Implication of the Whitefly, *Bemisia tabaci*, Collagen Protein in Begomoviruses Acquisition and Transmission

Vipin Singh Rana,^{1,2} Sonam Popli,^{1,3} Gunjan Kumar Saurav,¹ Harpreet Singh Raina,^{1,4} Rohit Jamwal,¹ Rahul Chaubey,⁵ V. V. Ramamurthy,⁵ K. Natarajan,⁶ and Raman Rajagopal^{1,†}

¹ Gut Biology Laboratory, Department of Zoology, University of Delhi, Delhi, India 110007

² Department of Veterinary Medicine, University of Maryland, College Park, MD, U.S.A.

³ Department of Medical Microbiology and Immunology, College of Medicine, University of Toledo, Toledo, OH, U.S.A.

⁴ Department of Zoology, Shri Guru Tegh Bahadur Khalsa College, University of Delhi, Delhi, India 110007

⁵ Division of Entomology, Indian Agricultural Research Institute, New Delhi, India 110012

⁶ Dr. B. R. Ambedkar Centre for Biomedical Research, University of Delhi, Delhi, India 110007

Accepted for publication 21 April 2019.

ABSTRACT

Begomoviruses are the largest group of plant viruses transmitted exclusively by the whitefly, *Bemisia tabaci* (Gennadius), in a persistent, circulative, and nonpropagative manner. Begomoviruses in association with *B. tabaci* cause enormous loss to world agricultural crops. Transmission, retention, and circulation of begomovirus in *B. tabaci* are facilitated by its interaction with several proteins of the insect and its endosymbionts. However, very few such proteins have been identified from *B. tabaci* that are involved in this specific interaction. Here, we have performed yeast two-hybrid assay between *B. tabaci* complementary DNA expression library and the coat protein (CP) of tomato leaf curl New Delhi virus (ToLCNDV) and cotton leaf curl

Rajasthan virus (CLCuV). Collagen was the common protein found to be interacting with both of the viruses. The collagen protein was found to be localized in gut layers of *B. tabaci*. Additionally, pull-down and dot-blot assays confirmed the association of endogenous collagen with ToLCNDV CP. Immunolocalization analysis also showed colocalization of ToLCNDV particles and collagen within insect gut. Finally, *B. tabaci* fed on anticollagen antibody and exhibited ~46% reduction in ToLCNDV transmission, suggesting a supportive role for collagen in virus transmission.

Keyword: possible candidate receptor

Bemisia tabaci (Gennadius; Hemiptera: Aleyrodidae) is known as the most economically important insect pest across the globe, causing huge losses to world agriculture by feeding and transmitting begomoviruses in food and cash crops (Brown et al. 1995; Navas-Castillo et al. 2011). It is a cryptic species complex consisting of >35 morphologically indistinguishable genetic groups, which differ in host range, insecticide resistance, endosymbionts that they harbor, and their efficiency in virus transmission (Firdaus et al. 2013; Gottlieb et al. 2010; Liu et al. 2012). The *Begomovirus* genus contains >320 definitive members that are either monopartite (single-component) or bipartite (A and B component) viruses infecting a wide range of dicotyledonous plants (Rojas et al. 2005; Zerbini et al. 2017). Tomato and cotton leaf curl diseases are prevalent plant viral diseases in India attributed to tomato leaf curl virus and cotton leaf curl virus, respectively. Tomato leaf curl New Delhi virus (ToLCNDV), a bipartite begomovirus, is the most predominant virus identified from northern India, infecting elite cultivars (Padidam et al. 1995; Pandey et al. 2009).

Begomoviruses are transmitted by *B. tabaci* in a persistent, circulative, and nonpropagative manner (Hogenhout et al. 2008;

Rubinstein and Czosnek 1997). During feeding, *B. tabaci* ingest geminivirus particles, which then enter the gut, are transported to the salivary gland via hemolymph, and are inoculated back into the plant during subsequent feeding (Cicero et al. 1995; Czosnek and Ghanim 2002; Ghanim et al. 2001). During this process, the virion particles cross several physical barriers, which are located on gut wall as well as on basement membrane of salivary glands. To cross these barriers, virus uses several insect proteins that are present on membrane as a receptor or coreceptor. In addition to these barriers, hemolymph also poses a hostile environment for virion particles, because it consists of plasma and several proteases. Endosymbionts of *B. tabaci* as well as other insect vectors have also been shown to play a major role in virus transmission by protecting virions from rapid degradation. For example, a GroEL protein of *Buchnera aphidicola* and endosymbiont of *Myzus persicae* interacts with luteovirus coat protein (CP) and protects virus particles from rapid proteolysis in the gut and hemolymph (Van den Heuvel et al. 1997), although another study questioned the role of GroEL–virion interaction in the virus transmission by aphid (Bouvaine et al. 2011). Similarly, in *B. tabaci*, GroEL protein secreted by endosymbionts *Hamiltonella* and *Arsenophonus* interacts with CPs of Tomato yellow leaf curl virus (TYLCV) and cotton leaf curl Rajasthan virus (CLCuV), respectively, and enhances their transmission efficiency (Gottlieb et al. 2010; Rana et al. 2012). CP of begomoviruses is the only known viral protein required for insect-mediated transmission (Briddon et al. 1990; Noris et al. 1998). Previous studies have also shown that changes in amino acid sequence of CP lead to altered vector specificity and transmission ability (Noris et al. 1998).

However, in contrast to *B. tabaci*, nonvector glasshouse whitefly *Trialeurodes vaporariorum* does not show the same mechanism of virus movement inside the body of insect. In this nonvector species, the ingested geminivirus particles were incapable of crossing the gut epithelium, suggesting the presence of some specific receptor

†Corresponding author: R. Rajagopal; zoorajagopal@gmail.com

Funding: This work was funded by the Indian Council of Agriculture Research–National Agricultural Science Fund (ICAR-NASF) and DST-PURSE. V. S. Rana was provided a research fellowship by UGC-DSKPDF.

V. S. Rana and S. Popli contributed equally to this work.

*The e-Xtra logo stands for “electronic extra” and indicates that two supplementary figures are published online.

The author(s) declare no conflict of interest.

(for begomoviruses) on the gut membrane in *B. tabaci* that allows viral particles to cross the epithelium using transcytosis mechanism. However, these receptors are absent in other nonvector whitefly species (Ohnishi et al. 2009). A study has shown that TYLCV-viruliferous *B. tabaci* was attracted more to infected plant rather than healthy one, probed it more quickly with a greater number of feeding bouts, and also, spent more time salivating into the sieve tube element (Moreno-Delafuente et al. 2013). This suggests that the viruses may alter the insect feeding behavior, which is favorable for their efficient transmission. The gene expression profile of *B. tabaci* on tomato yellow leaf curl China virus infection was analyzed, and results indicated that 1,606 genes associated with 157 biochemical pathways were expressed differentially in viruliferous and aviruliferous whiteflies (Luan et al. 2011). The intimate relationship between the virus and *B. tabaci* suggests an active molecular and cellular interaction between them.

In addition to the endosymbiont's GroEL protein, other host proteins, like 16-kDa small heat shock protein (named BtHSP16) (Ohnesorge and Bejarano 2009), HSP70 (Götz et al. 2012), midgut protein (MGP) (Rana et al. 2016), Cyclophilin B (Kanakala and Ghanim 2016), peptidoglycan (Wang et al. 2016), and thioredoxin-like protein (Saurav et al. 2019), have been identified to interact with begomoviruses in *B. tabaci*. Considering the complex and evolved nature of virus–vector interactions, there is a possibility of additional receptor proteins being synchronously involved in virus transmission pathway.

This study was conducted to identify and characterize the possible candidate receptors of *B. tabaci* against the CP from two Begomoviruses: ToLCNDV and CLCuV. A yeast two-hybrid complementary DNA (cDNA) expression library was prepared from *B. tabaci* and screened independently against the CPs of ToLCNDV and CLCuV. Using same strategy, our group had earlier identified and characterized a gut-specific receptor known as MGP, which exhibited a significant impact on viral transmission (Rana et al. 2016). Here, we demonstrate another protein, collagen, from *B. tabaci* that was found to interact with the CPs of both ToLCNDV and CLCuV in yeast two-hybrid assays. Additionally, we examined the localization of collagen protein in *B. tabaci* guts through western blots and confocal microscopy experiments. Interestingly, blocking of collagen protein by feeding *B. tabaci* on anticollagen antibody exhibited significantly reduced ToLCNDV transmission, suggesting a role for collagen protein in virus transmission by *B. tabaci*. In the future, this protein can also be used as a potential target for integrated pest management strategies.

MATERIALS AND METHODS

Yeast cDNA library preparation. To perform cDNA library screening, total RNA was isolated from *B. tabaci* and used for cDNA preparation. Prepared cDNA was ligated into pGADT7-Rec vector according to manufacturer's instructions (Clontech Laboratories, Inc.). The CP genes of CLCuV and ToLCNDV were cloned into pGBKT7 vector and used as bait individually (Rana et al. 2012, 2016). *Sma*I-linearized pGADT7-Rec vector and cDNA were cotransformed with bait vectors (pGBKT7-ToLCNDV CP or pGBKT7-CLCuV CP) individually into AH109 yeast cells.

Transformed AH109 cells were plated on selective media lacking leucine and tryptophan. Plates were incubated at 30°C for 3 to 4 days. Healthy colonies were restreaked on high-stringency synthetic 3 dropout media (3DO; Leu[−] Trp[−] His[−]) medium as well as 4DO (Leu[−] Trp[−] His[−] Ade[−]) medium. Additional β-galactosidase assays were also performed to revalidate the positive clones. Plasmid DNA from positive clones was rescued into bacterial strain DH5α and sequenced using the T7 primer at a sequencing facility (Macrogen). Sequences were analyzed for similarity to known sequence database (NCBI) using the BLAST algorithm. Isolated plasmids from positive clones were again retransformed into yeast AH109 and assayed for HIS3 and β-galactosidase activity.

5' rapid amplification of cDNA end (RACE) of cDNA from positive clone. Sequence analysis of the clones revealed the presence of 927-bp-long open reading frame, where the 5' end of the sequence was observed to be missing. To amplify the complete *B. tabaci* collagen gene, 5' RACE reaction was performed according to the manufacturer's instructions (Invitrogen). Primers used for 5' RACE are enlisted in Table 1.

Cloning and sequence analysis of collagen. Amplified 5' RACE products were cloned into pGEMT Easy T/A cloning vector and transformed into *Escherichia coli* strain DH5α. Plasmid DNA was isolated from five individual transformants and sequenced. Sequences were assembled using Bio-Edit software, whereas alignment was performed using Clustal-W with default settings. Web-based tools, like Conserved Domain Database (CDD) of NCBI, InterProScan database, and SignalP (ExpASY), were used for the analysis of collagen protein. The three-dimensional (3D) structure of collagen protein was analyzed using RaptorX. Furthermore, a phylogenetic tree was generated using the neighbor-joining algorithm and keeping bootstrap value 1,000 in MEGA7 to evaluate phylogenetic association of collagen protein among various organisms (Tamura et al. 2013).

Expression and purification of collagen protein. To amplify collagen gene, PCR was performed using gene-specific primers enlisted in Table 1. The cDNA prepared from total RNA of *B. tabaci* was used as a template for PCR. Amplified product was cloned into pGEMT-Easy Vector and sequenced. Collagen gene was further subcloned into pET32a expression vector (Novagen) between *Eco*R1 and *Sal*I restriction sites. Cloning was confirmed using restriction enzyme digestion with *Eco*R1 and *Sal*I. Sequence analysis was performed to confirm proper gene frame in pET32a expression vector and also, absence of any mutation in collagen gene.

Induction of collagen protein was achieved by transforming the expression construct into *E. coli* strain BL-21. Transformed culture was induced with 0.01 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 0.6 optical density at 600 nm (OD₆₀₀) and incubated at 16°C for 16 h. Because we were unable to solubilize the protein under native purification condition, the induced pellet was resuspended in buffer A (20 mM sodium phosphate [pH 7.5], 150 mM NaCl, 1 mM β-mercaptoethanol, 0.03% sodium dodecyl sulfate (SDS), and 1× protease inhibitor; GE Healthcare) and lysed by sonication. Lysate was centrifuged at high speed at 18,514 × g for 20 min at 4°C. Soluble fraction was incubated with Ni-NTA beads

TABLE 1. Primer sequences used for 5' rapid amplification of cDNA end of collagen gene and its cloning into expression vectors

Name	Sequence (5' → 3')	Melting temperature, °C
Collagen fwd-1	AGGACAAGGCCAAAAAGGAGAAC	55
Collagen fwd-2	TGTAAAAGGAGATATGGGTTTCCC	
Collagen rev-1	GTAGAAACAAGCTGATCCAACAG	
Collagen rev-2	ATGATTGCATCGTTGTAACCCAGC	55
Collagen rev-3	GATCCACATCCATTATCGGTAACC	
Collagen fwd primer (full gene)	GAATTCCATATGAGTTTTTCTTTCAAAATAAAGAGACC	
Collagen rev primer (full gene)	GACGTCGACTTAGTAGAAACAAGCTGATCCAAC	

(preequilibrated with buffer A) for 2 h at 4°C on Nutator. After binding, the beads were washed with washing buffer (20 mM sodium phosphate [pH 7.5], 150 mM NaCl, 1× protease inhibitor, 40 mM imidazole, and 1% Triton X-100 [vol/vol]) twice. Bound protein fractions were eluted with elution buffer (20 mM sodium phosphate [pH 7.5], 150 mM NaCl, 250 mM imidazole, and 1× protease inhibitor) containing 250 mM imidazole. Eluted fractions were resolved on SDS polyacrylamide gel electrophoresis (PAGE) and dialyzed against buffer B (50 mM Tris Cl [pH 7.5], 150 mM NaCl, 1× protease inhibitor, and 20% glycerol). The dialyzed protein samples were stored at −70°C until further use. glutathione S-transferase (GST)-tagged ToLCNDV CP was isolated and purified, and protein samples were stored at −70°C until further use as mentioned in Rana et al. (2016).

Anticollagen antibody generation, purification, and labeling. Anticollagen antibodies were produced using 6- to 8-week-old BALB/c mice. Purified collagen protein (100 µg) was injected subcutaneously into mice. Two mice were used for antibody production. Purified recombinant proteins were emulsified with Freund's complete adjuvant (Sigma-Aldrich) for the first immunization and Freund's incomplete adjuvant (Sigma-Aldrich) for the subsequent booster dosages. Antisera were prepared and stored at −20°C. Anticollagen antibodies immunoglobulin G (IgG) fraction was purified using Protein A Sepharose Fast Flow (GE Healthcare). Approximately 200 µl of antiserum was passed through a Protein A sepharose column preequilibrated with 10-bed volume of buffer A (0.05 M boric acid and 4.0 M NaCl [pH 9.0]). Thereafter, bead-bound protein was washed with 5-bed volumes of buffer A followed by elution with buffer B (0.05 M sodium phosphate, 0.05 M sodium citrate, and 0.3 M NaCl [pH 3.0]). Eluted fractions were immediately neutralized with 1 M Tris (pH 8.0) and analyzed on SDS PAGE. Purified collagen antibodies were labeled using the NHS-Rodamine Antibody labeling kit (Thermo Scientific). Anti-ToLCNDV CP rabbit polyclonal antibodies (a gift from V. G. Malathi, Indian Agricultural Research Institute) were also labeled with the DyLight 488 antibody labeling kit (Thermo Scientific) using the manufacturer's instructions. Labeled antibodies fractions were stored at −20°C in small aliquots.

Detection of collagen protein in *B. tabaci* protein extract. One hundred whiteflies were homogenized in 200 µl of homogenization buffer (10 mM Na₂HPO₄ [pH 7.5], 1 mM EDTA, 500 mM NaCl, and protease inhibitor cocktail; Sigma-Aldrich). After complete homogenization, 400 µl of additional homogenization buffer was added and centrifuged at 1,485 × g for 5 min. Supernatant was collected in a separate tube. Protein detection was performed by

harvesting the clarified supernatant along with SDS PAGE loading dye. Sample was resolved through 12% SDS PAGE and transferred onto nitrocellulose membrane. Anticollagen antibodies (1:2,000) were used as primary antibody, whereas antimouse horse radish peroxidase (HRP)-labeled antibodies (1:3,000) were used as secondary antibody. Bands were visualized using the enhanced chemiluminescence system (Millipore).

Localization of collagen protein in *B. tabaci*. To localize the collagen protein in *B. tabaci*, guts and ovaries were dissected from virus-free *B. tabaci* and fixed in 4% paraformaldehyde for 20 min. Gut and ovaries were incubated for 1 h with 2% Triton X-100 to make them permeable, and they were subsequently incubated with NHS-Rodamine-labeled anticollagen antibodies for another 1 h. After incubation, insect guts and ovaries were washed with phosphate buffered saline with tween 20 (phosphate buffered saline [PBS] plus 0.1% Tween 20) three times and mounted using Vectashield mounting medium containing 4', 6-diamidine-2'-phenyl-indole dihydrochloride (DAPI) (Vector Labs). All of the images were acquired with a Nikon A1 confocal microscope.

Pull-down assay. Purified GST-tagged ToLCNDV CP was incubated with *B. tabaci* protein lysate in binding buffer (10 mM Na₂HPO₄ [pH 7.5], 1 mM EDTA, 500 mM NaCl, and protease inhibitor cocktail; Sigma-Aldrich). As a negative control, GST tag alone was also incubated with *B. tabaci* protein lysate. In each reaction, binding buffer-equilibrated GST beads were added. The reactions were incubated in a shaker at 4°C for 2 h. After incubation, the beads were centrifuged at 1,485 × g at 4°C. Unbound fractions were separated, and beads were washed twice with washing buffer (0.01% Triton X-100 in binding buffer). Beads were mixed with 40 µl sample buffer and boiled at 100°C for 10 min. After incubation, 20 µl from each reaction were resolved in a 10% SDS PAGE and transferred to nitrocellulose membrane. Blot was developed using anticollagen and antimouse HRP-labeled antibodies.

Interaction of collagen and purified particles using dot-blot assay. ToLCNDV virion particles were isolated as previously mentioned by Bock et al. (1978). To perform the dot-blot assay, 250 ng of purified ToLCNDV virion particles were blotted on nitrocellulose membrane and blocked with 3% bovine serum albumin (BSA) for 1 h at room temperature. After blocking, membrane was incubated with *B. tabaci* protein lysate for another 1 h. Membrane was then washed three times with PBST (PBS plus 0.1% Tween 20) and treated with primary anticollagen antibodies followed by alkaline phosphatase (APS)-labeled secondary antimouse antibodies. Blots were developed using BCIP/NBT substrate (Sigma-Aldrich).

TABLE 2. List of interacting proteins from yeast two-hybrid library sequencing using Tomato leaf curl New Delhi virus (ToLCNDV) and cotton leaf curl Rajasthan virus (CLCuV) coat protein (CP) as bait

ToLCNDV CP	CLCuV CP
1. <i>Bemisia tabaci</i> collagen α-2(I) chain-like (LOC109037601), messenger RNA, Sequence ID: XM_019052360.1	1. <i>Bemisia tabaci</i> cytospin A-like (LOC109042747), transcript variant X3, messenger RNA, Sequence ID: XM_019059667.1
2. <i>Bemisia tabaci</i> DEAD-box ATP-dependent RNA helicase 20-like (LOC109043421), messenger RNA, Sequence ID: XM_019060616.1	2. <i>Bemisia tabaci</i> collagen α-2(I) chain-like (LOC109037601), messenger RNA, Sequence ID: XM_019052360.1
3. <i>Bemisia tabaci</i> dihydrolipoyl dehydrogenase, mitochondrial (LOC109032698), messenger RNA, Sequence ID: XM_019044958.1	3. <i>Bemisia tabaci</i> , MAGUK p55 subfamily member 5 (LOC109032964), messenger RNA, Sequence ID: XM_019045324.1
4. <i>Bemisia tabaci</i> DNA polymerase ζ-catalytic subunit (LOC109041849), transcript variant X2, messenger RNA, Sequence ID: XM_019058329.1	4. <i>Bemisia tabaci</i> ADP-ribosylation factor 1 (LOC109041963), transcript variant X2, messenger RNA, Sequence ID: XM_019058491.1
5. <i>Bemisia tabaci</i> cytochrome b-c1 complex subunit 6, mitochondrial (LOC109044620), messenger RNA, Sequence ID: XM_019062442.1	5. <i>Bemisia tabaci</i> BAHD acyltransferase DCR-like (LOC109036545), messenger RNA, Sequence ID: XM_019050838.1
6. <i>Bemisia tabaci</i> eukaryotic initiation factor 4A-I (LOC109034044), transcript variant X1, messenger RNA, Sequence ID: XM_019046967.1	6. <i>Bemisia tabaci</i> cytochrome P450 4g15-like (LOC109042606), transcript variant X3, messenger RNA, Sequence ID: XM_019059469.1
7. <i>Bemisia tabaci</i> ensconsin-like (LOC109042067), messenger RNA, Sequence ID: XM_019058629.1	7. <i>Bemisia tabaci</i> dihydrolipoyl dehydrogenase, mitochondrial (LOC109032698), messenger RNA, Sequence ID: XM_019044958.1

Colocalization of collagen and ToLCNDV virions. To perform the colocalization study, guts were dissected from ToLCNDV-infected *B. tabaci*, which was further fixed in 4% paraformaldehyde for 20 min followed by permeabilization with 2% Triton X-100 for 1 h. Permeabilized guts were incubated with NHS-Rodamine-labeled anticollagen antibodies for another 1 h. After incubation, insect guts were washed three times with PBST (PBS plus 0.1% Tween 20) and again incubated with DyLight 488-labeled anti-CP antibodies for colocalization experiment. Guts were again washed with PBST and mounted using Vectashield mounting medium containing DAPI (Vector Labs). All of the images were acquired with the Nikon A1 confocal microscope. Virus-free *B. tabaci* fed on DyLight 488-labeled anti-CP antibodies served as negative control.

Virus acquisition assay. To analyze the role of collagen protein in virus acquisition, *B. tabaci* were fed on anticollagen antibodies for 48 h and subsequently shifted to ToLCNDV-infected tomato plants or CLCuV-infected cotton plants for 24 h for virus acquisition. Thereafter, *B. tabaci* was removed from infected plants and stored in ethanol at -20°C . Total DNA was isolated from *B. tabaci* and quantified using a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific). The same amount of DNA was used in all sets as a template to check viruses acquired by *B. tabaci* using semiquantitative PCR with *cp* gene-specific primers. The amount of

virus in the virus-free *B. tabaci* (healthy individuals not fed with ToLCNDV or CLCuV); viruliferous *B. tabaci* (healthy *B. tabaci* fed on ToLCNDV- or CLCuV-infected plants); and *B. tabaci* fed with 15% sucrose (followed by feeding on ToLCNDV- or CLCuV-infected plants), antithioredoxin antibodies (followed by feeding on ToLCNDV- or CLCuV-infected plants; 1:100 dilution), and anticollagen antibodies (followed by feeding on ToLCNDV- or CLCuV-infected plants; 1:100 dilution) was normalized using β -actin quantification. The amount of ToLCNDV and CLCuV acquisition was quantified by measuring the intensities of PCR bands using C-DiGit Blot Scanner: 1.1.3 program (LI-COR Biosciences). The amount of ToLCNDV and CLCuV acquired by the viruliferous *B. tabaci* was considered to be 100% for additional quantification of other treatments. The virus-free *B. tabaci*, viruliferous *B. tabaci*, *B. tabaci* fed with 15% sucrose, antithioredoxin antibodies fed *B. tabaci* served as the control against *B. tabaci* fed with the anticollagen antibodies.

ToLCNDV transmission efficiency experiments. To assess the role of collagen in ToLCNDV transmission, newly emerged (5- to 8-day-old) whiteflies were fed on anticollagen antibodies at 1:100 dilution in 15% sucrose solution for 48 h. Subsequently, whiteflies were shifted to ToLCNDV-infected tomato plants for virus acquisition for an acquisition access period of 24 h (infectious

Fig. 1. Domain mapping of collagen. **A**, Representation of complete collagen protein template of 394 amino acids (open reading frame; ORF), indicating the N-terminal signal peptide (1 to 29 amino acids) immediately followed with collagen triple-helix repeat (31 to 119 amino acids), whereas the C terminus of the collagen protein shows the presence of fibrillar collagen C-terminal domain (165 to 394 amino acids). Protein was also analyzed on Phobius server for protein's transmembrane topology, which revealed the presence of noncytoplasmic domain (30 to 394 amino acids). **B**, Schematic diagram and three-dimensional structure of collagen protein. N represents the N terminus of collagen, and C represents the C terminus of collagen. The β -strands (yellow), α -helices ($\alpha 1$ and $\alpha 2$; pink), 3Helix (purple), and loop (blue) are shown. The three-dimensional structure was predicted using the Raptor X online web tool.

clone and ToLCNDV-infected tomato plants were collected from V. G. Malathi, IARI) (Jyothsna et al. 2013). *B. tabaci* fed only on ToLCNDV-infected plants and whiteflies fed with antithioredoxin tag antibodies served as controls. Each whitefly was transferred to a single healthy tomato plant (two- to three-leaf stage) and incubated in cages under greenhouse conditions for 3 days, after which the flies were removed. Tomato plants were analyzed for virus symptoms at 20 days postinfection. Detection of ToLCNDV was also confirmed in infected plants and whiteflies through PCR. Genomic DNA isolated from the tomato plant (Pandey et al. 2009) and *B. tabaci* (Singh et al. 2012) was used as a template for PCR with CP gene-specific primers. Four independent virus transmission experiments were performed.

Data analysis. All statistical analysis was performed using Statistica 12 (Statsoft Inc.). The means are represented as mean $\pm$ standard error. The virus transmission efficiency between different groups was compared by one-way analysis of variance with significance level of $P \leq 0.05$.

RESULTS

***B. tabaci* yeast cDNA expression library construction and screening against the CPs of ToLCNDV and CLCuV.** To identify begomovirus receptors, a gut cDNA library of *B. tabaci* was constructed and screened against the CPs of ToLCNDV and CLCuV using yeast two-hybrid assays. Several positive clones were isolated and sequenced (Table 2). A positive clone with sequence similarity to collagen was found to be interacted with the CPs of both viruses.

Sequence analysis of this clone revealed the presence of a 3' untranslated region (UTR) sequence, whereas the 5' sequence was absent. Using the 5' RACE technique, complete sequence of *collagen* gene was retrieved and analyzed with Bio-Edit software. Sequence

analysis suggested the presence of an open reading frame of 1,185 bp that encodes for 394 amino acids exhibiting molecular weight of 41.97 kDa. This gene contained the 5' UTR and 3' UTR sequences of 202 and 97 bp, respectively. This protein has been named collagen protein, and its complete sequence was submitted to the NCBI database under accession number MG837560.

Collagen protein exhibited similarity to COLF1 superfamily and contains an N-terminal signal peptide. The sequence of *collagen* gene was analyzed for sequence similarity in the NCBI database using the BLASTn 2.2.28 program against nucleotide collections. BLAST result indicated significant similarity (92% identity) with collagen α -2 (1) gene (accession number XM_D19052360.1) of *B. tabaci*, suggesting that the identified positive clone through the yeast two-hybrid screen in this study was a member of collagen super family proteins.

To further identify the conserved domains present in collagen protein, the amino acid sequence of collagen protein was searched for similarity against the CDD on NCBI. Two conserved domains, a fibrillar collagen C-terminal domain and a collagen triple-helix repeat, were found in collagen protein (Fig. 1A).

Furthermore, amino acid sequence of collagen was also analyzed on the InterProScan database version 4.8. The InterProScan result also revealed two conserved domains: fibrillar collagen C-terminal domain and a collagen triple-helix repeat. The software predicted a signal peptide residing at the N terminus of the collagen protein (1 to 29 amino acids), which could be required for protein trafficking (Supplementary Fig. S1). The location of cleavage site for signal peptide was found between the 29th and 30th amino acids by using SignalP (Supplementary Fig. S2). The complete schematic template drawn according to above-mentioned information is summarized in Figure 1A.

Fig. 2. Phylogenetic analysis of collagen protein. The phylogenetic tree of collagen protein of *Bemisia tabaci* and collagen sequences from other organisms was constructed using MEGA 7 with the neighbor-joining method. Numbers on each internal branch in the phylogenetic tree node represent the boot strap value from 1,000 replicates.

Additionally, the 3D structure of collagen protein was predicted using RaportX online web server. The structure revealed the presence of 12 β -sheets, α -helix, and 2 small 3- α helices in COLF1 domain, whereas no secondary structure was observed in the N terminus of collagen protein (Fig. 1B). The structures of N- and C-terminal domains are based on templates 3HQV (single type I

collagen molecule, *Rattus norvegicus*) and 4ac2A (crystal structure of human fibrillar procollagen type III C-propeptide trimer), respectively, which are available in the Protein Data Bank. To further analyze the phylogenetic relationship of *B. tabaci* collagen protein, a phylogenetic tree was constructed using the neighbor-joining method with MEGA 7.0 software (Fig. 2). Phylogenetic

Fig. 3. Heterologous expression and purification of collagen and thioredoxin protein. **A**, Induced and uninduced cultures of bacterial cells transformed with pET32a-collagen plasmid. Lane 1, Molecular weight marker (in kilodaltons). Lane 2, Total protein extract from induced bacterial cell culture expressing a prominent ~60-kDa band of overexpressed thioredoxin-tagged collagen. Lane 3, Total protein extract from uninduced bacterial culture. **B**, Purification of collagen. Lane 1, Molecular weight marker (in kilodaltons). Lane 2, Flow through. Lanes 3 to 5, Purified fractions of ~60-kDa thioredoxin-tagged collagen. **C**, Purification of glutathione S-transferase (GST)-tagged tomato leaf curl New Delhi virus (ToLCNDV) coat protein (CP). Lane 1, Molecular weight marker (in kilodaltons). Lane 2, Purified fraction of ToLCNDV CP. **D**, Detection of collagen in *Bemisia tabaci* protein extract. An ~45-kDa band of collagen was observed in the whitefly homogenate using anticollagen antibodies in soluble fraction of total protein extract.

analysis revealed clustering of collagen protein (isolated in this study) with the predicted collagen protein of *B. tabaci* and formation of a sister group with collagen proteins from *Bombyx mori*, *Harpegnathos saltator*, *Anopheles gambiae*, and *Danaus plexippus*.

Purified recombinant collagen protein was successfully used for anticollagen antibody production. After sequence analysis of the *collagen* gene, a new set of primers was designed for amplifying the complete gene. Amplified PCR product was excised, purified, and cloned directly into TA cloning pGEM-T Easy vector. Sequence analysis suggested the presence of complete open reading frame (ORF) of *collagen* gene in pGEM-T Easy vector. To express and purify the collagen protein, the *collagen* gene was subcloned into pET32a expression vector between *Eco*R1 and *Sal*I restriction sites. Induced protein was solubilized with detergent and purified using Ni-NTA beads (Fig. 3A and B), and it was used for antibody production in BALB/c mice. GST-tagged ToLCNDV CP (Fig. 3C) was purified as mentioned in Rana et al. (2016). Collagen protein was successfully detected in *B. tabaci* protein extracts (Fig. 3D). Western blot analysis revealed the presence of an ~45-kDa band in supernatant fraction of *B. tabaci* homogenate.

Collagen protein was localized in whole gut of *B. tabaci* using fluorophore-labeled antibodies. Localization experiments were performed using rhodamine-labeled, IgG-purified anticollagen antibodies. The expression of collagen protein was found confined mostly to the gut region (Fig. 4A to C). However, in

the case of ovary, no red signals were observed (Fig. 4D to F), thus clearly indicating that expression of collagen was confined to gut. Moreover, no such signals were observed by using anti-TRITC-labeled secondary antibody.

Collagen protein interacts with CLCuV and ToLCNDV CPs. To validate the collagen-CP interaction, we performed an in vivo yeast two-hybrid assay. Yeast cells were cotransformed with pGADT7-Rec-collagen (AD-collagen) and pGBKT7-ToLCNDV CP (BD-ToLCNDV CP) or AD-collagen and pGBKT7-CLCuV CP (BD-CLCuV CP) in separate experiments. Transformed yeast cells (experimental) were able to grow on 2DO and 4DO media, and they tested positive in β -galactosidase assay. Yeast cells cotransformed with pGADT7-Rec-T (AD-T) and pGBKT7-53 (BD-53; positive controls) mimicked the same result as experiments, whereas yeast cells cotransformed with pGADT7-Rec (AD) and BD-ToLCNDV CP or AD and BD-ToLCNDV CP (negative controls) grew on 2DO but not on 4DO, and they also did not show β -galactosidase activity (Fig. 5A). These results indicated specific interaction of collagen protein with CP of both ToLCNDV and CLCuV.

Furthermore, in vitro pull-down assays were performed to validate the interaction between purified GST-tagged ToLCNDV CP with endogenous collagen protein using *B. tabaci* protein extract (Fig. 5B). Presence of collagen protein band in bound fraction (in which collagen and GST-tagged ToLCNDV CP were allowed to interact) indicated the interaction of collagen with GST-tagged

Panel 1

Panel 2

Fig. 4. Confocal analysis of collagen expression in dissected organs of *Bemisia tabaci*. **A to C**, Localization of collagen protein in dissected gut of *B. tabaci*. **A**, Bright field. **B**, Collagen expression in dissected gut using anticollagen rhodamine-labeled immunoglobulin G (IgG) antibodies. **C**, Merged image from **A** and **B**. CA, caecae; FC, filter chamber; HG, hind gut; MG, midgut. **D to F**, Localization of collagen protein in dissected ovary of *B. tabaci*. **D**, Bright field. **E**, Collagen expression in dissected gut of flies using anticollagen rhodamine-labeled IgG antibodies. **F**, Merged image from **A** and **B**. All images were viewed at 20 $\times$ magnification.

ToLCNDV CP. Absence of collagen band in bound fraction when collagen and GST were allowed to interact indicated that collagen interacts with ToLCNDV CP but not with GST tag. These results indicated that ToLCNDV CP specifically interacts with endogenous collagen protein (Fig. 5B).

Dot-blot assays were also performed to check the interaction of purified ToLCNDV particles with endogenous collagen protein (Fig. 5C). When endogenous collagen was allowed to interact with blotted purified ToLCNDV particles, positive signal was observed. However, no signals were observed when mock virion particles

Fig. 5. Interaction of collagen protein with tomato leaf curl New Delhi virus (ToLCNDV) and cotton leaf curl Rajasthan virus (CLCuV). **A**, Yeast two-hybrid assay. Interaction of ToLCNDV and CLCuV coat protein (CP) with collagen protein in yeast. 1, Yeast cells were transformed with both plasmids constructs (AD and BD constructs), and all transformed cells were plated on a 2 dropout media (2DO) plate. 2, These cells were streaked on 4DO plates. 3, β-gal assay of the cells grown in 2DO plates was done. Positive control (AD-Rec-T plus BD-53) and experiments (AD-collagen plus BD-ToLCNDV CP and AD-collagen plus BD-CLCuV CP) showed β-gal activity and are indicated by green, yellow, and blue arrows, respectively. 4, Negative control for both is represented by red arrows. Schematic diagram representing the streaking pattern of the constructs. **B**, Pull-down assay. In vitro interaction of endogenous collagen with glutathione S-transferase (GST)-tagged ToLCNDV CP. In panel 1, protein bands were visualized using anticollagen antibodies (primary antibody), whereas in panel 2, protein bands were visualized using anti-GST antibodies (primary antibody). Lanes 1 and 3 are of unbound fraction, and lanes 2 and 4 are of bound fraction. Lane 5 is of lysate, whereas lane 6 denotes marker. In panel 1, presence of collagen band in lane 4 and absence in lane 3 indicated that GST-tagged ToLCNDV CP was interacting with collagen and hence, able to pull it down in bound fraction. Absence of collagen in lane 2 indicates that collagen did not interact with GST unspecifically. In panel 2, anti-GST tag antibody helped to develop bands of GST in lanes 1 and 2 and bands of GST-tagged ToLCNDV CP in lanes 3 and 4. **C**, Dot-blot assay. 1, Purified collagen protein incubated with mock virion particles showed no signal, suggesting negative interaction. Blot was developed using anticollagen antibody. 2, In positive control experiment, blotted ToLCNDV particles showed signals when blot was developed using anti-CP antibodies. 3, Endogenous collagen protein of *Bemisia tabaci* incubated with blotted ToLCNDV particles showed a signal, suggesting positive interaction of ToLCNDV particles and endogenous collagen (blot was developed using anticollagen antibody).

were allowed to interact with collagen protein (negative control). As a positive control, presence of ToLCNDV particles was also confirmed using anti-CP antibodies. These results indicated the specific interaction of endogenous collagen with purified ToLCNDV. Thus, pull-down and dot-blot assay experiments further confirmed the yeast two-hybrid screen results.

Colocalization of collagen and ToLCNDV virion particles.

To analyze the possible colocalization of ToLCNDV and collagen in viruliferous *B. tabaci*, guts were dissected and incubated with fluorophore-labeled anti-CP and anticollagen antibodies (Fig. 6). Samples were further analyzed on confocal microscope. Results showed the clear colocalization of collagen and ToLCNDV particles in the midgut loop region of *B. tabaci*. Colocalization of collagen and viral CP was seen as yellow patches in the midgut loop region; however, scanty colocalization signal was also observed in other parts of midgut. No signal of ToLCNDV was found in negative control, where virus-free whiteflies were fed with labeled anti-CP antibodies.

Collagen protein helps in virus acquisition. To evaluate the role of collagen protein in ToLCNDV acquisition by *B. tabaci*, virus-free flies were fed with anticollagen antibodies and shifted for 24 h to ToLCNDV-infected tomato plants for virus acquisition. The amount of virion acquired by *B. tabaci* after collagen antibody feeding was quantified through CP gene using semiquantitative PCR (Fig. 7A). The amount of virion present in *B. tabaci* was detected in virus-free *B. tabaci*, viruliferous *B. tabaci*, *B. tabaci* fed on 15% sucrose, *B. tabaci* fed on antithioredoxin tag antibodies, and *B. tabaci* fed on anticollagen antibodies. ToLCNDV was detected in all of the treatments as shown in Figure 7B, except in virus-free *B. tabaci* (100% in viruliferous *B. tabaci*, 94.5% in sucrose-fed *B. tabaci*, 74% in antithioredoxin tag antibodies fed *B. tabaci*, and 22.6% in anticollagen antibodies fed *B. tabaci*). Similar virus acquisition results were obtained with CLCuV when *B. tabaci* was fed on CLCuV-infected cotton plants. *B. tabaci* exhibited 0, 100, 75, 70, and 26% CLCuV acquisition efficiency with virus-free, viruliferous, sucrose-fed, antithioredoxin tag antibodies-fed, and

Fig. 6. Colocalization of collagen with tomato leaf curl New Delhi virus (ToLCNDV). Collagen and ToLCNDV were localized using specific fluorophore-labeled antibodies. DyLight 488 and Rodamine fluorophores were used for coat protein and collagen antibodies, respectively. Guts were dissected from whiteflies and incubated with fluorophore-labeled antibodies. Processed whiteflies were mounted on slides and analyzed in microscope. **A**, The localization of ToLCNDV and collagen in different parts of the gut of whiteflies. **a**, Distribution of collagen. **b**, Bright-field image showing different parts of gut. **c**, Colocalization of ToLCNDV and collagen. **d**, Distribution of ToLCNDV. All images were viewed at 20× magnification, and **B**, a selected area (in A, c) is viewed in 40× magnification. CA, caecae; FC, filter chamber; HG, hind gut; MG, midgut.

anticollagen antibodies-fed *B. tabaci*, respectively (Fig. 7B). These results indicated a reduced acquisition of ToLCNDV and CLCuV in case of anticollagen antibodies fed *B. tabaci* compared with viruliferous, sucrose-fed, and antithioredoxin tag antibodies-fed *B. tabaci*, suggesting the role of collagen protein in acquisition of both ToLCNDV and CLCuV.

Collagen plays an important role in virus transmission.

To evaluate the role of collagen in ToLCNDV transmission, *B. tabaci* fed with anticollagen antibodies was incubated on ToLCNDV-infected tomato plant for ToLCNDV acquisition. They were subsequently shifted onto healthy tomato plants (three- to four-leaf stage) for virus inoculation. *B. tabaci* fed on antithioredoxin tag antibodies instead of anticollagen antibodies and *B. tabaci* fed only on ToLCNDV-infected plants were used as controls.

Four independent experiments were conducted for checking virus transmission with 20 plants each. In total, 80 plants each were used for ToLCNDV transmission by *B. tabaci* fed with anticollagen antibodies or antithioredoxin tag or only fed on ToLCNDV-infected plants. *B. tabaci* for both the controls showed significantly higher transmission efficiency compared with the anticollagen antibodies fed *B. tabaci*. The ToLCNDV transmission was 53.75% for anticollagen antibodies fed *B. tabaci*, 96.25% for antithioredoxin fed *B. tabaci*, and 100% for only virus-fed *B. tabaci*. There was a significant difference in virus transmission between anticollagen antibodies fed *B. tabaci* and antithioredoxin antibodies fed *B. tabaci*

($F_{[1,6]} = 157.64$, $P < 0.001$) or only ToLCNDV-fed *B. tabaci* ($F_{[1,6]} = 373.36$, $P < 0.001$). However, there was no significant difference in ToLCNDV transmission between the two controls (antithioredoxin antibodies fed *B. tabaci* and only ToLCNDV-fed *B. tabaci*; $F_{[1,6]} = 2.45$, $P = 0.1682$) (Fig. 8).

DISCUSSION

Circulation of viruses within insect vectors involves complex interaction between the virus and vector components. In the last two decades, knowledge regarding the molecular interactions between insect vectors and their viruses has improved with the use of modern tools, like phage display, yeast two-hybrid, microarrays, etc. However, the molecular mechanisms of begomovirus-*B. tabaci* interaction remain largely unknown. To reveal this complex molecular association, a gut cDNA library of *B. tabaci* was constructed and screened against the CPs of ToLCNDV and CLCuV using yeast two-hybrid assays. Several positive clones were isolated and characterized. For additional work, a clone similar to collagen protein, which appeared in the library screens of both of the begomoviruses, was selected based on review of literature.

In preliminary work, the sequence analysis of collagen protein revealed the presence of collagen triple-helix repeat and COLF1 domain at N- and C-terminal regions, respectively. COLF1 domain (also known as C-propeptide domain) is highly conserved among

Fig. 7. Tomato leaf curl New Delhi virus (ToLCNDV) and cotton leaf curl Rajasthan virus (CLCuV) acquisition assays in anticollagen antibodies fed *Bemisia tabaci*. **A**, Semiquantitative PCR performed to estimate virus acquisition through amplification of coat protein (CP) DNA. ToLCNDV and CLCuV acquisition efficiencies were analyzed in panels 1 and 2 and panels 3 and 4, respectively. Five sets of whiteflies were used for virus acquisition experiment. Virus-free *B. tabaci* represents the whiteflies not fed on ToLCNDV- or CLCuV-infected plants. Viruliferous *B. tabaci* represents whiteflies fed on ToLCNDV- or CLCuV-infected plants. *B. tabaci* in other experimental sets was first fed with 15% sucrose, antithioredoxin tag antibodies, or anticollagen antibodies and then, fed on ToLCNDV- or CLCuV-infected plants. Panels 1 and 3 represent the amplification of CP gene of ToLCNDV and CLCuV from quantified genomic DNA of virus-free, viruliferous, 15% sucrose-fed, antithioredoxin tag antibodies-fed, and anticollagen antibodies-fed *B. tabaci* in lanes 1 to 5, respectively. Panels 2 and 4 represent the amplification of actin from quantified genomic DNA of virus-free, viruliferous, 15% sucrose-fed, antithioredoxin tag-fed, and anticollagen antibodies-fed *B. tabaci* in lanes 1 to 5, respectively. **B**, Representation of the percentage of ToLCNDV (yellow) and CLCuV (orange) acquisition by *B. tabaci* on various treatments. The amounts of ToLCNDV and CLCuV present in viruliferous *B. tabaci* were considered to be 100% for additional quantification of other treatments.

different fibrillar collagen proteins across many organisms (Exposito et al. 2010). Insect collagen fibrils are very variable; in flies and moths, they are very small in diameter and indistinctly banded. However, in other insects, such as locusts and cockroaches, the fibrils are much larger in diameter and display a distinct axial periodicity in sections (Ashhurst and Baily 1980). The formation of procollagen molecule from its three polypeptide chains is initiated by the association of COLF1 domain, which is a crucial step in nucleation and folding of the newly formed rod-like molecule (Bottomley et al. 2001; Boudko et al. 2012). Proteins having a collagen triple-helix repeat belong to the collagen superfamily; however, such structure has also been reported in a variety of noncollagenous proteins, such as bacteria-binding receptor MARCO, the complement component C1q, acetylcholinesterase, collectin-43 aggrexin, and several other proteins isolated from invertebrates and bacteria (Fields 2010).

As a next step, *collagen* gene was cloned, expressed, and purified using the bacterial expression system. Because collagen protein was solubilized and purified using SDS, it could not be used for in vitro protein–protein interaction studies. However, such solubilized collagen protein was used for antibody production in mice, which was further used to examine collagen protein localization in various organs of *B. tabaci*. Collagen protein was localized throughout the gut membrane, whereas no signals were observed in the head and ovary, suggesting a local or specific collagen protein expression in the gut epithelium layers, where it could be part of extracellular matrix in the basal lamina.

The basal lamina is a part of extracellular matrix secreted by underlying tissue that consists of collagen and laminin glycoproteins combined with a mixture of smaller molecules, including entactin, glutactin, heparin sulfate proteoglycans, fibronectin, and integrin depending on the organism. These compounds cross-link and stabilize the collagen laminin network (Pederson 1991). Several viruses, such as herpes simplex virus (Shieh et al. 1992), sindbis virus (Wang et al. 1992), and influenza viruses (Julkunen et al. 1984), interact with the basal lamina-associated molecules from a variety of host cells and tissues. In aphids, a high level of vector specificity is observed between luteoviruses and the accessory salivary gland (ASG) of their aphid vector. Transmission electron microscopy analysis of ASG suggested that basal lamina binds to luteoviruses and concentrates them (Peiffer et al. 1997). However, in contrary to aphid, we could not locate collagen protein in basal lamina of salivary gland in *B. tabaci* (data not shown). It may be owing to the absence of collagen in basal lamina or the presence of the other type of collagen, which could not be detected using our laboratory-raised collagen-specific antisera.

To confirm the association of collagen with ToLCNDV in *B. tabaci*, protein lysate prepared from *B. tabaci* was incubated with ToLCNDV CP. Lysate proteins bound to ToLCNDV CP were resolved on SDS PAGE and analyzed using western blot with anticollagen antibodies for detecting the presence of collagen protein. Our results showed collagen to be associated with ToLCNDV CP. Dot-blot assay also suggested positive interaction of virion particles with collagen protein. However, because of lack of direct physical protein–protein interaction information between purified collagen and ToLCNDV CP, it could not be confirmed whether collagen interacts with ToLCNDV CP directly or indirectly in association with some other protein(s), although yeast two-hybrid assay suggested that collagen protein interacts directly with the CP of ToLCNDV. Association of ToLCNDV with collagen was further confirmed using colocalization studies, which suggested that virion particles and collagen protein coexist in the midgut loop region. In earlier studies, midgut loop area had been considered as one of the crucial possible sites for virus binding and its entry into hemolymph (Hunter et al. 1998). Furthermore, colocalization of collagen protein with ToLCNDV in midgut loop suggests a substantial role for collagen protein during ToLCNDV binding and translocation to midgut epithelium cells. This kind of interaction between the collagen and virus CP may be important

for the begomoviruses to get docked over the cell surface via basal lamina and thus, subsequently move across the gut epithelial cells.

In previous studies, collagen protein exhibited physical interaction with other extracellular matrix proteins, like integrin (Jokinen et al. 2004), fibronectin (Kleinman et al. 1981), vitronectin (Gebb et al. 1986), and glycoproteins (Munakata et al. 1999), which act as a receptor for many insect and mammalian viruses, like echo virus (Jokinen et al. 2010), influenza A virus, para influenza virus, mumps virus (Julkunen et al. 1983), zika virus (Kim et al. 2017), dengue virus (Conde et al. 2016), bovine viral diarrhea virus (Iqbal et al. 2000), syncytial virus (Hallak et al. 2007), etc. These observations suggest that collagen could be involved in pathogen binding and internalization into cells via extracellular matrix proteins. These strategies may be followed by the ToLCNDV viral particles in associating with the basal membrane of gut cells and thus, facilitating their movement across the epithelial lining. This is further emphasized by the reduction in ToLCNDV and CLCuV acquisition and ToLCNDV transmission by whiteflies fed on anticollagen antibodies. ToLCNDV acquired by *B. tabaci* fed with anticollagen antibodies was as low as 22.6% compared with whiteflies fed with ToLCNDV (100%). Similarly, CLCuV acquired by *B. tabaci* fed with anticollagen antibodies was only 26% compared with whiteflies fed with CLCuV (100%).

Furthermore, *B. tabaci* fed with anticollagen antibodies exhibited reduction of 46.25% in ToLCNDV transmission with respect to positive control (ToLCNDV-fed healthy whiteflies). These observations suggest that anticollagen antibody bound to collagen protein in *B. tabaci* blocks the virus attachment to gut membrane, which further leads to the disruption of collagen–ToLCNDV complex. These observations could be the result of direct interference by anticollagen antibodies in collagen–ToLCNDV complex, or they may be because of the loss of binding capacity of mediator protein(s) to collagen protein, which acts as a bridge between ToLCNDV and collagen protein during virion attachment to gut wall.

Several studies have reported a role for collagen in insect metabolism. Using the lepidopteran model insect *Galleria mellonella*, it has been postulated that fragments of collagen IV, resulting from hydrolysis by microbial metalloproteinases, represent danger or alarm signals in insects (Altincicek and Vilcinskis 2007; Altincicek et al. 2009). Furthermore, there are

Fig. 8. Tomato leaf curl New Delhi virus (ToLCNDV) transmission efficiency of anticollagen antibodies fed *Bemisia tabaci*. There was a reduction in ToLCNDV transmission rate after feeding whiteflies on anticollagen antibodies, whereas antithioredoxin antibodies-fed whiteflies and only ToLCNDV-fed whiteflies showed 90 to 100% transmission efficiency. Different letters on the bars indicate significant difference between the three treatments with $P \leq 0.05$.

also reports indicating the role of collagen in insect immune responses. Therefore, it is crucial to explore additional interaction of viral CPs and collagen. However, it remains an interesting open question as to how viral CP hijacks collagen to escape from immune response while maintaining the viral titer for transmission.

ACKNOWLEDGMENTS

We thank V. G. Malathi for providing the ToLCNDV infectious clone. We also thank Ashok Kumar and Ravi Singh for excellent laboratory assistance.

LITERATURE CITED

- Altincicek, B., Berisha, A., Mukherjee, K., Spengler, B., Römpf, A., and Vilcinskis, A. 2009. Identification of collagen IV derived danger/alarms signals in insect immunity by nanoLC-FTICR MS. *Biol. Chem.* 390: 1303-1311.
- Altincicek, B., and Vilcinskis, A. 2007. Identification of immune-related genes from an apterygote insect, the firebrat *Thermobia domestica*. *Insect Biochem. Mol. Biol.* 37:726-731.
- Ashhurst, D. E., and Baily, A. J. 1980. Locust collagen: Morphological and biochemical characterization. *Eur. J. Biochem.* 103:75-83.
- Bock, K. R., Guthrie, E. J., and Meredith, G. 1978. Distribution, host range, properties and purification of cassava latent virus, a geminivirus. *Ann. Appl. Biol.* 90:361-367.
- Bottomley, M. J., Batten, M. R., Lumb, R. A., and Bulleid, N. J. 2001. Quality control in the endoplasmic reticulum: PDI mediates the ER retention of unassembled procollagen C-propeptides. *Curr. Biol.* 11:P1114-P1118.
- Boudko, S. P., Engel, J., and Bachinger, H. P. 2012. The crucial role of trimerization domains in collagen folding. *Int. J. Biochem. Cell Biol.* 44: 21-32.
- Bouvaire, S., Boonham, N., and Douglas, A. E. 2011. Interactions between a luteovirus and the GroEL chaperonin protein of the symbiotic bacterium *Buchnera aphidicola* of aphids. *J. Gen. Virol.* 92:1467-1474.
- Bridson, R. W., Pinner, M. S., Stanley, J., and Markham, P. G. 1990. Geminivirus coat protein gene replacement alters insect specificity. *Virology* 177:85-94.
- Brown, J. K., Frohlich, D., and Rosell, R. 1995. The sweetpotato/silverleaf whiteflies: Biotypes of *Bemisia tabaci* (Genn.), or a species complex? *Annu. Rev. Entomol.* 40:511-534.
- Cicero, J. M., Hiebert, E., and Webb, S. E. 1995. The alimentary canal of *Bemisia tabaci* and *Trialeurodes abutilonea* (Homoptera, Sternorrhynchi): Histology, ultrastructure and correlation to function. *Zoomorphology* 115: 31-39.
- Conde, J. N., Da Silva, E. M., Allonso, D., Coelho, D. R., Andrade, I. D., De Medeiros, L. N., Menezes, J. L., Barbosa, A. S., and Mohana-Borges, R. 2016. Inhibition of the membrane attack complex by dengue NS1 through interaction with vitronectin and terminal complement proteins. *J. Virol.* 90: 9570-9581.
- Czosnek, H., and Ghanim, M. 2002. Circulative pathway of begomoviruses in the whitefly vector *Bemisia tabaci*—Insights from studies with *Tomato yellow leaf curl virus*. *Ann. Appl. Biol.* 140:215-231.
- Exposito, J. Y., Valcourt, U., Cluzel, C., and Lethias, C. 2010. The fibrillar collagen family. *Int. J. Mol. Sci.* 11:407-426.
- Fields, G. B. 2010. Synthesis and biological applications of collagen-model triple-helical peptides. *Org. Biomol. Chem.* 8:1237-1258.
- Firdaus, S., Heusden, A., Hidayati, N., Supena, E., Mumm, R., Vos, R. H., Visser, R. F., and Vosman, B. 2013. Identification and QTL mapping of whitefly resistance components in *Solanum galapagense*. *Theor. Appl. Genet.* 126:1487-1501.
- Gebb, C., Hayman, E. G., Engvall, E., and Ruoslahti, E. 1986. Interaction of vitronectin with collagen. *J. Biol. Chem.* 35:16698-16703.
- Ghanim, M., Morin, S., and Czosnek, H. 2001. Rate of Tomato yellow leaf curl virus translocation in the circulative transmission pathway of its vector, the whitefly *Bemisia tabaci*. *Phytopathology* 91:188-196.
- Gottlieb, Y., Zchori-Fein, E., Mozes-Daube, N., Kontsedalov, S., Skoljac, M., Brumin, M., Sobol, I., Czosnek, H., Vavre, F., Fleury, F., and Ghanim, M. 2010. The transmission efficiency of tomato yellow leaf curl virus by the whitefly *Bemisia tabaci* is correlated with the presence of a specific symbiotic bacterium species. *J. Virol.* 84:9310-9317.
- Götz, M., Popovski, S., Kollenberg, M., Gorovitz, R., Brown, J. K., Cicero, J., Czosnek, H., Winter, S., and Ghanim, M. 2012. Implication of *Bemisia tabaci* heat shock protein 70 in Begomovirus-whitefly interactions. *J. Virol.* 86:13241-13252.
- Hallak, L. K., Kwilas, S. A., and Peebles, M. E. 2007. Interaction between respiratory syncytial virus and glycosaminoglycans, including heparan sulfate. *Methods Mol. Biol.* 379:15-34.
- Hogenhout, S. A., Ammar, E. D., Whitfield, A. E., and Redinbaugh, M. G. 2008. Insect Vector interactions with persistently transmitted viruses. *Annu. Rev. Phytopathol.* 46:327-359.
- Hunter, W., Hiebert, E., Webb, S., Tsai, J., and Polston, J. 1998. Location of geminiviruses in the whitefly *Bemisia tabaci* (Homoptera: Aleyrodidae). *Plant Dis.* 82:1147-1151.
- Iqbal, M., Flick-Smith, H., and McCauley, J. W. 2000. Interactions of bovine viral diarrhoea virus glycoprotein E(rns) with cell surface glycosaminoglycans. *J. Gen. Virol.* 81:451-459.
- Jokinen, J., Dadu, E., Nykvist, P., Kapyla, J., White, D. J., Ivaska, J., Vehviläinen, P., Reunanen, H., Larjava, H., Hakkinen, L., and Heino, J. 2004. Integrin-mediated cell adhesion to type I collagen fibrils. *J. Biol. Chem.* 279:31956-31963.
- Jokinen, J., White, D. J., Salmela, M., Huhtala, M., Kapyla, J. K., Puranen, J. S., Nissinen, L., Kankaanpää, P., Marjomäki, V., Hyypiä, T., Johnson, M. S., and Heino, J. 2010. Molecular mechanism of $\alpha 2 \beta 1$ integrin interaction with human echovirus 1. *EMBO J.* 29:196-208.
- Julkunen, I., Hautanen, A., and Keski-Oja, J. 1983. Interaction of viral envelope glycoproteins with fibronectin. *Infect. Immun.* 40:876-881.
- Julkunen, I., Vartio, T., and Keskiöja, J. 1984. Localization of viral-envelope glycoprotein binding sites in fibronectin. *Biochem. J.* 219:425-428.
- Jyothsna, P., Haq, Q. M., Singh, P., Sumiya, K. V., Praveen, S., Rawat, R., Briddon, R. W., and Malathi, V. G. 2013. Infection of Tomato leaf curl New Delhi virus (ToLCNDV), a bipartite begomovirus with betasatellites, results in enhanced level of helper virus components and antagonistic interaction between DNA B and betasatellites. *Appl. Microbiol. Biotechnol.* 97: 5457-5471.
- Kanakala, S., and Ghanim, M. 2016. Implication of the whitefly *Bemisia tabaci* cyclophilin B protein in the transmission of *Tomato yellow leaf curl virus*. *Front. Plant Sci.* 7:1702.
- Kim, S. Y., Zhao, J., Liu, X., Fraser, K., Lin, L., Zhang, X., Zhang, F., Dordick, J. S., and Linhardt, R. J. 2017. Interaction of Zika virus envelope protein with glycosaminoglycans. *Biochemistry* 56:1151-1162.
- Kleinman, H. K., Wilkes, C. M., and Martin, G. R. 1981. Interaction of fibronectin with collagen fibrils. *Biochemistry* 20:2325-2330.
- Liu, S. S., Colvin, J., and De Barro, P. J. 2012. Species concepts as applied to the whitefly *Bemisia tabaci* systematics: How many species are there? *J. Integr. Agric.* 11:176-186.
- Luan, J. B., Li, J. M., Varela, N., Wang, Y. L., Li, F. F., Bao, Y. Y., Zhang, C. X., Liu, S. S., and Wang, X. W. 2011. Global analysis of the transcriptional response of whitefly to *Tomato yellow leaf curl china virus* reveals their relationship of coevolved adaptations. *J. Virol.* 85:3330-3340.
- Moreno-Delafuente, A., Garzo, E., Moreno, A., and Fereres, A. 2013. A plant virus manipulates the behavior of its whitefly vector to enhance its transmission efficiency and spread. *PLoS One* 8:e61543.
- Munakata, H., Takagaki, K., Majima, M., and Endo, M. 1999. Interaction between collagens and glycosaminoglycans investigated using a surface plasmon resonance biosensor. *Glycobiology* 9:1023-1027.
- Navas-Castillo, J., Fiallo-Olive, E., and Sanchez-Campos, S. 2011. Emerging virus diseases transmitted by whiteflies. *Annu. Rev. Phytopathol.* 49: 219-248.
- Noris, E., Vaira, A. M., Caciagli, P., Masenga, V., Gronenborn, B., and Accotto, G. P. 1998. Amino acids in the capsid protein of *Tomato yellow leaf curl virus* that are crucial for systemic infection, particle formation, and insect transmission. *J. Virol.* 72:10050-10057.
- Ohnesorge, S., and Bejarano, E. R. 2009. Begomovirus coat protein interacts with a small heat shock protein of its transmission vector (*Bemisia tabaci*). *Insect Mol. Biol.* 18:693-703.
- Ohnishi, J., Kitamura, T., Terami, F., and Honda, K. 2009. A selective barrier in the midgut epithelial cell membrane of the nonvector whitefly *Trialeurodes vaporariorum* to *Tomato yellow leaf curl virus* uptake. *J. Gen. Plant Pathol.* 75:131-139.
- Padidam, M., Beachy, R. N., and Fauquet, C. M. 1995. Tomato leaf curl geminivirus from India has a bipartite genome and coat protein is not essential for infectivity. *J. Gen. Virol.* 76:25-35.
- Pandey, P., Choudhury, N. R., and Mukherjee, S. K. 2009. A geminiviral amplicon (VA) derived from tomato leaf curl virus (ToLCV) can replicate in a wide variety of plant species and also acts as a VIGS vector. *Virol. J.* 6: 152.
- Pederson, K. J. 1991. Invited review: Structure and composition of basement membranes and other basal lamina matrix systems in selected invertebrates. *Acta Zool.* 72:181-201.
- Peiffer, M. L., Gildow, F. E., and Gray, S. M. 1997. Two distinct mechanisms regulate luteovirus transmission efficiency and specificity at the aphid salivary gland. *J. Gen. Virol.* 78:495-503.
- Rana, V. S., Popli, S., Saurav, G. K., Raina, H. S., Chaubey, R., Ramamurthy, V. V., and Rajagopal, R. 2016. A *Bemisia tabaci* midgut protein interacts

- with begomoviruses and plays a role in virus transmission. *Cell. Microbiol.* 18:663-678.
- Rana, V. S., Singh, S. T., Priya, N. G., Kumar, J., and Rajagopal, R. 2012. Arsenophonus GroEL interacts with CLCuV and is localized in midgut and salivary gland of whitefly *B. tabaci*. *PLoS One* 7:e42168.
- Rojas, M. R., Hagen, C., Lucas, W. J., and Gilbertson, R. L. 2005. Exploiting chinks in the plant's armor: Evolution and emergence of geminiviruses. *Annu. Rev. Phytopathol.* 43:361-394.
- Rubinstein, G., and Czosnek, H. 1997. Long-term association of tomato yellow leaf curl virus with its whitefly vector *Bemisia tabaci*: Effect on the insect transmission capacity, longevity and fecundity. *J. Gen. Virol.* 78:2683-2689.
- Saurav, G. K., Rana, V. S., Popli, S., Daimei, G., and Rajagopal, R. 2019. A thioredoxin-like protein of *Bemisia tabaci* interacts with coat protein of begomoviruses. *Virus Genes* 55:356-367.
- Shieh, M. T., Wudunn, D., Montgomery, R. I., Esko, J. D., and Spear, P. G. 1992. Cell-surface receptors for Herpes-simplex virus are heparan-sulfate proteoglycans. *J. Cell Biol.* 116:1273-1281.
- Singh, S. T., Priya, N. G., Kumar, J., Rana, V. S., Ellango, R., Joshi, A., Priyadarshini, G., Asokan, R., and Rajagopal, R. 2012. Diversity and phylogenetic analysis of endosymbiotic bacteria from field caught *Bemisia tabaci* from different locations of North India based on 16S rDNA library screening. *Infect. Genet. Evol.* 12:411-419.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. 2013. MEGA 6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30:2725-2729.
- Van den Heuvel, J. F., Bruyère, A., Hogenhout, S. A., Ziegler-Graff, V., Brault, V., Verbeek, M., Van der Wilk, F., and Richards, K. 1997. The N-terminal region of the luteovirus readthrough domain determines virus binding to *Buchnera* GroEL and is essential for virus persistence in the aphid. *J. Virol.* 71:7258-7265.
- Wang, K. S., Kuhn, R. J., Strauss, E. G., Ou, S., and Strauss, J. H. 1992. High-affinity laminin receptor is a receptor for sindbis virus in mammalian cells. *J. Virol.* 66:4992-5001.
- Wang, Z. Z., Shi, M., Huang, Y. C., Wang, X. W., Stanley, D., and Chen, X. X. 2016. A peptidoglycan recognition protein acts in whitefly (*Bemisia tabaci*) immunity and involves in *Begomovirus* acquisition. *Sci. Rep.* 6:37806.
- Zerbini, F. M., Bridson, R. W., Idris, A., Martin, D. P., Moriones, E., Navas-Castillo, J., Rivera-Bustamante, R., Roumagnac, P., Varsani, A., and ICTV Report Consortium. 2017. ICTV virus taxonomy profile: Geminiviridae. *J. Gen. Virol.* 98:131-133.