEEATKRLYYRVER------ADN DESKLD-AWKAYRDSVAQA TPAALKMVHTWIRKEYIKDEEAAKVVAVIPHAA
PaVg SSDDPAEIVKSLAEEIKSDM-KKPAYIPERSTHAKLMMMRDIVRTMTAKQLQKATSLI----------HSESKHDL WIAYRDMVSES THPALEELSIWIISKKLSSEE AELLATLPRAV
BmVg DKKQNAMNAQKILQDIAQQL-QNPNNMPKSDFLSKFNILVRLIASMSTEQLSQTSRSI----------ETAKTSNNIIKSDMWMIFRD VTQA TLPAFKQIQSWIENKKIQEEEAAQVVVALPRTL
ApVg DKKQNSVNAQKLLQDKAQEL-QNPNNMPKSDFLSKLNILVRLLAAMTTEQLSQTSRTI----------GGDKSSNNNIKADMWMVYRDAVAQA TMPAFQQIKAWINSKKIQ EEAAQVVASLPSTL
AyVg DKKQNSVNAQKLLQDIAQQL-QNPNNMPKSDFLSKFNILVRLLAAMTTEQLSQTSRTI----------EAAKSSNNNIKADMWMVYRDAVAQT TMPAFQQIKAWINSKKIQ EEAAQVVASLPSTL
ScrVg DKKQNSMNAQKLLQEIAQQI-QNPNNMPKADFLSKFNILVHLIASMTSEQLSQTSRTI----------EAAKSSNNNIKADIWMVYRDAVTQA TLPAFQQIKTWIQSKKIQ EEAAQVVASLSTTL
LdVg DKKQNAMSAQKLLQDLAQQL-QNPSTMPQNEFLTKFNILVRLISTMTTDQLAQFSR I----------EVGKASNKIAKVDMWFIFRDAVIQA TQPAFQLIETWIKTKKIQ FLAAEVISSLAKTL






AfVg1
DrVg3
FhVg2
OlVg2
MaVgB
OaVg
CcVg
PpVg
DrVg1
FhVg1
OlVg1
MaVgA
SjVg
OmVg
AfVg2
GgVg2
LaVg
AtVg
XlVgA2
GgVg1
IuVg
AaVg
AgaVg1
PstVg1
PstVg2
PstVg3
RcVg
ArVg
PnVg
GnVg
AgrVg
BgVg
LmVg
PaVg2
PaVg
BmVg
ApVg
AyVg
ScrVg
LdVg



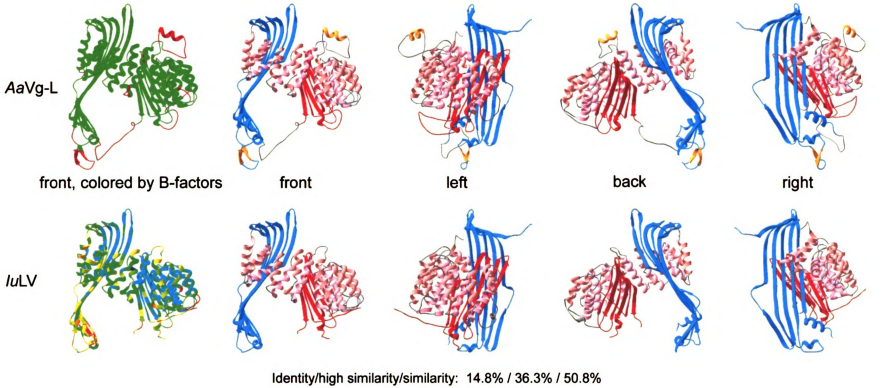


Fig. 25. The *Aa*Vg large subunit has the helical domain, C-sheet, and A-sheet. The top and bottom rows are ribbon representation of modeled residue 601-1301 of *Aa*Vg-L and its corresponding region (residue 297-950) on *Iu*LV. Diagrams in the first column are colored with B-factors, and a warmer color represents a region with higher uncertainty for the upper one or less rigidity for the lower one. In the rest ribbon diagrams, α helices in the outer and inner layers of the helical domain, the C-sheet, and the A-sheet are tan-, rose-, red-, and blue-colored respectively. One extra α helix and two β strands in *Aa*Vg-L are golden-colored.

the modeled *Aa*Vg-L is very similar to that of *Iu*LV. In the modeled helical domain, all 18 helices in the *Iu*LV are conserved in *Aa*Vg-L. Between two *Aa*Vg counterparts corresponding to the α9 and α10 on *Iu*LV, an *Aa*Vg-L insertion sequence was predicted to contain an extra α helix by the Swiss-Model server (Figure 25, golden-colored helix). The exact orientation of this extra helix is unsure because *Iu*LV lack this helix. It is possible that this extra helix joins one layer of helices by interacting with the AaVg-L counterparts to helices α9 and α10, but this is beyond the modeling capability of the Swiss-Model server. In the modeled N terminal part of the A-sheet, there is an insertion that was predicted to have two short anti-parallel β strands (Figure 25, golden-colored helix). The structure and location of this extra tiny β sheet look nice. Because most of the A-sheet was not predicted, it is still too early to check the surface EP of the AaVg-L.

DISCUSSION

The CLI of CRs in *Aa*VgR Has A More Strongly Negative Surface Than The CLII

All the thirteen modules from both clusters of CRs were modeled. Comparisons of surface EPs of modules from both clusters led to a prediction that the CLI has a more strongly negative EP than the CLII (Figure 13). The calcium atom was excluded during modeling. Because the five to six calcium-coordinating residues are highly conserved in each module, although the calcium atom will attenuate the negative surface potential somewhat, the trend of the surface EP of each module will not change much. In addition, the influence of calcium on the surface potential is roughly equal for each module, so calcium can be overlooked during module comparison.

The prediction from protein modeling of two clusters of CRs in *Aa*VgR is a good fit to the experimental results from ligand-binding assays. With regard to the *Aa*VgR, the protein modelings predicted that CRI-1, CRI-5, and CRII-3 would have the most strongly negatively charged surfaces; CRI-3, CRII-2, and CRII-4 would have strongly negative surfaces; and CRII-7 would have a somewhat strongly negative surface. To sum up, the CLI has a more strongly negative EP than the CLII of *Aa*VgR, and the surface charges of both clusters fall in the same scale. The results of saturation ligand-binding assays also showed that, in *Aa*VgR, the affinity of the CLI is approximately 100% higher than that of the CLII, while the affinity of the full-length *Aa*VgR is much higher than either cluster. In *Aa*VgR, the surface EP of CRII-5 is one of the four weakest CRs (the other three are those of CRII-1, CRII-6, and CRII-8) in the CLII. In *A. gambiae* VgR, counterparts of almost all *Aa*VgR CRs are present, except that of CRII-5. Nematode YPR (*Ce*YPR) lacks not only CRII-5 but also CRII-7 and CRII-8.

The Role of the *Aa*VgR EGF Homology Domain in *Aa*Vg Release

The decision to include one β propeller fold together with its three flanking EGF-like repeats C terminal to the CLI and CLII in currently constructed mini-receptors was based on two earlier reports. Davis *et. al.* found that a truncated LDLR lacking the EGF precursor homology domain showed markedly reduced affinity to LDL but not β-VLDL compared to native LDLR on the surface of transfected Chinese hamster ovary cells, and complete degradation in the transfected cells 4 hours after incubation with β-VLDL on the cell surface (Davis *et. al.* 1987). The EGF precursor homology domain was also shown to be important for *in vivo* binding of the LDLR to LDL on the cell surface, but

not for *in vitro* binding, because a truncated LDLR did bind LDL on a ligand blot (Davis *et. al.* 1987). Deletion analysis in the EGF precursor homology domain of LDLR by Esser *et. al.* also suggested its role in efficient binding of LDL (Esser *et. al.* 1988). Binding analyses of their mutant receptors showed that the first EGF-like repeat was required for binding of LDL, but not β-VLDL, while the second EGF-like repeat was not required for the ligand binding (Esser *et. al.* 1988).

Each modeled EGF-like repeat of *Aa*VgR has a surface EP varying from strongly positive to weakly negative (Figure 16). Although the bound calcium was not included during protein modeling, calcium will not make the actual surface EP of each EGF-like module more negative, and thus the conclusion that EGF-like repeats have no overall strongly negative surfaces is correct. Because until now the negative-positive electrostatic attraction is still thought to dominate the receptor-ligand interaction in the LDLR family, and because *Aa*VgR EGF-like repeats do not have strongly negative surfaces, EGF-like repeats in *Aa*VgR surely could not bind *Aa*Vg via electrostatic attraction, and quite possibly they do not bind *Aa*Vg at all.

The acid-dependent conformational change in the EGF-like and /or YWTD repeats of the LDLR has long been reported to alter the ligand binding properties of the LDLR, allowing dissociation from bound ligand after acidification of endocytic vesicle and recycling of LDLR (Davis *et. al.* 1987). Rudenko *et. al.* have resolved the extracellular domain of the LDLR at endosomal pH (pH5.3), and showed that both the CR4 and CR5 interact with the top face of the YWTD β propeller via hydrophobic contacts and salt bridges (Rudenko *et. al.* 2002).

The three YWTD β propellers of *Aa*VgR were modeled and surface EPs at pH7 calculated (Figure 17 and 18). The YWTD1 has five histidine residues on the top face and another five on the bottom face; the YWTD2 has five histidine residues on the β4 sheet with moderately negatively charged surface; and the YWTD3 also has five histidine residues on the nearly neutral top face. These gathered histidine residues on three YWTD propellers potentially could significantly change the surface charge at endosomal pH (pH5), even though the calculation of the surface EPs of the *Aa*VgR YWTD propellers at pH5 (rather than pH7) has not been performed yet. These histidine residues might possibly act as pH-sensitive switches that turn β propellers into false ligands of *Aa*VgR at acidic pH. The change in the surface charges of *Aa*VgR YWTD β propellers should be far more significant than the predicted limited change on the weakly negative top face of the human LDLR YWTD fold via merely two histidine residues (Jeon and Blacklow 2003, Figure 4).

Sorting Signals in The Cytoplasmic Tail of LDLR Family Members

The sorting signals in the cytoplasmic tails of membrane proteins mediate targeting of membrane proteins to the endosomes, the lysosomes, the apical plasma membrane, the basolateral plasma membrane, and the trans-Golgi network (Ohno *et. al.* 1995; Takeda *et. al.* 2003). Four sorting signals have been reported so far, which include the NPxY motif, the NPxY-like motif, the Yxxϕ motif, and the LL/LI motif. The internalization efficiency of the FxNPxY signal increases when the xNPx core sequence is flanked by two aromatic residues (F, H, and Y), including at least one Y (Paccaud *et. al.* 1993). Y at the first and/or the last position of this motif has a stronger propensity than

F in inducing a tight turn conformation and hence stronger internalization efficiency (Paccaud *et. al.* 1993). The Yxxϕ motif (ϕ is an residue with a bulky hydrophobic side chain, *e.g.* I, L, F) mediates not only internalization but also targeting to the trans-Golgi network (Bos *et. al.* 1993; Wong and Hong 1993). The LL/LI motif is an endocytosis motif present within many transmembrane cell surface proteins and functions as an internalization and lysosomal-targeting signal. The LL motif and a serine phosphorylation within the LRP tail were found to contribute to the receptor-mediated endocytosis (Li *et. al.* 2000 and 2001). The NPxY-like motif (where P is replaced with other residues) in the megalin cytoplasmic tail was recently found to be an apical sorting signal of LRP2, which is expressed on the apical plasma membrane of polarized epithelia cells (Takeda *et. al.* 2003). The sorting signals of membrane proteins can be recognized by clathrin adaptor protein complexes, which in turn associate with clathrin and other accessory molecules to generate clathrin coats and coated transport vesicles during the clathrin-coated and pit-mediated receptor internalization (Hirst and Robinson 1998). A set of cytoplasmic adaptor and scaffold proteins bind to the cytoplasmic tails of LDLR family members, which suggests the participation of quite a few family members in several signal transduction pathways. For example, The VLDLR and ApoER2 are involved in the reelin signaling pathway (Nimpf and Schneider 2002; Howell and Herz 2001; Herz *et. al.* 2000; Herz and Beffert 2000). The VLDLR is also involved in the urokinase-type plasminogen activator (uPA)-PAI-1 complex initiated cell signaling (Strickland *et. al.* 2002). LRP is involved in the signaling pathway initiated with the signaling protein, amyloid precursor protein (APP) (Herz *et. al.* 2000*A* and 2000*B*). The LRP is also involved in the midkine (a growth factor with migration - and survival-

promoting activites) pathway, in the regulation of neuronal calcium influx mediated by α_2M, and perhaps in the Ras-extracellular signal regulated kinase-mitogen activated protein (MAP) kinase pathway (Strickland *et. al.* 2002). The LRP5/6 function in the Wnt/Wingless signaling pathway initiated with Wnt proteins (a family of secreted cysteine-rich proteins) (Howell and Herz 2001).

Potential Receptor-Binding Sites on *Aa*Vg

Alignment among Vg sequences showed a varying number of basic residues around the corresponding region of the "receptor-binding motif" in several insect Vgs (Figure 20). The vicinity of this motif on *Aa*Vg has a roughly equal number of basic and acidic residues (lRgEDDqfiqvtKtqnfDRcDqR). Although the claimed "receptor-binding motif" on *O*aVg could be aligned with receptor-binding sites on ApoB and ApoE (Li *et. al.* 2003), a common pattern of these three motifs is not conserved among Vgs in investigated species (data not shown). Moreover, in *Oa*Vg, this motif is on a β strand in the N-sheet, while the receptor-binding motif on both ApoB and ApoE is on an α helix. Although the receptor-binding motif on ApoB and ApoE share high sequence homology, the protein sequence of ApoE shares no homology with that of Vg. Although the N terminal portion of ApoB-100 has a lipovitellin-like domain, which is homologous to Vg and LV, the receptor-binding site on ApoB is near the C terminus, which has no homology to Vg or LV (Segrest *et. al.* 2001, Figure 2). Searches for an amphipathic receptor-binding helix-homologous sequence on *Aa*Vg were also unsuccessful (data not shown).

The corresponding region of the claimed "receptor-binding site" on *Iu*LV has the positive surface patch comparable to most other areas on the N-sheet, and it is much weaker than the outer surface of the lipid-binding cavity (Figure 21). This cavity is composed primarily of the A-sheet, the C-sheet, and a small portion of the helical domain. Because the interface of a homodimer is formed mainly by the N-sheets and helical domains from two monomers, and most of the A-sheet is not supported by the monomer contact (Anderson *et. al.* 1998, Figure 6 and Table 2), the strongly positive surface on the A-sheet is exposed to solvent even in a dimer.

Earlier binding studies showed that *Aa*Vn-S has lower affinity to *Aa*VgR than to *Aa*Vn-L and that both subunits have lower affinity than the intact *Aa*Vg and *Aa*Vn (Dhadialla *et. al.* 1992; Sappington *et. al.* 1995). The weakened affinities of both *Aa*Vn-S and *Aa*Vn-L indicate that both subunits participate in binding to the *Aa*VgR through either more than one binding site or a large binding face spanning both subunits. The *Aa*Vg-S has a moderately positive surface, which supports the hypothesis that the force mediating *Aa*VgR-*Aa*Vg interaction is negative-positive charge interaction. The surface EP of *Aa*Vg-L was not checked because most of the A-sheet domain was not in the model.

An interesting observation is that there are many tyrosine-enriched sequences in *Aa*Vg. In *Iu*LV, tyrosine residues in the N-sheet mostly face inward, and tyrosine residues in the C-sheet and A-sheet domains mostly face toward the lipid-binding cavity (figure not shown). Thompson *et. al.* (Thompson and Banaszak 2002) reported 12 tyrosine residues on *Iu*LV that interacted with loaded lipids (Y90, Y187, Y194, Y644, Y669, Y800, Y807, Y742, Y888, Y1048, Y1382, Y1518). Dhadialla (Dhadialla *et. al.*

1992) reported that dephosphorylation of *Aa*Vg decreased its binding to *Aa*VgR to less than 20% (Dhadialla *et. al.* 1992, Figure 9). These evidences suggested that most tyrosine residues in *Aa*Vg might also interact with loaded lipids in the lipid-binding cavity and that some tyrosine residues may be involved in maintaining the receptor-binding site in the correct conformation and keeping it approachable.

CONCLUSIONS

Sequence analysis and protein modeling work in this thesis gave several predictions. First, both the CLI and CLII of CRs in the *Aa*VgR have predominantly negatively charged surfaces. CRI-1, CRI-5, and CRII-3 have most strongly negatively charged surfaces, CRI-3, CRII-2, and CRII-4 have strongly negative surfaces, and CRII-7 has a somewhat strongly negative surface. CRII-3, CRII-4, and CRII-7 are three modules with strongly negatively charged surfaces in CLIIs of both *Aa*VgR and *A. gambiae* VgR. The CLI of *Aa*VgR has more strongly negative EP than the CLII, and the negative surface charge of both clusters falls into the same scale. Second, *Aa*VgR has three YWTD β propellers with quite a few histidine residues on each propeller surface. These many surface histidine residues potentially could significantly change the surface charge of propellers at endosomal pH. Third, all modeled EGF-like repeats in *Aa*VgR lack strongly negatively charged surfaces, and thus they do not contribute to the negative-positive charge interaction between the *Aa*VgR and *Aa*Vg. Fourth, the modeled small subunit of *Aa*Vg has a moderately positively charged surface. The modeled portion of the large subunit of *Aa*Vg shows that it has one helical domain, one C-sheet domain, and one A-sheet domain. Fifth, the claimed "receptor-binding motif" on the N-sheet of tilapia Vg

is not basic residue-enriched in some Vg species and does not give a more positively charged surface patch than most other surface regions on the N-sheet of the modeled N-sheet of tilapia Vg; its counterparts in lamprey Vg and the *Aa*Vg small subunit are also the case. This motif in lamprey Vg is on the docking face of two monomers. Sixth, Insect VgR/YPR has either one or two kinds(s) of endocytosis signal(s).

Chapter 4

Summary and Future Research Perspectives

SUMMARY

Probably one of the most amazing biological phenomena known so far is the egg yolk accumulation in the mosquito oocytes during vitellogenesis, when the developing oocyte increases in size over 300-fold within 36 hours, mostly through the VgR-mediated endocytosis of mosquito Vg. Coupled with vitellogenesis, mosquitoes pass to human and animals numerous diseases, including dengue fever and malaria, which are threatening around half the world population. Therefore, investigation on the mosquito Vg-VgR interaction can elucidate the molecular mechanism of Vg internalization that is crucial to the mosquito-borne disease control.

*Aa*VgR is an ideal representative for the two-cluster LDLR subfamily. Unlike *A. gambiae* and *P. Americana* VgR, *Aa*VgR has eight contact modules in the CLII of CRs, and even better than *Drosophila* VgR, *Aa*VgR has an *O*-linked sugar domain. In addition, *A. aegypti* is the first model animal in elucidating the mechanism of receptor-mediated endocytosis. In this work, three mini-receptor genes were constructed and transfected into *Drosophila* cells. The saturation binding assay results showed that both the CLI and CLII of *Aa*VgR have one binding site each with moderate binding affinity and that they are responsible for the high affinity of the full-length *Aa*VgR to *Aa*Vg. The AaVgR has at least two binding sites, and both clusters supply strengthened binding of *Aa*Vg, probably in a synergistic way.

The CLI of *Aa*VgR is predicted to have, overall, a stronger negative surface than the CLII, which is a good fit to the results from the saturation binding assays. Both clusters have overall strongly negative surfaces, which also supports the conclusion from experimental results that two clusters supply strengthened binding of *Aa*Vg. Negatively

charged surfaces of both clusters highly support the hypothesis that the *Aa*VgR-*Aa*Vg interaction is mediated mainly by the negative-positive electrostatic attraction. In comparision with the CRs in both clusters, no EGF-like modules in *Aa*VgR have strongly negative surfaces. All three YWTD β propellers in *Aa*VgR have numerous histidine residues on their surfaces, implicating their critical role in regulating *Aa*Vg release and *Aa*VgR recycling through acid-induced change of surface charges in mosquitoes. The reported "receptor-binding motif" on tilapia Vg is not a real receptor-binding site. The modeled *Aa*Vg small subunit has a moderately positively charged surface. The *Aa*Vg large subunit model shows that it has a helical domain, a C-sheet domain, and an A-sheet domain. Evidence of the positively charged surface of AaVg supports the hypothesis of electrostatic complementarities from the ligand side. Finally, hypotheses were cast as a guidance to further studies.

FUTURE RESEARCH PERSPECTIVES

Investigation of individual CRs of AaVgR on binding AaVg: We now know that both the CLI and CLII in the *Aa*VgR bind *Aa*Vg. The next step would be to determine which module or combination of modules in each cluster is/are critical for Vg binding. Figure 26 shows the proposed future construction of mosquito mini-receptors. This design is based on two hypotheses. The first hypothesis is that the interaction between *Aa*Vg and *Aa*VgR is predominantly positive-negative electrostatic complementation. Based on this assumption, CRs with most negatively charged surfaces compose the core region of newly designed mini-receptors. The second hypothesis supposes that there are no intro-domain interactions between modules within the cluster and that each CR binds

Fig. 26. Proposed future design on mosquito minireceptors. Minireceptor constructs are shown as molecular surface EPs of the front and back faces (upper and lower halves of each pair) of component CRs, and were drawn in the same way as in figure 13. The left and right blocks show two series of minireceptors ($_{m}VgR_{I}$ and $_{m}VgR_{II}$) assembled with module(s) from either CLI or CLII of *Aa*VgR respectively. EGF-homology domains (including YWTD folds and its flanking EGF-like repeats) are excluded from this design.

CRI-1
CRI-2
CRI-3
CRI-4
CRI-5
mVgRI1-5
mVgRI135
mVgRII1-3
mVgRI1
mVgRI3
mVgRI5
CRII-1
CRII-2
CRII-3
CRII-4
CRII-5
CRII-6
CRII-7
CRII-8
mVgRII1-8
mVgRII2-4
mVgRII2-3
mVgRII3
mVgRII7
mVgRII3-4

ligand independently. Based on this assumption, mini-receptor $_{m}VgR_{I}135$ would be designed, with CRI-2 and CRI-4 kicked out except for their linker regions (Figure 26). This assumption is based on several observations. First, the crystal structure of LDLR shows that the ligand-binding domain is arranged in an extended way in the crystal form, with CR2 to CR7 separated by linkers, and modules do not interact with each other directly (*Rudenko et. al.* 2002). Second, Beglova *et. al.* observed in an NMR study that the linker connecting CR5 and CR6 is substantially flexible, and the covalent connection between CR5 and CR6 did not essentially change the intrinsic dynamic behavior of each repeat (Beglova *et. al.* 2001). Third, an NMR study of the CR1-CR2 pair of LDLR also demonstrated flexibility in this concatermer (Kurniawan *et. al.* 2000).

***Seeking more structural evidences in support of hypotheses*:** The direct evidence of intensified and/or expended positively charged surface areas on YWTD β propellers would be sought by calculating the surface EP at pH5. The docking interface between each YWTD β propeller and its contacting CR(s) would be made, with reference to Rudenko's X-ray result (Rudenko et. al. 2002). Modeling of the whole-length A-sheet domain of the *Aa*Vg large subunit needs to be finished, after which checking the surface EP of the *Aa*Vg-L will become possible.

***Finding residues on CRs critical for Vg binding*:** After we know which CR or combination of CRs is/are crucial for the Vg binding, a further study on residues critical to the Vg binding would be done by both biochemical and genetics methods. Mutations would be made on candidate residues and *in vitro* and *in vivo* binding assays would be performed. As a more powerful technology, surface plasmon resonance (SPR), would be used for several advantages over traditional binding assays. SPR can measure not only

affinity constants, but also reaction kinetic rate constants; SPR can monitor biomolecular interactions in real time without labeling requirements; SPR can also measure the amount of binding complex at equilibrium even in the presence of unbound ligand (Rich and Myszka, 2001). Finally, results from binding assays and theoretical modeling should be supported by 3-D structures of critical modules solved by solution nuclear magnetic resonance (NMR) or X-ray crystallography.

Bibliography

REFERENCES

Anderson, T. A., Levitt, D. G., and Banaszak, L. J. (1998). The structural basis of lipid interactions in lipovitellin, a soluble lipoprotein. *Structure.* **6**, 895-909.

Barr, P. J. (1991). Mammalian subtilisins: the long-sought dibasic processing endoproteases. *Cell* **66**, 1-3.

Beglova, N., North, C. L., and Blacklow, S. C. (2001). Backbone dynamics of a module pair from the ligand-binding domain of the LDL receptor. *Biochemistry* **40**, 2808-2815.

Boren, J., Lee, I., Zhu, W., Arnold, K., Taylor, S., and Innerarity, T. L. (1998). Identification of the low density lipoprotein receptor-binding site in apolipoprotein B100 and the modulation of its binding activity by the carboxyl terminus in familial defective apo-B100. *J.Clin.Invest* **101**, 1084-1093.

Bos, K., Wraight, C., and Stanley, K. K. (1993). TGN38 is maintained in the trans-Golgi network by a tyrosine-containing motif in the cytoplasmic domain. *EMBO J.* **12**, 2219-2228.

Brakch, N., Rholam, M., Boussetta, H., and Cohen, P. (1993). Role of beta-turn in proteolytic processing of peptide hormone precursors at dibasic sites. *Biochemistry* **32**, 4925-4930.

Brandes, C., Novak, S., Stockinger, W., Herz, J., Schneider, W. J., and Nimpf, J. (1997). Avian and murine LR8B and human apolipoprotein E receptor 2: differentially spliced products from corresponding genes. *Genomics* **42**, 185-191.

Brandes, C., Kahr, L., Stockinger, W., Hiesberger, T., Schneider, W. J., and Nimpf, J. (2001). Alternative splicing in the ligand binding domain of mouse ApoE receptor-2 produces receptor variants binding reelin but not alpha 2-macroglobulin. *J.Biol.Chem.* **276**, 22160-22169.

Byrne, B. M., Gruber, M., and Ab, G. (1989). The evolution of egg yolk proteins. *Prog.Biophys.Mol.Biol.* **53**, 33-69.

Chen, J. S. and Raikhel, A. S. (1996). Subunit cleavage of mosquito pro-vitellogenin by a subtilisin-like convertase. *Proc.Natl.Acad.Sci.U.S.A* **93**, 6186-6190.

Chen, J. S., Sappington, T. W., and Raikhel, A. S. (1997). Extensive sequence conservation among insect, nematode, and vertebrate vitellogenins reveals ancient common ancestry. *J.Mol.Evol.* **44**, 440-451.

Cheon, H. M., Seo, S. J., Sun, J., Sappington, T. W., and Raikhel, A. S. (2001). Molecular characterization of the VLDL receptor homolog mediating binding of lipophorin in oocyte of the mosquito Aedes aegypti. *Insect Biochem.Mol.Biol.* **31**, 753-760.

Cho, W. L., Deitsch, K. W., and Raikhel, A. S. (1991). An extraovarian protein accumulated in mosquito oocytes is a carboxypeptidase activated in embryos. *Proc.Natl.Acad.Sci.U.S.A* **88**, 10821-10824.

Cho, W. L., Tsao, S. M., Hays, A. R., Walter, R., Chen, J. S., Snigirevskaya, E. S., and Raikhel, A. S. (1999). Mosquito cathepsin B-like protease involved in embryonic degradation of vitellin is produced as a latent extraovarian precursor. *J.Biol.Chem.* **274**, 13311-13321.

Christie, R. H., Chung, H., Rebeck, G. W., Strickland, D., and Hyman, B. T. (1996). Expression of the very low-density lipoprotein receptor (VLDL-r), an apolipoprotein-E receptor, in the central nervous system and in Alzheimer's disease. *J.Neuropathol.Exp.Neurol.* **55**, 491-498.

Davis, C. G., Goldstein, J. L., Sudhof, T. C., Anderson, R. G., Russell, D. W., and Brown, M. S. (1987). Acid-dependent ligand dissociation and recycling of LDL receptor mediated by growth factor homology region. *Nature* **326**, 760-765.

Dhadialla, T. S. and Raikhel, A. S. (1990). Biosynthesis of mosquito vitellogenin. *J.Biol.Chem.* **265**, 9924-9933.

Dhadialla, T. S. and Raikhel, A. S. (1991). Binding of vitellogenin to membranes isolated from mosquito ovaries. *Arch.Insect Biochem.Physiol* **18**, 55-70.

Dhadialla, T. S., Hays, A. R., and Raikhel, A. S. (1992). Characterization of the Solubilized Mosquito Vitellogenin Receptor. *Insect Biochemistry and Molecular Biology* **22**, 803-816.

Dirlam, K. A., Gretch, D. G., LaCount, D. J., Sturley, S. L., and Attie, A. D. (1996). Expression and characterization of a truncated, soluble, low-density lipoprotein receptor. *Protein Expr.Purif.* **8**, 489-500.

Esser, V., Limbird, L. E., Brown, M. S., Goldstein, J. L., and Russell, D. W. (1988). Mutational analysis of the ligand binding domain of the low density lipoprotein receptor. *J.Biol.Chem.* **263**, 13282-13290.

Ford, P. S. and Van Heusden, M. C. (1994). Triglyceride-rich lipophorin in Aedes *aegypti* (Diptera: Culicidae). *J.Med.Entomol.* **31**, 435-441.

Grant, B. and Hirsh, D. (1999). Receptor-mediated endocytosis in the Caenorhabditis elegans oocyte. *Mol.Biol.Cell* **10**, 4311-4326.

Hagedorn, H. H., Turner, S., Hagedorn, E. A., Pontecorvo, D., Greenbaum, P., Pfeiffer, D., Wheelock, G., and Flanagan, T. R. (1977). Postemergence growth of the ovarian follicles of Aedes aegypti. *J.Insect Physiol* **23**, 203-206.

Hays, A. R. and Raikhel, A. S. (1990). A Novel Protein Produced by the Vitellogenic Fat-Body and Accumulated in Mosquito Oocytes. *Rouxs Archives of Developmental Biology* **199**, 114-121.

Herz, J., Gotthardt, M., and Willnow, T. E. (2000). Cellular signalling by lipoprotein receptors. *Curr.Opin.Lipidol.* **11**, 161-166.

Herz, J. and Beffert, U. (2000). Apolipoprotein E receptors: linking brain development and Alzheimer's disease. *Nat.Rev.Neurosci.* **1**, 51-58.

Hirst, J. and Robinson, M. S. (1998). Clathrin and adaptors. *Biochim.Biophys.Acta* **1404**, 173-193.

Howell, B. W. and Herz, J. (2001). The LDL receptor gene family: signaling functions during development. *Curr.Opin.Neurobiol.* **11**, 74-81.

Hussain, M. M., Strickland, D. K., and Bakillah, A. (1999). The mammalian low-density lipoprotein receptor family. *Annu.Rev.Nutr.* **19**, 141-172.

Jeon, H. and Blacklow, S. C. (2003). An Intramolecular Spin of the LDL Receptor beta Propeller. *Structure.(Camb.)* **11**, 133-136.

Koller, C. N., Dhadialla, T. S., and Raikhel, A. S. (1989). Selective Endocytosis of Vitellogenin by Oocytes of the Mosquito, Aedes-Aegypti - An Invitro Study. *Insect Biochemistry* **19**, 693-&.

Kounnas, M. Z., Morris, R. E., Thompson, M. R., FitzGerald, D. J., Strickland, D. K., and Saelinger, C. B. (1992). The alpha 2-macroglobulin receptor/low density lipoprotein receptor-related protein binds and internalizes Pseudomonas exotoxin A. *J.Biol.Chem.* **267**, 12420-12423.

Kurniawan, N. D., Atkins, A. R., Bieri, S., Brown, C. J., Brereton, I. M., Kroon, P. A., and Smith, R. (2000). NMR structure of a concatemer of the first and second ligand-binding modules of the human low-density lipoprotein receptor. *Protein Sci.* **9**, 1282-1293.

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. The first three ligand-binding repeats of VtgR interact with the amino-terminal region of Vtg. *J.Biol.Chem.* **278**, 2799-2806.

Mikhailenko, I., Battey, F. D., Migliorini, M., Ruiz, J. F., Argraves, K., Moayeri, M., and Strickland, D. K. (2001). Recognition of alpha 2-macroglobulin by the low density lipoprotein receptor-related protein requires the cooperation of two ligand binding cluster regions. *J.Biol.Chem.* **276**, 39484-39491.

Moestrup, S. K., Cui, S., Vorum, H., Bregengard, C., Bjorn, S. E., Norris, K., Gliemann, J., and Christensen, E. I. (1995). Evidence that epithelial glycoprotein 330/megalin mediates uptake of polybasic drugs. *J.Clin.Invest* **96**, 1404-1413.

Neels, J. G., van Den Berg, B. M., Lookene, A., Olivecrona, G., Pannekoek, H., and van Zonneveld, A. J. (1999). The second and fourth cluster of class A cysteine-rich repeats of the low density lipoprotein receptor-related protein share ligand-binding properties. *J.Biol.Chem.* **274**, 31305-31311.

Nielsen, K. L., Holtet, T. L., Etzerodt, M., Moestrup, S. K., Gliemann, J., Sottrup-Jensen, L., and Thogersen, H. C. (1996). Identification of residues in alpha-macroglobulins important for binding to the alpha2-macroglobulin receptor/Low density lipoprotein receptor-related protein. *J.Biol.Chem.* **271**, 12909-12912.

Nielsen, M. S., Brejning, J., Garcia, R., Zhang, H., Hayden, M. R., Vilaro, S., and Gliemann, J. (1997). Segments in the C-terminal folding domain of lipoprotein

lipase important for binding to the low density lipoprotein receptor-related protein and to heparan sulfate proteoglycans. *J.Biol.Chem.* **272**, 5821-5827.

Nimpf, J. and Schneider, W. J. (2000). From cholesterol transport to signal transduction: low density lipoprotein receptor, very low density lipoprotein receptor, and apolipoprotein E receptor-2. *Biochim.Biophys.Acta* **1529**, 287-298.

Norberg, B. and Haux, C. (1985). Induction, isolation and a characterization of the lipid content of plasma vitellogenin from two Salmo species: rainbow trout (Salmo gairdneri) and sea trout (Salmo trutta). *Comp Biochem.Physiol B* **81**, 869-876.

North, C. L. and Blacklow, S. C. (2000). Solution structure of the sixth LDL-A module of the LDL receptor. *Biochemistry* **39**, 2564-2571.

Ohlendorf, D. H., Barbarash, G. R., Trout, A., Kent, C., and Banaszak, L. J. (1977). Lipid and polypeptide components of the crystalline yolk system from Xenopus laevis. *J.Biol.Chem.* **252**, 7922-8001.

Ohno, H., Stewart, J., Fournier, M. C., Bosshart, H., Rhee, I., Miyatake, S., Saito, T., Gallusser, A., Kirchhausen, T., and Bonifacino, J. S. (1995). Interaction of tyrosine-based sorting signals with clathrin-associated proteins. *Science* **269**, 1872-1875.

Paccaud, J. P., Reith, W., Johansson, B., Magnusson, K. E., Mach, B., and Carpentier, J. L. (1993). Clathrin-coated pit-mediated receptor internalization. Role of internalization signals and receptor mobility. *J.Biol.Chem.* **268**, 23191-23196.

Pata, I., and Truve, E. (1998). Gel electrophoresis, transfer and hybridization of nucleic acids. In: V. Berzins, Editor, *Basic Cloning Procedures (Springer lab manual)*, Springer Verlag, Heidelberg, pp. 104-134.

Raag, R., Appelt, K., Xuong, N. H., and Banaszak, L. (1988). Structure of the lamprey yolk lipid-protein complex lipovitellin-phosvitin at 2.8 A resolution. *J.Mol.Biol.* **200**, 553-569.

Raikhel, A. S. (1987). The cell biology of mosquito vitellogenesis. *Mem.Inst.Oswaldo Cruz* **82 Suppl 3**, 93-101.

Raikhel, A. S. and Dhadialla, T. S. (1992). Accumulation of yolk proteins in insect oocytes. *Annu.Rev.Entomol.* **37**, 217-251.

Raikhel, A. S., Deitsch, K. W., and Sappington, T. W. (1997). Culture and analysis of the insect fat body. In: J. M. Crampton, C. B. Beard, and C. Louis, Editors, *Molecular Biology of Insect Disease Vectors: A Methods Manual, Chapman & Hall*, **chapter 42**, pp. 507-521

Rich, R. L. and Myszka, D. G. (2001). BIACORE J: a new platform for routine biomolecular interaction analysis. *J.Mol.Recognit.* **14**, 223-228.

Roehrkasten, A. and Ferenz, H. J. (1992). Role of the lysine and arginine residues of vitellogenin in high affinity binding to vitellogenin receptors in locust oocyte membranes. *Biochim.Biophys.Acta* **1133**, 160-166.

Rouille, Y., Duguay, S. J., Lund, K., Furuta, M., Gong, Q., Lipkind, G., Oliva, A. A., Jr., Chan, S. J., and Steiner, D. F. (1995). Proteolytic processing mechanisms in the biosynthesis of neuroendocrine peptides: the subtilisin-like proprotein convertases. *Front Neuroendocrinol.* **16**, 322-361.

Rudenko, G., Henry, L., Henderson, K., Ichtchenko, K., Brown, M. S., Goldstein, J. L., and Deisenhofer, J. (2002). Structure of the LDL receptor extracellular domain at endosomal pH. *Science* **298**, 2353-2358.

Sambrook, J., Russell, D.W. (2001). Molecular Cloning: a laboratory manual (third edition). *Cold Spring Harbor Lab. Press*, pp. 5.47-5.48

Sappington, T. W., Hays, A. R., and Raikhel, A. S. (1995). Mosquito vitellogenin receptor: purification, developmental and biochemical characterization. *Insect Biochem.Mol.Biol.* **25**, 807-817.

Sappington, T. W., Kokoza, V. A., Cho, W. L., and Raikhel, A. S. (1996). Molecular characterization of the mosquito vitellogenin receptor reveals unexpected high homology to the Drosophila yolk protein receptor. *Proc.Natl.Acad.Sci.U.S.A* **93**, 8934-8939.

Sappington, T. W. and Raikhel, A. S. (1998). Molecular characteristics of insect vitellogenins and vitellogenin receptors. *Insect Biochem.Mol.Biol.* **28**, 277-300.

Schneider, W. J. (1996). Vitellogenin receptors: oocyte-specific members of the low-density lipoprotein receptor supergene family. *Int.Rev.Cytol.* **166**, 103-137.

Schonbaum, C. P., Perrino, J. J., and Mahowald, A. P. (2000). Regulation of the vitellogenin receptor during Drosophila melanogaster oogenesis. *Mol.Biol.Cell* **11**, 511-521.

Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucl. Aci. Res.* **31**, 3381-3385.

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**, 1346-1367.

Shen, X., Steyrer, E., Retzek, H., Sanders, E. J., and Schneider, W. J. (1993). Chicken oocyte growth: receptor-mediated yolk deposition. *Cell Tissue Res.* **272**, 459-471.

Snigirevskaya, E. S., Sappington, T. W., and Raikhel, A. S. (1997). Internalization and recycling of vitellogenin receptor in the mosquito oocyte. *Cell Tissue Res.* **290**, 175-183.

Soulages, J. L. and Wells, M. A. (1994). Effect of diacylglycerol content on some physicochemical properties of the insect lipoprotein, lipophorin. Correlation with the binding of apolipophorin-III. *Biochemistry* **33**, 2356-2362.

Stefansson, S., Muhammad, S., Cheng, X. F., Battey, F. D., Strickland, D. K., and Lawrence, D. A. (1998). Plasminogen activator inhibitor-1 contains a cryptic high affinity binding site for the low density lipoprotein receptor-related protein. *J.Biol.Chem.* **273**, 6358-6366.

Strickland, D. K., Gonias, S. L., and Argraves, W. S. (2002). Diverse roles for the LDL receptor family. *Trends Endocrinol.Metab* **13**, 66-74.

Sun, J., Hiraoka, T., Dittmer, N. T., Cho, K. H., and Raikhel, A. S. (2000). Lipophorin as a yolk protein precursor in the mosquito, Aedes aegypti. *Insect Biochem.Mol.Biol.* **30**, 1161-1171.

Takeda, T., Yamazaki, H., and Farquhar, M. G. (2003). Identification of an apical sorting determinant in the cytoplasmic tail of megalin. *Am.J.Physiol Cell Physiol* **284**, C1105-C1113.

Thompson, J. R. and Banaszak, L. J. (2002). Lipid-protein interactions in lipovitellin. *Biochemistry* **41**, 9398-9409.

Tufail, M., Takeda, M. (2002). Vitellogenin receptor of the American cockroach: an oocyte-specific member of the low-density lipoprotein receptor gene superfamily. *NCBI Protein Sequences database*, **accession ID**, BAC02725

Van Driel, I. R., Davis, C. G., Goldstein, J. L., and Brown, M. S. (1987*A*). Self-association of the low density lipoprotein receptor mediated by the cytoplasmic domain. *J.Biol.Chem.* **262**, 16127-16134.

Van Driel, I. R., Goldstein, J. L., Sudhof, T. C., and Brown, M. S. (1987*B*). First cysteine-rich repeat in ligand-binding domain of low density lipoprotein receptor binds Ca2+ and monoclonal antibodies, but not lipoproteins. *J.Biol.Chem.* **262**, 17443-17449.

Wallace, R. A. (1985). Vitellogenesis and oocyte growth in nonmammalian vertebrates. *Dev.Biol.(N.Y.1985.)* **1**, 127-177.

Weisgraber, K. H. (1994). Apolipoprotein E: structure-function relationships. *Adv.Protein Chem.* **45**, 249-302.

Willnow, T. E., Orth, K., and Herz, J. (1994). Molecular dissection of ligand binding sites on the low density lipoprotein receptor-related protein. *J.Biol.Chem.* **269**, 15827-15832.

Willnow, T. E., Nykjaer, A., and Herz, J. (1999). Lipoprotein receptors: new roles for ancient proteins. *Nat.Cell Biol.* **1**, E157-E162.

Wilson, C., Wardell, M. R., Weisgraber, K. H., Mahley, R. W., and Agard, D. A. (1991). Three-dimensional structure of the LDL receptor-binding domain of human apolipoprotein E. *Science* **252**, 1817-1822.

Wong, S. H. and Hong, W. (1993). The SXYQRL sequence in the cytoplasmic domain of TGN38 plays a major role in trans-Golgi network localization. *J.Biol.Chem.* **268**, 22853-22862.

Zaiou, M., Arnold, K. S., Newhouse, Y. M., Innerarity, T. L., Weisgraber, K. H., Segall, M. L., Phillips, M. C., and Lund-Katz, S. (2000). Apolipoprotein E;-low density

lipoprotein receptor interaction. Influences of basic residue and amphipathic alpha-helix organization in the ligand. *J.Lipid Res.* **41**, 1087-1095.