Role of C terminal of Rep helicase in Rolling Circle Replication of *Tomato leaf curl virus*
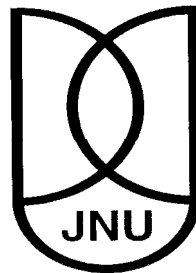
Thesis submitted to

Jawaharlal Nehru University, New Delhi

For the award of the degree of

DOCTOR OF PHILOSOPHY

JNU

RAJRANI RUHEL

School of Life Sciences
Jawaharlal Nehru University
New Delhi – 110 067

2016

SCHOOL OF LIFE SCIENCES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI -110067

Date: 19th July, 2016

CERTIFICATE

This is to certify that the work embodied in this thesis entitled **"Role of C terminal of Rep helicase in Rolling Circle Replication of *Tomato leaf curl virus*"** has been carried out at School of Life Sciences, Jawaharlal Nehru University, New Delhi, India. This work is original and has not been submitted so far in part or in full, for the award of any degree or diploma by any university.

Rajrani Ruhel

Student

Prof. Supriya Chakraborty

Supervisor

School of Life Sciences, JNU

SUPRIYA CHAKRABORTY, Ph.D.
Professor
School of Life Sciences
Jawaharlal Nehru University
New Delhi - 110067

Dean

School of Life Sciences, JNU

कार्यकारी डीन/Acting Dean
जीवन विज्ञान संस्थान/School of Life Sciences
जवाहरलाल नेहरू विश्वविद्यालय
Jawaharlal Nehru University
नई दिल्ली/New Delhi-110067

Dedication

To the glorious love of my father and my mother,

which I could never return back

Table of Contents

Acknowledgement

First of all, I would like to thank God and the support of my wonderful family in accomplishment of my goal of a PhD degree. The constant love, affection, patience and understanding from my Mama, Papa and my in-laws have been an incredible support for me throughout. I am indebted to my brother Raman and my husband Ravi for being there for me always. I love you all and I hope that you realize that I could not do any of this without you all.

*With due regards, I would like to express my sincere gratitude to my supervisor **Prof. Supriya Chakraborty**, for his kind support, guidance and affection, and for believing in me. I thank him for providing me moral support, for sharing my personal concerns and encouraging me throughout to do my best effort. This thesis could not have been written without his keen observation and analytical approach.*

I was lucky to study in a great working environment created by the staff, faculty and students of the School of life Sciences, Jawaharlal Nehru University. I would like to thank to the present dean, Prof. B. C. Tripathy and former deans for providing me the necessary facilities for carrying out my PhD work. I also thank the faculties of SLS who taught me in Pre-PhD course work and created a vibrant intellectual atmosphere.

I heartily thank the non-teaching staff of the school, especially Meenu ma'am, Talwar Ji, Shiny ma'am, Sunita ma'am and the administrative officer Dr. Sajjan Singh and others for the smooth functioning of the school office and for making official work simple for students.

I greatly acknowledge the technical staffs of the central instrumentation facility (CIF) at SLS, Mr. S.K. Mishra, Mr. Rajendra Meena, Mr. Amarchand, Mr. Suresh, Mr. Prankaj and other allied temporary staffs for ensuring the smooth functioning of the various instruments and other central facilities. I would also like to thank the Dr Manish and Dr. Gajendra at the AIRF, JNU for their help in circular dichorism (CD) studies.

It would not have been possible to pursue this PhD without the help and support of the people around me. I appreciate the help and support of my senior colleagues,

Abbreviations

A	absorbance
aa	amino acid
AAA+	ATPase associated with various cellular activities
ATP	adenosine triphosphate
bp	base pair
C	degree celsius
CD	circular dichorism
cm	centrimeter
CP	capsid protein or coat protein
cpm	count per minutes
CR	common region
CSR	complementary synthesis replication
CTAB	cetyl trimethyl ammonium bromide
Da	dalton
DF	dilution factor
DNA	deoxyribo-nucleic acid
dNTPs	deoxyribonucleotide tri phosphates
dpi	days post inoculation
ds	double stranded
EDTA	ethylenediamine tetra acetic acid
fmole	femtomole
g	gram
h	hour
IPTG	isopropyl-β-Dithiogalactopyranoside
IR	intergenic region
kb	kilo base
Kd	dissociation constant
kDa	kilo dalton
LB	Luria bertain
m	meter
MBP	maltose binding protein
mg	milligram
min	minutes
ml	milliliter
mM	millimolar
mm	millimeter
MP	movement protein
MP	movement protein
MW	molecular weight
ng	nanogram
NIG	NSP interacting GTPase
NLS	nucleus localization signal

nm	nanometer
nmol	nano mole
NMR	nuclear magnetic resonance
NSP	nuclear shuttle protein
nt	nucleotide
NTP	nucleotide triphosphate
OD	optical density
ORF	open reading frame
Ori	origin of replication
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
Pi	inorganic phosphate
pI	isoelectric point
RBR	retinoblastoma related protein
RCR	rolling circle replication
RDR	recombination dependent replication
REn	replication enhancer protein
Rep	replication initiator protein
RF	replicative form
RNA	ribonucleic acid
rpm	revolution per minute
SDM	site directed mutagenesis
SDW	single distilled water
sec	second
SF	superfamily of helicase
ss	single stranded
TAE	Tris Acetate EDTA
TBE	Tris Borate EDTA
TLC	thin layer chromatography
ToLC	*Tomato leaf curl virus*
ToLCGV	*Tomato leaf curl Gujarat virus*
ToLCNDV	*Tomato leaf curl New Delhi virus*
ToLCV	Tomato leaf curl disease
TrAP	transcription activator protein
TYLCV	*Tomato yellow leaf curl virus*
U	unit
μCi	micro curie
μg	microgram
μl	microlitre
w/v	weight/volume
wt	wild type

List of Figures

List of Tables

INTRODUCTION

Viruses are obligate parasites and infect almost all kingdoms of life which includes archae, algae, bacteria, fungi, protozoa, plants, vertebrates and invertebrates. More than 500 of them are known to infect many cultivable and non-cultivable plants worldwide (Kings *et al*., 2011). Crop plants most often encounter different biotic as well abiotic challenges in field. Among the biotic factors, viruses are known to cause substantial damage by causing devastated diseased condition in plants which affect the crop′s yield and productivity. Viruses are known to cause an increased predisposition to attack by other pathogens. In developing countries like India where agriculture forms an important sector of country′s economy and a large population relies on agriculture for their income and sustenance, the enormous crop loss caused by viruses had adversely affected the nation's economy as well.

The first report on isolation and characterization of the plant virus dates back to the year 1899, when the Dutch scientist, M. W. Beijerinck in his paper entitled "*Ueber ein Contagium vivum fluidum als Ursache der Fleckenkrankheit der Tabaksblitter*" demonstrated that the virulent soluble agent, *contagium vivum fluidum,* (contagious living fluid) being the causative agent of the disease in tobacco (Beijerinck, 1889; Johnson, 1942). He reported that this infectious agent is smaller than bacterium, could not be cultured and multiplies only within living plants. This causative agent is now referred to as *Tobacco mosaic virus* (TMV) that infects tobacco plant and produce mosaic symptom. Among many plant infecting viruses, begomoviruses like *Tomato yellow leaf curl virus* (TYLCV) and *African cassava mosaic virus* (ACMV) of the *Geminiviridae* family are known as major reason for the losses to crop production and have been considered to be among the top ten devastating plant viruses of the world (Rybicki, 2015).

Geminiviruses are circular ssDNA containing plant viruses that are encapsidated within the geminate virion (Zhang *et al*., 2001). The genome of the geminiviruses can be either monopartite or bipartite, comprising of either one or two ssDNA molecules, respectively. Thus, the encapsidated virion contains a single molecule of circular ssDNA genome in case of monopartite while for bipartite two virions each containing either of the ssDNA molecules (DNA-A or DNA-B) exist. The family *Geminiviridae* has been classified into seven genera based on the insect vector used for their transmission, their genome organization and their host range (Adams *et al*., 2013;

Varsani *et al*., 2014). Geminivirus infection causes modifications in the expression level of a several number of plant genes involved in regulation of biological processes like cell cycle, nucleotide metabolism, DNA repair and recombination (Ascencio-Ibanez *et al*., 2008). Their replication occurs within the infected plant cells through a dsDNA intermediate via rolling circle mode of replication (Lazarowitz and Shepherd, 1992). However, characterization of various DNA intermediates produced during replication indicates that geminivirus multiplication adopts recombination dependent mode of replication as well (Jeske *et al*., 2001).

The whitefly transmitted *Begomovirus* constitutes the largest genus of the family *Geminiviridae* (Kings *et al*., 2011). Tomato leaf curl disease (ToLCD) in plants results from the infection by begomoviruses. They lead to a debilitating state in cultivated tomato (*S. lycopersicum*) and results in extensive crop losses around the world. *Tomato leaf curl New Delhi virus* (ToLCNDV) and *Tomato leaf curl Gujarat virus* (ToLCGV) are two prevalent begomovirus species responsible for severe ToLCD in India (Chakraborty, 2008; Chakraborty *et al*., 2003; Padidam *et al*., 1995). ToLCNDV is a bipartite begomovirus. It contains DNA-A molecule which encodes for two (AV1 and AV2) and four ORFs (AC1, AC2, AC3 and AC4) from the virion-sense and complementary-sense strand, respectively. AV1 and AV2 ORFs code for the capsid protein or coat protein (CP) and the pre-coat protein, respectively, while AC1, AC2 and AC3 functions as replication-associated protein (Rep), the transcription activator protein (TrAP) and the replication enhancer protein (REn), respectively. AC4 is implicated in symptom production. The DNA-B molecule codes for BC1 and BV1 which function as movement protein (MP) and nuclear shuttle protein (NSP) respectively.

ToLCNDV Rep is a 361 amino acid protein with N-terminus (1-120 residues) having endonuclease, ligase and sequence specific DNA binding activity (Chatterji *et al*., 2000; Fontes *et al*., 1992; Orozco *et al*., 1997) while the ATPase and helicase activity is contributed by 120-361 amino acid residues in the C-terminal (Choudhury *et al*., 2006; Clerot and Bernardi, 2006; Desbiez *et al*., 1995). Amino acid residues 120-180 of Rep form the oligomerization domain which also is essential for interacting with many host factors such as proliferating cell nuclear antigen (PCNA), retinoblastoma related protein (RBR) and geminivirus Rep interacting kinase (GRIK) (Rizvi *et al*., 2015). Rep is multifunctional in nature and possesses modular functions, but the

coordination between these different modules is yet to be dissected. The NMR structure of the N-terminal of Rep protein is available and the roles of many of the N-terminal amino acids are also known (Campos-Olivas *et al*., 2002). Rep protein which is indispensable for viral replication belongs to AAA+ (ATPase associated with various cellular activities) family of ATPases and is grouped within the superfamily 3 (SF3) family of helicases (Gorbalenya and Koonin, 1993). Four conserved SF3 motifs of Rep helicase are namely A, B, B′, and C, all of which are located in approximately 100 amino acid residues stretch of C-terminal of Rep protein. AAA+ ATPase domain of the geminiviral Rep is found to be much reduced and lacks structural element such as conserved arginine finger (Clerot and Bernardi, 2006). Motif A and Motif B serve as the nucleoside triphosphate (NTP)-binding pocket and metal ion coordination site respectively, thus are required for ATP hydrolysis. Motif C is essential for interacting with the PO_4 at the gamma position of ATP and an 'apical' water molecule. Motif B′ is involved in ssDNA-binding (Hickman and Dyda, 2005).

Studies on different amino acids of SF3 signature motifs of the Rep protein may prove useful in understanding the mechanism employed by these plant virus encoded helicase during geminivirus replication. Keeping in view the progress made in geminivirus research and the importance of Rep protein in geminivirus replication, the present research initiative was undertaken to characterize the C-terminal of Rep in order to find out how nucleotide binding and hydrolysis drive geminivirus Rep helicase.

Keeping in view the progress made and the information available on role of Rep in geminivirus replication, the present study was proposed with the following objectives:-

1. To determine the efficiency of DNA unwinding by Rep using a variety of NTPs.

2. To identify critical amino acid residues within B' motif and β- hairpin loop required for various biochemical activities involved during RCR.

3. To study *in vivo* replicational efficiency of the wild type and mutants of Rep.

REVIEW
OF
LITERATURE

2.1 Tomato and tomato leaf curl disease (ToLCD)

Solanum lycopersicum, commonly known as tomato has originated in the Latin America. The spread of its cultivation throughout the world is considered result of Spanish colonization of America. Tomato is a fruit of the flowering plant and is consumed as an edible fruit as well as used as vegetable. It has a high water content and low sugar, fat and cholesterol content. It contains ample of minerals like potassium, magnesium, calcium and phosphorus. It is also a rich source of vitamin A, vitamin C and antioxidants like lycopene and glutathione. Worldwide, India is the the second major tomato producing country in the world after China. **Table 2.1** below highlights the major tomato cultivating countries in the world (FAOSTAT, 2015; http://faostat3.fao.org/home/E).

Table 2.1 Statistics of major tomato cultivating countries of the world for the year 2013.

Country	Area harvested (Ha)	Yield Hg/Ha	Production tonnes
China, mainland	984603	514565	50664255
India	880000	207125	18227000
United States of America	149977	838432	12574550
Turkey	311000	380064	11820000
Egypt	212946	400750	8533803

In India, the major tomato producers are Karnataka, Andhra Pradesh and Madhya Pradesh (NHB database, 2014; http://nhb.gov.in/default.aspx) (**Figure 2.1**).
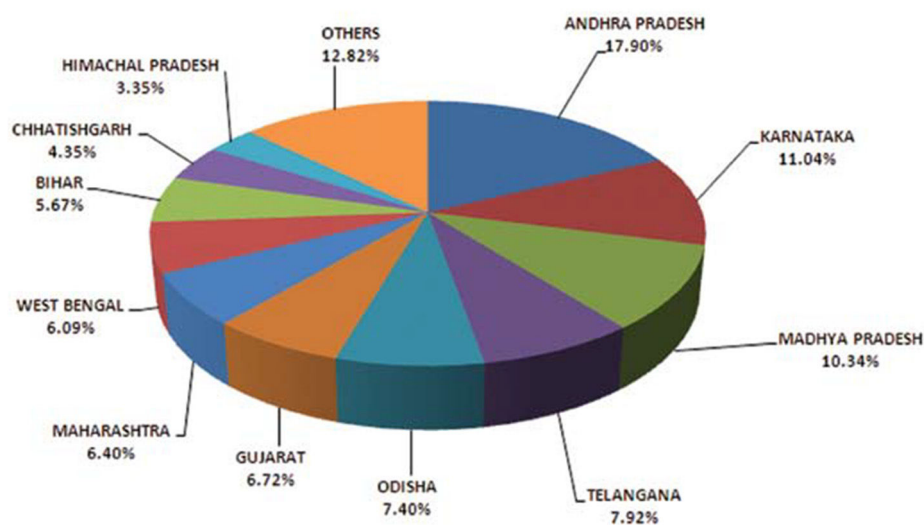
Figure 2.1 Pie chart showing the leading tomato producing states in India (National Horticultural Board-NHB database, 2014).

Tomatoes are cultivated in more than 100 countries all over the world. They are subjected to various biotic and abiotic stresses which result in drastic reduction in its production in subtropical and tropical regions across the world. Tomato crop is attacked by fungal, bacterial and viral pathogens. Diseases of tomato include early-blight, late-blight, powdery mildew, bacterial leaf spot, bacterial canker, fusarium wilt, tomato spotted wilt, tomato big bud, tomato bunch top, tomato mosaic, tomato leaf curl etc. But the major serious threat are viuses, mainly from the begomoviruses which are reported to cause upto 100% crop losses in some cases (Saikia, 1989).

Tomato production have been severly hampered by mainly two viruses, namely *Tomato yellow leaf curl virus* (TYLCV) and *Tomato leaf curl virus* (TLCV). Depending upon the kind of symptoms found on tomato plants, disease caused by begomovirus is called as either 'leaf curl' or 'yellow leaf curl' disease. The viruses causing these diseases are distinguished by their genomic analysis and are collectively known as tomato leaf curl viruses (ToLCVs). ToLCVs are more diverse than TYLCVs. Tomato leaf curl disease are reported from India, Pakistan, Bangladesh, Australia, Taiwan, Oman and Panama while tomato yellow leaf curl disease incidences are known from Asia, America, Africa and European countries. ToLCV are known to cause great damage to tomato cultivation and production in the Middle East (Shirazi *et al.*, 2014), Africa (Bock *et al.*, 1981; Patil and Fauquet, 2009; Stanley and Gay, 1983) as well as in the Western Hemisphere (Accotto et al., 2000; Polston and Anderson, 1997). The symptoms of a typical leaf curl disease includes reduced leaf size, upward or downward leaf curling, stunted growth and reduced fruiting (**Figure 2.2**). ToLCV infection at an early stage of plant development may lead to more severe symptoms, partial/complete sterility and results in much higher yield loss. Apart from tomato, ToLCVs can also infect other member of the *Solanaceae* family such as *Solanum tuberosum, Capsicum annuum, Nicotiana tabacum, Nicotiana benthamiana*. It can also infect weed species and other non-solanaceae species like *Vigna unguiculata and Luffa cylindrica* (Ramappa *et al.*, 1998).

ToLCV has emerged a serious problem for tomato cultivating regions in the Indian subcontinent (Chakraborty, 2008; Chakraborty *et al.*, 2003). The ToLCV species are diverse, transmitted by the whitefly-*Bemisia tabaci* Genn. (Harrison A. D., 1991; Padidam *et al.*, 1995) and are spread in all the parts of India. The first report of

ToLCD in India was by Vasudeva (1948) while the first sequence for ToLCV was made available by Srivastava et al. (1995).

Figure 2.2 Leaf curl diseases induced by begomoviruses infecting various host plants such as (a) tomato, (b) chilli, (c) maize, (d) cabbage, (e) cotton and (f) tobacco.

Table 2.2 Begomovirus species causing ToLCD on tomatoes reported in India.

Monopartite species	Bipartite species
Tomato leaf curl Kerala virus (ToLCKeV)	*Tomato leaf curl New Delhi virus* (ToLCNDV)
*Croton yellow vein mosaic virus (*CYVMV)	*Tomato leaf curl Palampur virus* (ToLCPalV)
Tomato leaf curl Rajasthan virus (ToLCRaV)	
Tomato leaf curl Pune virus (ToLCPV)	
Tomato leaf curl Lucknow virus (ToLCLuV)	
Tomato leaf curl Bangalore virus (ToLCBaV)	
Tomato leaf curl Karnataka virus(ToLCKaV)	
Tomato leaf curl Joydebpur virus (ToLCJoV)	
Tomato leaf curl Gujarat virus (ToLCGuV)	
Cotton *leaf curl Burewala virus (*CLCuBuV*)*	
Papaya leaf curl virus(PaLCuV)	
*Tomato leaf curl Pakistan virus (*ToLCPkV)	
*Tomato leaf curl Bangladesh virus (*ToLBdV)	

In India, as many as 15 begomoviruses infecting tomatoes has been identified which cause leaf curl disease in them (**Table 2.2**). It comprises of both monopartite as well as bipartite viruses (Jyothsna *et al*., 2013b; Kanakala *et al.*, 2013; Tiwari *et al.*, 2013). Recombinations are known to occur among the Indian ToLCV begomoviruses and mixed infections are also found to be frequent in occurrence (Chatchawankanphanich and Maxwell, 2002; Kirthi and Savithri, 2003; Muniyappa *et al.*, 2000).Certain isolates of ToLCV are known to cause mild symptoms and also there are reports showing its presence in cucurbitaceous crops (Phaneendra *et al*., 2012). ToLCJoV is a typical monopartite virus where DNA-A alone is highly infectious (Tiwari *et al.*, 2013). Satellite molecules are found to be associated with both the monopartite as well as bipartite viruses. Association of a helper virus with different betasatellites in different locations is known. For instance monopartite ToLCBaV is known to be associated with ToLCBaB in Karnataka and with TYLVTHB in Maharashtra (Jyothsna *et al.*, 2013a). ToLCNDV and ToLCPaV are the two bipartite Indian ToLCV species known. ToLCNDV distribution is reported all over the country and has a wider host range. This devastating pathogen is found to be associated with many betasatellites (Jyothsna *et al*., 2013b).

2.2 Geminivirus

The family *Geminiviridae* comprises of more than 300 species (215 approved and 103 tentative species) and constitutes one of the largest families of plant viruses (Kings *et al.*, 2011). These plant viruses affect a wide range of essential and economically important crops such as maize, beans, chilli, tomato, cassava, legumes and even cash crop like cotton (Mansoor *et al.*, 2006; Mansoor *et al.*, 2003). Geminivirus infection leads to reprogramming of the host cellular machinery to create an environment suitable for viral infection and its establishment within the host (Hanley-Bowdoin *et al*., 2013). Because of its high dependence on the host biosynthetic machinery, geminivirus serves as a suitable model system for study of host DNA replication and gene expression. Moreover, the circular genome which is single stranded DNA molecule of small size of approximately (2.7 - 3 kb) makes them an ideal choice for use as vectors for expression of foreign gene in plants.

2.2.1 General features and basic genome organization of *Geminiviridae* family

The viral particles associated with the Maize streak disease and Beety curly top disease showed a distinctive twinned quasi-isomeric morphology (Bock *et al.*, 1974; Mumford, 1974). The name geminivirus was coined from the latin word "Gemini" meaning "twins" as they contain twined icosahedral capsid (Goodman, 1977). The virion particles are ~18 nm in diameter and 30 nm long. The virion constitutes of two incomplete icosahedral capsid (T=1) joined together to generate a structure with 22 pentameric capsomers and 110 identical protein subunits (Zhang *et al.*, 2001). Subsequently, the studies on infectious clones of *African cassava mosaic virus* demonstrated bipartite nature of its genome and that both the DNA molecules are required for its effective host infectivity (Stanley and Gay, 1983).

The small genome of geminivirus utilize bidirectional mode of transcription and have overlapping genes in different frames for its efficient usage. The genome of geminiviruses can be either monopartite, if consists of only one DNA molecule, or bipartite if it contains two molecules (Kings *et al.*, 2011). DNA-A and DNA-B molecules share no sequence similarity except for a highly conserved region (~200 nt) known as common region. Different ORFs diverge in opposite direction from this conserved common region (Lazarowitz and Shepherd, 1992). The 5′ intergenic region contains the viral origin of replication and RNA polymerase II promoters for the expression of Rep gene and virion sense genes (AV1 and AV2). This region also comprises of cis elements that mediate viral replication (Hanley-Bowdoin *et al.*, 2000). This region consists of inverted repeat sequences that forms a characteristic hairpin secondary structure comprising of a GC rich stem while an AT rich loop region. The AT rich loop region contains an invariant nonamer sequence 5'-TAATATT↓AC-3'. Within this invariant nonancleotide resides T7-A8 site that is in replication needed for the cleavage and joining of the viral DNA (Laufs *et al.*, 1995).

2.2.2 Classification of the *Geminiviridae* family

On the basis of the insect vector, host range, and genome structure the *Geminiviridae* family is divided into seven genera (Adams *et al.*, 2013; Varsani *et al.,* 2014). The genome organisation of the genera belonging to family geminiviridae is shown in **Figure 2.3**. Characteristic features of the well-studied genera are briefly mentioned in the table below.

Table 2.3 Classification of the *Geminiviridae* family.

Genus	Genome	Vector	Type species	Host Range
Mastrevirus	Monopartite	Leafhopper	*Maize streak virus*	Mono/ Dicots
Curtovirus	Monopartite	Leafhopper	*Beet curly top virus*	Dicots
Topocuvirus	Monopartite	Treehopper	*Tomato pseudo curly top virus*	Dicots
Begomovirus	Mono or Bipartite	Whitefly	*Bean golden mosaic virus*	Dicots
Becurtovirus	Monopartite	Leafhopper	*Beet curly top Iran virus*	Dicots
Turncurtovirus	Monopartite	Leafhopper	*Turnip curly top virus*	Dicots
Eragrovirus	Monopartite	Unknown vector	*Eragrostis curvula streak virus*	Dicots

2.2.2.1 Genus Mastrevirus

Mastreviruses (type species: *Maize streak virus* - MSV) are leafhopper transmitted viruses. They have monopartite genome. They infect mainly monocotyledonous plant species. However, few mastreviruses *i.e. Tobacco yellow dwarf virus* (TYDV) and *Bean yellow dwarf virus* (BeYDV) are reported to infect dicotyledonous plants (Liu *et al.*, 1998). But both TYDV and BeYDV are found to be serologically not related to *Wheat dwarf virus* (WDV) and *Digitaria streak virus* (DSV) that are monocotyledonous infecting mastreviruses (Liu *et al.*, 1998; Padidam *et al.*, 1995). It is known to be the only genera that contain introns and utilizes transcript splicing. Unlike other genera, mastreviruses regulate expression of virion and complementary sense genes by posttranscriptional splicing event (Rojas *et al.*, 2005). There are total four genes on the virion- and complementary- sense. DNA strands that are separated by two noncoding regions designated as the long intergenic region (LIR) and the short intergenic region (SIR) (Mullineaux *et al.*, 1984). A small complementary-sense DNA is found annealed to the ssDNA genome of *Chloris striate mosaic virus* (CSMV), *Digitaria streak virus* (DSV), MSV, *Tobacco yellow dwarf virus* (TYDV) and *wheat dwarf virus* (WDV) which is also encapsidated within the virion particle. This small DNA is believed to prime the complementary sense DNA synthesis to produce the dsDNA replicative form and is associated with the SIR (Donson *et al.*, 1984).

Figure 2.3 Genome organizations of the different genera of the family *Geminiviridae*. Arrows represent open reading frames. LIR, long intergenic region; SIR, short intergenic region; CR, common region; rep, replication-associated protein; ren, replication enhancer; trap, transactivator protein; ss, silencing suppressor; tgs, transcriptional gene silencing; sd, symptom determinant; cp, capsid protein; mp, movement protein; reg, regulatory gene; NSP, nuclear shuttle protein. Adapted from (Varsani *et al*., 2014).

2.2.2.2 Genus Curtovirus

Like mastreviruses, curtoviruses (type species: *Beet curly top virus* - BCTV) are monopartite and leafhoppers-transmitted viruses but they infect dicot plants only. They infect a wide range of dicot plants in both the eastern as well as the western hemisphere. A typical curtovirus genome codes for seven ORFs and contains a single intergenic region (IR) within which lies the origin of replication (Hormuzdi and Bisaro, 1993; Hormuzdi and Bisaro, 1995). The complementary sense and virion strand encodes four (C1, C2, C3, C4) and three (V1, V2, V3) ORFs, respectively. V1 codes for coat protein that is also required for insect transmission and V2 is movement protein (MP). V2 also regulates the ssDNA and dsDNA accumulation in

10

infected cells. The complementary-sense strand codes for C1 and C3 ORFs which is required for initiation and enhancement of viral replication, respectively. C2 protein acts as pathogenicity factor while C4 protein are involved in cell cycle regulation.

2.2.2.3 Genus Topocuvirus

Topocuviruses possess a single genome component. *Tomato pseudo-curly top virus* (TPCTV) is the type species of this genus and is also the only member. They are transmitted by the treehopper, *Micrutalis malleifera* and infect only dicotyledonous plants. They have been reported only in the America (New World, NW). Analysis of TPCTV genome has shown common feature of mastrevirus and begomovirus and are believed to have emerged as a result of recombination event between a mastrevirus and a begomovirus (Briddon *et al.*, 1996).

2.2.2.4 Genus Begomovirus

The members of the genus begomovirus (type species: *Bean golden mosaic virus* - BGMV) are diverse, economically important and have a wide geographical distribution. They are transmitted by the whitefly and infect dicotyledonous plants. It forms the largest genus of family *Geminiviridae* and contains more than 200 species (Kings *et al.*, 2011). The symptoms associated with begomovirus infection include stunted plant growth, mosaic, vein yellowing, leaf-curling and leaf yellowing. Unlike other genera of the family, members of this genus can have either monopartite or bipartite genome. Bipartite begomoviruses consist of about equal size (~2.5- 2.6 kb) two DNA components, known as DNA-A and DNA-B. Single DNA component of monopartite begomoviruses is homologous to the DNA-A molecule of bipartite viruses (Kings *et al.*, 2011; Rojas *et al.*, 2005).

The genome organisation of begomovirus is discussed in detail in section 2.4. Based on phylogenetic analysis and their genome organisation, begomoviruses have been broadly catergorised into two groups: the old world viruses and the new world viruses. While the bipartite begomoviruses are widely distributed and found in both NW and OW, the monopartite begomoviruses are native to the OW (Nawaz-ul-Rehman *et al.*, 2009). NW begomoviruses differ from OW viruses in their genetic composition, as they possess another gene AV2 in bipartite and V2 in monopartite begomoviruses which codes for pre-coat protein.

2.2.2.5 Genus Becurtovirus

Recently, three new genera are included in the family *Geminiviridae,* which are *Becurtovirus*, *Turncurtovirus* and *Eragrovirus* (Adams *et al.*, 2013; Varsani *et al.,* 2014).

Beet curly top Iran virus (BCTIRV) is the type species of the genus *Becurtovirus* and is transmitted by leaf hopper (Yazdi *et al.*, 2008). This genus comprises of only two species namely, *Beet curly top Iran virus* and *Spinach curly top Arizona virus*. Both contains an unusual ''TAAGATTCC'' nonanucleotide sequence at the origin of virion-sense strand replication instead of the ''TAATATTAC'' nonanucleotide found in most genera of geminiviruses (Yazdi *et al.*, 2008). BCTIRV showed similarity to curtovirus in terms of the biological symptoms and wide host range (Varsani *et al.*, 2014). The BCTIRV coat protein shows high amino acid similarity to that of curtoviruses, which is the reason of its transmission by the same leafhopper vector. However, surprisingly, BCTIRV genome shows little nucleotide identity to curtoviruses. The genome organisation shows greater resemblance to that of mastreviruses. Like mastreviruses, BCTIRV genome contains two intergenic regions, SIR and LIR. The BCTIRV Rep also shows more phylogenetic correlation to that of *Mastrevirus*. The nucleotide analysis suggests their evolution from the recombination event between *Curtovirus* and *Mastrevirus*.

2.2.2.6 Genus Eragrovirus

Eragrovirus genus contains only one species member, *Eragrostis curvula streak virus* (ECSV) which is also the type species (Varsani *et al.*, 2014). ECSV infects monocotyledonous plants belonging to *Eragrostis curvula* species in South Africa (Varsani *et al.*, 2009). Six isolates of ECSVs are known. Like becurtovirus, they also bear an unusual ''TAAGATTCC'' nonanucleotide sequence at their origin of virion-sense strand replication.

2.2.2.7 Genus Turncurtovirus

The genus *Turncurtovirus* constitutes only *Turnip curly top virus* (TCTV) species (Varsani *et al.*, 2014). However, as many as 20 isolates of TCTV are isolated from *Brassica rapa* and *Raphanus sativus*. They are transmitted by leafhopper. Like mastreviruses, begomoviruses, curtoviruses and topocuviruses, they possess the same

''TAATATTAC'' nonanucleotide found in the virion-strand origins of replication. TCTV Rep shows the highest similarity to curtoviruses.

2.3 Geminivirus infection cycle

Geminivirus infection occurs in terminally differentiated plant cells where they induce the host gene expression and modify the plant cell cycle machinery to complete its life cycle inside the host (Hanley-Bowdoin *et al*., 2013; Nagar *et al.*, 1995). As a consequence of whitefly feeding upon plants, the viral genomic ssDNA (+) molecule is introduced into the plant cells (Horns and Jeske, 1991). The viral replication takes place in the nuclei of the infected plant cell via rolling circle replication mechanism (Saunders *et al.*, 1991; Stenger *et al*., 1991). The replication proceeds through an intermediate dsDNA molecule which is also referred to as "replicative form" (Kammann *et al*., 1991; Saunders *et al.*, 1992). AC1 (replication initiator protein) is the only viral ORF that is indispensable for replication. It induces the accumulation of PCNA in the infected differentiated plant cells (Nagar *et al*., 1995). Various host factors/proteins are involved to accomplish the viral life cycle inside the plant.

2.3.1 Rolling circle replication

Various extrachromosomal elements such as bacterial plasmid, ssDNA of bacteriophage and RNA viriod replicate their genome via rolling circle replication (Branch *et al.*, 1988; Gros *et al*., 1987; Koepsel *et al.*, 1985). The RCR mode of replication is initiated after nick at specific site "Ori", origin of replication. This generates 3′ - OH which serves as primer during DNA synthesis. The elongation of DNA synthesis in phage produces multiple single stranded linear copies of DNA which is called as concatomer. However, in replication of ssDNA phage genome and plasmids, after each round of replication a unit length single stranded molecule is formed which after cleavage is ligated to produce a circularized DNA molecule.

Number of evidences supported the idea of geminivirus adopting rolling circle mechanism for its replication. Various DNA forms that are formed upon *African cassava mosaic virus* infection clearly indicated rolling circle mechanism as the mode of geminivirus replication (Saunders *et al.*, 1991). Apart from that homology studies on the geminiviral Rep and the replication initiator proteins of the bacteriophages and eubacterial plasmid families had shown presence of the signature motifs namely motif

I, motif II and Motif III in the amino terminal half of Rep that aids in nicking and joining during the rolling circle replication (Ilyina and Koonin, 1992). Replication of geminivirus is best characterized using the members belonging to genus *Begomovirus* as depicted in **Figure 2.4**.
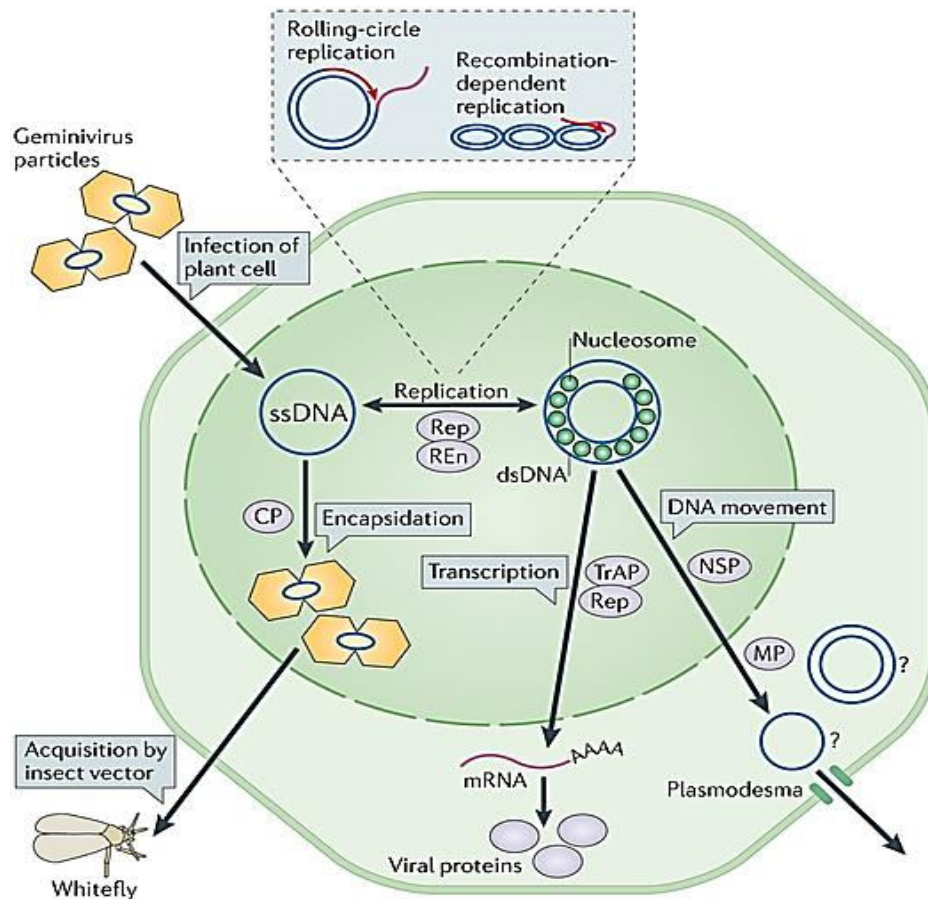
Figure 2.4 Geminivirus infection cycle inside the plant cell Adapted from Hanley-Bowdoin *et al.* (2013).

The geminivirus replication via rolling circle mechanism is completed in three stages:

- In the first stage viral ssDNA (+ strand) is converted to dsDNA intermediate replicative form with the help of host factors which is depicted in **Figure 2.5**. As a result of feeding by the insect vectors ssDNA genome is introduced into the plant cell. The viral genome enters the nucleus as part of ssDNA-CP complex. The entry to nucleus is mediated by nuclear localization signal present in coat protein (Guerra-Peraza *et al*., 2005). Within the host nucleus complementary strand synthesis occurs in which ssDNA molecule is converted into a covalently closed dsDNA replicative form (RF) by host DNA polymerase. This RF serves as

template for viral replication as well as its bidirectional transcription. In case of mastreviruses, an oligonucleotide complementary to small intergenic region (SIR) is found packed within the virion which acts as primer for DNA synthesis (Donson *et al.*, 1984). Instead in begomoviruses such a primer is generated within the large intergenic region by host RNA polymerase (Saunders *et al*., 1992). The dsDNA replicative form also interacts with cellular histone to form viral minichromosomes (Pilartz and Jeske, 1992). Viral ssDNA genome is more vulnerable than dsDNA and therefore more prone to damage. Recently, RAD51D, a paralog of RAD51 has been suggested to play a role during the complementary strand replication (CSR) of geminivirus to produce RF molecules (Richter *et al.*, 2016b). RAD51D is a key player in RAD51-independent single strand annealing (SSA) recombination pathway in somatic cells of *A. thaliana* and the possible roles
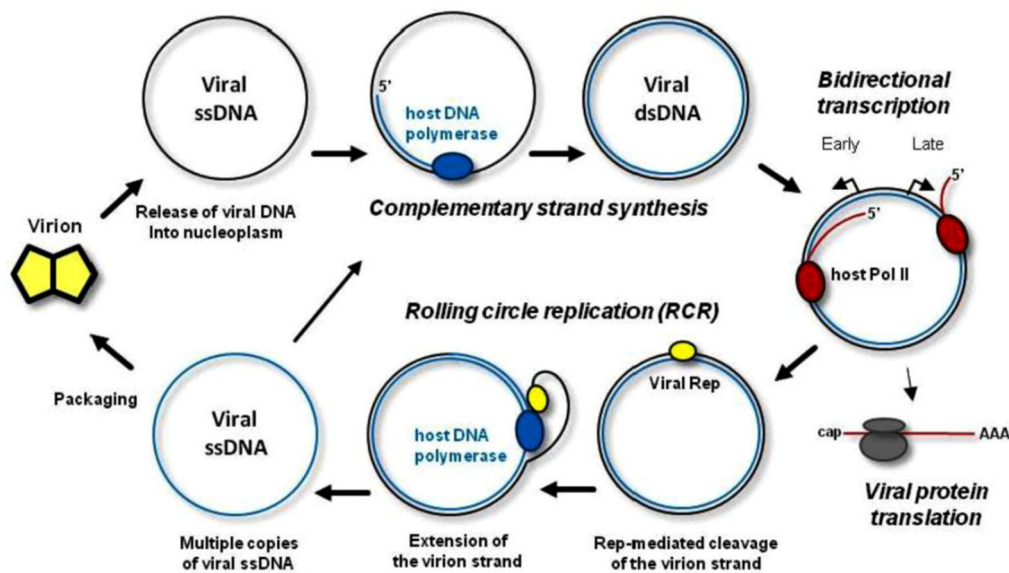
Figure 2.5 Model for Rolling circle replication during geminivirus DNA synthesis. During the complementary strand synthesis, the ssDNA is converted into dsDNA intermediate by host DNA polymerase (shown as blue oval). Rep protein is produced by translation by host polymerase II (shown in red oval) from this dsDNA. This dsDNA molecule is further nicked by Rep protein (shown in yellow oval) and further DNA is synthesized with the help of host DNA polymerase. (Adapted from Richter and Serra *et al.,* 2016).

of SSA in complementary strand synthesis of the geminiviral replication have been suggested (Serra *et al*., 2013). In *A. thaliana*, four translesion synthesis (TLS) polymerases have been reported namely, Polη, Polζ, Polκ and Rev1. Out of these, Polζ is believed to carry out complementary strand replication of geminiviral DNA

(Richter *et al*., 2016a). Supporting this, it has been found that TLS polymerases are constitutively expressed in differentiated plant cell where geminivirus replicates.

- During the second stage of RCR, more RF DNAs are generated from RF. In this step, Rep (AC1) produces a nick on the (+) strand of the RF molecule. This nick is created at a conserved, specific nonamer sequence (TAATATT↓AC) present in the loop region of the stem-loop structure present in the common region of the circular DNA molecule (Laufs *et al*., 1995). Various cis elements are identified in the common region that affects the process of replication **Figure 2.6**. Origin of replication on the TGMV (+)-strand possesses binding sites for two transcription factors, one at the TATA-box and another at G-box. These sites are not required for viral replication. In contrast to this, other two elements namely, AG- and CA-motifs are required for replication probably by binding to host factors. Following the cleavage, Rep protein recruits the host DNA polymerase which extends the 3'-end of the cleaved virion strand while Rep remains covalently bound to the 5' terminus via a phosphotyrosine linkage. Rep is also known to interact with several host protein during the course of replication. For example,
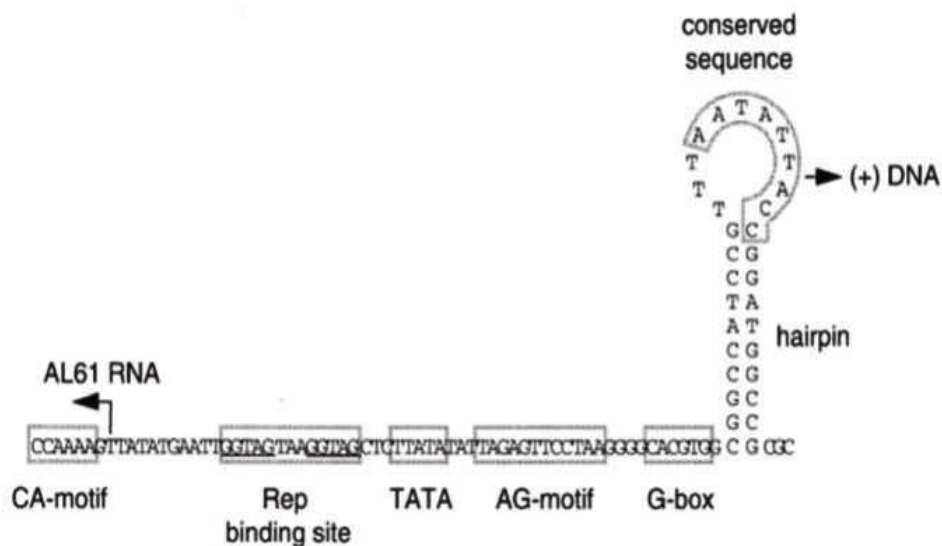
Figure 2.6 The Plus-strand origin of Replication and AC1 promoter in TGMV. The initiation site and direction of replication and transcription are marked (adapted from Hanley-Bowdoin *et al*., 2000).

it interacts with host RFC, RPA70, MCM and PCNA (Rizvi *et al*., 2015). Rep acts as ATPase and is believed to act as a replicative helicase (Choudhury *et al*., 2006; Clerot and Bernardi, 2006). Rep interacts with another viral

protein REn (encoded by AC3) to enhance replication (Pasumarthy *et al.*, 2010). Later during this step when the origin of replication is regenerated, another nick is introduced by Rep and then Rep is transferred to the new 5' terminus. Rep acts as ligase finally to produce a circular ssDNA molecule.

- During the third stage of RCR the newly synthesized ssDNA molecules are accumulated, encapsidated and new viral progenies are produced. During this final step, there is a shift from the production of dsDNA molecules to accumulation of ssDNA molecules. Various AC2 (TrAP) and AV2 mutants were found to result in reduction in ssDNA levels and a corresponding increase in level of dsDNA molecules (Hayes and Buck, 1989; Hormuzdi and Bisaro, 1993; Sunter *et al.*, 1990). Thus, these viral factors are believed to control this transition step during replication. Both these viral proteins are implicated in the inhibition of minus strand synthesis and as a result of which ssDNA molecules produced are diverted from the replication pool towards encapsidation and viral assembly.

Further the host's transport machinery is utilized and exploited to spread the virus throughout the plant (**Figure 2.7**). Viral infection is established throughout the plant with the help of the nuclear shuttle protein (NSP encoded by BV1) and the movement protein (MP encoded by) of virus. Viral movement within the host occurs at two different levels: a) short distance cell-to-cell movement through plasmodesmata and b) long distance movement to the distal parts of the plant which occurs through the vascular system. NSP binds to the newly synthesized ssDNA viral genome and transports them from nucleus to cytoplasm (Pascal *et al.*, 1994). NSP interacting GTPase (NIG) interacts with NSP and help in the release of the complex of NSP-DNA from the nuclear export machinery. Subsequently cell-to-cell movement of viral DNA is carried out with the help of movement protein. In case for monopartite viruses, movement of the viral DNA is carried out by coat protein (CP) and movement protein (MP). CP can bind to viral DNA and contains the NLS (Guerra-Peraza *et al.*, 2005). Movement of the viral genome in and out the nucleus is required for its replication and spread respectively, which occurs via recognition by the nuclear transport machinery components.
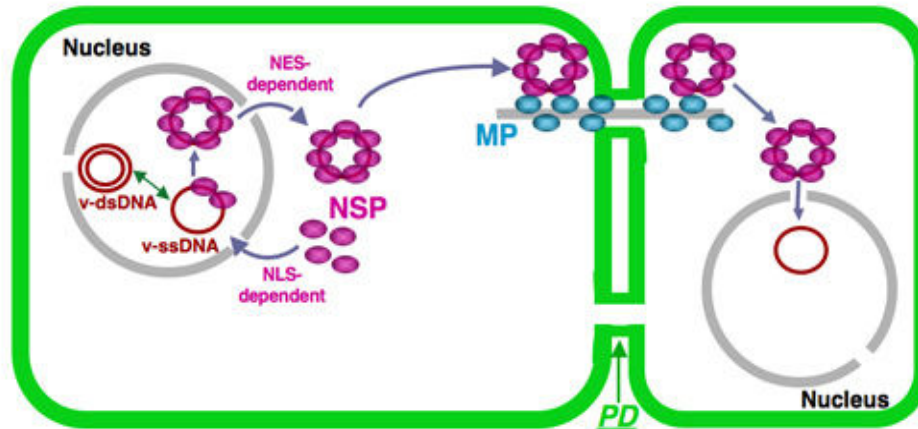
Figure 2.7 Model for role of NSP and MP in geminivirus movement. The movement of the viral DNA (shown in red) is assisted by NSP protein depicted as pink ovals which help in transport of the viral DNA from nucleus to cytoplasm while the movement of the viral DNA to other cell is carried out by MP depicted as blue ovals (http://www.plantpath.cornell.edu/labs/lazarowitz/Research.html).

CP docks the viral genome at the nuclear pore and is believed to enter into the nucleus interacting with the importin alpha component of nuclear import machinery. Host's Histone H3 interacts with both Viral NSP and MP suggesting its possible participation in viral movement competent complex (Zhou *et al*., 2011). *Arabidopsis* calcium sensor protein synaptogmin A (SYTA) which also regulates endocytosis is found to be involved in MP mediated cell-to-cell movement of viral genome. It helps in targeting the MP-DNA complex to plasmodesmata via endocytosis (Lewis and Lazarowitz, 2010) as well by altering plasmodesmata permeability (Uchiyama *et al*., 2014).

2.3.2 Recombination-dependent replication

In infected plant cells, reservoir of geminiviral molecules results from not only rolling circle replication, but recombination and repair pathways which are error-prone processes. Recombination and repair pathways are believed to be the major forces to overcome plant defense as well as for its evolution. An increased homologous recombination of transgenes upon geminivirus infection has been found and it has been observed that this recombination occurs in tissue specific manner (Richter *et al*., 2014). Recombination-dependent replication have been suggested as an important part of viral replication in case of incomplete replication due to DNA damage or low processivity of polymerases (Alberter *et al*., 2005; Jeske *et al*., 2001; Preiss and Jeske, 2003; Ruschhaupt *et al*., 2013).
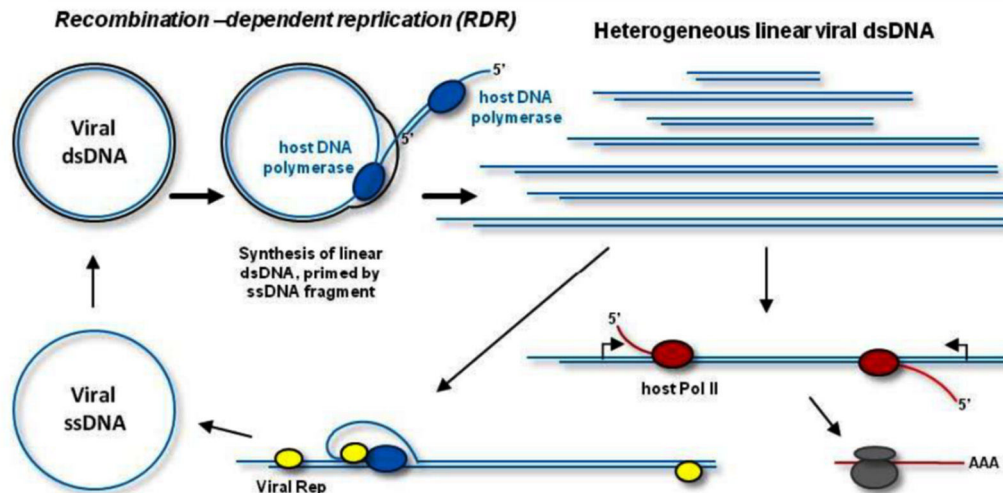
Figure 2.8 Recombination dependent replication in geminiviruses. RDR results in a heterogeneous pool of linear dsDNA from which Rep protein is produced by host polymerase II (shown in blue oval). Rep protein is produced by translation by host polymerase II (shown in red oval) from this dsDNA. This linear dsDNA is further nicked by Rep protein (shown in yellow oval) and further DNA is synthesized with the help of host DNA polymerase. (Adapted from Pooggin, 2013)

Invasion of a homologous region of the circular dsDNA molecule by a viral ssDNA fragment occurs with the assistance of host recombination proteins. This marks the initiation of RDR. After this, the invaded ssDNA is extended on the viral template strand by the host DNA polymerase. RDR produces a heterogeneous pool of linear dsDNAs that accumulate at high levels upon geminivirus infection (**Figure 2.8**). The priming during RDR does not involve Rep activity. Instead Rep helps in release of the ssDNA genome fragment from the heterogeneous linear dsDNA molecule which can again enter the replication cycle. The long linear dsDNAs containing more than one origin of replication can be transcribed to generate viral mRNAs by host RNA Polymerase II. Rep protein produced after translation starts replication of the long linear dsDNA bearing more than one origin of replication (Pooggin, 2013).

2.4 Begomovirus

Begomovirus is a genera comprising of whitefly transmitted geminiviruses that infects dicot plants in tropical and subtropical regions (Brown, 1994; Brown, 2007; Brown, 2009; Varma and Malathi, 2003). They infect various economically important cultivated plants such as legumes and fibres as well as noncultivated plants. Weeds are found to harbour many begomoviruses and they serve as depository of plant virions for their subsequent spread to other plants by whiteflies. They exhibit diversity in their genomic components which can be either monopartite or bipartite.

Begomoviruses have a high rate of mutation, pseudorecombination and recombination. Rapidly increasing spread of begomovirus worldwide is also due to expansion of geographical range of B-biotype of *B.tabaci* that is invasive and polyphagus (Seal *et al.*, 2006). There are several reports of acquisition of new DNA components and satellite molecules. Different subgenomic molecules such as beta satellite, alpha satellite and defective interfering molecules often are associated with their genomes. Thus, they constitute a diverse group of plant viruses and are most successfully emerging pathogens with the highest number of species among other genus of the family *Geminiviridae* (Kings *et al.*, 2011). On the basis of geographical and phylogenetic studies begomoviruses are caterogised into New World (NW) viruses and Old world (OW) viruses.

2.4.1 Genome organization of begomoviruses

2.4.1.1 Bipartite begomoviruses

A bipartite begomovirus has two DNA components, namely DNA-A and DNA-B molecule each of which is ~2.5-2.6 kb in size. DNA-A molecule of bipartite virus perform functions associated with its replication, gene expression control, encapsidation and overcoming host defence. DNA-B molecule is required for viral intracellular as well as intercellular movement inside the host (Hanley-Bowdoin *et al.,* 2000). DNA-A and DNA-B exhibits less sequence similarity except the conserved common region that shows more than 85% sequence similarity (Eagle *et al*., 1994). DNA-B requires DNA-A for its replication while DNA-A is dependent on DNA-B for systemic movement. However this obligatory relationship between DNA-A and DNA-B is not always the case. For incidence *Tomato yellow leaf curl Thailand virus* (TYLCTHV) produce symptoms on *N. benthamania* in absence of cognate DNA-B, however coinfection with DNA-B results in an increased symptom severity (Rochester *et al.*, 1990). Bipartite begomoviruses are found in both the new world as well as the old world. The DNA-A molecule codes for five and six ORFs in new world and old world begomoviruses, respectively (Hanley-Bowdoin *et al.*, 2000; Harrison and Robinson, 1999). It contains AC1, AC2, AC3 and AC4 ORFs in the complementary sense strand while AV1 in the virion sense strand. An additional ORF, AV2 encoding pre-coat protein is present in the case of bipartite old world

begomoviruses. The DNA-B codes for BC1 and BV1 in the complementary sense and virion sense strand, respectively (Sanderfoot *et al.*, 1996).

2.4.1.2 Monopartite begomoviruses

As early as 1990 all identified begomoviruses were found to be bipartite. Reports on infectivity of TYLCSV and TYLCV demonstrated existence of single genomic component containing begomoviruses (Kheyr-Pour *et al.*, 1991; Navot *et al.*, 1991). In the subsequent years, many monopartite begomoviruses have been identified. Monopartite begomoviruses are native to Old world. However, recently *Tomato Leaf deformation virus* (ToLDeV) in Peru has been identified which is not associated with DNA-B molecule and the biological properties of its genotype exhibit similarities to the Old world monopartite begomoviruses. It is believed to emerged from a New world bipartite DNA-A progenitor through convergent evolution and recombination events (Sánchez-Campos *et al.*, 2013). The monopartite begomoviruses genome contains six genes and displays homology to DNA-A component of Old world bipartite begomoviruses (Navot *et al.*, 1991).

2.4.2 Functions of begomovirus encoded proteins

The important functions of various proteins encoded by begomoviruses have been summarised in **Table 2.4**.

Table 2.4 Function of the viral proteins encoded by begomoviruses.

Viral ORF(s)	Protein & its Functions
DNA-A AV1	Coat protein, viral genome packaging
AV2	Suppress transcriptional and post-transcriptional gene silencing (TGS and PTGS), Virus movement
AC1	Replication associated protein, required for viral replication
AC2	Transcription activator protein, Suppress transcriptional and post-transcriptional gene silencing (TGS and PTGS), Activation of late viral gene transcription
AC3	Replication enhancer protein, interact with Rep and enhance Rep mediated virus replication
AC4	Enhancement of pathogenesis, symptom induction
DNA-B BV1	Nuclear shuttle protein (NSP), Virus trafficking
BC1	Movement of viral protein

2.4.2.1 Coat protein

AV1 codes for coat protein which encapsidate the viral genomic DNA. It is an oligomeric/multimeric protein where the N-terminal amino acids of one CP molecule interact with C-terminal residues of another CP molecule (Hallan and Gafni, 2001). It has a critical role in determining vector specificity and insect transmission (Briddon *et al*., 1990) as well as in viral ssDNA accumulation (Harrison *et al*., 2002; Wartig *et al.*, 1997a). TYLCV and ACMV CP possess nuclear localization signal (NLS) at its N-terminal (Kunik *et al*., 1998). However, TYLCV coat protein bears nuclear export signal (NES) in the C-terminal. Thus it believed that CP functions as nuclear shuttle protein in monopartite TYLCV begomoviruses while for bipartite begomoviruses acts as nuclear import protein. Coat protein interacts with HSP16, vector whitefly encoded Heat shock protein (Ohnesorge and Bejarano, 2009) and GroEL proteins encoded by whitefly gut endosymbiont (Morin *et al.,* 2000). These interactions might provide stability to the viral genomic DNA during its transfer into the haemolymph of whitefly. It might also be responsible for limiting the copy number of viral DNA by inhibiting Rep endonuclease activity (Yadava *et al.*, 2010).

2.4.2.2 Pre-coat protein

AV2 encodes for pre-coat protein/movement protein. The new world begomoviruses do not contain AV2 (Kings *et al.*, 2011; Rybicki, 1994). Recently it has been found to be suppressor of RNA-silencing (Yadava *et al.*, 2010; Zrachya *et al.*, 2007). It has been also involved in virus movement. TYLCV AV2 mutations prevented the viral systemic infection in tomato, suggesting its role in movement (Harrison and Robinson, 1999; Sharma and Ikegami, 2009; Wartig *et al*., 1997b). It is also proposed that TYLCV AV2 increases the viral DNA transport from nucleus to cell periphery (Rojas *et al.*, 2001).

2.4.2.3 Replication intiator protein

AC1 encodes for replication initiator protein/ replication associated protein (Rep) which is indispensible for viral DNA replication (Elmer, 1988). It is a multifunctional protein. It has a high oligomeric state which is required for many of its functions. It binds to DNA in a site specific manner at the iterons present in the intergenic region (IR). Thereafter it initiates viral replication by creating a nick at the conserved

nonanucleotide sequence. It also possesses ligase and ATP-dependent topoisomerase I activity (Laufs *et al.*, 1995; Pant *et al.*, 2001). The Rep protein exhibits ATPase and helicase activities *(Choudhury et al.*, 2006; Clerot and Bernardi, 2006; George *et al.*, 2014). Role of Rep protein in viral replication and other functions of Rep protein during viral infection along with its interaction with various host factors have been discussed in detail in the following section.

2.4.2.4 Transcription activator protein

AC2 ORF encodes for a protein that functions as a transcriptional activator. It is required for AV1, AV2 and BV1 gene expression (Gopal *et al.*, 2007; Hanley-Bowdoin *et al.*, 2000; Harrison and Robinson, 1999; Pandey *et al.*, 2009; Trinks *et al.*, 2005). The N terminal of TrAP is basic and contains NLS while C terminal possesses an acidic transactivation domain. The middle Cys-His rich region is required for nonspecific DNA binding. TrAP protein has been reported to function as suppressor of RNA silencing in both monopartite and bipartite begomovirus (Gopal *et al.*, 2007; Kon *et al.*, 2007; Trinks *et al.*, 2005; Vanitharani *et al.*, 2004; Voinnet *et al.*, 1999; Wang *et al.*, 2003). ACMV AC2 was the first to be shown as suppressor of PTGS. AC2 suppressor activity was demonstrated to be independent of transcription activation activity in case of ICMV (Vanitharani *et al.*, 2004). AC2 inactivates adenosine kinase (ADK) (Wang *et al.*, 2003) and leads to suppression of gene silencing. It also interacts with karyopherin α of tomato (Chandran *et al.*, 2012).

2.4.2.5 Replication enhancer protein

ORF AC3 encodes for protein named as the replication enhancer (REn) as it increases viral replication (Hanley-Bowdoin *et al.,* 2000; Harrison and Robinson, 1999; Sunter *et al.,* 1990). Mutation in AC3 ORF leads to reduced ssDNA and dsDNA accumulation (Sunter *et al.*, 1990). REn protein can oligomerize and bind to viral Rep protein. TGMV REn interacts with few host cell cycle proteins like maize retinoblastoma homologue (RBR1) and *Arabidopsis* PCNA (Castillo *et al.*, 2003; Hanley-Bowdoin *et al.*, 2004; Settlage *et al.*, 2001).

2.4.2.6 AC4 encoded protein

AC4 encodes for a small protein that function as a symptom determinant (Fondong *et al.*, 2007; Gopal *et al.*, 2007; Pandey *et al.*, 2009; Wartig *et al.*, 1997b) TLCV AC4

mutants produce less severe symptoms compared to plants agroinoculated with functional AC4 gene (Rigden *et al.*, 1994). ACMV AC4 has been shown to supress PTGS (Vanitharani *et al.,* 2004).

2.4.2.7 Nuclear shuttle protein

Movement of viral DNA between the nucleus and the cytoplasm is carried out by viral nuclear shuttle protein encoded by BV1 (Hanley-Bowdoin *et al.*, 2000; Noueiry *et al.*, 1994; Sanderfoot and Lazarowitz, 1995). It binds to the newly synthesized ssDNA molecule and transports them between nucleus and cytoplasm.

2.4.2.8 Movement protein

BC1 encodes for movement protein (MP) which helps in cell-to-cell movement of viral DNA (Jeffrey *et al.*, 1996; Lazarowitz and Beachy, 1999).

2.4.3 DNA satellite molecules associated with begomoviruses

Satellites are described as nucleic acids or viruses that are indispensably dependent for their replication on the helper virus (Fauquet *et al.*, 2005). The first satellite identified was a RNA molecule found associated with the nepovirus *Tobacco ringspot virus* (Schneider, 1969).
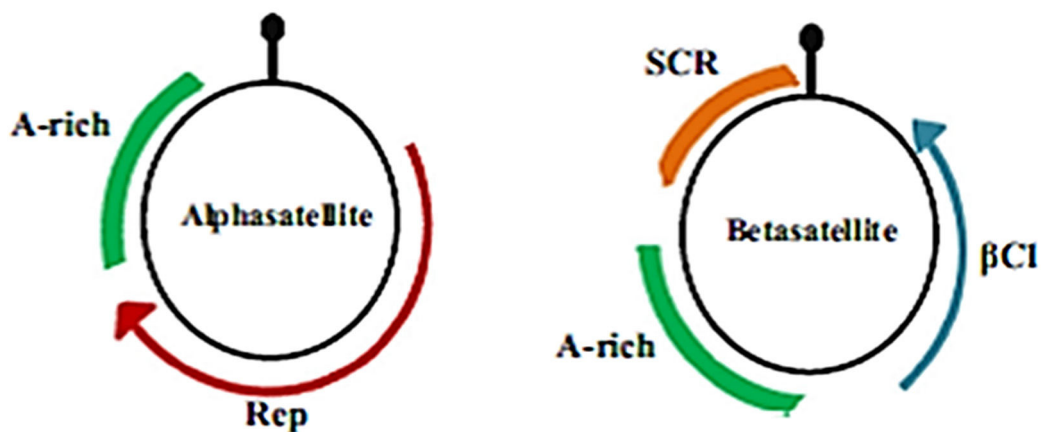
Figure 2.9 Genome organization of alphasatellite and betasatellite DNA molecule (~1.4 Kb) showing the key conserved domains. (Adapted from Rizvi *et al.*, 2015)

Thereafter, several groups of plant viruses have been reported to be associated with various satellite RNAs (Fauquet *et al.*, 2005), which interferes with the replication of the helper viruses and results in attenuated symptoms. According to studies few

satellites may intensify disease symptoms that are induced by the helper virus, while certain results in novel symptoms which were usually not produced during the helper virus infection (Roossinck *et al*., 1992).

The first DNA satellite reported in begomovirus was found associated with *Tomato leaf curl virus* (ToLCV) from Australia (Dry *et al*., 1997). It was 682 nt circular ssDNA whose replication and encapsidation depends on ToLCV. However, it was found to have no effect on either symptoms caused by ToLCV or the viral replication. Except for the nonanucleotide TAATATTAC sequence present in the stem loop, the satellite DNA showed very less DNA sequence similarity to its helper virus (Akbar Behjatnia *et al*., 1998).

Ageratum yellow vein virus (AYVV) (Saunders and Stanley, 1999; Saunders *et al*., 2000), a monopartite begomovirus isolated from *Ageratum conyzoides* (goat weed) when re-introduced in goat weed failed to produce the typical yellow vein symptoms associated with the infection which suggested the involvement of another molecule/factor that may have role in pathogenicity in the host. However, it was found to be infectious in *Nicotiana benthamiana*. Subsequent studies led to identification of small circular ssDNA molecule having sequence of unknown origin but with the AYVV origin of replication (Stanley *et al*., 1997).

Subsequently many begomoviruses with monopartite genome were found to be associated with satellite molecule referred to as betasatellite which are circular ssDNA, ~1.3 kb in size which is half the size of the helper DNA molecule. According to several reports many betasatellites are necessary for symptom development that are associated with the infection/ disease (Briddon *et al*., 2001; Briddon *et al.,* 2003; Jose and Usha, 2003; Saunders *et al.,* 2000; Zhou *et al.,* 2003). However, all betasatellites obligatory require a helper virus for its replication, local as well as systemic spread and transmission by white fly (Briddon *et al*., 2003). The betasatellite molecule comprises of one ORF (*βC1*) in addition to a ~240 nt long A-rich region and a~220 nt satellite-conserved region (SCR) (**Figure 2.9**). The *βC1* has been found to suppress the jasmonic acid signaling (Yang *et al.,* 2008).

The begomovirus are also often associated with another kind of circular ssDNA satellite molecule, initially referred to as DNA-1 (Briddon *et al.,* 2004; Mansoor *et al.,* 1999; Saunders and Stanley, 1999), but now renamed to alphasatellites (Mubin *et al.,*

2009). Alphasatellites contain a conserved hairpin structure and adenine rich (A-rich) region (**Figure 2.9**). They encode a single protein Rep that exhibits high nucleotide identity with the Nanovirus encoded Rep protein. Members of family *Nanoviridae* also have a circular ssDNA genome (Gronenborn, 2004). Alphasatellites require helper begomoviruses for insect transmission and movement in host but are capable of autonomous replication. Unlike betasatellites, alphasatellites possess TAGTATTTAC nonanucleotide sequence in their stem loop which is typically present in viruses of *Nanoviridae* family.

2.5 Replication initiator protein

Functional studies on various mutants of the entire ORFs of *Tomato golden mosaic virus* concluded that the Rep is the only viral protein absolutely necessary for its replication (Elmer, 1988). In addition it is also involved in the process of transcription and regulates the expression of certain viral genes. Rep is a multifunctional protein as is evident from the presence of various motifs in the protein (**Figure 2.10**) and its ability to interact with various host factors (Rizvi *et al.*, 2015).
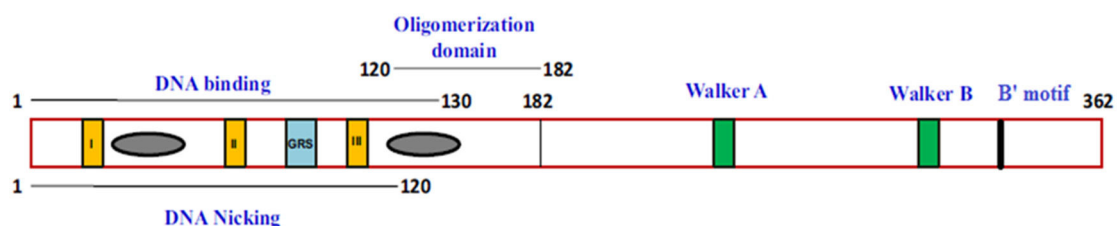
Figure 2.10 Diagrammatic representation of replication initiator protein (Rep) of geminivirus. The orange boxes represent the motif I, II and III; the grey oval represents the α-helices; the green boxes represent the Walker A and Walker B; the blue box represent the geminivirus Rep sequence motif. The position of the B' motif has also been shown at the C-terminal (modified from Rizvi *et al.* 2015).

2.5.1 Role of Rep in geminivirus replication

Geminivirus Rep is highly conserved protein. It exhibits no similarity with known polymerases but instead showed remarkable similarity with the replication initiator proteins (Rep proteins) of eubacterial plasmids due to the presence of three conserved motifs with similar spacing between them, namely motifs I, II and III at their N-termini (Koonin and Ilyina, 1992). These motifs are needed in initiation as well as termination of DNA synthesis during RCR. Rep functions as a site and strand specific endonuclease. Motif I (FLTY) and motif II (HLH) are required for

specific dsDNA binding and metal-binding respectively, while the catalytic site for endonuclease activity comprises of motif III (YxxKD/E). The hydroxyl group of the Y (tyrosine) residue of motif III forms a covalent bond with the 5'-PO_4 of the cleaved DNA strand.

The Rep protein has been demonstrated to initiate the process of replication by introducing a nick in the plus strand within the highly conserved nonanucleotide sequence 5'TAATATTAC3′ (Laufs *et al*., 1995). For rolling-circle replication, Rep binds in the CR region in sequence specific manner and mapping results have shown the cleavage occurs at the phoshodiester bond between the seventh and eighth residues of the invariant nonamer 5'TAATATT↓AC3′ (**Figure 2.11**). The 3'-hydroxyl end thus generated is utilized for the RCR while to the 5'-phosphate end of the cleaved DNA Rep remains covalently bound. After a complete cycle of replication, new origin sequence is generated that is again nicked and subsequently the nascent 3' end of DNA is ligated to the previously generated 5'end by Rep, thus resolving the nascent viral single strand into genome-sized units. The 211amino acids in the N-terminal of TYLCV AC1 contribute for origin cleavage and ligation (Heyraud-Nitschke *et al*., 1995). Apart from RCR motifs (I-III), helix 1 and helix 2 located between motif I and motif II also contribute to Rep DNA binding and endonuclease activity (Orozco and Hanley-Bowdoin, 1998). Mutational studies in TGMV Rep had shown that DNA binding domain spans from 1 to 130 amino acid region and overlaps with the oligomerization domain (120-180 amino acid residues) indicating that DNA binding is oligomerization dependent. However, oligomerization is not required for DNA cleavage and ligation activity which is conferred by 1-120 amino acids region of Rep. Geminivirus Rep Sequence (GRS) is another motif showing high extent of conservation among all geminivirus Rep proteins (Nash *et al*., 2011). GRS comprise of an uncharacterized sequence constituting of two clusters of amino acids between motif II and motif III and has been found to be required for the replication initiation as the GRS mutations resulted in impaired DNA cleavage activity.

2.5.2 Role of Rep in transcription regulation

The presence of Rep-binding site between TATA box and the transcription start site suggested Rep role in transcription process (Sunter *et al.*, 1993). Repression of early gene transcription by its own gene product has been known in the life cycle of many

DNA viruses like SV40. During the process of transcription, Rep mediated repression of its own promoter occurs via binding to the conserved iteron sequences in the CR
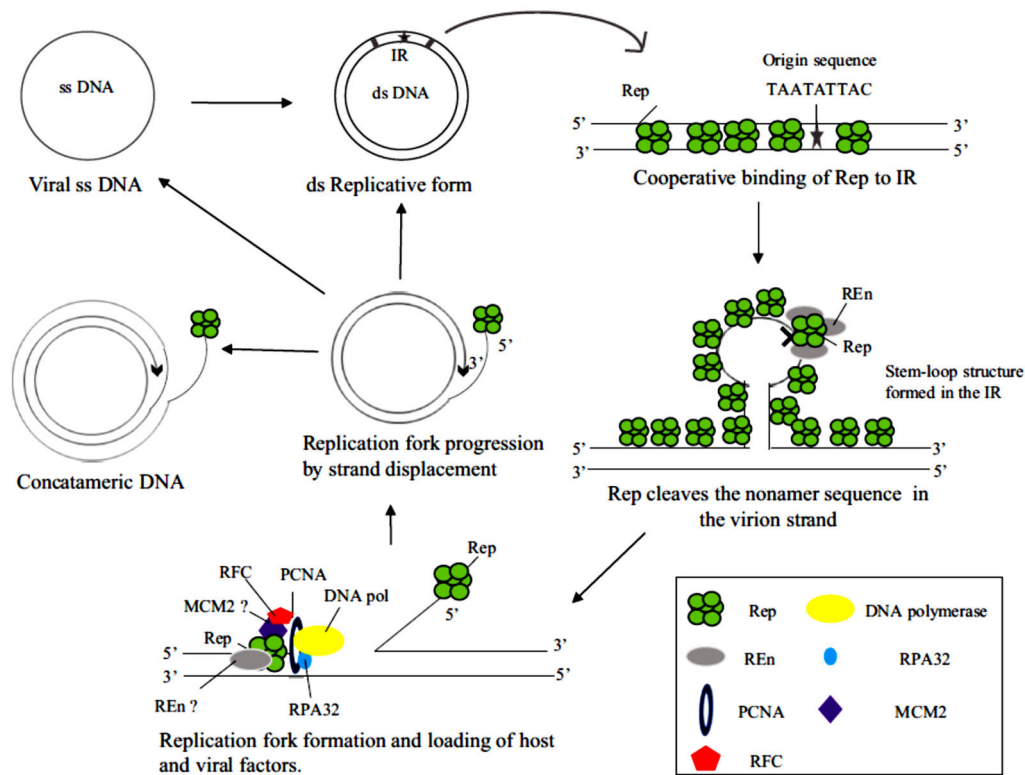
Figure 2.11 Model for Rep mediated rolling-circle replication of geminivirus DNA (Adapted from Rivzi *et al.*, 2015).

region of the viral genome (Eagle *et al.*, 1994). Initiation of replication by Rep iteron binding and autoregulation due to Rep iteron binding are suggested to be exclusive and independent events, as the transcription repression activity was still retained in replication impaired Rep mutants and the ability of the presence of Rep binding site into the heterologous promoter to confer AC1 mediated repression. This autoregulation is also required for the expression of two downstream genes AC3 and AC2 since the transcription site for the both are present in the coding region of Rep (Shung and Sunter, 2007). This enhanced expression of AC2 gene brings about the suppression of plant defence responses and the expression of late viral gene in later phase of geminiviral infection. *Tomato yellow leaf curl Sardinia virus* (TYLCSV) Rep protein contains a highly conserved RGG sequence at amino acid 124–126 which when mutated inhibits its autoregulation and its subcellular localization (Sardo *et al.*, 2011).

2.5.3 Role of Rep in suppression of gene silencing

Wheat dwarf virus Rep and RepA are the first mastreviral protein reported to have suppression of RNA silencing activity (Liu *et al.*, 2014). They are demonstrated to inhibit the PTGS of ssGFP as well as inverted repeat dsGFP constructs, although they exhibited weaker silencing suppressor activity relative to the *Tobacco etch virus* HC-pro protein.

Methylation of DNA (specifically at cytosine base) leads to repression of gene transcription while methylation of histone can lead to either activation or repression of gene transcription. The geminiviral genome is highly methylated *in vivo* (Raja *et al.*, 2008). Out of the various replicative intermediates namely open circular, covalently closed circular and heterogeneous linear DNA, it is the viral heterogeneous linear dsDNA found to be preferentially methylated (Paprotka *et al.,* 2011). Conserved hairpin and AC1 binding region of the viral DNA are found to be the frequent sites for cytosine methylation within which resides the early and late gene promoter and replication origin, thus suggesting the implication of methylation in the process of viral transcription and replication. *In vitro* DNA methylation found to have negative impact on the replication of the viral genome in tobacco protoplasts (Brough *et al.*, 1992; Ermak *et al.*, 1993). Several studies indicated that plant utilizes both the cytosine as well as histone methylation machinery against the invading geminivirus which viral RNA silencing suppressor proteins (RSS) counter by inhibiting the host silencing machinery (Burgyan and Havelda, 2011; Raja *et al.*, 2010). Supporting this notion it was found that *Arabidopsis* plants deficient in cytosine methyltransferases and histone methyltransferases as well plants mutants for methylation cycle components results in more pronounced symptoms on geminivirus infection and the viral DNA isolated from them had significantly reduced methylation level (Raja *et al.*, 2008). In addition both active and repressive H3K4 and H3K9 marks are associated with viral genome. All these findings indicated the role of host mediated viral genome methylation during the infection process.

Recently, it was demonstrated that geminivirus infection results in Rep mediated reduction in transcript of methyl cycle enzymes, MET1 and CMT3 in *N. benthamiana* which are responsible for maintenance of symmetric methylation (Rodriguez-Negrete *et al.*, 2013). Expression of Rep leads to reduced DNA methylation of *A. thaliana*

transgene as well as of the host loci whose expression is regulated by symmetric CG methylation suggesting its role in suppression of TGS.

2.5.4 Interaction Rep with host factors

Rep interacts with several host factors required for replication and plant defence.

1. **Plant Retinoblastoma related proteins (RBR):** Rb family of protein regulate the progression of cell cycle. Expression of a homolog of Rb (RBR) of maize results in reduced viral DNA replication in wheat cells. WDV RepA contains LXCXE motif which is required to physically interact with RBR protein (Xie *et al.*, 1995). This interaction seems critical for replication as the mutants incapable of this interaction abolished the WDV replication in cultured wheat cells. Maize Retinoblastoma-related proteins, RBR1 and RBR2 interact with both Rep as well as D-type cyclin (Ach *et al.*, 1997). However, TGMV Rep lacks LXCXE motif and interacts with RBR protein with a distinct and novel helix4 motif (Arguello-Astorga *et al.*, 2004). Helix 4 motif consists of charged amino acids residues flanking a hydrophobic core which are implicated in interaction with RBR and also in viral replication. Rep was able to re-replicate in fission yeast. It was surprising as fission yeast lacks RB homolog. But the alternative interaction with cyclins through RXL motif has been proposed that regulates replication (Hipp *et al.*, 2014).

2. **Proliferating cell nuclear antigen (PCNA):** Geminivirus infection induces the expression of PCNA in the mature infected cells (Nagar *et al.*, 2002). Tomato PCNA has been shown to interact with TYLCSV Rep protein as well as with REn protein (Castillo *et al.*, 2003). PCNA binds to IMYMV Rep protein. A 134 to 183 amino acid stretch of IMYMV Rep is required for the interaction on the other hand amino acid residues of PCNA involved in the interaction are dispersed throughout PCNA (Bagewadi *et al.*, 2004). This interaction downregulates the endonuclease and ATPase function of the Rep protein.

3. **Replication factor C (RFC):** Replication factor C is a multimeric protein that loads of the PCNA onto the DNA during replication process. Wheat large subunit of replication factor C complex (TmRFC-1) binds to WDV Rep

protein in the DNA/Rep/TmRFC-1 complexes which resembles the pre-initiation complex, thus assists in the further assembly of elongation complex for the viral replication (Luque *et al.*, 2002).

4. **Replication Protein A-32 (RPA-32):** RPA is a heterotrimeric protein that binds to ssDNA and in addition to replication, is involved in repair and recombination. RPA32 subunit interacts with C-terminal of MYMIV Rep protein, downregulates the endonuclease activity while upregulating its ATPase activity (Singh *et al.*, 2007). Thus, it has been believed that the interaction with RPA32 in addition, might limit the replication initiation and drives to elongation phase of RCR.

5. **Recombination enzymes:** RAD51 and RAD54 are two repair and recombination proteins which interact with Rep protein (Kaliappan *et al.*, 2012; Suyal *et al.*, 2013). They might play a critical role in case of replicational stress by stabilizing the replication fork. The N-terminal of RAD54 binds to the oligomerization domain of MYMIV-Rep and enhances its nicking, ATPase as well as helicase activities (Kaliappan *et al.*, 2012). However surprisingly, studies on rad54 mutant showed no effect on the CSR, RCR or RDR even though it physically interacts with Rep (Richter *et al.*, 2015).

6. **Sumoylation conjugating enzyme-1:** The N-terminal of Rep binds to sumoylation conjugating enzyme (SCE-1) (Castillo *et al.*, 2004). The K68 and K96 amino acid residues in the N-terminal of Rep are found to interact with SCE-1 and when mutated abolished the interaction and reduced the viral accumulation in infected plants (Sanchez-Duran *et al.*, 2011).

7. **Histones:** Geminiviral genomic DNA has been shown to assemble as minichromosomes (Pilartz and Jeske, 1992). TGMV Rep protein interacts with Histone-3 which suggests probable implication of this interaction in replication and transcription process (Kong and Hanley-Bowdoin, 2002). It has been hypothesized that Rep recruitment on the viral genome and its interaction with H3 may help in removal of nucleosomal block and thus helps in its efficient transcription and replication. Rep protein also interacts with a kinesin motor protein (GRIMP) that is involved in mitosis process (Kong and

31

Hanley-Bowdoin, 2002). Apart from that it also interacts with a kinase, Geminivirus rep interacting kinase (GRIK). These interactions might inhibit the cell from entry into the mitotic phase.

8. **NAC-domain containing proteins:** GRAB1 and GRAB2 belong to NAC-domain containing protein family which are involved in plant development. They have a unique C-terminal which contain negatively charged residues while the N-terminal is conserved, interact with WDV Rep to inhibit the replication (Xie *et al.*, 1999).

2.6 Helicases

Helicases are translocases that are capable of unwinding the complementary strands of the double-stranded nucleic acid (NA) required to facilitate various cellular functions. Helicase function as an essential and integral component of various larger protein complexes which are involved in diverse steps of all nucleic acid transactions. DNA helicase facilitates DNA metabolic processes such as repair, replication and recombination while RNA helicase targets transcription, RNA splicing, RNA editing, RNA transport, RNA degradation, ribosome biogenesis and translation machinery. Since helicases are involved in multitude of cellular processes, these enzymes are ubiquitously essential and therefore are evolutionarily conserved (Patel and Picha, 2000; Singleton and Wigley, 2002). The importance of this enzyme is reflected from the fact that due to its absence or defect various pathological states can arise (Andressoo *et al.*, 2005; Ellis *et al.*, 1995; Modrich, 1994). Thus, helicase are important players in assuring integrity and stability of genome.

In helicases, translocation is brought about by NTP-dependent conformational changes. Apart from this, enzymes having high processivity and capable of displacing other proteins on NAs, function as translocases which are part of complexes such as chromatin remodelling, ribosomal biogenesis etc (Saha *et al.*, 2006; Veaute *et al.*, 2005). The complexity of the helicase is further enhanced by the presence of different modular accessory domain which provides additional activities such as oligomerization, protein-protein interaction, substrate specificity etc (Hickman and Dyda, 2005).

On the basis of the conserved amino acid sequences that are shared by distantly related members, helicases have been classified by Gorbalenya and Koonin (1993) into six superfamilies (SF1-SF6). Members of the SF3-6 are hexamers and have been discussed in the following section.

2.6.1 Hexameric helicases

Hexameric helicases are ring shaped structures exhibiting two-tiered arrangement. They have either Rec A or an AAA+ like core domain which have NTP binding site at the interface between two monomers. The NTP interacts with conserved motifs that line the binding pocket as well as with arginine-finger that is contributed by the adjacent monomeric subunit. Ring shaped helicases require NTP-dependent loading partner for its loading onto the template (Abbate *et al.*, 2004). However, some can self-assemble around the DNA. Unwinding is an intricate process where NTP binding, NTP hydrolysis, conformational changes and translocation are closely interrelated (Lohman and Bjornson, 1996). The mechanism that couples the NTP binding and hydrolysis to DNA translocation has been unclear.

Structures of many hexameric helicases demonstrate that only one strand of dsDNA passes through the central channel of the hexamer. The DNA-binding hairpin loops that contain basic amino acids are projected into the central channel from each subunit and form a spiral 'staircase' tracking the backbone of the DNA sequentially. Height of the hairpin loops exhibit correlation with different ATP ligated states, like with ATP, ADP and apo (empty) states. As each subunit sequentially progresses through different ATP ligated configuration the DNA-binding hairpin loops travels from the top to the bottom position of the staircase (Singleton *et al.*, 2000). The movement of these highly flexible loop is driven by both NTP binding and hydrolysis.

2.6.2 Mechanism of NTP hydrolysis and strand separation in hexameric helicase

NTP-binding sites of the hexameric helicases exhibited strong negative cooperativity among them (Bujalowski and Klonowska, 1993). The mechanism of NTP-hydrolysis has been a subject of considerable debate. Four principal mechanisms have been proposed – three-site sequential model (F1-ATPase type), all-sites sequential model, stochastic model and concerted model. The three-site sequential model, also called as

F1-ATPase-like model was proposed after significant similarities between F1-ATPase and SF4/SF5 helicases were reported (Abrahams *et al.*, 1994; Hingorani *et al.*, 1997). Initial findings from the studies of T7 bacteriophage helicase suggested that there are three to four high affinity sites while the other sites have low affinity for NTP (Hingorani and Patel, 1996; Stitt and Xu, 1998). According to this model there are three potential NTP hydrolyzing sites at alternating subunit interfaces. Although initial biochemical studies with T7 bacteriophage helicase favoured three-site sequential model, but this model is no longer supported. In all-sites sequential model, all the sites are proposed to be equivalent in terms of their catalytic potential. Structural reports also strongly accept this model for the SF3 helicases, Rho and T7gp4 (Adelman *et al.*, 2006; Liao *et al.*, 2005). The only report that supports the stochastic model is proposed for ClpX which belongs to AAA+ family (Martin *et al.*, 2005). The salient feature of this model is that each subunit of the hexameric helicase independently catalyzes the NTP hydrolysis and thereafter translocate on the DNA substrate. The concerted model has been proposed for SV40 large-T antigen in which all the NTP binding site binds and hydrolyzes NTP simultaneously and are converted into NDP bound forms that further return back to apo (empty) form (Gai *et al.*, 2004; Huang *et al.*, 1998).

The Bovine papillomavirus E1 protein, SV40 large T-antigen and AAV Rep belongs to SF3 and are members of the AAA+ (ATPases associated with cellular activities) family (Neuwald *et al.*, 1999). These hexameric helicases unwind the dsDNA with 3′- 5′ polarity (Ahnert and Patel, 1997; Sedman and Stenlund, 1998). Hexameric replicative helicases of bacteriophage T7 gene product 4 (T7gp4) and *E. coli* DnaB belongs to the SF4 (Ahnert and Patel, 1997). According to the strand exclusion model for the strand separation, one of the strand moves through the central channel while the other complementary strand is excluded. This mechanism provides a way to check immediate and spontaneous reannealing of the separated DNA strands. It had been reported that both T7gp4 and DnaB encircle only one of the two DNA strands (Egelman *et al.*, 1995; Kaplan and O'Donnell, 2002). The strand displacement assays also suggested that the ring encircles one of the strands while the other strand is sterically excluded. The crystal structures of the SV40 T-antigen (Tag) had demonstrated that channel diameters in different NTP ligated state cannot permit the passage of 20Å diameter of the dsDNA (Li *et al.*, 2003).

In addition to this in E1 helicase, the six oligomerization domains form a rigid collar which is supposed to act as a processivity factor, making unwinding process more efficient by keeping the two strands of DNA topologically apart (Jeruzalmi *et al.*, 2002). Other models for the mechanism of strand separation also had been postulated for the hexameric helicases such as torsional mechanism and ploughshare mechanism (Ahnert and Patel, 1997; Hacker and Johnson, 1997; Kaplan, 2000; Kaplan *et al*., 2003).

Different helicase superfamilies adopt a wide range of mechanisms although there are several common features between them. Understanding on functions of hexameric helicase have considerably expanded due to increased available information from genetic and biochemical approaches, crystallography and NMR. The mechanisms associated with the unwinding process by these ring helicases lacks understanding. The studies of hexameric helicases will also provide insight into the mechanisms of AAA+ proteins, which are involved in diverse cellular activities.

MATERIALS AND METHODS

3.1. Materials

3.1.1 Viral clones

The monomeric clone of ToLCNDV DNA-A (Accession no. HM007113) in pUC18, full-length infectious clones of ToLCNDV DNA-A (Accession no. HM007113) and ToLCGV DNA- B (Accession no AY190291) in pCAMBIA 2300 were available in our laboratory (Molecular Virology laboratory, School of Life Sciences, Jawaharlal Nehru University). The monomeric clone of ToLCNDV DNA-A in pUC18 at *Kpn* I site was used as template of cloning of C-terminal of Rep in pET28a(+) vector. The pET28a(+)-RepC clone was used for introducing site directed mutations in the Rep protein. While for construction of infectious clones of ToLCNDV mutants, site directed mutations were performed on the monomeric clones of ToLCNDV DNA-A clone into pUC18 available in our laboratory.

3.1.2 Plant species, microbial strains and vectors

N. benthamiana was obtained from Central Tobacco Research Institute (CTRI), Rajamundary.

Table 3.1 Different microbial strains used in this study

S.No.	Strains	Genotype	Purpose
1.	*E.coli* strain DH10β	*araD139 Δ(ara-leu)7697 fhuA lacX74 galK (φ80 Δ(lacZ)M15) mcrA galU recA1 endA1 nupG rpsL (StrR) Δ(mrr-hsdRMS-mcrBC)*	Cloning of viral AC1 ORF (wild-type and mutant), Cloning of mutant viral genome
2.	*E.coli* strain DH5α	*fhuA2Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	Cloning of viral AC1 ORF (wild-type and mutant), Cloning of mutant viral genome
3.	*E.coli* strain BL21(DE3)	*F⁻ompT gal dcmlonhsdS_B(r_B⁻m_B⁻) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*	Expression of Wild-type and mutants of ORF AC1
4.	*A.tumefaciens* strain EHA105	genotype C58 pTiBo542; T-region::aph, Km(S); A281 derivative harboring pEHA101, T-DNA replaced with nptII, elimination of T-DNA boundaries unconfirmed, super-virulent	For agroinoculation of plants with infectious dimeric viral constructs

Table 3.2 Different plasmid vectors used in this study

S.No.	Vector	Source	Purpose
1.	pMALc2X	New England Biolabs, UK	Overexpression of AC1 ORF
2	pJET1.2	MBI Fermentas, USA	Cloning of PCR products with blunt ends
3	pCAMBIA 2300	Cambia	Construction of infectious viral clones
4	pET28a(+)	Novagen	Overexpression of AC1 ORF

3.1.3 Chemicals

All the chemicals used in this study were of high quality and purity. The common chemicals were obtained from Merck, Himedia, Qualigens, Fischer, Sigma (USA), Invitrogen and Millipore (USA) unless otherwise stated.

Trizma base, ethylenediaminetetraacetic acid (EDTA), sodium dodecyl sulfate (SDS), Bromophenol blue, xylene cyanol, antibiotics (kanamycin and ampicillin), isopropyl β-D-1-thiogalactopyranoside (IPTG), triton-X, lysozyme, agarose, bovine serum albumin (BSA), sodium chloride (NaCl), acrylamide, bis-acrylamide, ammonium persulfate (APS), Tetramethylethylenediamine (TEMED), Bradford reagent, β-mercaptoethanol, dithiothreitol (DTT), maltose, ATP, lithium chloride, potassium chloride, ammonium molybdate, ascorbic acid, imidazole and other chemicals were procured from Sigma, USA. Glycerol, sodium bicarbonate ($NaHCO_3$), sodium hydroxide (NaOH), bacto-tryptone, yeast extracts, agar, ethanol, glacial acetic acid, chloroform, dextrose, hydrochloric acid (HCl), sulphuric acid (H_2SO_4), isoamyl alcohol, isopropanol, monopotassium phosphate (KH_2PO_4), dipotassium phosphate (K_2HPO_4), magnesium chloride ($MgCl_2$), calcium chloride ($CaCl_2$), monosodium phosphate (NaH_2PO_4), disodium phosphate (Na_2HPO_4), sodium acetate (CH_3COONa) and potassium acetate (CH_3COOK) were purchased from Merck.

3.1 4 Enzymes and Kits

The enzymes and kits used in the present study were procured from the following sources as mentioned below.

Table 3.3 Enzymes and kits used in this study

S. No.	Enzymes and Kits	Source /Supplier company
1.	Restriction enzymes	Fermentas,USA; Bangalore Genei, India and New England Biolabs, UK
2.	T4 DNA ligase	Fermentas, USA
3.	Prestained protein molecular weight ladder	Fermentas, USA
4	Unstained protein molecular weight ladder	BioRad
5	Factor Xa	New England Biolabs, UK
6.	Antarctica Phosphatase	New England Biolabs, UK
7.	DNA ladder	Fermentas ,Bangalore Genei and New England Biolabs
8	Taq DNA polymerase	Chromous,India; Bio tools, USA
9	Pfu DNA Polymerase	New England Biolabs, UK
10	Genomic DNA isolation kit	Qiagen, Germany
11.	MinEluteTMGel extraction Kit	Qiagen, Germany
12	Montage PCR purification kit	Millipore,USA
13.	Nucleotide removal kit	Qiagen, Germany
14.	pTZ57R/T cloning kit	Fermentas,USA
15.	pJET 1.2 Blunt Cloning Vector	Thermo Fisher Scientific, USA
16	Quick-change site-directed mutagenesis kit	Stratagene, USA
17	Klenow fragment	Fermentas, USA

3.1 5 Antibodies, filters and membranes

Anti-His monoclonal antibody raised in mouse, anti-mouse IgG coupled to peroxidase (raised in goat) were purchased from Sigma. Membrane filters of 0.22 μm pore size were purchased from Millipore, USA. Other filter papers were purchased from Whatman, USA. Nylon and PVDF membrane was procured from Millipore, USA.

3.1.6 Chromatography resins

Amylose resin was procured from New England Biolab (NEB). Ni-NTA Agarose, DEAE Agarose and Sepharose CL-4B were obtained from Sigma, USA. Heparin sepharose and Q- sepharose were purchased from GE healthcare, UK.

3.1.7 Oligonucleotides

M13 MP18 circular ssDNA was purchased from Bangalore Genei. All the primers and oligonucleotides used in this study were synthesized from Sigma-Aldrich, Bangalore, India. The primers were diluted according to the company's instruction and stocks of 100 μM were prepared for both forward and reverse primers. Further 1:10 times dilution i.e. working of 10 μM was prepared for further use in PCR.

Table 3.4 List of the primers used in the present study

Primer Name	Length	Sequence (5′ to 3′)
RepF FP	32 nt	CATGCCATGGCTCCGCCACGTCGTTTTAGAAT
RepC FP	31 nt	CATGCCATGGGACGATCTGCTCGTGGTGGTC
RepF/RepC RP	31 nt	CGGCTCGAGACTCGCCTCCTGCGAATGCTCT
K227A FP	39 nt	GGGGATAGTAGAACGGGCGCCACAATGTGGGCTCGATGC
K227A RP	39 nt	GCATCGAGCCCACATTGTGGCGCCCGTTCTACTATCCCC
D261A FP	39 nt	GCCTGGTACAACGTCATTGCCGACGTTGATCCCCACTAT
D261A RP	39 nt	ATAGTGGGGATCAACGTCGGCAATGACGTTGTACCAGGC
K272A FP	39 nt	CACTATCTAAAGCACTTTGCCGAATTCATGGGGGCCCAG
K272A RP	39 nt	CTGGGCCCCCATGAATTCGGCAAAGTGCTTTAGATAGTG
R279A FP	39 nt	GAATTCATGGGGGCCCAGGCCGACTGGCAAAGCAACACG
R279A RP	39 nt	CGTGTTGCTTTGCCAGTCGGCCTGGGCCCCCATGAATTC
D280A FP	39 nt	TTCATGGGGGCCCAGCGTGCCTGGCAAAGCAACACGAAG
D280A RP	39 nt	CTTCGTGTTGCTTTGCCAGGCACGCTGGGCCCCCATGAA
Y287A FP	39 nt	TGGCAAAGCAACACGAAGGCCGGAAAGCCAGTCATGATT
Y287A RP	39 nt	AATCATGACTGGCTTTCCGGCCTTCGTGTTGCTTTGCCA
K289A FP	39 nt	AGCAACACGAAGTACGGAGCCCCGGTCATGATTAAAGGT
K289A RP	39 nt	ACCTTTAATCATGACCGGGGCTCCGTACTTCGTGTTGCT
P290A FP	39 nt	AACACGAAGTACGGAAAGGCCGTCATGATTAAAGGTGGA
P290A RP	39 nt	TCCACCTTTAATCATGACGGCCTTTCCGTACTTCGTGTT

Table 3.5 List of the oligonucleitides used in the present study along with their purpose

Name	Length	Sequence (5' to 3')	Purpose
M13FWD17	17 nt	GTTTTCCCAGTCACGAC	For DNA binding
M13FWD23	23 nt	CCCAGTCACGACGTTGTAAAACG	For helicase assay

3.2 Cloning of full-length and C-terminal region of Rep of *Tomato leaf curl New Delhi virus*

Full-length open reading frame (ORF) of AC1 was amplified from the monomeric clone of ToLCNDV DNA-A in pUC18 vector using Phusion enzyme and the blunt end PCR product was ligated into pJET1.2 vector. Further it was subcloned into pET28a(+) vector at *Nco*I and *Xho*I site. Using the same strategy and the restriction site, 726 bp fragment of the AC1 fragment corresponding to the C-terminal region spanning from amino acids 120 to 361 of Rep was PCR amplified from the monomeric clone of ToLCNDV DNA-A in pUC18 was cloned in pJET1.2 vector and finally into pET28a(+) vector.

3.2.1 Cloning of full-length and C-terminal region of Rep into pJET1.2 vector

3.2.1.1 PCR amplification by Phusion polymerase

Polymerase Chain Reaction (PCR) is an *in vitro* process used for amplification of specific sequences of DNA using small oligonucleotide called primers that can hybridize to opposite strands of the DNA. DNA amplification is achieved in the following steps: template denaturation, primer annealing and the extension of the annealed primers using the DNA polymerase enzyme using the protocol as described by Sambrook and Russell (2001).

Table 3.6 General composition of reaction used for PCR amplification

Composition of PCR reaction mixture	Final concentration
Reaction Buffer	1X (2.5 μl from 10X stock)
MgCl$_2$	0.5 μl (25 mM)
Forward and Reverse primer	0.1 μl+0.1 μl (from 100 mM stock)
Template	Plasmid DNA (20 ng) or plant genomic DNA (1 μg)
Phusion polymerase/Taq polymerase	1.5U-1U
Sterile distilled water	to adjust the ready volume 25 μl

The region corresponding to the full-length and C-terminal of the Rep was amplified from the monomeric clone of ToLCNDV DNA-A in pUC18 by PCR. On completion of the PCR reaction, amplification products were analyzed by agarose gel electrophoresis.

Table 3.7 PCR conditions used for the gene amplification

	Condition	No. of cycles
Initial denaturation	98ºC, 1.30 min	1
Denaturation	98ºC, 20 sec	
Annealing	60ºC, 30 sec	25
Polymerization	72ºC, 30 sec	
Final extension	72ºC, 10 min	1
Final storage	4ºC	1

Then the reaction sample was run on 1% agarose gel electrophoresis along with DNA marker of appropriate size, and was subsequently purified with a MinEluteTMGel extraction Kit (Qiagen, Germany) according to the manufacturer's instructions and ligated with pJET1.2 vector.

3.2.1.2 Ligation conditions to pJET1.2 vector

pJET1.2 vectors were used for directly cloning blunt end PCR products produced through amplification of high fidelity DNA polymerase. Ligation of the amplified product into the pJET vector takes less time and yield > 99% recombinant clones. The pJET vector contains a lethal restriction enzyme gene which is disrupted after ligation of a blunt end DNA insert at the cloning site. This results in colony formation from only bacterial cells harbouring recombinant plasmids. The pJET vector lacking the DNA insert expresses a lethal restriction enzyme on recircularizing that result in killing of the host *E. coli* cell after transformation.

Before setting up ligation of insert and vector, 2.0 µl of the insert DNA were analyzed on 1% agarose gel to estimate their concentration. The purified Blunt end PCR amplified AC1 gene and the pJET1.2 vector was ligated by using T4 DNA ligase and incubated for 16 h at 16^0C. The ng equivalent of vector DNA and insert DNA for 100 fmole [1 mole $\equiv 6.023 \times 10^{23}$ molecules and for a 1000 bp (1.0 kb) DNA, 1000×660 gm $\equiv 1$ mole] were calculated accordingly. One hundred fmole each of the vector and insert fragments were ligated in 1X T4 DNA ligase buffer with 1.0 µl T4 DNA ligase (MBI, Fermentas) in a final reaction volume of 20 µl at 16^0C for 16-18 h as per the manufacturer's instruction (MBI Fermentas). Only vector and T$_4$ DNA ligase in 1X T4 DNA ligase buffer were used as negative control/background for ligation

reaction. After completion of ligation, the reaction products were stored at -20^0C till further use for transformation.

Table 3.8 Ligation reaction of digested DNA fragment in vector

Ingredients	Volume (µl)
10X T4 DNA ligase buffer	1X
5U/µl T4 DNA ligase	5U
100 mM ATP	5 mM
vector and insert	In 1:3 molar ratio
SDW	Make volume upto 20 µl

3.2.1.3 Restriction digestion of pJET-RepF and pJET-RepC clones

The presence the RepF and RepC fragment insert in pJET vector was further confirmed by double digestion of the clone with *Nco*I and *Xho*I restriction enzymes (since the forward and reverse primer contains *Nco*I and *Xho*I sites, respectively).

The reaction mixture incubated at 37^0C for 4-5 h. Then the samples were loaded along with DNA marker on 1% agarose gel. The presence of band of approximately 1.1 kb and 700 bp confirmed the presence of the AC1 full length and C-terminal of the ORF, respectively. The digested fragment was further eluted and ligated into *Nco*I and *Xho*I digested pET28a (+) vector.

Table 3.9 Reaction for restriction digestion of plasmid DNA

Composition of restriction digestion Reaction mixture	Volume µl (Final concentration)
Restriction enzyme buffer	2 µl from 10X stock (1X)
Template DNA	Plasmid-2 to 6 µl/Viral DNA- 10 to15 µl (100 ng DNA/µl)
BSA (if required)	0.5 µl (0.1 mg/ml)
Restriction enzyme	0.5 µl to 1 µl (1 U/µg)
Sterile distilled water	To adjust the final volume 20 µl

3.2.2 Cloning of full-length and C-terminal region of Rep into pET28a (+) vector

3.2.2.1 *E. coli* competent cell preparation

Competent cells of *E. coli* strain DH10β and *E. coli* strain BL21 (DE3) were prepared as per protocol described by Mandel and Higa (1970) with some modifications. Strain of *E. coli* was streaked onto the luria agar plate. The plate was then incubated for overnight at 37^0C. Five ml seed culture was grown using single colony from the plate and further incubated overnight at 37^0C on shaker at 220 rpm. One ml of the culture was then inoculated into the 50ml LB media and incubated at 37^0C till OD at 600 nm reaches approximately 0.5. Culture was then cooled on ice for 15 min. The cooled culture was centrifuged at 6000 rpm for 5 min at 4^0C. After centrifugation, the supernatant was discarded and pellet was resuspended in 30 ml of ice cold sterile solution consisting of 6 ml of 0.1 M $CaCl_2$ and 24 ml of 0.1 M $MgCl_2$. The resuspended cells were incubated on ice for 30 min and then were centrifuged at 6000 rpm for 5 min at 4^0C. The pelletized cells were finally resuspended in 225 μl of 100% glycerol and 1275 μl of 100 mM $CaCl_2$. The cell suspension was gently mixed and aliquoted (75 μl) keeping on ice and immediately stored at -80^0C.

After overnight incubation at 16^oC, the ligated sample was transformed into *E.coli* DH10β competent cells.

3.2.2.2 Transformation of competent *E. coli* cells

Five microlitre of ligation mixture or 20 ng of the plasmid DNA was added to the competent cells and then mixed by gentle tapping and was kept on ice for 30 min. Heat shock was given to the cells in the culture at 42^0C for 2 min. Cells were immediately kept back on ice for 5 min. One ml of fresh LB without any antibiotic selection was added to the culture. Culture was then incubated at 37^0C for 1 h with shaking. The cell suspension was spun at 6000 rpm for 5 min at room temperature. Excess of LB was then discarded by simply inverting the micro centrifuge tube in a discard box inside the laminar. Pellet in the remaining LB media was mixed properly by using the pipette. The transformed cells were then plated on the antibiotic containing plate. Plate was then kept at 37^0C overnight. A single transformed colony was picked by a toothpick, which was scratched on the master plate and then inoculated into the test tube containing 2 ml of media supplemented with desired

antibiotic. For pJET clones 100 µg/ml ampicillin while for pET28a(+) clones 50 µg/ml kanamycin was used. The tubes and plate were incubated overnight at 37^0C. Next day plasmid DNA isolation was performed.

3.2.2.3 Plasmid DNA isolation from bacterial cells

Plasmid DNA from *E. coli* DH10β cells was isolated by alkaline lysis method as described by Birnboim and Doly (1979) with some modifications.

E. coli cells from 2 ml of overnight grown culture at 37^0C containing appropriate antibiotic were pelleted down at 13,000 rpm for 2 min at room temperature and the supernatant was discarded. Two hundred µl of solution-I [50 mM glucose, 25 mM Tris-HCl (pH 8.0), 10 mM EDTA pH 8.0, sterilized by autoclaving and stored at 4ºC] was added to the pellet and the cells were resuspended by vortexing. Then 200 µl of solution II (0.2 N NaOH, 1% SDS) was added and kept immediately on ice for 5 min, the solution was gently mixed by inverting the tube for 4 to 5 times. Subsequently, 200µl ice cold solution-III (for 2 ml: 1.2 ml of 5 M potassium acetate, 0.23 ml glacial acetic acid and 0.57 ml H_2O, pre-chilled) was added to the above mixture. The micro centrifuge tubes are again inverted 4 to 5 times while on ice for 10 min. Then equal volume of CI (24:1, chloroform: isoamylaclohol) mixture was added to the micro centrifuge tube. The mixture was then centrifuged at room temperature for 10 min at 13,000 rpm. The upper aqueous phase was transferred in a fresh micro centrifuge tube after the centrifugation and 0.8 volume of isopropanol was added to that. The tube was kept for 30 min at 4^0C and centrifuged at 13,000 rpm for 25 min at 4^0C. The supernatant was discarded and the pellet was resuspended in 500 µl of 70% ethanol by gentle tapping to loosen the pellet. The solution was spun down at 13,000 rpm for 10 min at 4^0C. Pellet was dried completely by vacuum centrifugation (Thermo Fisher, USA). Subsequently, the pellet was resuspended in 30 µl of sterile distilled water.

The quality of the DNA was checked by running the 2 µl of DNA along with the 1X DNA loading dye on agarose gel electrophoresis (Bio-Rad). The plasmid DNA solution was further stored at -20^0C for further use.

3.2.2.4 Ligation with pET28a(+) vector

Two µl of both the digested insert and vector DNA after gel elution was analyzed on 1% agarose gel to estimate its concentration prior to setting up ligation reaction. The

purified insert of either RepF or RepC and the pET28a(+) vector was ligated by using T4 DNA ligase and incubated for 16 h at 16^0C.

The 10 µl ligation mixtures of (pET28a +RepF) and (pET28a +RepC) were transformed into competent *E. coli* DH10β cells, plated onto 50 µg/ml kanamycin containing plates and incubated overnight at 37^0C. A single transformed colony was picked by a sterile toothpick, which was streaked on the master plate containing 50 µg/ml of kanamycin and then inoculated into the test tube containing 2 ml of kanamycin (50 µg/ml) containing luria broth media.

3.2.2.5 Confirmation of pET28a(+)-RepF and pET28a(+)-RepC clones

a. Mobility check analysis of the ligated clones

Three µl of the isolated plasmid DNA was electrophoresed on agarose gel along with the pET28a(+) control DNA which do not possess the insert. The vector possessing either the RepF or the RepC insert will run slower compared to the only pET28a(+). In this way the clones containing the insert were initially screened.

b. PCR amplification

To confirm the presence the RepF and RepC insert in pET28a vector, the selected clones were used as template for PCR. PCR was performed using the specific RepF and RepC primers along with positive control (AC1 template) and one negative control pET28a(+) plasmid DNA.Then the reaction sample was run on 1 % agarose gel electrophoresis along with DNA marker of appropriate size. The presence of band of approximately 1.1 kb and 700 bp confirmed the presence of the gene.

c. Restriction digestion confirmation

The presence of the insert of ~1.1 kb and ~700 bp size in pET28a(+)vector was further confirmed by digestion of clone plasmid DNA with *Nco*I and *Xho*I. Then the samples were loaded on 1% agarose gel, along with DNA marker. The presence of band of approximately 1.1 kb and 700 bp confirmed the presence of the desired insert.

3.3 Site directed mutagenesis of the AC1 ORF in pET28a(+) vector

Overlapping primers for site directed mutagenesis were designed that contained the desired mutation and was complementary to the template DNA around the mutation

site so that it can hybridize with the target DNA. Designed primers were synthesized at Sigma-Aldrich, Bangalore, India. The primers were diluted according to the manufacturer's recommendations and stocks of 100 µM were prepared for both forward and reverse primers. Further 1:10 times dilution i.e. working of 10 µM concentration was prepared.

The pET28a(+)-RepC clone was used as a template for site-directed mutagenesis to introduce mutations at the B' motif of the Rep protein using the overlapping primers mentioned in the primer list **Table 3.5**.

Table 3.10 Composition of PCR reaction used for the site directed mutagenesis in plasmid DNA

Ingredients	Volume (µl)
5x High Fidelity buffer	1X
10 mM dNTPs	1 mM
25 mM MgCl₂	300 nM
10 µM Forward primer	200 nM
10 µM Reverse primer	200 nM
10 U/µl Phusion /*Pfu* Turbo	5U
DMSO	1-5%
Template	50 ng
Sterile distilled water	Make volume upto 50 µl

Table 3.11 PCR condition used for site directed mutagenesis in plasmid DNA

	Condition	No. of cycles
Initial denaturation	95ºC, 4 min	1
Denaturation	95ºC, 45 sec	
Annealing	55ºC, 1 min	18
Polymerization	68ºC, 15 min	
Final Extension	68ºC, 30 min	1
Final storage	4ºC	1

Then 2.5 µl of the reaction was loaded on 1% agarose gel along with DNA marker of appropriate size and pET28a(+) as negative control. Following the reaction, 1 µl of *Dpn* I enzyme was added and incubated at 37^{0}C for 4 h. *Dpn* I cleaves only at the methylated sites, so it degrades the template plasmid but not the PCR product. The final reaction was transformed into DH10β competent cells and plated on 50 µg/ml

kanamycin containing plate. Finally a colony was picked, miniprep was done. All the mutants were sequenced to confirm the desired mutation and to rule out any error introduced during PCR amplifications.

3.4 Expression and purification of RepF and RepC in *E. coli* bacterial cells

The confirmed plasmids of pET-RepF and pET-RepC were transformed into *E. coli* BL21competent cells using the protocol as described in before in **section 3.2.2.2.**

3.4.1 Induction and overexpression profile of RepF and RepC proteins

The expression of Rep protein was optimized by varying different variables such as IPTG concentration, temperature, and time in *E. coli* strain BL21. One of the pET28a(+)-RepC transformed *E. coli* BL21 colony was inoculated in 50ml LB media containing kanamycin (50 µg/ml) and in one flask pET28a(+) vector transformed *E. coli* BL21colony was inoculated. The cultures were grown overnight at 37^0C at 220 rpm. One litre of secondary inoculation was done using 1% of the primary culture and the culture was allowed to grow at 37^0C for 3 h i.e. until the OD reached to 0.6. One ml culture was taken immediately before induction (non-induced control) and the cells were pelleted, and resuspended in lysis buffer. Samples were kept at -20^0C until needed for SDS-PAGE. Induction of expression of the gene was done by adding IPTG of 0.2 mM, 0.5 mM and 1 mM concentration in different culture flasks. Culture was grown at varied temperatures for expression standardization (37^0C and 16^0C) after attaining 0.5 at OD_{600}. The collected culture was pelleted at 6000 rpm for 7 min. Pellets were then stored at -80^0C until SDS-PAGE was performed.

3.4.2 Protein extraction and purification of RepF and RepC proteins

Cells were grown at 37^0C to an OD_{600} of 0.6. IPTG was added to final concentration of 0.2 mM and incubated with constant shaking at 220 rpm for 16 h at 16^0C. Pellets were resuspended in buffer containing 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 2 mM β-mercaptoethanol, 1 mM PMSF and 10% glycerol. Pellets were carefully dissolved to avoid frothing. Subsequently, lysozyme (0.1 mg/ml) was added to the dissolved pellets and freeze thaw cycles (liquid nitrogen - 37^0C cycle) were carried out until the lysate turns viscous (indicating proper lysis) and were incubated on ice for 30 min. The samples were sonicated for six cycles of 20 sec each at amplitude of 20%, with 50 seconds break after every cycle. Lysate was kept on ice at all the times.

The lysate was centrifuged at 13000 rpm at 4^0C for 30 min. The supernatant was filtered to avoid clogging of column and transferred to the fresh oakridges.

His-tagged protein was purified on Ni-NTA agarose beads (Sigma). Fresh Ni-NTA agarose (3-4 ml) was loaded in the Sigma column of total 10 ml bed volume capacity. Packed column was checked for any bubble as it will interfere with the protein binding. If there is any bubble, the column was agitated to remove it. The column was then washed thoroughly with 50-60 ml of water to remove any ethanol present. Column was never allowed to dry out during the whole process as it affects the binding capacity of the Ni-NTA to the His tagged protein.

Binding and elution was done as described by Gu *et al*. (1994) with some modifications. The column was kept in cold room and all the process of protein binding and purification was carried out in cold room. The extracted protein was passed through the column for 2-3 times. The column was then washed with 10 ml of lysis buffer. The column was then washed with 20 ml of wash buffer. The column was then washed with 10 ml of elution buffer. Fractions of 1 ml were collected in the micro centrifuge tube tubes.

3.4.3 Polyacrylamide gel electrophoresis

Appropriate volume of solution containing the desired concentration of acrylamide was prepared for the resolving gel as outlined in the **Table 3.12** below. The solution was mixed well after adding TEMED and quickly poured between the plates. One ml water was layered on the top of gel and kept aside for polymerization. Water was decanted once polymerization was over. Stacking gel solution (4%) was prepared and poured directly onto the surface of the polymerized resolving gel (10%). Comb was inserted immediately into the stacking gel solution and kept aside for polymerization. After polymerization, comb was removed and gel was fixed on the electrophoresis apparatus (Bio-Rad) and 1X SDS-running buffer was added to the top and bottom reservoirs of the electrophoresis apparatus.

The protein samples were prepared by mixing with that volume of 6X loading dye buffer such that the final working concentration becomes 1X. The mixture was boiled for 10 min in water bath. Samples were loaded and run at 8 V/cm (90 V) till dye front

moved up to the resolving gel. Voltage was increased to 15 V/cm (120 V) till the dye reached the bottom of the gel (Sambrook and Russell, 2001).

The gel was removed from the plates and transferred to 15 times the gel volume of staining solution and kept at room temperature with gentle shaking. The stained gel was transferred from the staining solution to the excess volume of destaining solution and kept at room temperature with shaking. Two-three changes of destaining solution made protein bands clearly visible with no background stain.

For performing western blotting a parallel SDS-PAGE gel was run and was not stained. The detail of the western blotting is provided in the preceding section.

Table 3.12 Composition of resolving and stacking gel solutions

	Resolving gel (10%)	**Stacking gel (5%)**
Constituents	Volume (ml)	Volume (ml)
H_2O	3.896	2.75
30% Acrylamide	3.3	0.67
1.5M Tris (pH-8.8)	2.6	-
1M Tris (pH-6.8)	-	0.5
10% SDS	0.1	0.04
10% APS	0.1	0.04
TEMED	0.004	0.004

3.4.4 Western blotting

Western blotting was done according to Towbin *et al*. (1979) with modification. A piece of PVDF membrane was kept in methanol for 30 min. Electro blotting was performed in the presence of transfer buffer at a constant voltage of 17 V for 30 min. The membrane was placed in blocking solution consisting of phosphate buffer saline (PBS) and 0.5% non skimmed milk powder for 2 h with gentle shaking at room temperaure. Membrane was washed with PBS + 0.1% Tween 20 (PBST) for 5 times. Primary antibody was diluted to 1:10,000 dilution in PBS containing 1% BSA and then added to membrane. The incubation was continued at room temperature for 1 h with gentle shaking. Membrane was again washed with PBS+0.1% Tween 20 for 5 min each 2 to 3 times to remove unbound antibody. The membrane was then

incubated with horseradish peroxidase conjugated secondary antibody diluted to 1:5000 dilution in PBS at room temperature for 1 h. The membrane was again washed with PBST to remove unbound fraction. Membrane was incubated with developing solution (1% DAB + PBS + H_2O_2) in dark for 10 to 15 min. The reaction stopped by rinsing the blot in water.

3.4.5 Preparation of dialysis tubing

The dialysis tubes were cut into pieces of convenient length (10-20 cm). Gloves were used while handling the dialysis bag. The tubing was boiled for 10 min. in a large volume of 2 % (w/v) sodium bicarbonate ($NaHCO_3$) and 1 mM EDTA (pH 8.0). The tubing was then rinsed thoroughly with distilled H_2O and was then boiled for 10 min in 1 mM EDTA (pH 8.0). The tubing was allowed to cool. Tubing was washed inside and out with distilled H_2O prior to dialysis.

3.4.6 Concentration of several elution fractions

Fresh CentriCon column (Millipore, cutoff 10 kD) was used for concentration of the protein. The column was equilibrated with lysis buffer. Elution fractions of Ni-NTA purified protein was pooled in the column. The column was centrifuged at 5000 rpm at 4^0C. The protein was concentrated to a final volume of ~ 0.5 ml and the final concentration was determined by measuring OD at 280 nm.

3.4.7 Quantification of protein concentration

OD of the protein sample (A) was taken at 280 nm. Concentration of the protein (c) in mg/ml was calculated using following formula: $c = A \times DF \times (MW.)/(\varepsilon_{280} \times l)$, where DF= dilution factor, ε_{280}= extinction coefficient and l= pathlength in cm. Using Expasy tool ProtParam (http://web.expasy.org/protparam/), the molecular weight (MW) of the His-RepF and His-RepC protein were found to be 42190.6 Da and 28643.2 Da, respectively, while the extinction coefficient were 54235 M^{-1} cm^{-1} and 43890 M^{-1} cm^{-1} respectively.

The protein concentration was also determined by using the Bradford reagent Sigma, USA. This assay was performed according to the manufacturer's recommendation in a 96-well plate.

3.5 Biochemical assays performed for wild-type and mutant RepC ptoteins

3.5.1 ATPase assay

In a standard reaction of 10 µl, the indicated amount of proteins were incubated along with 0.2 µCi of [γ-P^{32}] ATP in a buffer containing 20 mM Tris-HCl (pH 8.0), 1 mM ATP, 1 mM MgCl$_2$, 100 mM KCl, 8 mM DTT and 80 µg/ml BSA at 30^{0}C for 30 min. The reaction was terminated by addition of EDTA and 1 µl of the reaction mixture was spotted on polyethyleneimine-cellulose strips (Sigma-Aldrich, USA) and air-dried. Ascending chromatography was performed using 0.5 M LiCl and 1 M HCOOH as the running solvent. The TLC plate was air dried and exposed to imaging plate. The plate was read by phosphoimager (GE Amersham).

For nonradioactive calorimetric ATPase assay, the indicated amount of RepC protein was incubated 30^{0}C for 45 min with 2 mM ATP in 100 µl reaction mixture containing 1X ATPase buffer [50 mM Tris-HCl (pH 8) and 5 mM MgCl$_2$]. After stopping the reaction by adding 1 ml of stop solution (0.5% SDS, 2% H$_2$SO$_4$, and 0.5% ammonium molybdate), 10 µl of freshly prepared coloring reagent (10% ascorbic acid) was added and absorbance value at 750 nm was taken using a UV- spectrophotometer. ATPase activity was calculated from the standard curve for 0-250 nmol of inorganic phosphate. A$_{750}$ values could be converted to amounts of phosphate ion released per reaction by comparison with a standard curve of K$_2$HPO$_4$.

3.5.2 Fluorescence based ATPγs and ssDNA binding assay

Fluorescence measurements were carried out at 25^{0}C using Cary-Varian spectrofluorimeter with a spectral bandpass of 5 and 10 nm for excitation and emission, respectively. The binding of ATPγS and DNA (with ATPγS saturated protein) was measured by monitoring the intrinsic tryptophan fluorescence of either RepC or the mutants having concentration of 1.5-1.85 µM at 350 nm upon excitation at 295 nm. RepC was incubated with with a final concentration of 100 µM of ATPγS for 15 min prior to fluorescence titration with DNA. All fluorescence measurements were repeated at least thrice in buffer containing 50 mM Tris.SO$_4$ (pH 7.5), 1 mM MgSO$_4$, and 5 mM β-mercaptoethanol. Measured fluorescence spectra were corrected for buffer background, inner filter effects and dilution. All binding data were analyzed using a one-site saturation model and equation $\Delta F/F_o = B_{max} [L]/ (K_d + [L])$ where

F_o= Initial fluorescence, ΔF_c is the fluorescence intensity corrected for dilution as well as inner filter effects, $\Delta F = F_o - F_c$ is the difference between the initial fluorescence intensity and that obtained after titration with ligand, B_{max} is the maximal binding, [L] is the ligand concentration, and K_d is the dissociation constant for the interaction.

3.5.3 Helicase assay

For the helicase assay, the dsDNA substrate was prepared by annealing a ^{32}P- labelled 23 base oligonucleotide to circular M13mp18 ssDNA. This oligonucleotide has 17 nt annealed to the M13 ssDNA with a 6 nt 3′ overhang. Helicase assay was performed at 37°C in buffer containing 50 mM Tris-Cl pH 8.0, 10 mM MgCl$_2$, 2.5 mM DTT, 0.1 mg/ml BSA, 1 % glycerol, 1 mM ATP and 10 nM of substrates with indicated amount of either wild-type or mutant proteins, in a volume of 10 μl. After 30 min, the reaction was stopped by addition of (final concentrations) 8% glycerol, 0.48% SDS, 20 mm EDTA, and 0.02% bromophenol blue. The samples were subsequently loaded on 4 % native polyacrylamide gels and were runned using 0.5X TBE buffer at 16 V/cm for 40 min. The gels were dried on whatman paper and exposed to phosphorimager screen (GE Amersham).

3.5.4 Circular dichroism

Circular dichroism (CD) spectrum measurements of the wild-type and mutant RepC proteins were obtained at 25^{0}C in a Quartz cell with 1 mm light path with Chirascan Model of Applied Photosystem CD spectrophotometer (Greenfield, 2006). An average of 10 scans was corrected for buffer contributions. The secondary structure content of the proteins was predicted by using K2D and CDNN software.

3.6 Homology modelling of ToLCNDV RepC

The NCBI blast search against the protein data bank for ToLCNDV RepC had shown very less sequence similarity amongst known structures. Therefore, conserved motifs of ToLCNDV RepC were searched against PROSITE and PRINTS databases as the NCBI blast search against the protein data bank displayed less similarity at sequence level amongst the structures already known.

From the PHYRE analysis, E1 papilloma (2GXA- member of SF3 helicases) was ranked top structure for structural prediction of ToLCNDV. Three-dimensional model

of ToLCNDV RepC was generated using MODELLER package (Eswar *et al.*, 2007). The models were built using an alignment of region 304-577 from the crystal structure of Papillomavirus E1 and 104-361 residues of ToLCNDV Rep. The model was selected on the basis of their lowest DOPE score and highest GA341 assessment score and further validated by PROCHECK.

Furthermore, to obtain a hexameric state, modelled structure was superimposed on the crystal structure of Papillomavirus E1 (2GXA). Using the available hexameric ADP bound crystal structure of the template papillomavirus E1 protein, the ADP molecules were added at the active site of ToLCNDV Rep model. The molecular minimization simulations were done with AMBER molecular dynamics package using AMBER99SB force field and steepest descent algorithm to remove close van der Waals contacts, followed by conjugate gradient minimization until the energy was stable in sequential repetitions. All hydrogen atoms were included in the calculation.

3.7 Construction of infectious dimer of ToLCNDVs

3.7.1 Generation of mutations in the monomeric ToLCNDV clone

The monomeric clone of *Tomato leaf curl New Delhi virus* DNA-A (ToLCNDV DNA-A, Ac.no. HM007113) in pUC18 vector was already available in our laboratory. The pUC18- ToLCNDV monomeric clone was used as template to generate mutants in full length viral genome by SDM. Using the pUC18- ToLCNDV monomeric clone, mutations were introduced in full length viral genome at RepC region in the following positions (K227A, D261A, K272A, R279A, D280A, Y287A, K289A and P290A) by *Pfu Turbo* enzyme.The PCR amplified product was restriction digested with 1 µl of the *Dpn* I restriction enzyme (10 U/µl) by incubation at 37^0C for 4 h. Subsequently the reaction products were transformed into competent *E.coli* DH5α cells.

3.7.2 Cloning of tandem repeat ToLCNDV constructs in pCAMBIA2300 vector

The partial viral fragment of ToLCNDV genome was cloned in pCAMBIA2300 vector at *Kpn*I and *Xba*I site**.** In this partial clone, then the full length ~2.7 kb genome fragment containing mutations within the AC1 gene were mobilized from the monomeric clones. The tandem orientation was then confirmed by *Xba*I restriction digestion of the genome. The strategy used for the cloning has been outlined in the **Figure 3.1**.
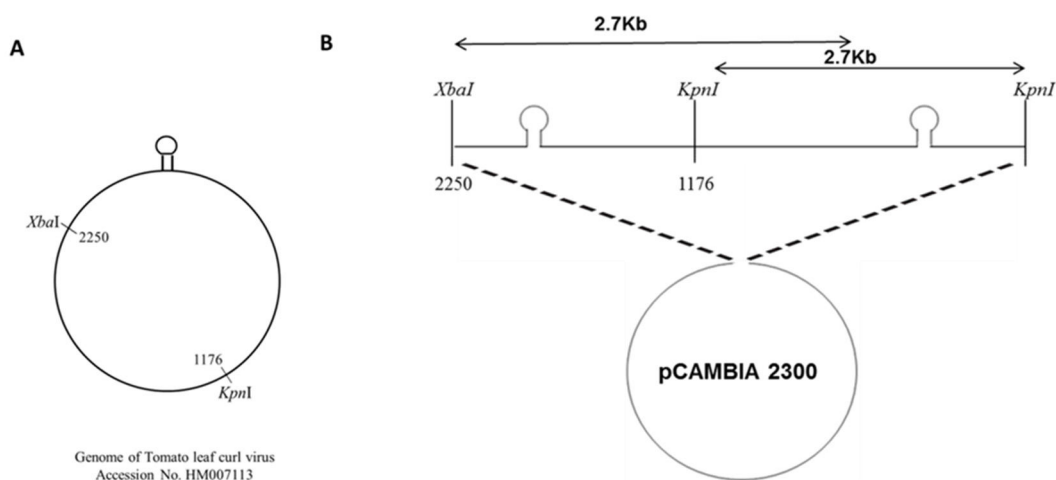
Figure 3.1 Generation of partial tandem repeat constructs (A) Schematic diagram showing the presence of *Xba*I and *Kpn*I restriction site in the genome of *Tomato leaf curl virus* (Accession No. HM007113) which are used for cloning of partial tandem repeat construct harbouring either wild type or mutant AC1 gene. (B) Strategy used for cloning of the partial tandem repeat constructs in pCAMBIA 2300. The full length ToLCNDV genomw was ligated at *Kpn*I site to the partial clone containing the *Xba*I and *Kpn*I digested fragment in pCAMBIA 2300.

3.7.3 Preparation of competent *Agrobacterium tumefaciens* EHA105 cells

Agrobacterium strain EHA105 was grown in 5 ml LB media containing 1% sterilized glucose and 30 µg/ml rifampicin antibiotic at 28^0C for 36-48 h. Two ml of grown culture was further inoculated into 50 ml LB media in a 250 ml conical flask. The culture was allowed to grow at 28^0C and 220 rpm until the 0.6 OD is reached. Culture was chilled for 5 min and 30 ml of culture was transferred to a centrifuge tube. Culture was harvested at 4^0C, 7000 rpm for 5 min. Supernatant was discarded, and pellet was resuspended in 10 ml of ice-cold 0.15 M NaCl and was allowed for 15 min incubation on ice. Again it was centrifuged at 7000 rpm for 5min and pellet was resuspended in 1ml of ice-cold 20 mM $CaCl_2$. After resuspension, 100 µl of competent cells was aliquoted. Competent cells were kept in liquid nitrogen for 5 min, immediately stored at -80^0C.

3.7.4 Transformation of competent *Agrobacterium tumefaciens* cells

Agrobacterium tumefaciens strain EHA105 was transformed using cold shock method, the competent cells were taken and kept on ice for two hours for thawing. One µg DNA (respective dimer in pCAMBIA 2300) was added and kept on ice for 30 min. After this Cold shock was given by dipping the cells in liquid nitrogen for 1 min and then thawed at room temperature. Cells were allowed to recover with

30 µg/ml rifampicin added LB media for 5 h at 28^0C incubator shaker. Transformed cells were plated on LB agar plates containing 30 µg/ml rifampicin and 50 µg/ml kanamycin, and were allowed to grow at 28^0C for 48 h. The colonies were inoculated into LB media containing 30 µg/ml rifampicin and 50 µg/ml kanamycin media and incubated at 28^0C, 220 rpm incubator shaker for 48 h. The plasmid was isolated and the PCR confirmed clones were further processed.

3.8 Plant Inoculation

Agrobacterium tumefaciens mediated inoculation of test plants was carried out following Chattopadhyay *et al.* (2008). For this, *Agrobacterium tumefaciens* strain EHA105 harbouring viral construct was inoculated in 100ml LB media supplemented with kanamycin (50 µg/ml) and rifampicin (30 µg/ml) and grown at 28^0C for 36 h at 200 rpm. Bacterial cells were pelletized at 6000 rpm for 15 min and pellet was resuspended in SDW. The OD of each inoculum was brought to 1.0 and the equimolar concentration of *Agrobacterium* transformed cells containing DNA-A and DNA-B tandem repeat constructs were used to inoculate plants. Inoculation of was carried out by pricking at the petiole and stem with a fine needle. Five-leaf stage *N. benthamiana* plants (two week old seedlings, one week after re-transplantation) are inoculated by pricking at the joint of third and second leaf from the apex and 2 µl of bacterial culture was added. All inoculated plants along with the plant inoculated with empty construct (mock) were allowed to grow at 28^0C in glass house chamber maintaining 16 h day and 8 h night period. Agroinoculated plants were kept under insect-free greenhouse condition. Symptom observation and assessment was done till 30 dpi. Inoculated plants were scored for symptom severity. Symptom analysis and scoring was performed. Samples were collected from inoculated plants at 14 dpi for further analysis.

3.9 Plant genomic DNA isolation

Total DNA was isolated from systemically infected leaves of *Nicotiana benthamiana*, as described by Dellaporta *et al.* (1983). The frozen samples were ground in liquid nitrogen at first and the finely grounded powders were added into extraction buffer containing 100 mM Tris-HCl pH 8, 50 mM EDTA pH 8, 50 mM NaCl, 10 mM β-mercaptoethanol. After addition of SDS, samples were incubated at 65^0C for 15 min followed by precipitation by 5 M sodium acetate. Supernatant was collected and

purified with chloroform isoamylalcohol (24:1) solution. Total DNA was precipitated after addition of isopropanol, followed by centrifugation at 14000g for 30 min. DNA pellet was washed with 70% ethanol and was air dried before finally dissolved into 200 μl of sterile double distilled water.

Presence of ToLCNDV wild-type and its mutant's genomic components (DNA-A and DNA-B) were tested in inoculated *N. benthamiana* by PCR. Total genomic DNA isolated from systematic leaves of either from infected or mock inoculated plants used template. Detection of viral genome was carried out using primers specific for AC1 and BC1 ORFs.

3.10 Southern hybridization

Southern hybridization was carried out using following steps:
1. *Separation of nucleic acid and transfer of DNA:* Viral DNA accumulation was detected in the inoculated plants using Southern hybridization (Southern, 1975). Southern hybridization was performed according to the protocol described in Sambrook and Russell (2001). For southern blot analysis, 8.0 μg of total DNA isolated from ToLCNDV DNA-A and DNA-B inoculated plants was electrophoresed in 0.8% agarose gel by applying voltage not more than 50 V to allow smooth and clear separation of viral DNA bands. Depurination is usually practiced during the transfer of high molecular weight DNA. Gel was treated with 0.25 M HCl for depurination for 10 min which causes the partial hydrolysis of the DNA before alkali denaturation and helps in the transfer of DNA. It was washed with deionised water and was treated with denaturation solution (0.5 M NaOH and 1.5 M NaCl) for 45 min with gentle agitation. The denatured gel was treated with neutralization solution consisting of 0.5 M Tris-HCl and 1.5 M NaCl twice for 30 min and 15 min, respectively. While the gel was in neutralizing solution, a glass plate was placed over a 10X SSC buffer containing tank and sufficient volume of 10X SSC [(20X SSC (3 M NaCl, 0.3 M Na-Citrate, pH 7.0)] was added. A filter paper (larger than size of the gel, should cover maximum area of glass plate) was placed in such a way so that both the sides were completely dipped in tank buffer and above then filter paper was saturated with 10 X SSC. After neutralization, the gel was placed on the top of the filter paper in inverted position and also no air bubbles remained trapped under the gel. Parafilm was kept over areas of filter paper not covered by gel to avoid direct

contact of filter paper and membrane. Pre-soaked nylon membrane of the same size of the gel was layered on the gel so that no air bubble remained trapped under the membrane. Three pieces of pre-soaked 3 mm whatmann filter papers were kept over the membrane. Blotting papers (paper towels) were placed on the filter paper and with a flat plate load on the top. The DNA was kept for transfer, overnight (Sambrook and Russell, 2001).

After 16-20 h, the blot was carefully removed with the gel and the marker positions were carefully marked. The blot was then placed under UV light for 1 min. for crosslinking of DNA with the membrane. The membrane was then placed between two layers of 3 mm whatman filters, wrapped properly and stored until use.

2. *Probe labelling*: Random probe labelling method was used (Feinberg and Vogelstein, 1983). The reaction mixture was prepared as follows: Random hexamer primer- 0.5 μl (0.2 μg/μl), DNA template- 4 μl (30 ng), sterile distilled water - 17 μl, which were mixed in a 1.5 ml eppendorf tube and boiled at 100ºC for 5 min, snap-chilled on ice for 3-4 min, then added Klenow buffer- 3 μl (10X), Klenow fragment - 1 μl (5 U/μl), α-[^{32}P]- 3 μl, and the reaction was incubated at 37^0C for 1 h. After the random labelling step was completed, free nucleotides were removed from labelled DNA by nucleotide removal kit (Qiagen, Germany) following manufacturer's instruction. The labelled DNA was denatured by boiling at 100^0C for 10 min and snap-chilled on ice for 5 min, and then added in the running pre hybridization solution and allowed to hybridize.

3. *Pre hybridization and hybridization of DNA bound membrane*: Prehybridization was done at 65ºC for 4-6 h using pre-hybridization buffer in a hybridizing oven. The blots were hybridized with denatured α-[^{32}P]-dCTP-labelled DNA probe at 106 cpm/ml, by adding it in pre-hybridization buffer. Hybridization was allowed for 12-16 h at 65ºC.

4. *Washing of blots*: Membranes were washed sequentially as follows: 2X SSC, 0.1% SDS; 1X SSC, 0.1% SDS; 0.5X SSC, 0.1% SDS; 0.2X SSC, 0.1% SDS for 10 min each at 65ºC until the non-specific counts had substantially reduced.

5. *Scanning and phosphomoimager analysis*: Membranes were air-dried and exposed to imaging plate. The image was read by Phosphoimager (GE Amersham).

RESULTS

4.1 Cloning, expression and purification of His-RepF and His-RepC protein

For overexpression of full-length and C-terminal of Rep protein, 1086bp full-length and 726 bp C-terminal fragment of geminivirus ToLCNDV ORF was cloned in pJET1.2 and subcloned in pET28a(+) expression vector at *Nco*I and *Xho*I restriction sites. The clones were confirmed by PCR amplification and also by restriction digestion. The cloned full-length AC1 ORF in pET-28a(+) showed an amplicon of ~1.1 kb by specific primers (**Figure 4.1 A, Lane 4 & 5)** and a fall out of same size fragment (~1.1 kb) with *Nco*I and *Xho*I restriction enzymes (**Figure 4.1 B, Lane 3**). While the cloned C-terminal AC1 fragment in pET-28a(+) showed an amplicon of ~700 bp by RepC specific primers (**Figure 4.1 A, Lane 1 & 2)** and a corresponding release of ~700 bp with *Nco*I and *Xho*I restriction enzymes (**Figure 4.1 B, Lane 1**).
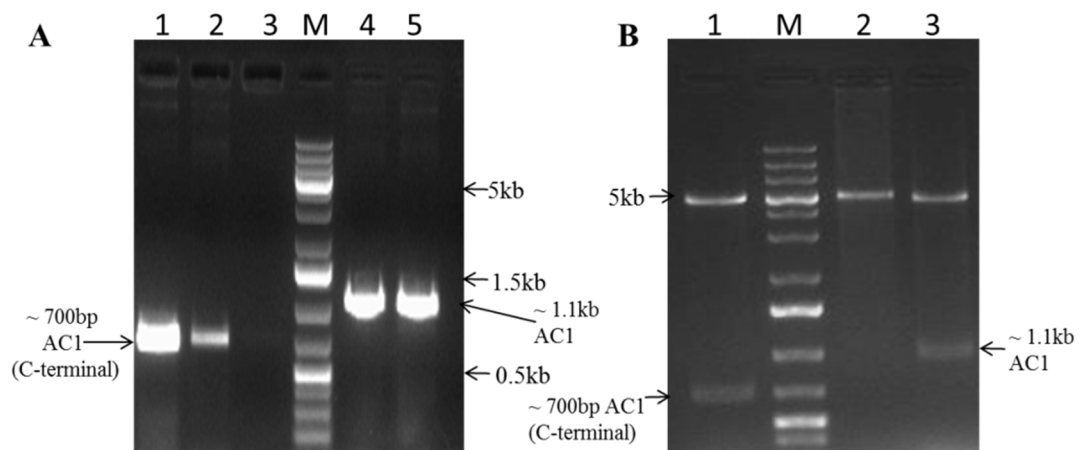
Figure. 4.1 Cloning of full-length and C-terminal of AC1 ORF in pET28a (+) vector (A) Agarose gel showing PCR using specific primers for AC1 and C terminal of AC1 ORF. Lane 1 & 2 and lane 4 & 5 showing the PCR amplicon of ~700 bp and ~1.1 kb corresponding to full-length and C-terminal clone of AC1 ORF, while pET28a(+) used as negative control in lane 3 showing no amplification. Lane M indicates 1 kb DNA ladder. (B) Agarose gel showing digestion of full-length and c-terminal clone of AC1 ORF in pET28a (+) vector with *Nco*I and *Xho*I. Lane 1 and Lane 3 showing ~700 bp and ~1.1 kb fragment release from C-terminal and full-length AC1 ORF clone in pET28a (+) vector. Lane M indicates 1 kb DNA ladder. The positions of three reference bands in the DNA marker (5 kb, 1.5 kb and 0.5 kb) are indicated.

The sequencing confirmed clones were transformed into *E. coli* BL21 (DE3) cells and overexpression conditions were standardized to select the IPTG concentration and temperature for optimum expression of the RepF (full-length Rep) protein and RepC (C-terminal Rep) protein. The overexpression of the ToLCNDV RepF and RepC in the total bacterial cell extract was monitored by SDS-PAGE. RepF and RepC proteins were successfully induced with 1 mM IPTG at various temperatures (37^0C, 28^0C and

16^0C) (**Figure 4.2 A & B**). The presence of expected molecular weight protein of ~ 43 kD for His-RepF and ~29 kD for His-RepC in the induced bacterial cultures was clearly observed from SDS-PAGE analysis which was absent in the uninduced bacterial cultures.
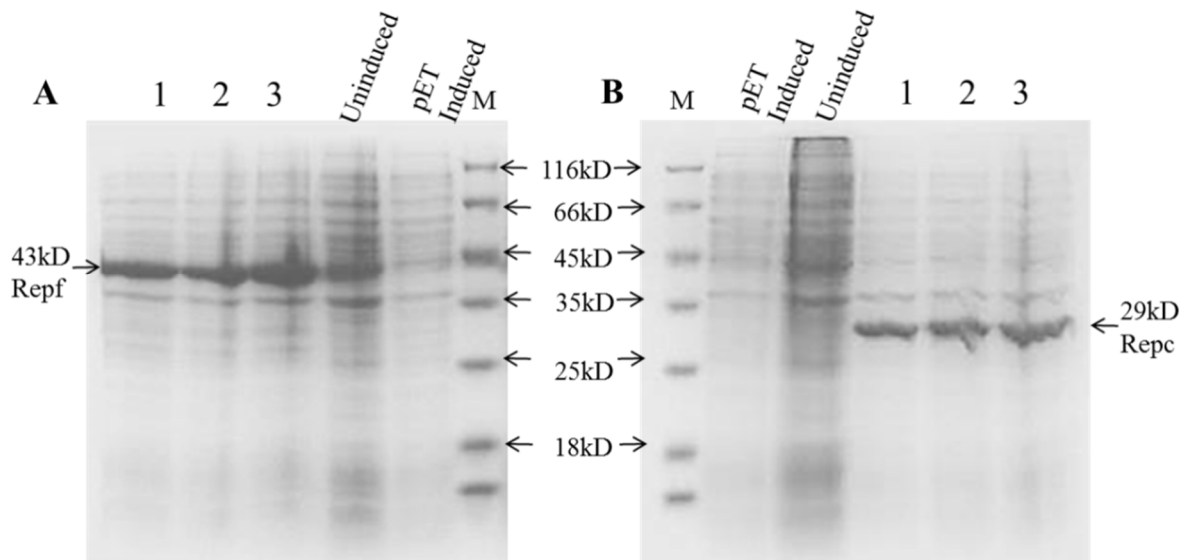
Figure. 4.2 Overexpression profile of His-RepF and His-RepC in *E. coli* BL21 cells. (A) SDS-PAGE (10%) showing the presence of His-RepF (~43 kD) in *E.coli* BL21 cells by inducing with 1mM IPTG at 37^0C (lane 1), 28^0C (lane 2) and 16^0C (lane 3). (B) SDS-PAGE showing the overexpression of His-RepC (~29 kD) in *E. coli* BL21 cells by inducing with 1mM IPTG at 37^0C (lane 1), 28^0C (lane 2) and 16^0C (lane 3). pET vector induced with 1 mM and uninduced *E. coli* BL21 culture were used as controls in the experiment. Lane M indicates the protein ladder. Corresponding molecular mass of the protein marker are indicated in kilodaltons (kD).

For the protein purification under native conditions, mobilization of the His-RepF and His-RepC in soluble fraction is necessary. Therefore, the cells after induction with different concentration of IPTG (0.1 mM, 0.5 mM and 1 mM) at 37^0C and 16^0C were further lysed by sonication and after centrifugation at 10,000 rpm for 30 min, the soluble fractions were loaded onto the (10%) SDS-PAGE along with the insoluble fractions. At 37^0C both RepF and RepC were largely present in the insoluble fraction (**Figure 4.3 A & B**). The efforts to increase the amount of His-RepF protein in the soluble fraction have proved very difficult. SDS-PAGE analysis of the His-RepF in the soluble and inclusion bodies fractions from the total bacterial lysate showed that very little difference in change in solubility of the His-RepF protein when induced with different IPTG (0.1 mM, 0.5 mM and 1 mM) concentration at 37^0C and 16^0C (**Figure 4.3 A & C**). However, the amount of the RepC protein in soluble fraction

was greatly increased at 16^0C which was evident from the SDS-PAGE showing reduced His-RepC in inclusion bodies and increased in the soluble fraction at 16^0C. It was found that about 30% of the His-RepC protein was present in the soluble fraction while 70% was detected in the insoluble fraction (**Figure 4.3 B & D**). Thus, induction with 0.1 mM IPTG at 16°C for 16 h was selected for optimum expression for His-RepC in the *E. coli* BL21 cells based on the supernatant-pellet profile. For His-RepF induction was also performed at 16°C for 16 h with 0.1 mM IPTG, however, the purification was carried out with larger amount of secondary culture.
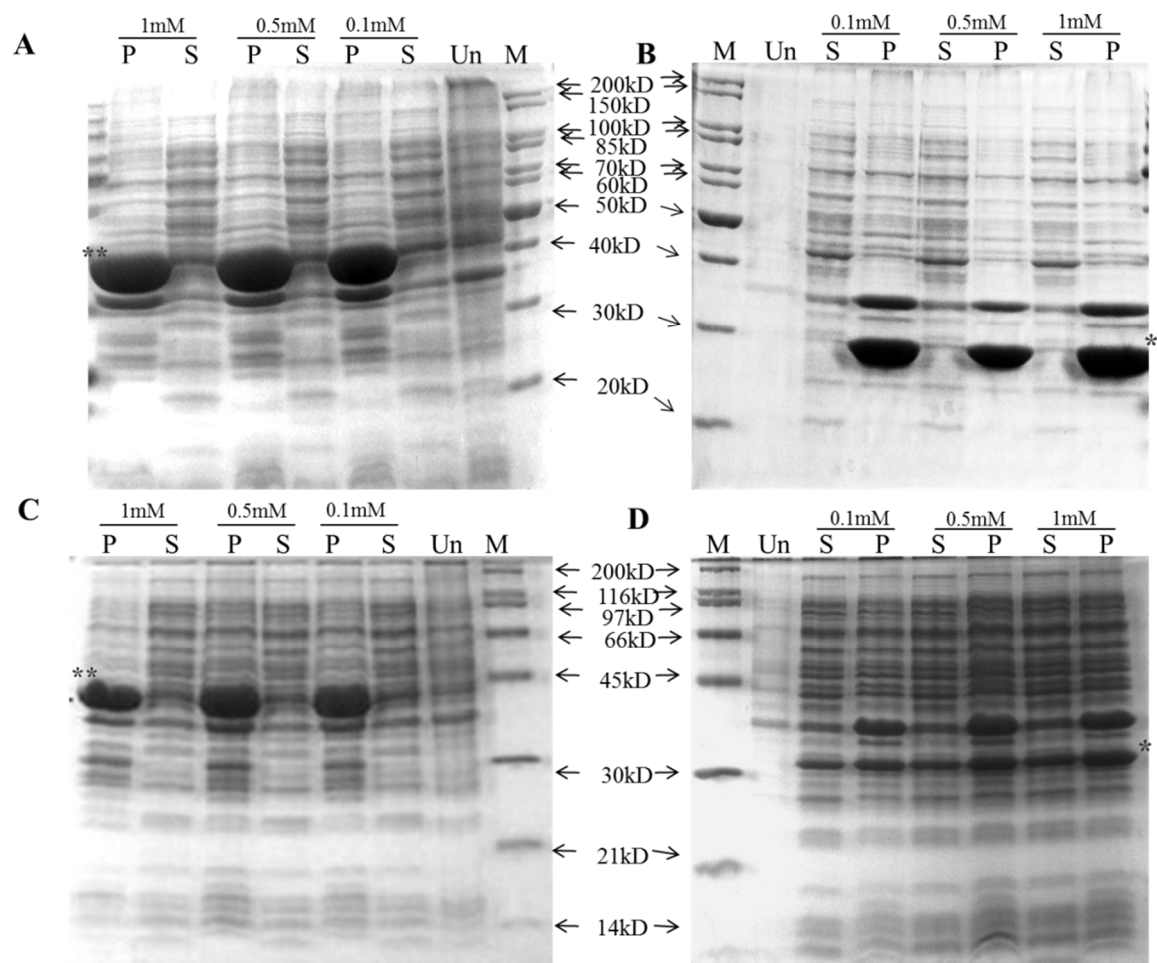
Figure. 4.3 Analysis of the soluble and insoluble fractions of total bacterial lysate overexpressing His-RepF and His-RepC in *E.coli* BL21 cells. (A & B) SDS-PAGE showing the supernatant and pellet profile of His-RepF and His-RepC, respectively at 37^0C. (C & D) SDS-PAGE showing the supernatant and pellet profile of His-RepF and His-RepC, respectively at 16^0C. M-protein marker, Un-uninduced, P-pellet, S-supernant. Double asterisk (**) indicates the RepF protein (~43 kD) while single asterisk (*) indicates the RepC (~29 kD) protein. Corresponding molecular mass of the protein marker are indicated in kilodaltons (kD).

The *E. coli* BL21 bacterial cell lysate harbouring the induced His-RepF (with 0.1 mM IPTG) was passed through the Ni-NTA agarose matrix to immobilize the fusion protein. To determine the optimum imidazole concentration for efficient elution of the RepF protein, gradient elution using increasing concentration of imidazole was carried out. The SDS-PAGE analysis of the collected protein elutes along with the uninduced and induced soluble bacterial fraction was performed. As observed on the SDS-PAGE, the His-RepF started to get elute with 250 mM imidazole (**Figure 4.4 A**). Similar elution profile was obtained for RepC (**Figure 4.4 B**). More RepF and RepC proteins were obtained by repeating purification more than once. The purified RepF and RepC protein were dialysed to remove imidazole and were concentrated using centricon column. The protein concentration was determined by measuring OD at 280 nm using spectrophotometer as well as by bradford assay.
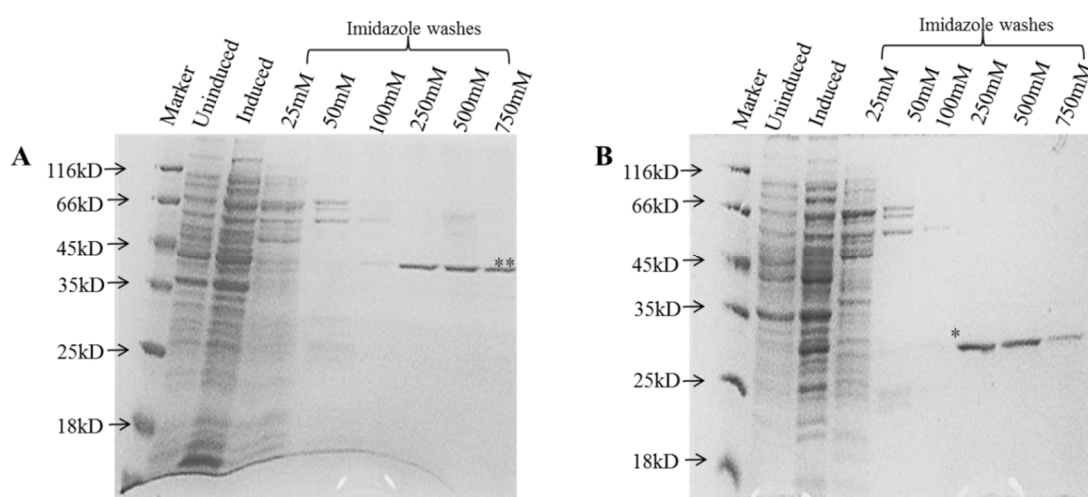
Figure 4.4 Purification profile of His-RepF and His-RepC protein. (A) SDS-PAGE showing gradient elution of His-RepF protein with different imidazole concentration (25mM-750 mM) as indicated. (B) SDS-PAGE showing gradient elution of His-RepC protein with different imidazole concentration (25-750 mM) as indicated. M-Marker; uninduced culture showed the absence of the overexpressed RepF and RepC protein; soluble fraction of the induced culture showed overexpressed RepF and RepC protein respectively. Double asterisk (**) indicates the RepF protein (~43 kD) while single asterisk (*) indicates the RepC protein (~29 kD). Corresponding molecular mass of the protein marker are indicated in kilodaltons (kD).

The purified RepF and RepC protein were confirmed through MALDI analysis of the trypsin digested peptide fragments of the respective proteins from SDS-PAGE gel. **Figure 4.5 A** shows the western blotting for the detection of purified His-RepC as well as His-RepC protein in the induced total bacterial lysate with anti-His antibody while no signal was obtained in the uninduced bacterial culture. SDS-PAGE (10%)

analysis of purified RepC and RepF proteins showed that Ni-NTA chromatography could yield almost purified RepC and RepF proteins (**Figure 4.5 B**).
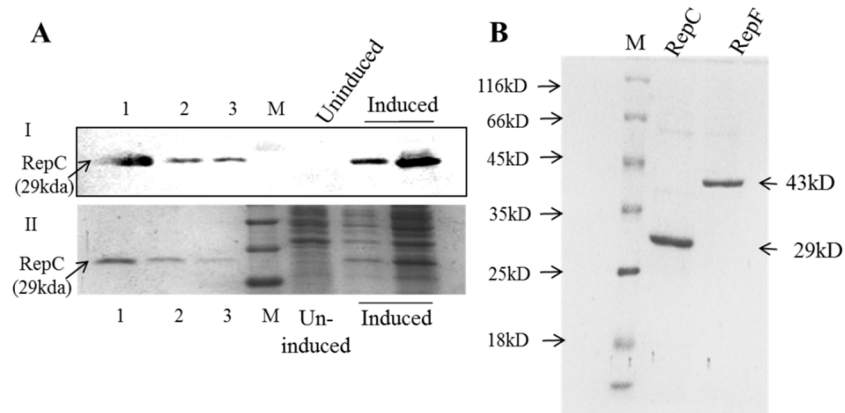
Figure 4.5 Purification of RepC and RepF (A) Detection of His tagged RepC using anti-His antibody (I) Western blot and (II) the corresponding coomassie brilliant blue stained SDS-PAGE. Lane 1-3 contains decreasing amount of purified His-RepC; M-Marker; His-RepC was detected using anti-His antibody in induced culture; while uninduced culture was used as negative control. (B) SDS-PAGE showing the purified, dialysed and concentrated RepC and RepF protein. Corresponding molecular mass of the protein marker are indicated in kilodaltons (kD).

4.2 Characterization of ATPase activity of RepC

The C-terminal of the Rep protein of the geminivirus contains the ATPase activity. In order to determine the ATP hydrolysis activity, the purified RepC and RepF proteins (~300 ng) were incubated with radiolabelled [γ-P^{32}]ATP at 30^0C for 30 min and
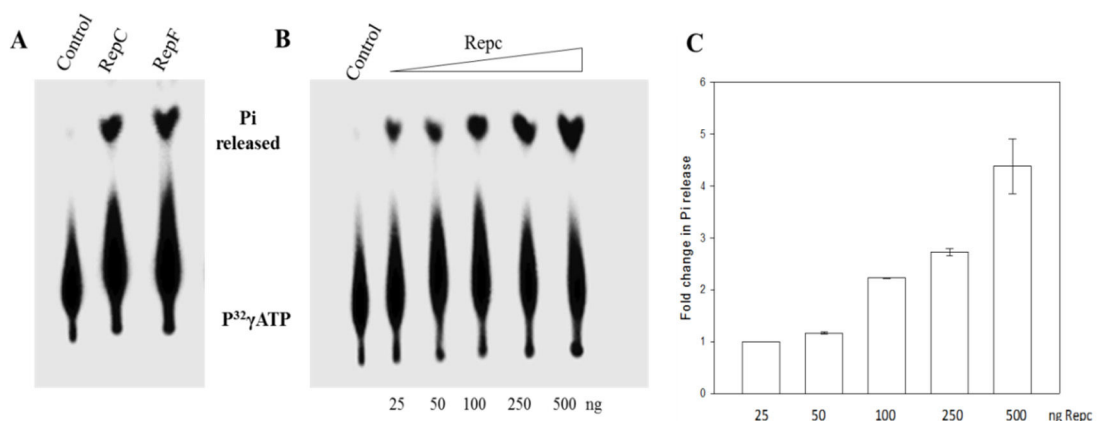
Figure 4.6 ATPase activity of RepC and RepF (A) Autoradiograph showing the ATP hydrolysis (ATPase) activity of His-RepC and His-RepF. In ATPase assay, 0.2μCi of [γ-P^{32}] ATP (6000 Ci/mmol) was incubated with 300ng of RepC and RepF protein for 30 min at 30ºC. Position of P^{32} labelled ATP and released Pi is indicated. (B) Autoradiograph showing the ATPase activity at increasing concentration of wild-type RepC protein. Amount of the protein (in ng) used is indicated below each lane (C) the quantified plot of graph (B) showing the fold change in Pi release by increasing amount of RepC protein.
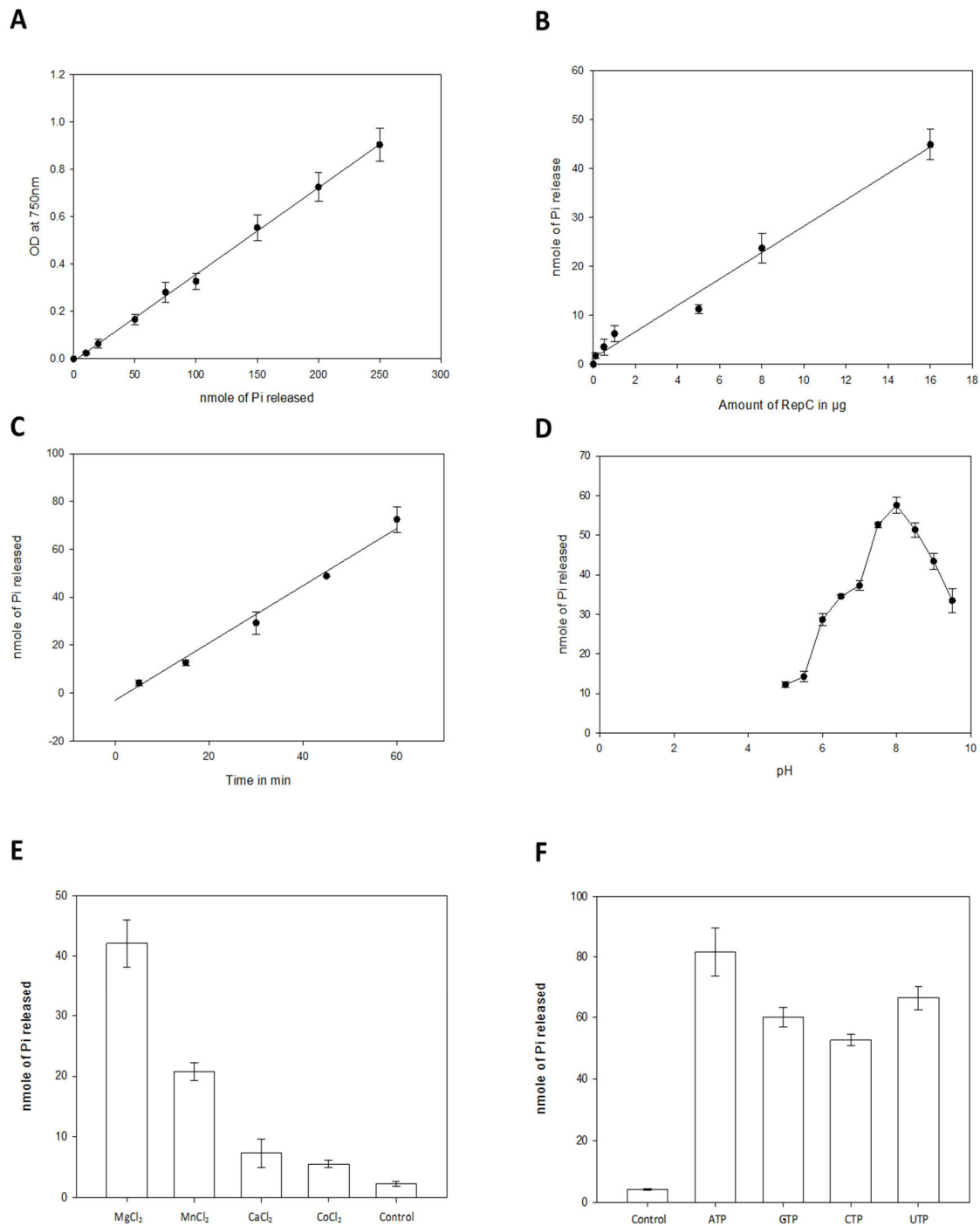
Figure 4.7 Biochemical characterization of the RepC mediated ATPase activity. (A) Standard plot using K_2HPO_4 for 0-250 nmol of inorganic phosphate. (B)ATPase activity of RepC at increasing concentrations of purified RepC protein (0.25 µg-16 µg) in the presence of 1 mM ATP. (C) ATPase activity of RepC (10 µg) for varying lengths of time as indicated in the plotted graph. (D) pH profile of RepC (10 µg) ATPase activity. The ATPase assay was performed in the pH ranges of 5-6.0 (by using 50 mM MES) and 6.5-9.5 (by using 50 mM Tris-HCl). (E) Divalent cation dependent ATPase activity of RepC (10 µg) in presence of 5 mM each of $MgCl_2$, $MnCl_2$, $CaCl_2$ and $CoCl_2$. The control reaction contains only RepC protein. (F) Substrate specificity of RepC. The ATPase assay was performed at 30^0C for 45 min in the presence of 5 mM each NTP used in independent sets of the assay mixture. The control reaction contains only RepC protein.

the ATPase activity was measured in terms of the amount of the Pi liberated from $[\gamma\text{-}P^{32}]$ATP. After the reaction, products were separated on PEI-TLC plates by ascending chromatography and was followed by autoradiography It was evident from the autoradiograph that both the purified His-RepF and His-RepC proteins were biologically active (**Figure 4.6 A**). Almost equal amount of released Pi from $[\gamma\text{-}P^{32}]$ATP was observed in case of RepC and RepF which indicates that both the proteins exhibited comparable ATPase activity. The amount of the Pi release from $[\gamma\text{-}P^{32}]$ATP was found to increase with increasing amount of the RepC protein thus, showing that the activity is protein concentration dependent (**Figure 4.6 B & C**).

4.3 RepC exhibits general Ribonucleoside tri-phosphatase (NTPase) activity

The ATPase activity of RepC protein was also monitored by another assay where ATP hydrolysis was measured colorimetrically as a function of Pi release. During the course of the hydrolysis reaction, a reduced phosphomolybdate complex was formed which is blue in colour and can be easily quantified by measuring OD at 750 nm.

The amount of Pi release was estimated from the standard curve for 0-100 nmol of Pi plotted from the K_2HPO_4 (**Figure 4.7 A**). The rate of ATP hydrolysis was found to be linear function of RepC protein concentration (**Figure 4.7 B**). Time course analysis of ATPase activity of RepC showed that the ATP hydrolysis was time dependent. It was found to be linear over the time period of 60 min (**Figure 4.7 C**). The pH dependence on RepC mediated ATPase activity revealed that the activity was optimum over a range of 7.5-8.5 which is in agreement with the activity of the RepC protein since the pI of the RepC protein was 6.0 (**Figure 4.7 D**). ATPase activity of the RepC protein was highly dependent on the presence of divalent cations in order of $Mg^{2+}>Mn^{2+}>Ca^{2+}=Co^{2+}$ (**Figure 4.7 E**). The ability of RepC protein to utilize other NTPs (i.e. GTP, CTP and UTP) was also studied using this assay. The hydrolysis was found to be maximum for ATP comparative to other NTPs (**Figure 4.7 F**). The NTP hydrolysis was observed in the following order ATP>GTP=UTP>CTP.

To validate the above observations, the effect of the presence of different NTPs (unlabelled) on the process of ATP hydrolysis of the radiolabelled $[\gamma\text{-}P^{32}]$ATP by RepC protein was carried out. It was found that 5mM of various unlabelled NTPs

(both deoxy and ribose forms) can inhibit the RepC mediated ATPase activity, thus suggesting that they can act as competitors of the ATP hydrolysis reaction carried out by RepC (**Figure 4.8)**. However, the observed extent of inhibition of ATPase activity of RepC by different NTPs was not same.
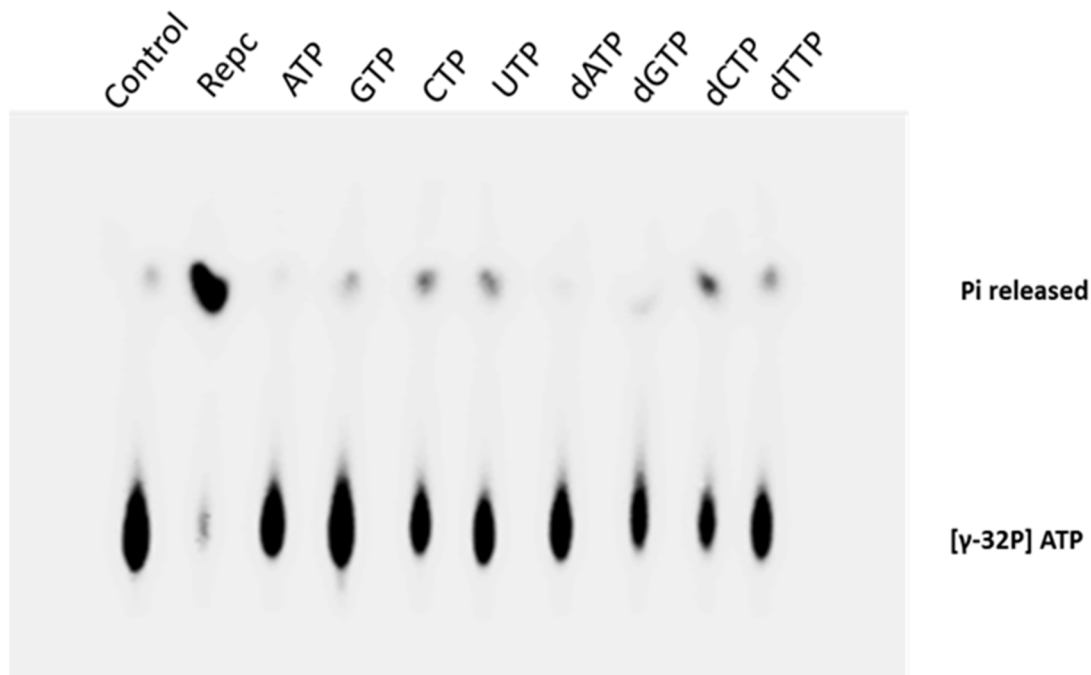
Figure 4.8 ATPase activity of wild-type RepC protein in presence of different NTPs. In ATPase assay, 0.2 µCi of [γ-P^{32}] ATP (6000 Ci/mmol) was incubated in a buffer containing 20mM Tris-HCl (pH 8.0), 1 mM MgCl$_2$, 100 mM KCl, 8mM DTT and 80 µg/ml BSA at 30^{0}C for 60 min with 500 ng of RepC protein either in the absence or presence of 5 mM NTPs in a total reaction volume of 10 µl. One µl of the reaction mixture was spotted on a polyethyleneimine (PEI) thin layer chromatography (TLC) plate and air dried and autoradiographed**.** Positions of P^{32} labelled ATP and released Pi are indicated.

To make it more evident, per cent inhibition of the ATPase activity of the wild-type RepC protein with increasing concentration (50nM – 2mM) of cold NTPs was measured which also revealed that the degree of inhibition varies **(Figure 4.9 A & B)**. It was found that ATP irrespective of its deoxyribose form (dATP) or ribose form (rATP) exhibited far greater inhibition to ATPase activity of RepC as compared to that of other dNTPs or rNTPs **(Figure 4.9 C & D)**.
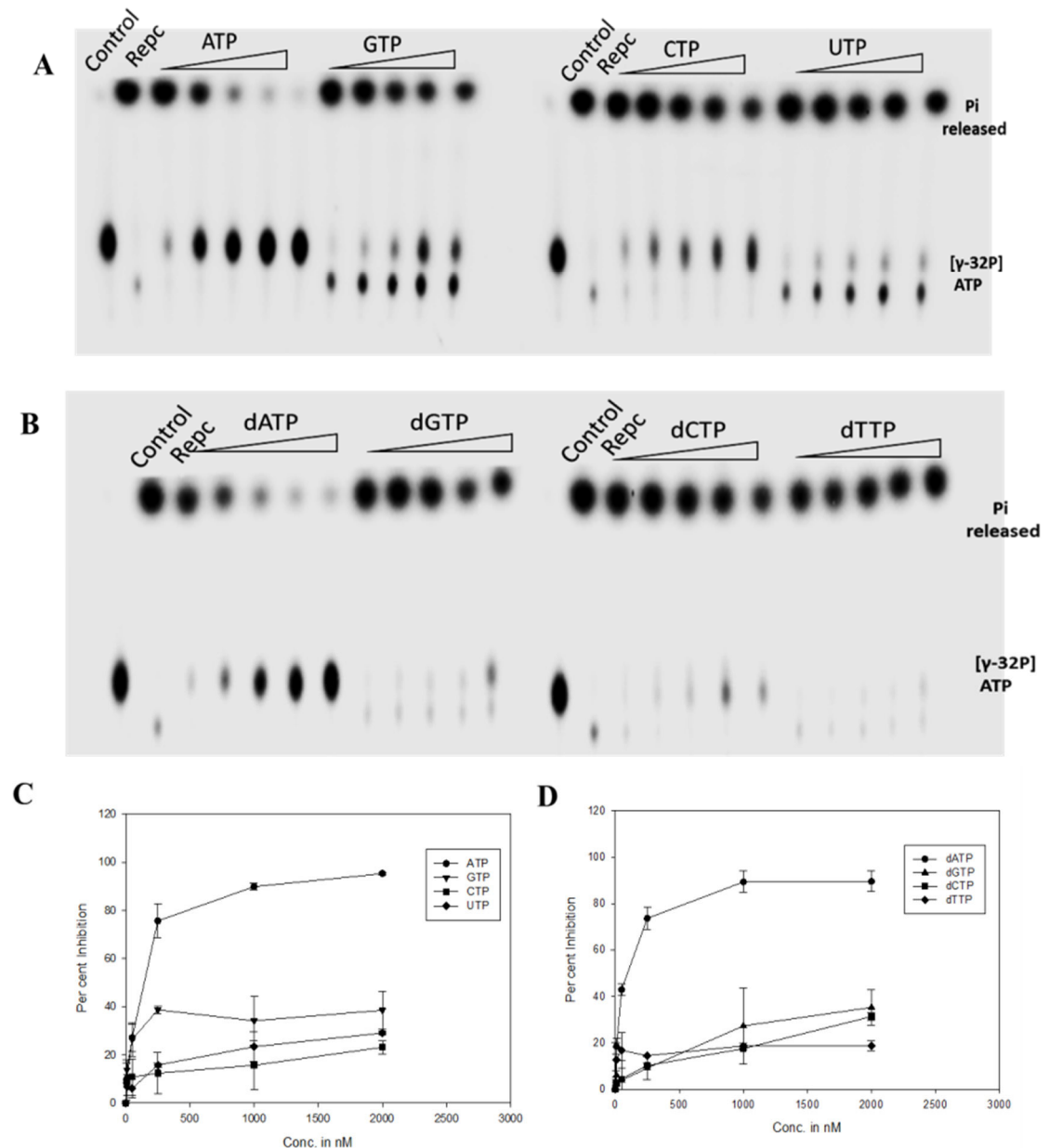
Figure 4.9 ATPase activity of wild-type RepC protein in presence of different concentration of various NTPs. (A) and (B) In ATPAse assay, 0.2 µCi of [γ-^{32}P] ATP (6000 Ci/mmol) was incubated at 30^{0}C for 60 min with 500ng of RepC protein in the presence of increasing concentration of (50-2000 nM) rNTPs and dNTPs, respectively. Positions of P^{32} labelled ATP and released Pi are indicated. (C) and (D) Graph showing the per cent inhibition of radiolabelled-ATP hydrolysis in the presence of increasing concentration of rNTPs and dNTPs respectively.

4.4 Binding efficiency of RepC protein with various NTPs and with ssDNA in presence of different NTPs

To further understand the NTP binding of wild-type RepC protein, fluorescence based binding studies of RepC protein with various NTPs were performed. Wild-type RepC

(~1.5 μM) was titrated with ATPγS and other NTPs i.e. ATP, GTP, UTP and CTP along with dNTPs as substrate and their binding affinities were determined by measuring the change in intrinsic fluorescence of RepC protein. Unlike ATP, ATPγS cannot be hydrolyzed but can bind to Rep just like ATP. All the values were corrected for background and dilution. The data obtained was analyzed and graphical representation was prepared that indicated the change in fluorescence to the initial fluorescence i.e. DelF/F$_o$ observed on increasing the substrate concentration. All NTPs had the similar Kd value which is the measure of the binding affinity to the RepC protein (**Table 4.1**). The affinities of RepC for ssDNA binding in the presence of different NTPs were also found to be similar (**Table 4.2**).

Table 4.1 NTP binding of wild-type ToLCNDV RepC protein using fluorimetric based experiments.

Ligand	Kd	R^2
ATPγS	$(1.4016\pm0.199) \times 10^{-6}$ M	0.9715
GTP	$(1.8866\pm0.3858) \times 10^{-6}$ M	0.9472
UTP	$(1.7223\pm0.3812) \times 10^{-6}$ M	0.9374
CTP	$(2.0532 \pm0.1220) \times 10^{-6}$ M	0.9951
ATP	$(1.1748\pm0.3768) \times 10^{-6}$ M	0.9603
dATP	$(0.8427\pm0.1431) \times 10^{-6}$ M	0.9623
dGTP	$(0.9977\pm0.2541) \times 10^{-6}$ M	0.9149
dCTP	$(0.7467\pm0.2277) \times 10^{-6}$ M	0.9417
dTTP	$(0.8637\pm0.2276) \times 10^{-6}$ M	0.9096

The error bar represents the standard deviation from two independent experiments.

Table 4.2 DNA binding of wild-type ToLCNDV RepC protein in the presence of various NTPs using fluorimetric based experiments.

Ligand	Kd	R^2
ATPγS+ DNA	$(37.875 \pm6.14)\times 10^{-9}$ M	0.9753
GTP+ DNA	$(40.1430\pm5.5313)\times 10^{-9}$ M	0.9829
UTP+ DNA	$(38.4877\pm4.3425)\times 10^{-9}$ M	0.9890
CTP+ DNA	$(35.3771\pm3.7672)\times 10^{-9}$ M	0.9780

The error bar represents the standard deviation from two independent experiments.

4.5 Secondary structure content of the wild-type RepC protein

Using the Sable tool (www.sable.cchmc.org), the amino acid sequence of the ToLCNDV RepC was analysed for its secondary content and solvent accessibility (**Figure 4.10 A**). The results showed that the Rep protein is highly solvent accessible and contains no transmembrane domain. The Rep protein was found to contain mainly random coil. The CD spectrum of the wild-type His-RepC protein also showed that purified RepC protein is structured and contains mainly loop region (**Figure 4.10 B***).* The CD spectra was found to change upon ATP binding and DNA binding in the presence of ATP.
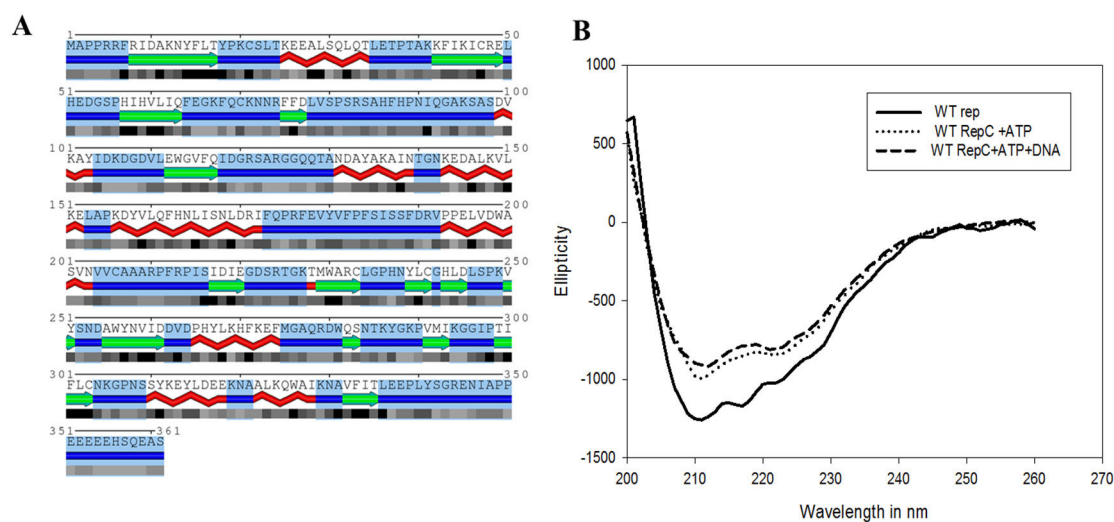
Figure 4.10 Secondary structure content of the ToLCNDV Rep protein. (A) Schematic diagram showing the α-helix (in red), β-sheet (in green) and random coil (higlighted in blue) regions present in the full-length ToLCNDV RepC protein depicted by Sable tool. With different shades of black is depicted probability of the solvent accessibility of different residues of the Rep protein (B) CD spectra of the wild-type RepC (~0.1 mg/ml final concentration) in the absence or presence of ATP and DNA (+ATP) in 10 mM phosphate buffer (pH 8.0), 50 mM NaCl and 10% glycerol.

4.6 Double-stranded DNA unwinding activity of RepC

The RepC protein is an unwinding enzyme. *In vitro* unwinding assay was carried out using a radiolabelled forked dsDNA helicase substrate or probe. The helicase activity of RepC was determined using a partial duplex DNA substrate that consists of 5' P^{32}-labelled 23-mer oligonucleotide partially annealed to the M13 circular ssDNA. This 23-mer oligonucleotide has 17 bp complementary region and 6 nt 3' overhang.
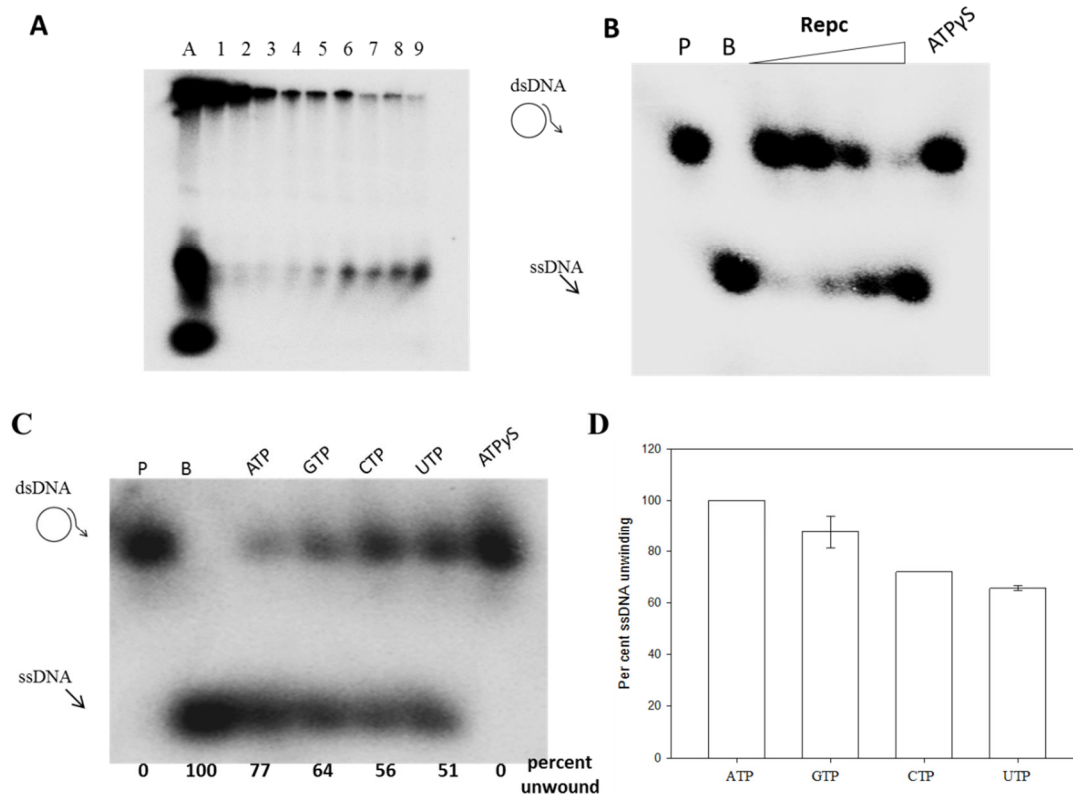
Figure 4.11 DNA unwinding activity of wild-type RepC protein. (A) Autoradiograph showing the dsDNA helicase substrate preparation. Single stranded circular M13 DNA along with ssDNA was boiled and further left for slow cooling to produce annealed product which was purified through gel filtration using sepharose CL-4B beads. Lane A- annealed product, lane 1-9 elutes collected after gel filtration using sepharose CL-4B beads. (B) Autoradiograph showing the unwinding reactions carried out by wild-type RepC as a function of protein concentration (1X-10X; X=0.1 µg). The lane marked P is the annealed dsDNA helicase probe while the lane marked B is the annealed dsDNA helicase substrate heated to 100°C for 5 min before loading onto the gel. Lane P and Lane B serve as control for the experiment. (C) Autoradiograph showing the unwinding reactions carried out by (~0.8 µg) wild-type RepC in the presence of 1 mM NTPs. The reaction was carried out for 30 min at 25°C using wild-type RepC with 1 mM of the indicated nucleotide. (D) Plot of the helicase activities quantified from the autoradiograph as shown in panel C.

The annealed product was then subjected to further purification to remove the free unannealed ssDNA using gel filtration Sepharose CL-4B beads. Following the purification different elutes were runned on 4% native polyacrylamide gel (**Figure 4.11 A**). Elutes which did not contain any free ssDNA was used as substrate for the helicase assay. The dsDNA helicase substrate was incubated with increasing concentration of RepC and reaction was stopped by adding 1X loading dye following which it was loaded onto 4% native polyacrylamide gel. It is noteworthy that unwinding assay of the Ni-NTA purified ToLCNDV RepC protein resulted in degradation of the radiolabelled ssDNA being released from the dsDNA helicase

substrate indicating the presence of nuclease contamination in the protein preparation. Therefore, the Ni-NTA purified RepC proteins were further purified using heparin chromatography and Q-sepharose chromatography. With the increasing concentration of this purified RepC protein, an increase in the amount of the radiolabelled ssDNA from the dsDNA helicase substrate in the presence of 1 mM ATP was observed (**Figure 4.11 B**). Apart from that no ssDNA was found to be released from helicase substrate by RepC protein in the presence of ATPγS which is non-hydrolyzable form of ATP suggesting that unwinding process being driven by ATP hydrolysis. Thus, ToLCNDV RepC protein can unwind the dsDNA utilizing the energy from ATP hydrolysis. To further determine the unwinding of dsDNA by RepC in the presence of other NTPs, same experiment was performed in the presence of 1 mM of GTP, CTP and UTP. The dsDNA unwinding by RepC was found to be driven by hydrolysis of ATP and other NTPs as well (**Figure 4.11 C**). While in the presence of ATP the unwinding activity of dsDNA by ~0.8 μg RepC protein was found to be 77%, the unwinding activity percent was reduced to ~60-50% in the presence of other NTPs i.e. GTP, CTP and UTP.

4.7 Site-directed mutagenesis of RepC, expression and purification of RepC mutants

Apart from the ATP binding and hydrolysis, the C-terminal of the geminivirus Rep protein is also involved in DNA binding and DNA unwinding activities, and Rep is classified as a member of the SF3 family of helicases (Choudhury *et al*., 2006; Clerot and Bernardi, 2006; George *et al*., 2014). The B' motif was predicted previously in the modelled structure of ToLCGV Rep generated by using Bovine papilloma virus E1 protein as template (George *et al.*, 2014). As an extrapolation to that the present study has been carried out, but was focussed on ToLCNDV Rep protein. The amino acid sequence analysis of the ToLCNDV Rep protein also showed that it contained all the classic features present in SF3 helicases like Walker A, Walker B and C motif. In addition to that B' motif was also predicted in the ToLCNDV Rep protein.

The multiple sequence alignment of helicase motifs of geminivirus Reps with other members of SF3 family showed various conserved amino acids within Walker A, Walker B, B' motif and C motif (**Figure 4.12 A**). The B' motif of Rep showed a much greater degree of conservation among members of various genera of the

Geminiviridae family (**Figure 4.12 B)**. The conserved nature of the B' motif in geminivirus Rep suggested critical role of these amino acids of B' motif in DNA unwinding. Various conserved amino acids of the B' motif of ToLCNDV Rep i. e. K272, R279, D280, K289 and P290 were targeted and their roles were further investigated by mutating them to alanine (**Figure 4.12 B)**.
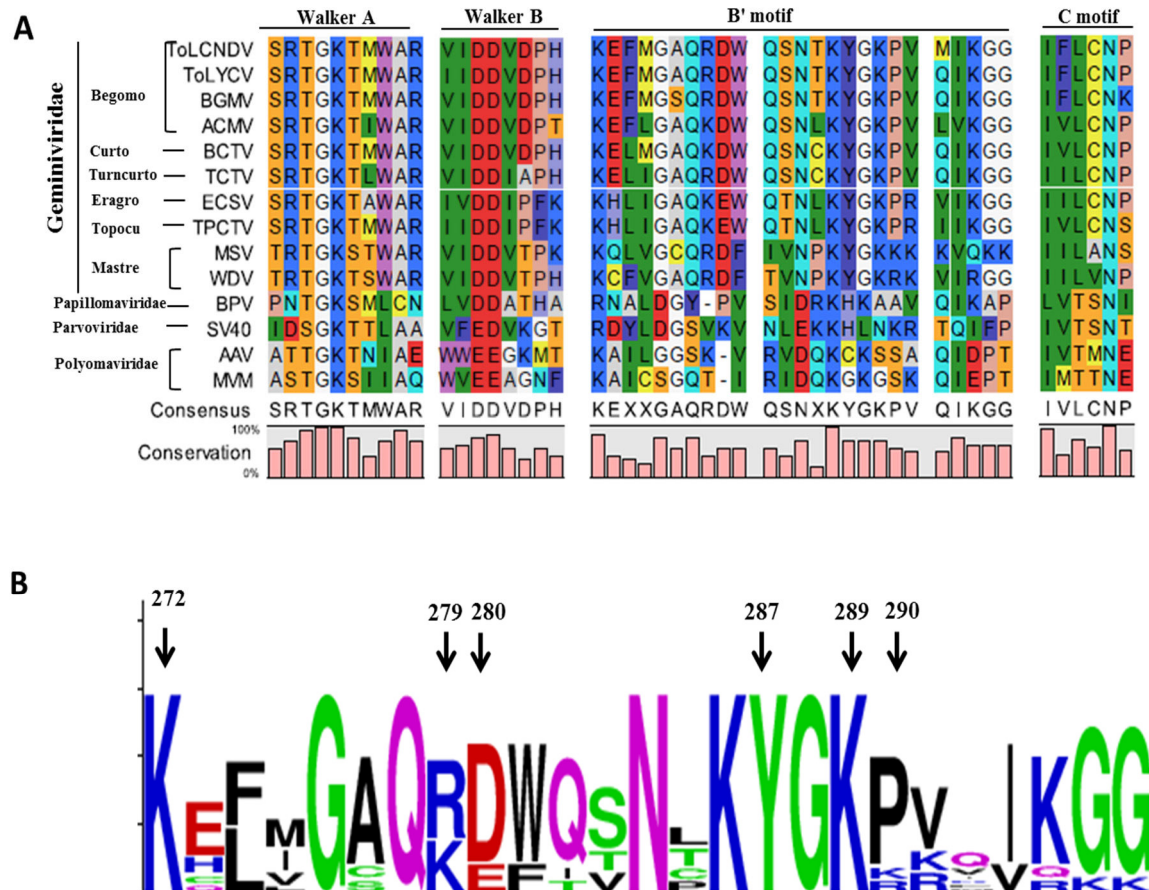
Figure 4.12 Analysis of the conserved signature SF3 motifs in ToLCNDV Rep protein (A) Multiple alignment of Walker A, Walker B, B motif and C motif of replication initiator proteins from the members of the SF3 family from ssDNA and dsDNA viruses using CLC Workbench. *Tomato leaf curl virus*, ToLCNDV-Rep; *Tomato leaf yellow curl virus*, ToLYCV; *Bean golden mosaic virus*, BGMV; *African cassava mosaic virus*, ACMV-Rep; *Beet curly top virus*, BCTV-Rep; Turnip curly top virus, TCTV-Rep; Tomato pseudo curly top virus, TPCTV-Rep; Maize streak virus, MSV-Rep; Wheat dwarf virus, WDV-Rep; Bovine papillomavirus,BPV-E1; Simian virus 40, SV40 Large-T Antigen; Adeno-associated virus2, AAV2-Rep78; Minute virus of mice, MVM-NS. (B) Sequence conservation of B' motif among various representative members of the geminivirus family (Tomato leaf curl virus, ToLCNDV-Rep; *Tomato leaf yellow curl virus*, ToLYCV; *Bean golden mosaic virus*, BGMV; *African cassava mosaic virus*, ACMV-Rep; *Beet curly top virus*, BCTV-Rep; *Turnip curly top virus*, TCTV-Rep; *Tomato pseudo curly top virus*, TPCTV-Rep; *Maize streak virus*, MSV-Rep; *Wheat dwarf virus*, WDV-Rep;) depicted using WebLogo (Crooks *et al*., 2004). With the arrow are indicated the various amino acids of ToLCNDV RepC investigated.
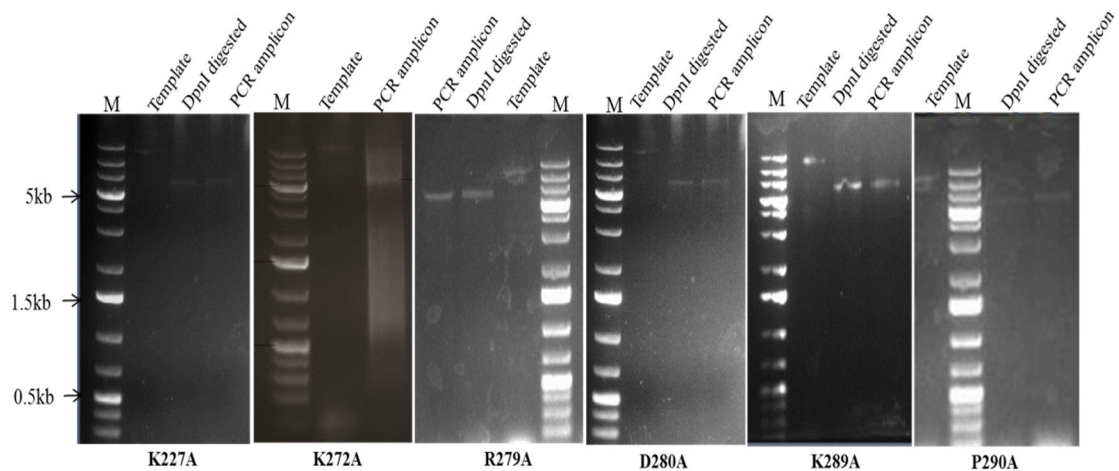
Figure 4.13 Site directed mutagenesis PCR of wild-type pET28a(+)-RepC plasmid using *Pfu* Turbo enzyme and overlapping primers. Various RepC mutations, as mentioned below the corresponding gel were generated. The gels show the PCR products (before and after *Dpn*I digestion) along with the template used [pET28a(+)-RepC plasmid] for each mutation. Lanes with pET28a(+)-RepC template, PCR product (showing size corresponding to pET vector i.e. ~6 kb) and *Dpn*I digested PCR product are indicated above for each gel. Lane M is 1 kb DNA ladder.

In vitro site-specific RepC mutation were created in double stranded pET28a(+)-RepC plasmid using the overlapping primers that contained the desired mutation. Alanine substitution mutagenesis PCR was performed by using high fidelity *Pfu* Turbo enzyme. The PCR amplified product was subjected to *Dpn*I digestion to remove the parental methylated and hemimethylated plasmid. PCR amplicons and *Dpn*I treated amplicons for each site directed mutagenesis were visualized by agarose gel electrophoresis in which bands corresponding to the size of pET28a-RepC plasmid (~6kb) were found in few cases indicating amplification of the complete plasmid (**Figure 4.13**). In some case smear was also observed indicating the presence of partially amplified PCR products. All the RepC mutants were further verified by sequencing.

The confirmed RepC mutant plasmid clones were transformed into *E. coli* BL21 cells for their overexpression followed by their subsequent purification. K227A, D261A, K272A, R279A, D280A, K289A and P290A RepC mutants were induced with 0.1mM IPTG, at 16^0C for 16 h as that used for the wild-type RepC. The SDS-PAGE analysis of the induced total bacterial lysate containing mutant RepC proteins along with lysate of the wild-type RepC confirmed the successful overexpression of the
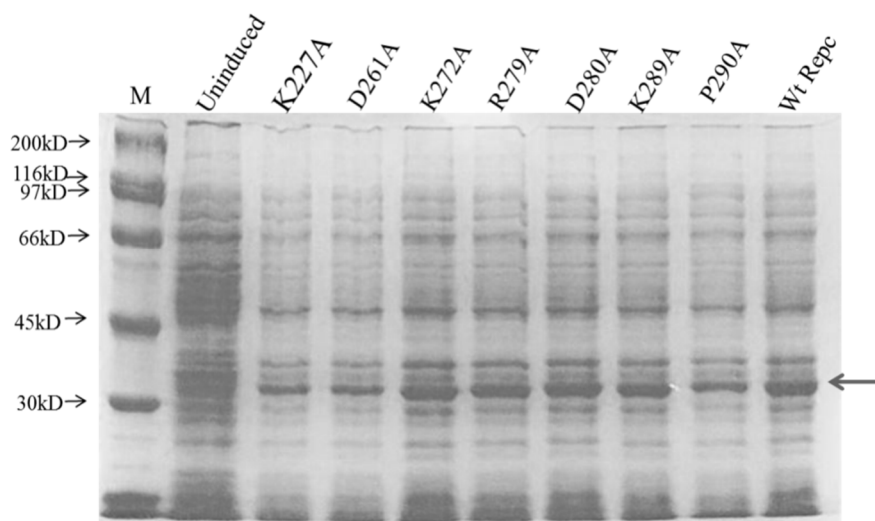
Figure 4.14 Overexpression profile of RepC mutants in *E. coli* BL21 DE3 cells. SDS-PAGE (10%) showing the overexpression of various RepC mutant proteins along with the wild-type RepC protein (Wt RepC) in *E.coli* BL21 cells by inducing with 0.1mM IPTG at 16°C for 16 h. RepC mutant proteins are indicated above each lane. Corresponding molecular mass of the protein marker are indicated in kilodaltons (kD). The arrow is indicated corresponding to the ~29kD protein band of either wild-type or mutant RepC.
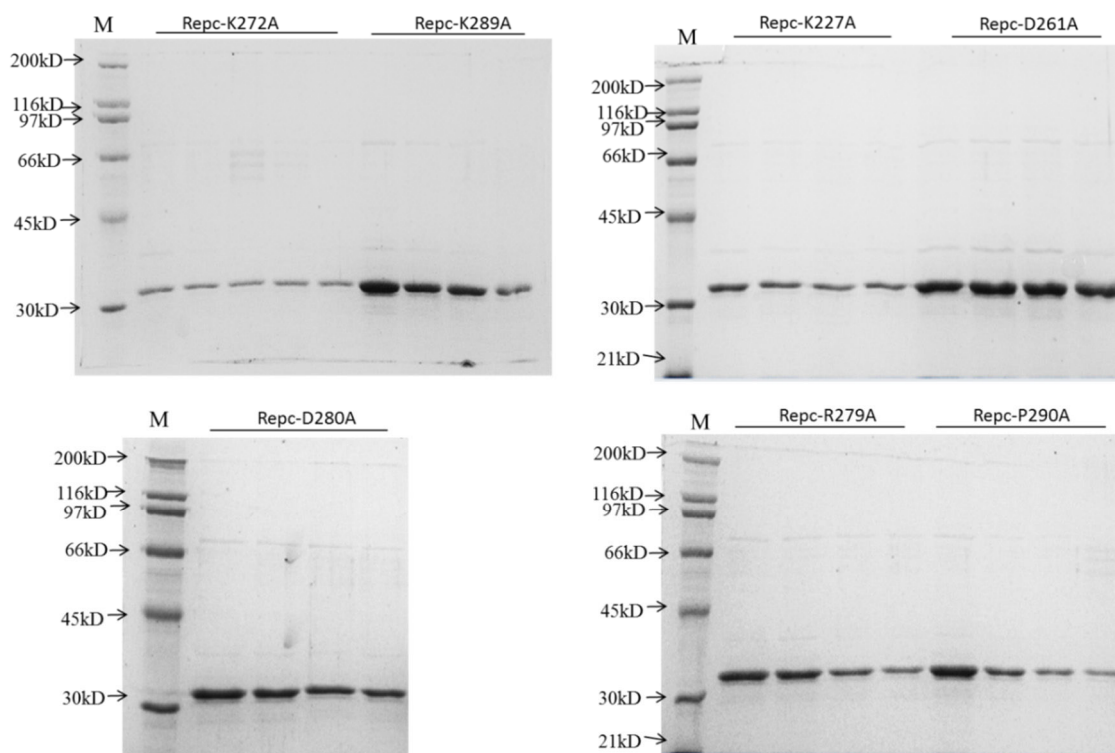
Figure 4.15 Purification profile of various His-RepC mutant proteins. SDS-PAGE showing the Ni-NTA eluted His-RepC mutant proteins (indicated above the gel) with 250mM imidazole of the estimated molecular weight of ~29 kD size. Corresponding molecular mass of the protein marker are indicated in kilodaltons (kD).

mutant RepC proteins (**Figure 4.14**). However, Y287A RepC protein failed to overexpress in *E. coli* BL21 cells and thus could not be purified. **Figure 4.15** shows the purity of various His-RepC mutants after Ni-NTA purification. The purified mutant proteins were further dialysed, concentrated and quantified.

4.8 Effect of RepC mutations on structural integrity and hydrolysis of ATP

Mutation of amino acid in a protein may result in alteration in secondary structure of the protein which may further lead to functional defect in protein. Structural integrity of the wild-type and mutants were tested by circular dichorism (CD) spectroscopy. The analysis of the CD spectrum of a protein gives the content of various secondary structure i.e. α-helix, β-sheet and random coil present in the protein. This also helps in assessing the structural change on ligand binding. To monitor the effect of RepC mutations on the structure of the RepC protein, CD was performed on the purified His tagged RepC protein along with various mutants. The CD spectra of purified RepC mutant proteins (K227A, D261A, K272A, R279A, D280A, K289A and P290A) were found to be similar to that of the wild-type RepC protein (**Figure 4.16 A**).
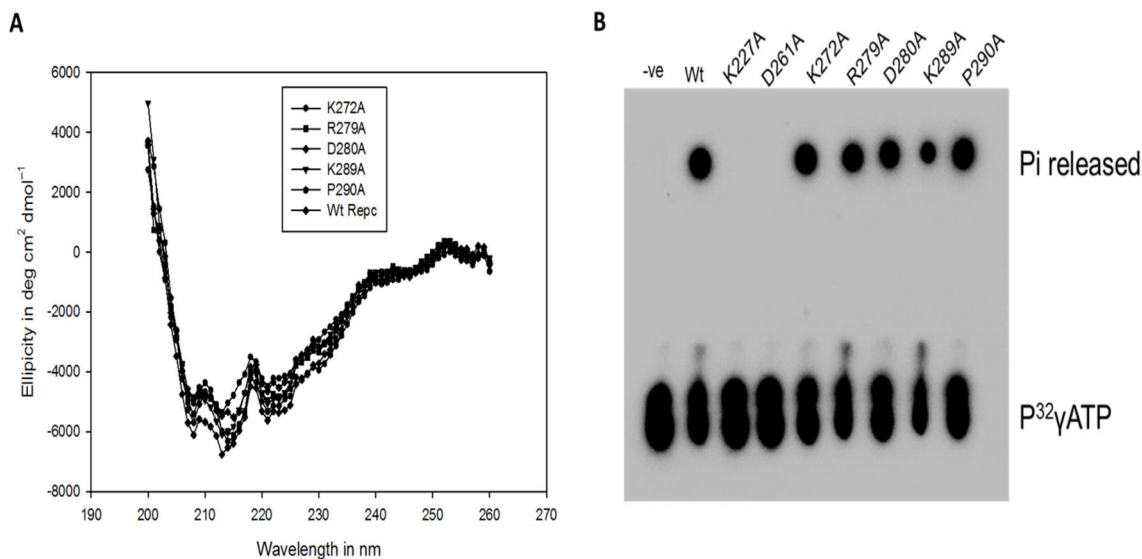
Figure 4.16 Effect of mutations on the overall structure of RepC protein. (A) CD spectra of the wild-type RepC along with all the mutants (~0.1 mg/ml concentration) in 10 mM phosphate buffer (pH 8.0), 50 mM NaCl and 10% glycerol (B) Autoradiograph of ATPase activity of wild-type and mutant RepC. In ATPase assay, 0.2μCi of [γ-^{32}P] ATP (6000 Ci/mmol) was incubated with 300 ng of either wild-type or mutant RepC proteins. Positions of P^{32} labelled ATP and released Pi are indicated. The respective mutations are indicated above the blot.

This indicates that the mutations result in no conformational alteration in RepC protein and structural integrity was maintained. The ATP hydrolysis activity of these B' motif mutants were found to be retained similar to that of wild-type RepC protein, thus indicating no effect of mutations on the RepC mediated ATPase activity. While K227A and D261A being Walker A and Walker B mutations respectively were ATPase defective as had been reported previously (**Figure 4.16 B**).

4.9 Effect of B' motif mutations on the helicase activity of ToLCNDV RepC protein

The dsDNA unwinding activity of ToLCNDV RepC was previously studied using dsDNA substrate consisting of 5' γP^{32}-labelled 23-mer oligonucleotide that was partially annealed to M13 ss circular DNA as dicussed in **section 4.6**.
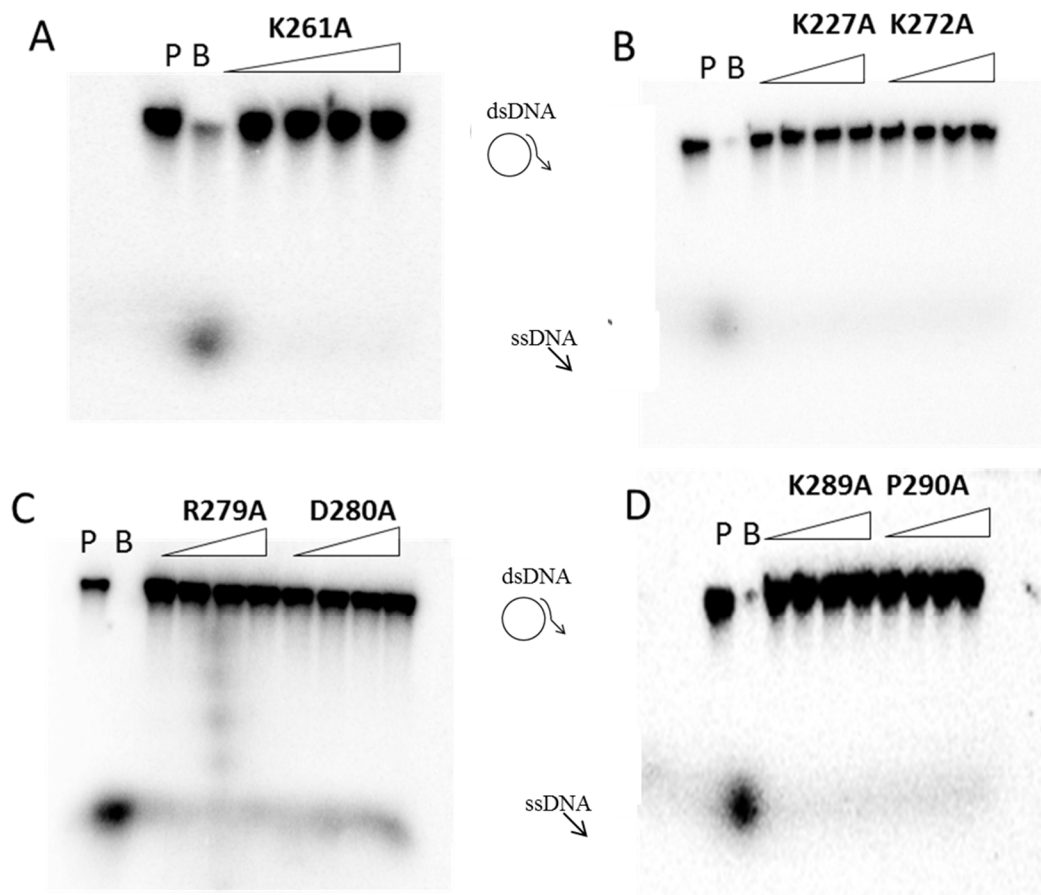
Figure 4.17 dsDNA unwinding activity of various RepC mutant proteins. (A-D) The autoradiograph of gels of the unwinding reactions carried out by increasing concentration of mutant RepC as a function of protein concentration (range from 0.1-1 μg). The reaction mixtures in the lanes marked P is the annealed dsDNA helicase probe while the lane marked B is the annealed dsDNA helicase substrate heated to 100°C for 5 min before loading onto the gel.

No detectable amount of ssDNA from dsDNA helicase substrate by K227A, D261A, K272A, K289A and P290A RepC mutants were observed, thus confirming abolishment of the helicase activity of the RepC proteins. The R279A and D280A mutant proteins showed drastic and significant decrease in DNA unwinding activity (**Figure 4.17**). This suggests that these residues play critical role in dsDNA unwinding process.

4.10 Binding studies of ToLCNDV mutant RepC proteins to ATPγS and ssDNA by fluorescence spectroscopy

The binding of the RepC mutant proteins with ATPγS as well as with the ssDNA in the presence of ATPγS was investigated with the help of fluorescence binding studies. The tryptophan amino acid of the RepC protein (wild-type/mutant) was used as intrinsic fluorophore. The emission spectra of the wild-type RepC showed a peak at 350nm. The change in intensity was observed on addition of ligand, i.e. ATPγS and DNA (+ATPγS) and from that the Kd value was determined, which is the reflection of the binding affinity of the RepC protein to the corresponding ligand.

The Kd values of the mutant RepC protein with ATPγS and DNA (+ATPγS) are provided in **Table 4.3** and **Table 4.4** respectively. The Kd value of ATPγS to R279A, D280A and P290A were 1.38±0.09 µM, 2.25±1.27 µM and 1.34±0.45 µM, respectively, thus they possess similar affinity to ATPγS as that of the wild-type protein (Kd=1.40±0.20 µM). Affinity to ssDNA in the presence of ATPγS of R279A, D280A and P290A mutant proteins were found to be similar as that of wild-type ToLCNDV RepC protein.

In addition to this, the binding affinity of K272A and K289A to ATPγS and DNA in the presence of ATPγS were previously determined using the same experiment set up for ToLCGV (George *et al*., 2014) (**Table 4 5**). Binding affinity of the wild-type ToLCGV RepC and of the ToLCGV RepC K272A mutant to ATPγS were 0.95±0.07 µM and 1.27±0.09 µM, respectively. While the ssDNA binding affinity of wild-type ToLCGV RepC and of ToLCGV K272A mutant were 21.6±3 nM and 54.9±7.8 nM, respectively. This indicated that K272 is involved in ssDNA binding and not in ATP binding. Another ToLCGV RepC mutant K289A showed drastic effect on DNA binding and the binding cannot be studied by fluorescence spectroscopic assay while the affinity to ATPγS was same as that of the wild-type protein.

Table 4.3 ATPγS binding of wild-type and mutant ToLCNDV RepC proteins using fluorimetric binding experiments.

Ligand	Kd	R^2
Wild-type	1.4016±0.199 µM	0.9715
R279A	1.37765±0.093126 µ M	0.9671
D280A	2.25295 ± 1.26692 µM	0.9750
P290A	1.3394±0.452973 µM	0.9734

The error bar represents the standard deviation from two independent experiments.

Table 4.4 DNA binding of wild-type and mutant ToLCNDV RepC proteins in the presence of ATPγS using fluorimetric binding experiments.

Ligand	Kd	R^2
Wild-type	37.875 ±6.14 nM	0.9715
R279A	37.4139 ± 7.8503 n M	0.9628
D280A	42.2717 ± 18.6736 nM	0.9305
P290A	55.5735 ± 7.9629 nM	0.9854

The error bar represents the standard deviation from two independent experiments.

Table 4.5 ATPγS and DNA binding of wild-type ToLCGV RepC and K272A RepC mutant using fluorimetric binding experiments.

Protein	Ligand	Kd	R^2
Wild-type	ATPγS	$(0.95 \pm 0.068) \times 10^{-6}$ M	0.99
	DNA (+ATPγS)	$(21.56 \pm 2.99) \times 10^{-9}$ M	0.98
K272A	ATPγS	$(1.27 \pm 0.09) \times 10^{-6}$ M	0.99
	DNA (+ATPγS)	$(54.98 \pm 7.75) \times 10^{-9}$ M	0.98

The error bar represents the standard deviation from two independent experiments.

4.11 Effect of Rep mutations on ToLCNDV replication *in planta*

The monomeric clone of ToLCNDV DNA-A containing ~2.7 kb ToLCNDV DNA-A viral genome was cloned in pUC18 vector at *Kpn*I site. Site-directed mutations (at K227, D261, K272, R279, D280, Y287, K289 and P290) in the C-terminal of AC1 ORF were generated in this monomeric clone and the mutations were confirmed by sequencing. The strategy used for the generation of the infectious clones has been

discussed in materials and methods **section 3.7.** The partial tandem repeat of ToLCNDV containing desired mutation in the AC1 ORF were confirmed by digestion with *Xba*I showing a release of full-length genome corresponding to the fragment of 2739 bp (**Figure 4.18**).
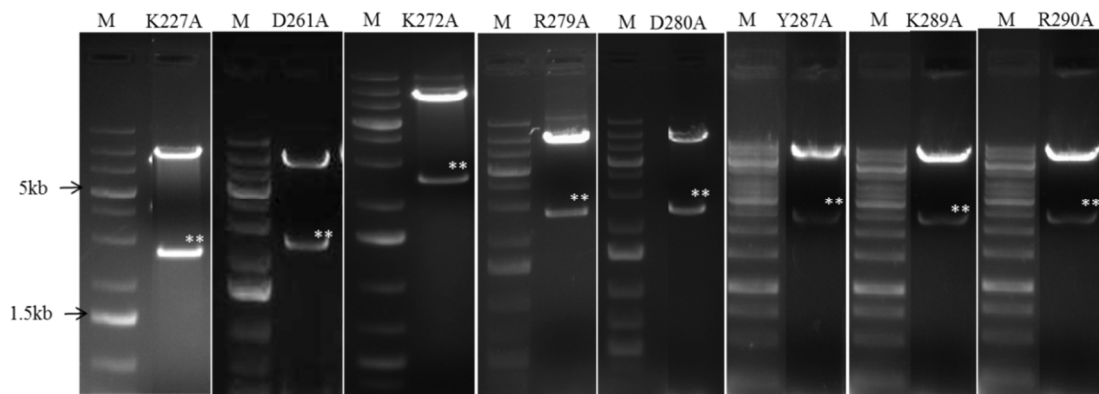
Figure 4.18 Construction of partial tandem repeat construct of ToLCNDV containing mutations in the AC1 ORF. Agarose gel electrophoresis of digestion of partial tandem repeat of ToLCNDV containing desired mutations in AC1 gene with *Xba*I restriction enzyme showing release of ~2.7 kb fragment. Asterisk shows the released 2.7 kb fragment after *Xba*I digestion. Mutations are indicated at the top of each lane.

Figure 4.19 Phenotype of representative *N. benthamiana* plants inoculated with either infectious wild-type or mutant ToLCNDV virus at 14 dpi.

The plasmid of the confirmed infectious clones of ToLCNDV DNA-A of each mutant was then transformed into *Agrobacterium tumefaciens* strain EHA-105 along with mock (pCAMBIA 2300 vector only) and wild-type ToLCNDV DNA-A. Three weeks old *N. benthamiana* plants were agro-inoculated with these constructs along with ToLCGV DNA-B and the severity of the symptoms were observed till 30 dpi.

The symptoms were noticed only in the plants inoculated with wild-type ToLCNDV (DNA-A + DNA-B combination), while the plants inoculated with infectious clones of Rep mutants were found to be symptomless. **Figure 4.19** shows symptom development on *N. benthamiana* plants after inoculation with infectious clones of ToLCNDV DNA-A containing either wild-type or mutant AC1 in the presence of ToLCGV DNA-B. ToLCNDV DNA-A inoculated *N. benthamiana* plants along with ToLCGV DNA-B resulted in stunted growth and in addition typical leaf curling was observed at 14 dpi which was not observed for ToLCNDV mutants. Although in case of inoculation with ToLCNDV DNA-A containing R279A, K289A and P290A mutations, the plant height was slightly affected but no apparent symptoms were observed.
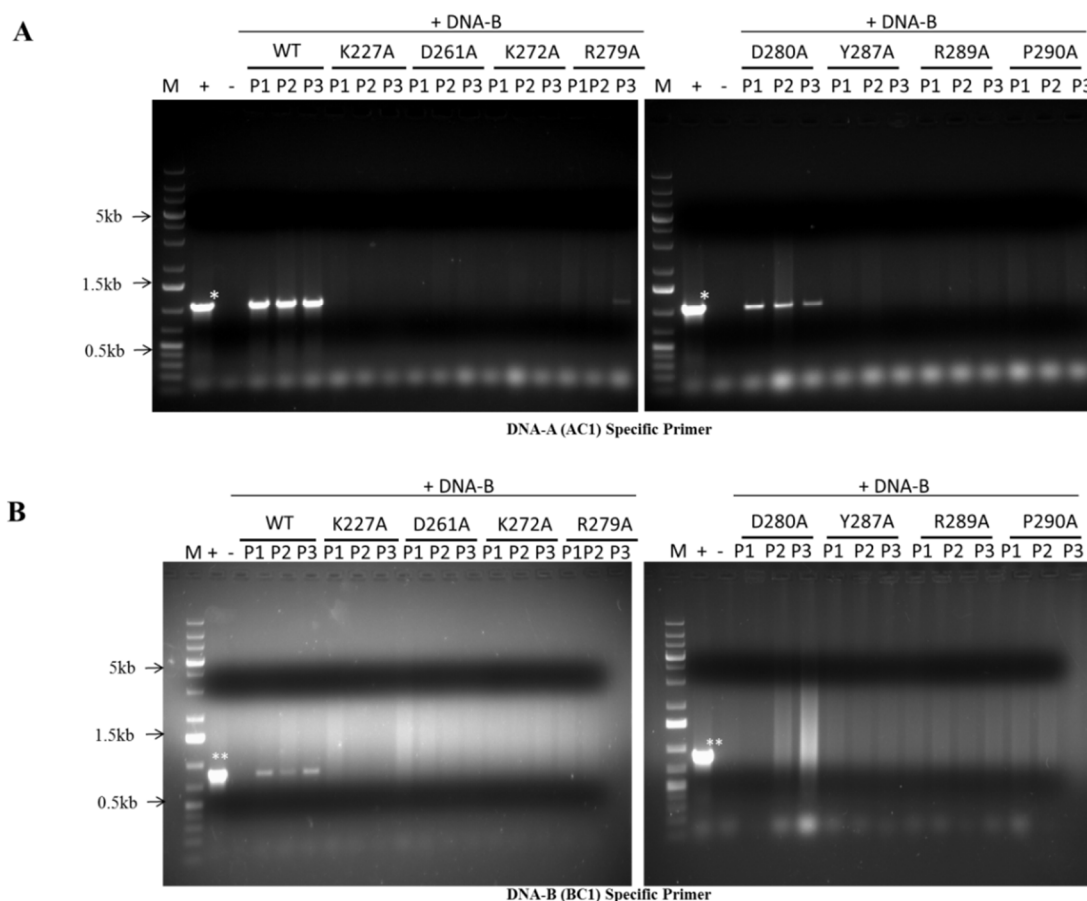
Figure 4.20 Detection of viral DNA isolated from systemic leaves of *N. benthamiana* at 14dpi by PCR. (A) using AC1 specific primers (B) using BC1 specific primers. P1, P2 & P3 are the three plants that are inoculated with various ToLCNDV combinations as indicated above the gel; M- DNA ladder; (+) positive control- monomeric ToLCNDV DNA-A plasmid; (-) negative control- mock plant inoculated with only pCAMBIA vector. The presence of ~1.1 kb fragment corresponding to AC1 ORF in positive control is indicated with single asterisk (*).The presence of ~800 bp fragment corresponding to AC1 gene in positive control is indicated with double asterisk (**). The size of the three reference bands (in kb) of the DNA marker are indicated for each gel.
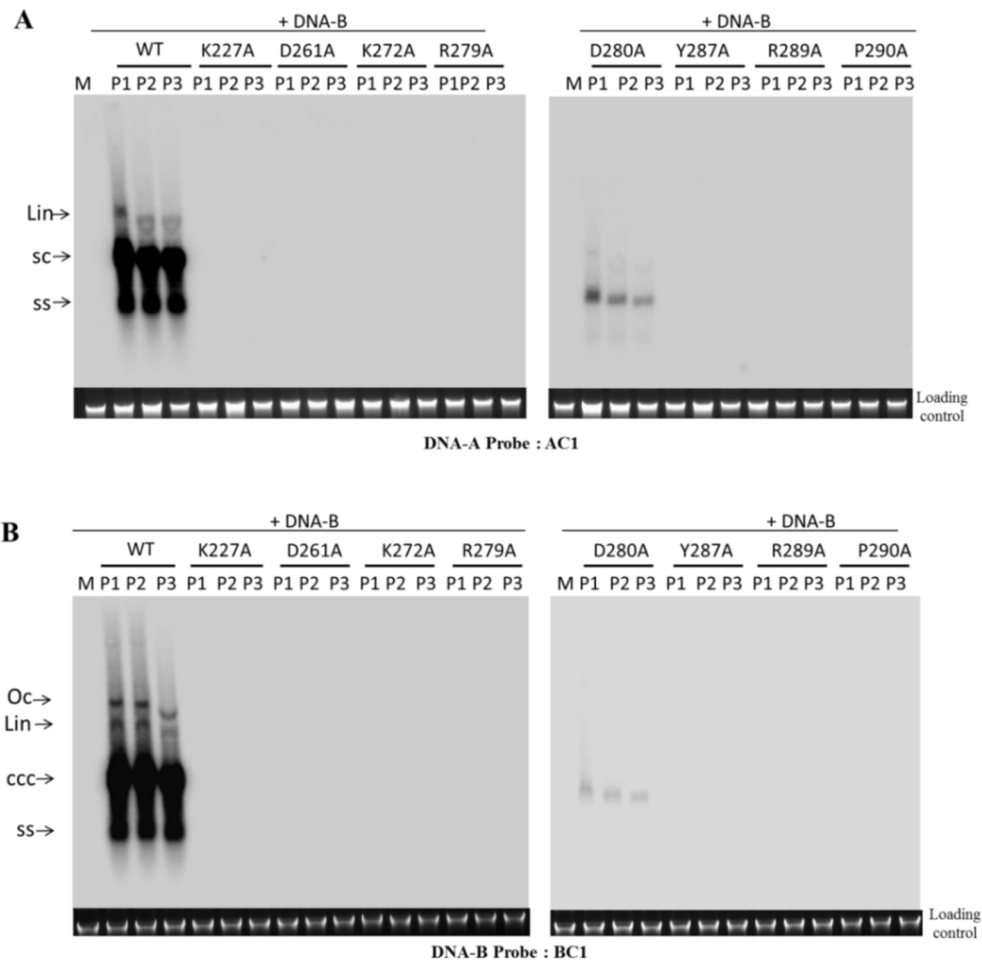
Figure 4.21 Southern blot showing relative level of ToLCNDV DNA-A and DNA-B in *N. benthamiana* plants inoculated with either wild-type or mutants of ToLCNDV along with DNA-B infectious clone. OC- open circular; lin- linear, ccc- covalently closed circular and ss, single stranded represent the different forms of replicated viral DNA.

To check the viral DNA accumulation in the inoculated plants, the total DNA from these plants were isolated and were subjected to PCR using AC1 specific and BC1 specific primers for determining the accumulation of viral DNA-A and DNA-B molecule, respectively. Monomeric ToLCNDV DNA-A and ToLCGV DNA-B plasmid were used as positive control in the experiment showing the presence of amplicon of ~1.1 kb of the AC1 ORF and ~800 bp of the BC1 ORF, respectively, after PCR. Plants inoculated with wild-type ToLCNDV DNA-A infcetious construct along ToLCGV DNA-B showed accumulation of both ToLCNDV DNA-A and ToLCGV DNA-B molecule (**Figure 4.20 A & B**). Neither of the ToLCNDV mutants except D280A mutant showed accumulation of ToLCNDV DNA-A in *N. benthamiana* plants. However, it is found that accumulation of ToLCNDV DNA-A

and ToLCGV DNA-B in D280A mutant inoculated plants were significantly reduced relative to that in the wild-type ToLCNDV virus inoculated plants. There was also no accumulation of DNA-B molecule in the plants inoculated with AC1 mutants of ToLCNDV virus (**Figure 4.20 B**). All these results indicated that due to mutations in the B′ motif of Rep, the viral replication was affected.

The southern blot analysis of relative levels of DNA accumulation in *N. benthamiana* plants inoculated with various mutants ToLCNDV viruses also showed similar results. D280A mutation in the Rep containing ToLCNDV virus inoculated plants showed drastically reduced accumulation of both DNA-A and DNA-B molecules relative to wild-type ToLCNDV inoculated plants (**Figure 4.21 A & B**). No accumulation of DNA-A and DNA-B molecules were detected by Southern blot analysis in the *N. benthamiana* plants inoculated with other ToLCNDV mutants.

The biochemical activities associated with various Rep mutant proteins (Walker A, Walker B and B' motif) along with wild-type Rep have been summarised in the **Table 4.6**.

Table 4.6 Summary of the biochemical activities of wild-type and mutant RepC proteins

	ATP hydrolysis	ATPγS binding	DNA binding	Helicase activity	Replication (*in planta*)
Wt RepC	++	++	+++	+++	+++
K227A	-	ND	ND	-	-
D261A	-	ND	ND	-	-
K272A	++	++*	-*	-	-
R279A	++	++	++	-	-
D280A	++	++	++	-	-
Y287A	++	++	++	-	-
K289A	++	++*	-*	-	-
P290A	++	++	++	-	-

 *Data from ToLCGV RepC (George *et al.,* 2014); ND- not determined.

Walker A (K227) and Walker B (D261) Rep mutants have been used as control as their implications on geminivirus Rep mediated ATP hydrolysis, helicase and rolling circle replication have been previously studied (Choudhury *et al*., 2006). The loss of helicase activity of the Rep protein on B' motif mutations while the ATP binding and hydrolysis activity being unaffected evidently prove their role in the process of DNA unwinding.

4.12 Structural modelling of helicase domain of ToLCNDV Rep

ToLCV Rep protein belongs to the SF3 family of helicases. But it shares very less sequence similarity to members of SF3 family. In addition to that, the structural information with respect to the helicase activity of the Rep protein is lacking. The modelled structure of the ToLCV Rep protein was generated using Bovine Papilloma Virus-E1 structure as a template.
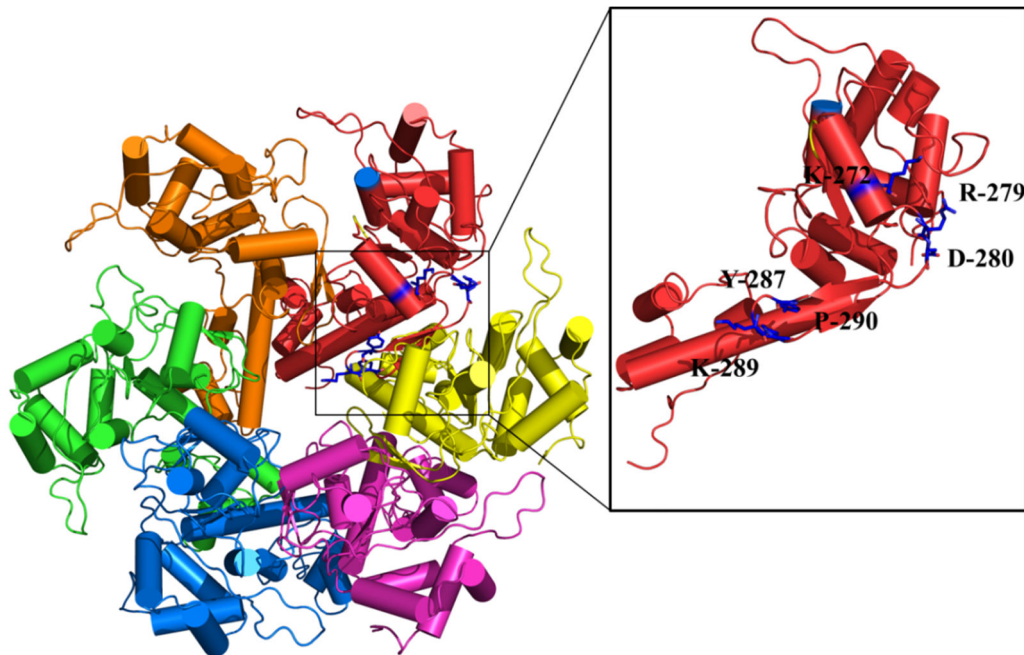
Figure 4.22 Homology model of the ToLCNDV RepC helicase domain based on the Bovine papillomavirus-E1 helicase domain as a template. Cartoon representation of the helicase hexamer where the six subunits are shown in different colours. Locations of residues relevant for this study are depicted in the magnified picture (in boxes) from a single subunit.

The modelled structure showed a degree of structural conservation for both the protein, despite as low as 16% amino acid similarity between them. **Figure 4.22** shows the location of the conserved residues of B' motif present in the C terminal of the Rep protein model that have been generated using Bovine Papilloma Virus -E1structure as a template. K272, R279 and D280 residues of the B motif was found at the interface of the subunits in the hexameric model while the Y287, K289 and P290 were found to be into the central cavity through which the DNA passes. The position of the conserved amino acids of the B' motif suggests their direct/indirect role in DNA binding during the unwinding/ translocation process.

DISCUSSION

Crop cultivating areas including tomatoes face serious challenge and significant losses from the leaf curl disease caused by begomoviruses which have a high rate of mutation and recombination. The Rep protein is involved in the rolling circle replication of the geminiviruses (Saunders *et al*., 1991). This protein also interacts and redirects various host factors to create an environment suitable for the viral propagation and spread within the host cells. Hence, Rep has been considered a target for establishing antiviral state and broad spectrum resistance in crops (Chatterji *et al*., 2001; Noris *et al*., 1996; Reyes *et al.*, 2013).

Helicase activity has been found to be associated with many viral encoded proteins required for their life cycle completion such as in picornaviruses (Norder *et al.*, 2011; Xia *et al.,* 2015), in papillomaviruses (Masterson *et al*., 1998) and in polyomaviruses (An *et al.*, 2012). Helicase domains are frequently found to be part of various non-structural proteins, which have highly disordered regions. Replication initiator protein is the master viral protein in the life cycle of geminiviruses that infect plants. It is an early viral protein which is required almost during all the stages of replication except for the complementary strand synthesis. It acts as an ATPase protein which is required during the process of viral replication via rolling circle replication (Desbiez *et al.,* 1995). It has been classified as a member of the SF3 family of helicases and was demonstrated to possess 5' to 3' dsDNA unwinding activity (Choudhury *et al*., 2006; Clerot and Bernardi, 2006). The analysis of the secondary structure alignment of geminivirus RepC with other SF3 members shows that it contains AAA+ fold which is relatively reduced and do not contain "arginine finger". Thus, it is required to explore the precise role of various conserved motifs present in this reduced AAA+ fold and the mechanism adopted for the DNA unwinding during viral replication.

Solubility issue associated with various recombinant proteins pose a serious challenge in their functional characterization. Successful purification of MBP tagged Rep protein had been performed previously where the presence of MBP tag had resulted in enhanced solubility of the protein (Clerot and Bernardi, 2006). However, in few cases MBP tagged Rep protein was found to be highly insoluble instead (Girish *et al.*, 2006) and hence as a result of which the protein was purified from the inclusion bodies.For the present study, the ToLCGV RepC was overexpressed as MBP tagged protein as the MBP tag is believed to increase the expression, solubility and subsequent

purification of the protein. Due to the need of expensive and commercially limited factor Xa enzyme which was used for the tag cleavage of the MBP-RepC protein, the study was switched to ToLCNDV RepC which shows a relatively increased solubility at optimized growth conditions (0.1 mM IPTG, 16^0C for 16 h) as His tag recombinant protein. The analysis of the ToLCNDV Rep shows that on the basis of the amino acid content the protein is highly soluble and no transmembrane domains was found (using the Expasy Protparam tool). The CD spectra and the Sable tool used for the secondary structure content prediction indicate that the Rep protein comprises of mainly random coiled loop which makes upto ~60% of the total secondary content. The Rep protein was suggested to be a disordered protein by PrDOS (http://prdos.hgc.jp/cgi-bin/top.cgi) tool which may explain the associated intrinsic structural flexibility and thus, the difficulty in context to its purification and crystallization observed in many cases.

The NTP binding site in proteins are found to be evolutionary conserved and comprises of two essential and highly conserved motifs namely, the Walker A and Walker B motifs. ToLCNDV RepC was found to significantly hydrolyse other NTPs (i.e. GTP, CTP and UTP) in addition to ATP. However, Rep mediated hydrolysis was found to be maximum for ATP relative to others. Similar observations are found for many helicases like chikungunya virus (CHIKV) non-structural protein 2 (nsP2) (Das *et al.*, 2014), hepatitis E virus (HEV) ORF1 (Karpe and Lole, 2010) and equine arteritis virus (EAV) nsp10 (Seybert *et al.*, 2000) that shows no strong discrimination for the ATP, GTP, CTP or UTP.

Further in order to understand the NTP binding pocket of the ToLCNDV Rep, NTP binding and their binding affinity to the pocket were looked into. The inhibition of ToLCNDV RepC mediated [γ-P^{32}]ATP hydrolysis activity was found by various other cold labelled NTPs both deoxy and ribose forms. dATP and ATP are better competitors for the NTP binding site, thus indicating that the nucleotide binding site seems to be somehow more fit for ATP. For the inhibition of ATP hydrolysis, NTP need to competitively bind to the same active site meant for ATP, thus blocking the Pi release from the radio labelled ATP. The affinity of various NTPs binding to the ToLCNDV RepC lies in the same range as estimated by fluorescence based studies supporting the notion that various NTP can bind and therefore inhibit the ATP hydrolysis. The rate of inhibition varies even though the affinities are similar giving

rise to the possibility that the subtle local conformational changes within the nucleotide binding pocket upon different NTP binding may be distinct during their hydrolysis. NTP binding and hydrolysis are known to induce conformational change in the proteins. Members of the SF3 family are hexameric in nature and possess either RecA or AAA+ like core domain with the nucleotide binding site located at the interface of two adjacent subunits. ATP interacts with the conserved motifs lining the binding site and with the conserved arginine residue of the adjacent subunit (Frick and Lam, 2006). In addition to that, upon ATP hydrolysis a relative motion between the core domains of the neighbouring monomeric subunits occurs. To note is the fact that geminivirus Rep lack the arginine finger thus, may adopt alternate mechanism for the conformational change transduction. The cycle of ATP binding and hydrolysis are associated with distinct conformational states of protein and the process of unwinding is suggested to be influenced by the nucleotide mediated conformation changes (Erzberger and Berger, 2006). However the extent of the conformation change is not yet clear and generalised.

In general, the energy released by any NTP which is ~10 kcal/mol can be utilized for DNA unwinding as ~1.6 kcal/mol energy investment is sufficient for separation of an average base pair (Von Hippel and Delagoutte, 2001). The process of DNA unwinding is driven by utilizing the energy released by NTP hydrolysis. To examine whether the hydrolysis energy of different nucleotides could be used for RepC mediated DNA unwinding, a helicase assay was performed in the presence of various nucleotides. Interestingly, RepC utilizes ATP and GTP efficiently for driving DNA unwinding. From the present study, it seems that the NTP binding pocket of the ToLCNDV RepC do not discriminate either the ribose or the deoxyribose form of nucleotide binding. The CD analysis indicates change in the conformation of RepC protein upon ATP binding.

The present study on Replication initiator protein was carried out with the aim to understand the precise role of various amino acids in the helicase domain of Rep in the geminivirus replication. For this, a 3D structure model of ToLCNDV-Rep was generated based on the already available BPV-E1 protein structure. Although very less sequence similarity exists between them (~16%), the overall *in silico* model generated showed homology to the BPV-E1 structure and displayed superimposition of conserved structural elements of SF3 motifs. From the hexameric structural model

of ToLCNDV RepC, conserved B' motif and β-hairpin loop was predicted from the superimposition of the BPV-E1 structure. Interestingly the predicted B' motif displayed a higher degree of conservation among Rep proteins of *Geminiviridae* family. B' motif contains various positively charged amino acids and is suggested to play multiple roles during the unwinding process mediated by BPV E1 protein, SV40 and AAV 78 protein (Yoon-Robarts *et al.*, 2004). Like in other members of SF3 helicases, B′ motif is located similarly between Walker B and motif C in geminivirus Rep. The binding of ssDNA to the lysine present within the B' motif in *E. coli* RecA induces conformational change in the catalytic residues of the Walker B and motif C (Story *et al.*, 1992). Few residues of B' motif are shown to directly interact with ATP (Story *et al.*, 1992), while some are required for DNA binding (Castella *et al.*, 2006; Shen *et al.*, 2005). In SV40 large T antigen, the B' motif has also been implicated in positioning DNA binding β-hairpin during the course of DNA unwinding which is induced by ATP hydrolysis at the interface of the neighbouring subunits (Gai *et al.*, 2004). In addition to these roles, there are some residues within the B' motif that serve as coupler, linking the ATP hydrolysis activity to the DNA binding during unwinding (Singleton *et al.*, 2000).

Highly conserved lysine residue of the Walker A motif and the acidic residue present in the Walker B motif is found to be essential for NTP binding (Hall and Matson, 1999). The lysine and aspartate residues present in the Walker A and Walker B motifs respectively of the geminivirus Rep protein had previously been shown to be required for ATPase and are found critical for the viral replication (Choudhury *et al.*, 2006). Consistent with this, the K227A (Walker A) and D261A (Walker B) mutants of ToLCNDV Rep were found to be defective in ATP hydrolysis and subsequently the role of these mutations in the process of DNA unwinding and replication had been implicated. Various biochemical activities which include ATP binding and hydrolysis, ssDNA binding, dsDNA unwinding and *in planta* replication efficiency of RepC mutant proteins (K272A, R279A, D280A, K289A and P290A) for the B' motif were studied. The similar CD spectra recorded for the wild-type and mutant RepC proteins rule out the effect of the mutations on the structure of the Rep protein and their subsequent implication on the functions related to unwinding. Apart from that the ATPase activity of these mutant proteins comparable to that of wild-type levels also suggests that the mutant proteins are structured. ATP hydrolysis and other *in vitro*

biochemical assay could not be performed for Y287A owing to its lack of overexpression in *E.coli* BL21 cells. The unwinding activities of K272A, R279A, D280A, K289A and P290A were impaired while the ATP binding and hydrolysis remain unaffected. The loss of unwinding activity of K272A and K289A might be due to loss of ssDNA binding (**Table 4.5**). The ssDNA binding affinity of ToLCGV K272A was decreased while the binding affinity of ToLCGV K289A could not be determined as the change in fluorescence intensity upon DNA binding was very less, indicating insignificant binding to DNA (George *et al.*, 2014). Although R279A, D280A and P290A had similar DNA binding affinity as that of wild-type RepC, they were found to be defective in unwinding activity. This suggests that they may possess indirect role in DNA unwinding process. Rep being essential for geminivirus replication in the plants, the subsequent effect of all these mutations on the viral replication was observed on *N. benthamiana* plants. The inability of the mutants to support the viral replication *in planta* was consistent with the biochemical analysis of these mutants.

The B' motif of ToLCNDV Rep is positioned near to the ATP binding pocket at the interface of the subunits as well as near the central cavity as indicated from the *in silico* hexameric modelled structure. Thus these residues may potentially directly interact with the ssDNA and/or probably couple the ATP binding to the ssDNA binding. The Y287, K289 and P290 are present on the β-hairpin loop lining the central cavity. The K289 is located at the tip of the β- hairpin structure located in the central channel of the hexameric protein, and may thus directly interact with ssDNA during geminivirus. This provides an explanation of the observed loss of DNA binding in case of K289A protein. In addition, the Y287 residue may be essential for base stacking interaction with DNA passing through the central channel of the hexameric protein. Although the biochemical activities associated with Y287A protein could not be determined, the residue plays a critical role in viral replication since the mutation abolished the viral replication *in planta*. The role of the various positively charged and aromatic residues present in the β- hairpin loop in ssDNA binding has been previously reported for SV40 large T antigen (Shen *et al.*, 2005). Apart from this, role of positively charged residues of the β- hairpin loop has been demonstrated in AAV2 Rep and BPV-E1 as well (Castella *et al.*, 2006; Yoon-Robarts *et al.*, 2004). P290 residue may be required for maintaining the structure or the proper

orientation of the β- hairpin loop protruding into the DNA binding cavity, thus mutating this residue might result in loss of DNA unwinding. *In silico* RepC model shows that K272, R279 and D280 amino acids reside at the subunit interface where the ATP binding pocket is present. Even though not present near the central cavity, these residues when mutated do exert deficiency/ defect in DNA unwinding and subsequent viral replication. For K272A the unwinding activity was completely abolished while R279A and D280A exhibited greatly reduced helicase activity. K272A and R279A loss of unwinding activity leads to inability of these mutants to support replication of viruses in plants as expected. However, in D280A very mild unwinding activity and correspondingly a drastically reduced viral accumulation in *N. benthamiana* plants could be observed, but the failure of symptom development suggest that the level of viral replication was not enough to support the viral infectivity in plants. Thus, D280 may play a supporting/secondary role but is required for the process of viral replication. Based on the biochemical results obtained and the location of these residues away from the central pore, it may be suggested that these residue may show movement upon ATP mediated conformational change or may be required for positioning of the β-hairpin loop in response to the ATP hydrolysis during the DNA unwinding. Thus various residues of the B' motif present in the C-terminal of the geminivirus Rep is pivotal for the DNA unwinding and therefore for viral replication in plant cell.

Rep protein is a helicase and the loss of this activity has been found to be non-negotiable for the viral propagation and replication within plant cell. From studies on various helicases belonging to different helicase superfamily, it has been found that despite variability in the mechanism they follow, several critical residues involved in helicase activity are conserved to much extent in the tertiary structure arrangement. Thus, exploring such residues in the helicase domain of the geminivirus Rep protein will serve dual purpose. Firstly, it will provide insight into the mechanism adopted by the geminivirus virus helicase, which happens to be the only plant viral protein belonging to SF3 family of helicases. Secondly, understanding the biological relevance of geminivirus helicase domain in the modular structure of Rep might help in providing improved approach for generating broad spectrum resistance against plant viruses.

Based on the structural informations gathered, drug designing can be used to screened and identify small-molecule helicase inhibitors that may target any of the following step during DNA unwinding by geminivirus Rep: NTP binding, NTP hydrolysis (via stabilizing ADP-bound state), nucleic acid binding, coupling NTP hydrolysis to translocation (via restricting or blocking the domain movement or altering the interface cleft) or unwinding of nucleic acid. Oligomerization has been found to have implication on the unwinding activity of hexameric helicases. Thus, small molecules capable for disrupting the oligomerization of helicase or by inhibiting the crucial interaction with other host protein in larger complex may also create an antiviral state.

SUMMARY

The present study was carried out with the aim to investigate the role of the C-terminal of the replication initiator protein (Rep) of *Tomato leaf curl New Delhi virus* (ToLCNDV) in the intricate process of replication. The truncated ToLCNDV Rep protein showed no significant change in its ATPase activity as compared to the wild-type protein and the solubility of the truncated RepC protein was found to be enhanced than the wild-type. Therefore the present was taken up further.

The significant findings from the present study are as follows:

✓ Both the full-length (RepF) and C-terminal (RepC) of ToLCNDV Rep were successfully overexpressed and purified from the *E. coli* cells as His tagged recombinant protein. Both the proteins were biologically functional.

✓ ToLCNDV RepC can hydrolyse various NTPs and affinity to different NTPs was observed in similar range.

✓ RepC can efficiently use ATP and other NTPs for DNA unwinding.

✓ Based on the *in silico* model of helicase domain of Rep generated using BPV- E1 protein as template, B′ motif was identified.

✓ Role of eight conserved amino acids of B′ motif in ATP Hydrolysis, DNA unwinding and *in vivo* replicational efficiency of Rep were studied.

✓ Most of the Rep mutants (K272A, R279A, D280A, K289A, P290A) retained ATPase activity comparable to the wild-type except K227A and K261A.

✓ RepC mutations caused no alteration on the structural integrity of the protein.

✓ K272A mutation abolishes DNA unwinding due to reduced ssDNA binding of Rep. Apart from K272, the R279 and D280 residues present near the interface of the subunits of the hexameric protein model were also found to have implication on the DNA unwinding and geminivirus replication.

✓ Residues Y287, K289 and P290 were part of the β-hairpin loop lining the central channel and were critical for the unwinding of dsDNA during the course of replication.

✓ Although, Y287A RepC mutant could not be purified, *in planta* infectivity showed importance of this residue in ToLCNDV replication.

✓ *In planta* infectivity showed all the amino acids targeted are critical for unwinding of DNA and thus influences viral replication in permissive host.

REFERENCES

Abbate, E. A., Berger, J. M., Botchan, M. R., 2004. The X-ray structure of the papillomavirus helicase in complex with its molecular matchmaker E2. Genes Dev 18, 1981-96.

Abrahams, J. P., Leslie, A. G., Lutter, R., Walker, J. E., 1994. Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria. Nature 370, 621-8.

Accotto, G. P., Navas-Castillo, J., Noris, E., Moriones, E., Louro, D., 2000. Typing of tomato yellow leaf curl viruses in Europe. Eur J Plant Pathol 106, 179-186.

Ach, R. A., Durfee, T., Miller, A. B., Taranto, P., Hanley-Bowdoin, L., Zambryski, P. C., Gruissem, W., 1997. RRB1 and RRB2 encode maize retinoblastoma-related proteins that interact with a plant D-type cyclin and geminivirus replication protein. Mol Cell Biol 17, 5077-86.

Adams, M. J., King, A. M., Carstens, E. B., 2013. Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2013). Arch Virol 158, 2023-30.

Adelman, J. L., Jeong, Y. J., Liao, J. C., Patel, G., Kim, D. E., Oster, G., Patel, S. S., 2006. Mechanochemistry of transcription termination factor Rho. Mol Cell 22, 611-21.

Ahnert, P., Patel, S. S., 1997. Asymmetric interactions of hexameric bacteriophage T7 DNA helicase with the 5'- and 3'-tails of the forked DNA substrate. J Biol Chem 272, 32267-73.

Akbar Behjatnia, S. A., Dry, I. B., Ali Rezaian, M., 1998. Identification of the replication-associated protein binding domain within the intergenic region of tomato leaf curl geminivirus. Nucleic Acids Res 26, 925-31.

Alberter, B., Ali Rezaian, M., Jeske, H., 2005. Replicative intermediates of Tomato leaf curl virus and its satellite DNAs. Virology 331, 441-8.

An, P., Saenz Robles, M. T., Pipas, J. M., 2012. Large T antigens of polyomaviruses: amazing molecular machines. Annu Rev Microbiol 66, 213-36.

Andressoo, J. O., Hoeijmakers, J. H., de Waard, H., 2005. Nucleotide excision repair and its connection with cancer and ageing. Adv Exp Med Biol 570, 45-83.

Arguello-Astorga, G., Lopez-Ochoa, L., Kong, L. J., Orozco, B. M., Settlage, S. B., Hanley-Bowdoin, L., 2004. A novel motif in geminivirus replication proteins interacts with the plant retinoblastoma-related protein. J Virol 78, 4817-26.

Ascencio-Ibanez, J. T., Sozzani, R., Lee, T. J., Chu, T. M., Wolfinger, R. D., Cella, R., Hanley-Bowdoin, L., 2008. Global analysis of Arabidopsis gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection. Plant Physiol 148, 436-54.

Bagewadi, B., Chen, S., Lal, S. K., Choudhury, N. R., Mukherjee, S. K., 2004. PCNA interacts with Indian mung bean yellow mosaic virus rep and downregulates Rep activity. J Virol 78, 11890-903.

Beijerinck, M. W., 1889. Ueber ein Contagium vivum fluidum als Ursache der Fleckenkrankheit der Tabaksblatter. Verhandelingen der Koninkyke akademie Wettenschappen te Amsterdam 65, 3-21.

Birnboim, H. C., Doly, J., 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7, 1513-23.

Bock, K. R., Guthrie, E. J., Woods, R. D., 1974. Purification of maize streak virus and its relationship to viruses associated with streak diseases of sugar cane and Panicum maximum. Ann Appl Biol 77, 289-296.

Bock, K. R., Guthrie, E. J., Figueiredo, G., 1981. A strain of cassava latent virus occurring in coastal districts of Kenya. Ann Appl Biol 99, 151-159.

Branch, A. D., Benenfeld, B. J., Robertson, H. D., 1988. Evidence for a single rolling circle in the replication of potato spindle tuber viroid. Proc Natl Acad Sci USA 85, 9128-32.

Briddon, R., Pinner, M., Stanley, J., Markham, P., 1990. Geminivirus coat protein gene replacement alters insect specificity. Virology 177, 85-94.

Briddon, R. W., Bedford, I. D., Tsai, J. H., Markham, P. G., 1996. Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, tomato pseudo-curly top virus, suggests a recombinant origin. Virology 219, 387-94.

Briddon, R. W., Mansoor, S., Bedford, I. D., Pinner, M. S., Saunders, K., Stanley, J., Zafar, Y., Malik, K. A., Markham, P. G., 2001. Identification of dna components required for induction of cotton leaf curl disease. Virology 285, 234-43.

Briddon, R. W., Bull, S. E., Amin, I., Mansoor, S., Bedford, I. D., Rishi, N., Siwatch, S. S., Zafar, Y., Abdel-Salam, A. M., Markham, P. G., 2004. Diversity of DNA 1: a satellite-like molecule associated with monopartite begomovirus-DNA beta complexes. Virology 324, 462-74.

Briddon, R. W., Bull, S. E., Amin, I., Idris, A. M., Mansoor, S., Bedford, I. D., Dhawan, P., Rishi, N., Siwatch, S. S., Abdel-Salam, A. M., Brown, J. K., Zafar, Y., Markham, P. G., 2003. Diversity of DNA beta, a satellite molecule associated with some monopartite begomoviruses. Virology 312, 106-21.

Brough, C. L., Gardiner, W. E., Inamdar, N. M., Zhang, X. Y., Ehrlich, M., Bisaro, D. M., 1992. DNA methylation inhibits propagation of tomato golden mosaic virus DNA in transfected protoplasts. Plant Mol Biol 18, 703-12.

Brown, J. K., 1994. Current status of Bemisia tabaci as a plant pest and virus vector in agroecosystems worldwide. FAO Plant Protect B 42, 3-32.

Brown, J. K., 2007. The Bemisia tabaci complex: genetic and phenotypic variability drives begomovirus spread and virus diversification. Plant Dis 1, 25-56.

Brown, J. K., 2009. Phylogenetic biology of the Bemisia tabaci sibling species group. Bemisia: bionomics and management of a global pest. Springer, pp. 31-67.

Bujalowski, W., Klonowska, M. M., 1993. Negative cooperativity in the binding of nucleotides to Escherichia coli replicative helicase DnaB protein. Interactions with fluorescent nucleotide analogs. Biochemistry 32, 5888-900.

Burgyan, J., Havelda, Z., 2011. Viral suppressors of RNA silencing. Trends Plant Sci 16, 265-72.

Campos-Olivas, R., Louis, J. M., Clerot, D., Gronenborn, B., Gronenborn, A. M., 2002. The structure of a replication initiator unites diverse aspects of nucleic acid metabolism. Proc Natl Acad Sci USA 99, 10310-5.

Castella, S., Bingham, G., Sanders, C. M., 2006. Common determinants in DNA melting and helicase-catalysed DNA unwinding by papillomavirus replication protein E1. Nucleic Acids Res 34, 3008-19.

Castillo, A. G., Kong, L. J., Hanley-Bowdoin, L., Bejarano, E. R., 2004. Interaction between a geminivirus replication protein and the plant sumoylation system. J Virol 78, 2758-69.

Castillo, A. G., Collinet, D., Deret, S., Kashoggi, A., Bejarano, E. R., 2003. Dual interaction of plant PCNA with geminivirus replication accessory protein (Ren) and viral replication protein (Rep). Virology 312, 381-94.

Chakraborty, s., 2008. Tomato Leaf Curl Viruses from India. In: Regenmortel, B. W. J. M. a. M. H. V. V., (Ed.), Encyclopedia of Virology. Elsevier, Oxford, pp. 124-133.

Chakraborty, S., Pandey, P. K., Banerjee, M. K., Kalloo, G., Fauquet, C. M., 2003. Tomato leaf curl Gujarat virus, a New Begomovirus Species Causing a Severe Leaf Curl Disease of Tomato in Varanasi, India. Phytopathology 93, 1485-95.

Chandran, S. A., Levy, Y., Mett, A., Belausov, E., Ramakrishnan, U., Gafni, Y., 2012. Mapping of functional region conferring nuclear localization and karyopherin α-binding activity of the C2 protein of bhendi yellow vein mosaic virus. J Gen Virol 93, 1367-1374.

Chatchawankanphanich, O., Maxwell, D. P., 2002. Tomato leaf curl Karnataka virus from Bangalore, India, Appears to be a Recombinant Begomovirus. Phytopathology 92, 637-45.

Chatterji, A., Beachy, R. N., Fauquet, C. M., 2001. Expression of the oligomerization domain of the replication-associated protein (Rep) of Tomato leaf curl New Delhi virus interferes with DNA accumulation of heterologous geminiviruses. J Biol Chem 276, 25631-8.

Chatterji, A., Chatterji, U., Beachy, R. N., Fauquet, C. M., 2000. Sequence parameters that determine specificity of binding of the replication-associated protein to its cognate site in two strains of tomato leaf curl virus-New Delhi. Virology 273, 341-50.

Chattopadhyay, B., Singh, A. K., Yadav, T., Fauquet, C. M., Sarin, N. B., Chakraborty, S., 2008. Infectivity of the cloned components of a begomovirus: DNA beta complex causing chilli leaf curl disease in India. Arch Virol 153, 533-9.

Choudhury, N. R., Malik, P. S., Singh, D. K., Islam, M. N., Kaliappan, K., Mukherjee, S. K., 2006. The oligomeric Rep protein of Mungbean yellow mosaic India virus (MYMIV) is a likely replicative helicase. Nucleic Acids Res 34, 6362-77.

Clerot, D., Bernardi, F., 2006. DNA helicase activity is associated with the replication initiator protein rep of tomato yellow leaf curl geminivirus. J Virol 80, 11322-30.

Crooks, G. E., Hon, G., Chandonia, J.-M., Brenner, S. E., 2004. WebLogo: a sequence logo generator. Genome Res 14, 1188-1190.

Das, P. K., Merits, A., Lulla, A., 2014. Functional cross-talk between distant domains of chikungunya virus non-structural protein 2 is decisive for its RNA-modulating activity. J Biol Chem 289, 5635-53.

Dellaporta, S. L., Jonathan, W., Hicks, J. B., 1983. A plant DNA minipreparation: version II. Plant Mol Biol Report, 19-21.

Desbiez, C., David, C., Mettouchi, A., Laufs, J., Gronenborn, B., 1995. Rep protein of tomato yellow leaf curl geminivirus has an ATPase activity required for viral DNA replication. Proc Natl Acad Sci USA 92, 5640-4.

Donson, J., Morris-Krsinich, B. A., Mullineaux, P. M., Boulton, M. I., Davies, J. W., 1984. A putative primer for second-strand DNA synthesis of maize streak virus is virion-associated. EMBO J 3, 3069-73.

Dry, I. B., Krake, L. R., Rigden, J. E., Rezaian, M. A., 1997. A novel subviral agent associated with a geminivirus: the first report of a DNA satellite. Proc Natl Acad Sci USA 94, 7088-93.

Eagle, P. A., Orozco, B. M., Hanley-Bowdoin, L., 1994. A DNA sequence required for geminivirus replication also mediates transcriptional regulation. Plant Cell 6, 1157-70.

Egelman, E. H., Yu, X., Wild, R., Hingorani, M. M., Patel, S. S., 1995. Bacteriophage T7 helicase/primase proteins form rings around single-stranded DNA that suggest a general structure for hexameric helicases. Proc Natl Acad Sci USA 92, 3869-73.

Ellis, N. A., Groden, J., Ye, T. Z., Straughen, J., Lennon, D. J., Ciocci, S., Proytcheva, M., German, J., 1995. The Bloom's syndrome gene product is homologous to RecQ helicases. Cell 83, 655-66.

Elmer, J. S., Leslie Brand, Garry Sunter, William E. Gardiner, David M. Bisaro, and Stephen G. Rogers, 1988. Genetic analysis of the tomato golden mosaic virus II. The product of the AL1 coding sequence is required for replication. Nucleic Acids Res 16, 7043-7060.

Ermak, G., Paszkowski, U., Wohlmuth, M., Mittelsten Scheid, O., Paszkowski, J., 1993. Cytosine methylation inhibits replication of African cassava mosaic virus by two distinct mechanisms. Nucleic Acids Res 21, 3445-50.

Erzberger, J. P., Berger, J. M., 2006. Evolutionary relationships and structural mechanisms of AAA+ proteins. Annu Rev Biophys Biomol Struct 35, 93-114.

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U., Sali, A., 2007. Comparative protein structure modeling using MODELLER. Curr Protoc Protein Sci Chapter 2, Unit 2 9.

Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., Ball, L. A., 2005. Virus taxonomy: VIIIth report of the International Committee on Taxonomy of Viruses. Academic Press.

Feinberg, A. P., Vogelstein, B., 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132, 6-13.

Fondong, V. N., Reddy, R. C., Lu, C., Hankoua, B., Felton, C., Czymmek, K., Achenjang, F., 2007. The consensus N-myristoylation motif of a geminivirus AC4 protein is required for membrane binding and pathogenicity. MPMI 20, 380-391.

Fontes, E. P., Luckow, V. A., Hanley-Bowdoin, L., 1992. A geminivirus replication protein is a sequence-specific DNA binding protein. Plant Cell 4, 597-608.

Frick, D. N., Lam, A. M., 2006. Understanding helicases as a means of virus control. Curr Pharm Des 12, 1315-38.

Gai, D., Zhao, R., Li, D., Finkielstein, C. V., Chen, X. S., 2004. Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. Cell 119, 47-60.

George, B., Ruhel, R., Mazumder, M., Sharma, V. K., Jain, S. K., Gourinath, S., Chakraborty, S., 2014. Mutational analysis of the helicase domain of a replication initiator protein reveals critical roles of Lys 272 of the B' motif and Lys 289 of the beta-hairpin loop in geminivirus replication. J Gen Virol 95, 1591-602.

Girish, K. R., Palanivelu, S., Kumar, P. D., Usha, R., 2006. Refolding, purification and characterization of replication-initiator protein from soybean-infecting geminivirus. J Virol Methods 136, 154-9.

Goodman, R. M., 1977. Single-stranded DNA genome in a whitefly-transmitted plant virus. Virology 83, 171-9.

Gopal, P., Kumar, P. P., Sinilal, B., Jose, J., Yadunandam, A. K., Usha, R., 2007. Differential roles of C4 and βC1 in mediating suppression of post-transcriptional gene silencing: evidence for transactivation by the C2 of Bhendi yellow vein mosaic virus, a monopartite begomovirus. Virus Res 123, 9-18.

Gorbalenya, A. E., Koonin, E. V., 1993. Helicases: amino acid sequence comparisons and structure-function relationships. Curr Opin Struc Biol 3, 419-429.

Greenfield, N. J., 2006. Using circular dichroism spectra to estimate protein secondary structure. Nat Protoc 1, 2876-90.

Gronenborn, B., 2004. Nanoviruses: genome organisation and protein function. Vet Microbiol 98, 103-9.

Gros, M. F., te Riele, H., Ehrlich, S. D., 1987. Rolling circle replication of single-stranded DNA plasmid pC194. EMBO J 6, 3863-9.

Gu, J., Stephenson, C. G., Iadarola, M. J., 1994. Recombinant proteins attached to a nickel-NTA column: use in affinity purification of antibodies. Biotechniques 17, 257, 260, 262.

Guerra-Peraza, O., Kirk, D., Seltzer, V., Veluthambi, K., Schmit, A. C., Hohn, T., Herzog, E., 2005. Coat proteins of Rice tungro bacilliform virus and Mungbean yellow mosaic virus contain multiple nuclear-localization signals and interact with importin alpha. J Gen Virol 86, 1815-26.

Hacker, K. J., Johnson, K. A., 1997. A hexameric helicase encircles one DNA strand and excludes the other during DNA unwinding. Biochemistry 36, 14080-7.

Hall, M. C., Matson, S. W., 1999. Helicase motifs: the engine that powers DNA unwinding. Mol Microbiol 34, 867-77.

Hallan, V., Gafni, Y., 2001. Tomato yellow leaf curl virus (TYLCV) capsid protein (CP) subunit interactions: implications for viral assembly. Arch Virol 146, 1765-1773.

Hanley-Bowdoin, L., Bejarano, E. R., Robertson, D., Mansoor, S., 2013. Geminiviruses: masters at redirecting and reprogramming plant processes. Nat Rev Microbiol 11, 777-88.

Hanley-Bowdoin, L., Settlage, S. B., Orozco, B. M., Nagar, S., Robertson, D., 2000. Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. Crit Rev Biochem Mol Biol 35, 105-40.

Hanley-Bowdoin, L., Settlage, S. B., Robertson, D., 2004. Reprogramming plant gene expression: a prerequisite to geminivirus DNA replication. Mol Plant Pathol 5, 149-156.

Harrison A. D., M. V., Swanson M. M. and Roberts D. J., 1991. Recognition and differentiation of seven whitefly transmitted Geminiviruses from India and their relationship to African cassava mosaic and Thailand mungbean yellow mosaic viruses. Ann Appl Biol 118, 299-308.

Harrison, B., Robinson, D., 1999. Natural genomic and antigenic variation in whitefly-transmitted geminiviruses (begomoviruses). Annu Rev Phytopathol 37, 369-398.

Harrison, B., Swanson, M., Fargette, D., 2002. Begomovirus coat protein: serology, variation and functions. Physiol Mol Plant Path 60, 257-271.

Hayes, R. J., Buck, K. W., 1989. Replication of tomato golden mosaic virus DNA B in transgenic plants expressing open reading frames (ORFs) of DNA A: requirement of ORF AL2 for production of single-stranded DNA. Nucleic Acids Res 17, 10213-22.

Heyraud-Nitschke, F., Schumacher, S., Laufs, J., Schaefer, S., Schell, J., Gronenborn, B., 1995. Determination of the origin cleavage and joining domain of geminivirus Rep proteins. Nucleic Acids Res 23, 910-6.

Hickman, A. B., Dyda, F., 2005. Binding and unwinding: SF3 viral helicases. Curr Opin Struct Biol 15, 77-85.

Hingorani, M. M., Patel, S. S., 1996. Cooperative interactions of nucleotide ligands are linked to oligomerization and DNA binding in bacteriophage T7 gene 4 helicases. Biochemistry 35, 2218-28.

Hingorani, M. M., Washington, M. T., Moore, K. C., Patel, S. S., 1997. The dTTPase mechanism of T7 DNA helicase resembles the binding change mechanism of the F1-ATPase. Proc Natl Acad Sci USA 94, 5012-7.

Hipp, K., Rau, P., Schafer, B., Gronenborn, B., Jeske, H., 2014. The RXL motif of the African cassava mosaic virus Rep protein is necessary for rereplication of yeast DNA and viral infection in plants. Virology 462-463, 189-98.

Hormuzdi, S. G., Bisaro, D. M., 1993. Genetic analysis of beet curly top virus: evidence for three virion sense genes involved in movement and regulation of single- and double-stranded DNA levels. Virology 193, 900-9.

Hormuzdi, S. G., Bisaro, D. M., 1995. Genetic analysis of beet curly top virus: examination of the roles of L2 and L3 genes in viral pathogenesis. Virology 206, 1044-54.

Horns, T., Jeske, H., 1991. Localization of abutilon mosaic virus (AbMV) DNA within leaf tissue by in situ hybridization. Virology 181, 580-8.

Huang, S. G., Weisshart, K., Fanning, E., 1998. Characterization of the nucleotide binding properties of SV40 T antigen using fluorescent 3'(2')-O-(2,4,6-trinitrophenyl)adenine nucleotide analogues. Biochemistry 37, 15336-44.

Ilyina, T. V., Koonin, E. V., 1992. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. Nucleic Acids Res 20, 3279-85.

Jeffrey, J. L., Pooma, W., Petty, I. T., 1996. Genetic requirements for local and systemic movement of tomato golden mosaic virus in infected plants. Virology 223, 208-218.

Jeruzalmi, D., O'Donnell, M., Kuriyan, J., 2002. Clamp loaders and sliding clamps. Curr Opin Struct Biol 12, 217-24.

Jeske, H., Lutgemeier, M., Preiss, W., 2001. DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. EMBO J 20, 6158-67.

Johnson, J., 1942. Translators preface and translations with short biographies of Mayer, Ivanovsky, Beijerinck and Bauer. Phytopathological Classics 7, 33-54.

Jose, J., Usha, R., 2003. Bhendi yellow vein mosaic disease in India is caused by association of a DNA Beta satellite with a begomovirus. Virology 305, 310-7.

Jyothsna, P., Rawat, R., Malathi, V. G., 2013a. Predominance of tomato leaf curl Gujarat virus as a monopartite begomovirus: association with tomato yellow leaf curl Thailand betasatellite. Arch Virol 158, 217-24.

Jyothsna, P., Haq, Q. M., Singh, P., Sumiya, K. V., Praveen, S., Rawat, R., Briddon, R. W., Malathi, V. G., 2013b. 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, 5457-71.

Kaliappan, K., Choudhury, N. R., Suyal, G., Mukherjee, S. K., 2012. A novel role for RAD54: this host protein modulates geminiviral DNA replication. FASEB J 26, 1142-60.

Kammann, M., Schalk, H. J., Matzeit, V., Schaefer, S., Schell, J., Gronenborn, B., 1991. DNA replication of wheat dwarf virus, a geminivirus, requires two cis-acting signals. Virology 184, 786-90.

Kanakala, S., Jyothsna, P., Shukla, R., Tiwari, N., Veer, B. S., Swarnalatha, P., Krishnareddy, M., Malathi, V. G., 2013. Asymmetric synergism and heteroencapsidation between two bipartite begomoviruses, tomato leaf curl New Delhi virus and tomato leaf curl Palampur virus. Virus Res 174, 126-36.

Kaplan, D. L., 2000. The 3'-tail of a forked-duplex sterically determines whether one or two DNA strands pass through the central channel of a replication-fork helicase. J Mol Biol 301, 285-99.

Kaplan, D. L., O'Donnell, M., 2002. DnaB drives DNA branch migration and dislodges proteins while encircling two DNA strands. Mol Cell 10, 647-57.

Kaplan, D. L., Davey, M. J., O'Donnell, M., 2003. Mcm4,6,7 uses a "pump in ring" mechanism to unwind DNA by steric exclusion and actively translocate along a duplex. J Biol Chem 278, 49171-82.

Karpe, Y. A., Lole, K. S., 2010. NTPase and 5' to 3' RNA duplex-unwinding activities of the hepatitis E virus helicase domain. J Virol 84, 3595-602.

Kheyr-Pour, A., Bendahmane, M., Matzeit, V., Accotto, G. P., Crespi, S., Gronenborn, B., 1991. Tomato yellow leaf curl virus from sardinia is a whitefly-transmitted monoparatite geminivirus. Nucleic Acids Res 19, 6763-6769.

Kings, A., Adams, M., Carstens, E., Lefkowitz, E., 2011. Virus taxonomy. Ninth report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego.

Kirthi, N., Savithri, H. S., 2003. A conserved zinc finger motif in the coat protein of Tomato leaf curl Bangalore virus is responsible for binding to ssDNA. Arch Virol 148, 2369-80.

Koepsel, R. R., Murray, R. W., Rosenblum, W. D., Khan, S. A., 1985. The replication initiator protein of plasmid pT181 has sequence-specific endonuclease and topoisomerase-like activities. Proc Natl Acad Sci USA 82, 6845-9.

Kon, T., Sharma, P., Ikegami, M., 2007. Suppressor of RNA silencing encoded by the monopartite tomato leaf curl Java begomovirus. Arch virol 152, 1273-1282.

Kong, L. J., Hanley-Bowdoin, L., 2002. A geminivirus replication protein interacts with a protein kinase and a motor protein that display different expression patterns during plant development and infection. Plant Cell 14, 1817-32.

Koonin, E. V., Ilyina, T. V., 1992. Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA replication initiator proteins. J Gen Virol 73 (Pt 10), 2763-6.

Kunik, T., Palanichelvam, K., Czosnek, H., Citovsky, V., Gafni, Y., 1998. Nuclear import of the capsid protein of tomato yellow leaf curl virus (TYLCV) in plant and insect cells. Plant J 13, 393-399.

Laufs, J., Traut, W., Heyraud, F., Matzeit, V., Rogers, S. G., Schell, J., Gronenborn, B., 1995. In vitro cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. Proc Natl Acad Sci USA 92, 3879-83.

Lazarowitz, S. G., Shepherd, R. J., 1992 Geminiviruses: genome structure and gene function. Crit Rev Plant Sci 11, 327-349.

Lazarowitz, S. G., Beachy, R. N., 1999. Viral movement proteins as probes for intracellular and intercellular trafficking in plants. Plant Cell 11, 535-548.

Lewis, J. D., Lazarowitz, S. G., 2010. Arabidopsis synaptotagmin SYTA regulates endocytosis and virus movement protein cell-to-cell transport. Proc Natl Acad Sci USA 107, 2491-6.

Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J. A., Fanning, E., Jochimiak, A., Szakonyi, G., Chen, X. S., 2003. Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. Nature 423, 512-8.

Liao, J. C., Jeong, Y. J., Kim, D. E., Patel, S. S., Oster, G., 2005. Mechanochemistry of t7 DNA helicase. J Mol Biol 350, 452-75.

Liu, L., Davies, J. W., Stanley, J., 1998. Mutational analysis of bean yellow dwarf virus, a geminivirus of the genus Mastrevirus that is adapted to dicotyledonous plants. J Gen Virol 79 (Pt 9), 2265-74.

Liu, Y., Jin, W., Wang, L., Wang, X., 2014. Replication-associated proteins encoded by Wheat dwarf virus act as RNA silencing suppressors. Virus Res 190, 34-9.

Lohman, T. M., Bjornson, K. P., 1996. Mechanisms of helicase-catalyzed DNA unwinding. Annu Rev Biochem 65, 169-214.

Luque, A., Sanz-Burgos, A. P., Ramirez-Parra, E., Castellano, M. M., Gutierrez, C., 2002. Interaction of geminivirus Rep protein with replication factor C and its potential role during geminivirus DNA replication. Virology 302, 83-94.

Mandel, M., Higa, A., 1970. Calcium-dependent bacteriophage DNA infection. J Mol Biol 53, 159-62.

Mansoor, S., Zafar, Y., Briddon, R. W., 2006. Geminivirus disease complexes: the threat is spreading. Trends Plant Sci 11, 209-12.

Mansoor, S., Briddon, R. W., Zafar, Y., Stanley, J., 2003. Geminivirus disease complexes: an emerging threat. Trends Plant Sci 8, 128-34.

Mansoor, S., Khan, S. H., Bashir, A., Saeed, M., Zafar, Y., Malik, K. A., Briddon, R., Stanley, J., Markham, P. G., 1999. Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. Virology 259, 190-9.

Martin, A., Baker, T. A., Sauer, R. T., 2005. Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. Nature 437, 1115-20.

Masterson, P. J., Stanley, M. A., Lewis, A. P., Romanos, M. A., 1998. A C-terminal helicase domain of the human papillomavirus E1 protein binds E2 and the DNA polymerase alpha-primase p68 subunit. J Virol 72, 7407-19.

Modrich, P., 1994. Mismatch repair, genetic stability, and cancer. Science 266, 1959-60.

Morin, S., Ghanim, M., Sobol, I., Czosnek, H., 2000. The GroEL protein of the whitefly Bemisia tabaci interacts with the coat protein of transmissible and nontransmissible begomoviruses in the yeast two-hybrid system. Virology 276, 404-416.

Mubin, M., Briddon, R. W., Mansoor, S., 2009. Diverse and recombinant DNA betasatellites are associated with a begomovirus disease complex of Digera arvensis, a weed host. Virus Res 142, 208-12.

Mullineaux, P. M., Donson, J., Morris-Krsinich, B. A., Boulton, M. I., Davies, J. W., 1984. The nucleotide sequence of maize streak virus DNA. EMBO J 3, 3063-8.

Mumford, D. L., 1974. Purification of curly top virus. Phytopathology 64.

Muniyappa, V., Venkatesh, H. M., Ramappa, H. K., Kulkarni, R. S., Zeidan, M., Tarba, C. Y., Ghanim, M., Czosnek, H., 2000. Tomato leaf curl virus from

Bangalore (ToLCV-Ban4): sequence comparison with Indian ToLCV isolates, detection in plants and insects, and vector relationships. Arch Virol 145, 1583-98.

Nagar, S., Hanley-Bowdoin, L., Robertson, D., 2002. Host DNA replication is induced by geminivirus infection of differentiated plant cells. Plant Cell 14, 2995-3007.

Nagar, S., Pedersen, T. J., Carrick, K. M., Hanley-Bowdoin, L., Robertson, D., 1995. A geminivirus induces expression of a host DNA synthesis protein in terminally differentiated plant cells. Plant Cell 7, 705-19.

Nash, T. E., Dallas, M. B., Reyes, M. I., Buhrman, G. K., Ascencio-Ibanez, J. T., Hanley-Bowdoin, L., 2011. Functional analysis of a novel motif conserved across geminivirus Rep proteins. J Virol 85, 1182-92.

Navot, N., Pichersky, E., Zeidan, M., Zamir, D., Czosnek, H., 1991. Tomato yellow leaf curl virus: A whitefly-transmitted geminivirus with a single genomic component. Virology 185, 151-161.

Nawaz-ul-Rehman, M. S., Mansoor, S., Briddon, R. W., Fauquet, C. M., 2009. Maintenance of an old world betasatellite by a new world helper begomovirus and possible rapid adaptation of the betasatellite. J Virol 83, 9347-55.

Neuwald, A. F., Aravind, L., Spouge, J. L., Koonin, E. V., 1999. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res 9, 27-43.

Norder, H., De Palma, A. M., Selisko, B., Costenaro, L., Papageorgiou, N., Arnan, C., Coutard, B., Lantez, V., De Lamballerie, X., Baronti, C., Sola, M., Tan, J., Neyts, J., Canard, B., Coll, M., Gorbalenya, A. E., Hilgenfeld, R., 2011. Picornavirus non-structural proteins as targets for new anti-virals with broad activity. Antiviral Res 89, 204-18.

Noris, E., Accotto, G. P., Tavazza, R., Brunetti, A., Crespi, S., Tavazza, M., 1996. Resistance to tomato yellow leaf curl geminivirus in Nicotiana benthamiana plants transformed with a truncated viral C1 gene. Virology 224, 130-8.

Noueiry, A. O., Lucas, W. J., Gilbertson, R. L., 1994. Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport. Cell 76, 925-932.

Ohnesorge, S., Bejarano, E., 2009. Begomovirus coat protein interacts with a small heat-shock protein of its transmission vector (Bemisia tabaci). Insect mol biol 18, 693-703.

Orozco, B. M., Hanley-Bowdoin, L., 1998. Conserved sequence and structural motifs contribute to the DNA binding and cleavage activities of a geminivirus replication protein. J Biol Chem 273, 24448-56.

Orozco, B. M., Miller, A. B., Settlage, S. B., Hanley-Bowdoin, L., 1997. Functional domains of a geminivirus replication protein. J Biol Chem 272, 9840-6.

Padidam, M., Beachy, R. N., Fauquet, C. M., 1995. Classification and identification of geminiviruses using sequence comparisons. J Gen Virol 76 (Pt 2), 249-63,.

Pandey, P., Choudhury, N. R., 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.

Pant, V., Gupta, D., Choudhury, N. R., Malathi, V., Varma, A., Mukherjee, S., 2001. Molecular characterization of the Rep protein of the blackgram isolate of Indian mungbean yellow mosaic virus. Journal of General Virology 82, 2559-2567.

Paprotka, T., Deuschle, K., Metzler, V., Jeske, H., 2011. Conformation-selective methylation of geminivirus DNA. J Virol 85, 12001-12.

Pascal, E., Sanderfoot, A. A., Ward, B. M., Medville, R., Turgeon, R., Lazarowitz, S. G., 1994. The geminivirus BR1 movement protein binds single-stranded DNA and localizes to the cell nucleus. Plant Cell 6, 995-1006.

Pasumarthy, K. K., Choudhury, N. R., Mukherjee, S. K., 2010. Tomato leaf curl Kerala virus (ToLCKeV) AC3 protein forms a higher order oligomer and enhances ATPase activity of replication initiator protein (Rep/AC1). Virol J 7, 128.

Patel, S. S., Picha, K. M., 2000. Structure and function of hexameric helicases. Annu Rev Biochem 69, 651-97.

Patil, B. L., Fauquet, C. M., 2009. Cassava mosaic geminiviruses: actual knowledge and perspectives. Mol Plant Pathol 10, 685-701.

Phaneendra, C., Rao, K. R., Jain, R. K., Mandal, B., 2012. Tomato leaf curl New Delhi virus is Associated With Pumpkin Leaf Curl: A New Disease in Northern India. Indian J Virol 23, 42-5.

Pilartz, M., Jeske, H., 1992. Abutilon mosaic geminivirus double-stranded DNA is packed into minichromosomes. Virology 189, 800-2.

Polston, J. E., Anderson, P. K., 1997. The emergence of whitefly-transmitted geminiviruses in tomato in the western hemisphere. Plant Disease 81, 1358-1369.

Pooggin, M. M., 2013. How can plant DNA viruses evade siRNA-directed DNA methylation and silencing? Int J Mol Sci 14, 15233-59.

References

Preiss, W., Jeske, H., 2003. Multitasking in replication is common among geminiviruses. J Virol 77, 2972-80.

Raja, P., Wolf, J. N., Bisaro, D. M., 2010. RNA silencing directed against geminiviruses: post-transcriptional and epigenetic components. Biochim Biophys Acta 1799, 337-51.

Raja, P., Sanville, B. C., Buchmann, R. C., Bisaro, D. M., 2008. Viral genome methylation as an epigenetic defense against geminiviruses. J Virol 82, 8997-9007.

Ramappa, H. K., Muniyappa, V., Colvin, J., 1998. The contribution of tomato and alternative host plants to tomato leaf curl virus inoculum pressure in different areas of South India. Ann Appl Biol 133, 187-198.

Reyes, M. I., Nash, T. E., Dallas, M. M., Ascencio-Ibanez, J. T., Hanley-Bowdoin, L., 2013. Peptide aptamers that bind to geminivirus replication proteins confer a resistance phenotype to tomato yellow leaf curl virus and tomato mottle virus infection in tomato. J Virol 87, 9691-706.

Richter, K. S., Kleinow, T., Jeske, H., 2014. Somatic homologous recombination in plants is promoted by a geminivirus in a tissue-selective manner. Virology 452-453, 287-96.

Richter, K. S., Ende, L., Jeske, H., 2015. Rad54 is not essential for any geminiviral replication mode in planta. Plant Mol Biol 87, 193-202.

Richter, K. S., Gotz, M., Winter, S., Jeske, H., 2016a. The contribution of translesion synthesis polymerases on geminiviral replication. Virology 488, 137-48.

Richter, K. S., Serra, H., White, C. I., Jeske, H., 2016b. The recombination mediator RAD51D promotes geminiviral infection. Virology 493, 113-27.

Rigden, J. E., Krake, L. R., Rezaian, M. A., Dry, I. B., 1994. ORF C4 of tomato leaf curl geminivirus is a determinant of symptom severity. Virology 204, 847-850.

Rizvi, I., Choudhury, N. R., Tuteja, N., 2015. Insights into the functional characteristics of geminivirus rolling-circle replication initiator protein and its interaction with host factors affecting viral DNA replication. Arch Virol 160, 375-87.

Rochester, D. E., Kositratana, W., Beachy, R. N., 1990. Systemic movement and symptom production following agroinoculation with a single DNA of tomato yellow leaf curl geminivirus (Thailand). Virology 178, 520-6.

Rodriguez-Negrete, E., Lozano-Duran, R., Piedra-Aguilera, A., Cruzado, L., Bejarano, E. R., Castillo, A. G., 2013. Geminivirus Rep protein interferes with

the plant DNA methylation machinery and suppresses transcriptional gene silencing. New Phytol 199, 464-75.

Rojas, M. R., Hagen, C., Lucas, W. J., Gilbertson, R. L., 2005. Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. Annu Rev Phytopathol 43, 361-94.

Rojas, M. R., Jiang, H., Salati, R., Xoconostle-Cázares, B., Sudarshana, M., Lucas, W. J., Gilbertson, R. L., 2001. Functional analysis of proteins involved in movement of the monopartite begomovirus, Tomato yellow leaf curl virus. Virology 291, 110-125.

Roossinck, M. J., Sleat, D., Palukaitis, P., 1992. Satellite RNAs of plant viruses: structures and biological effects. Microbiol Rev 56, 265-79.

Ruschhaupt, M., Martin, D. P., Lakay, F., Bezuidenhout, M., Rybicki, E. P., Jeske, H., Shepherd, D. N., 2013. Replication modes of Maize streak virus mutants lacking RepA or the RepA-pRBR interaction motif. Virology 442, 173-9.

Rybicki, E., 1994. A phylogenetic and evolutionary justification for three genera of Geminiviridae. Archives of Virology 139, 49-77.

Rybicki, E. P., 2015. A Top Ten list for economically important plant viruses. Arch Virol 160, 17-20.

Saha, A., Wittmeyer, J., Cairns, B. R., 2006. Chromatin remodelling: the industrial revolution of DNA around histones. Nat Rev Mol Cell Biol 7, 437-47.

Saikia, A. K., V. Muniyappa, 1989. Epidemiology and control of tomato leaf curl virus in Southern India. Trop Agric 66, 350-354.

Sambrook, J., Russell, D. W., 2001. Molecular cloning : a laboratory manual 3rd. ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Sánchez-Campos, S., Martínez-Ayala, A., Márquez-Martín, B., Aragón-Caballero, L., Navas-Castillo, J., Moriones, E., 2013. Fulfilling Koch's postulates confirms the monopartite nature of tomato leaf deformation virus: A begomovirus native to the New World. Virus Res 173, 286-293.

Sanchez-Duran, M. A., Dallas, M. B., Ascencio-Ibanez, J. T., Reyes, M. I., Arroyo-Mateos, M., Ruiz-Albert, J., Hanley-Bowdoin, L., Bejarano, E. R., 2011. Interaction between geminivirus replication protein and the SUMO-conjugating enzyme is required for viral infection. J Virol 85, 9789-800.

Sanderfoot, A. A., Lazarowitz, S. G., 1995. Cooperation in viral movement: the geminivirus BL1 movement protein interacts with BR1 and redirects it from the nucleus to the cell periphery. Plant Cell 7, 1185-1194.

Sanderfoot, A. A., Ingham, D. J., Lazarowitz, S. G., 1996. A Viral Movement Protein as a Nuclear Shuttle (The Geminivirus BR1 Movement Protein Contains Domains Essential for Interaction with BL1 and Nuclear Localization). Plant Physiol 110, 23-33.

Sardo, L., Lucioli, A., Tavazza, M., Masenga, V., Tavazza, R., Accotto, G. P., Noris, E., 2011. An RGG sequence in the replication-associated protein (Rep) of Tomato yellow leaf curl Sardinia virus is involved in transcriptional repression and severely impacts resistance in Rep-expressing plants. J Gen Virol 92, 204-9.

Saunders, K., Stanley, J., 1999. A nanovirus-like DNA component associated with yellow vein disease of Ageratum conyzoides: evidence for interfamilial recombination between plant DNA viruses. Virology 264, 142-52.

Saunders, K., Lucy, A., Stanley, J., 1991. DNA forms of the geminivirus African cassava mosaic virus consistent with a rolling circle mechanism of replication. Nucleic Acids Res 19, 2325-30.

Saunders, K., Lucy, A., Stanley, J., 1992. RNA-primed complementary-sense DNA synthesis of the geminivirus African cassava mosaic virus. Nucleic Acids Res 20, 6311-5.

Saunders, K., Bedford, I. D., Briddon, R. W., Markham, P. G., Wong, S. M., Stanley, J., 2000. A unique virus complex causes Ageratum yellow vein disease. Proc Natl Acad Sci USA 97, 6890-5.

Schneider, I. R., 1969. Satellite-like particle of tobacco ringspot virus that resembles tobacco ringspot virus. Science 166, 1627-9.

Seal, S., VandenBosch, F., Jeger, M., 2006. Factors influencing begomovirus evolution and their increasing global significance: implications for sustainable control. Crit Rev Plant Sci 25, 23-46.

Sedman, J., Stenlund, A., 1998. The papillomavirus E1 protein forms a DNA-dependent hexameric complex with ATPase and DNA helicase activities. J Virol 72, 6893-7.

Serra, H., Da Ines, O., Degroote, F., Gallego, M. E., White, C. I., 2013. Roles of XRCC2, RAD51B and RAD51D in RAD51-independent SSA recombination. PLoS Genet 9, e1003971.

Settlage, S. B., Miller, A. B., Gruissem, W., Hanley-Bowdoin, L., 2001. Dual interaction of a geminivirus replication accessory factor with a viral replication protein and a plant cell cycle regulator. Virology 279, 570-576.

Seybert, A., Hegyi, A., Siddell, S. G., Ziebuhr, J., 2000. The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5'-to-3' polarity. RNA 6, 1056-68.

Sharma, P., Ikegami, M., 2009. Characterization of signals that dictate nuclear/nucleolar and cytoplasmic shuttling of the capsid protein of Tomato leaf curl Java virus associated with DNAβ satellite. Virus Res 144, 145-153.

Shen, J., Gai, D., Patrick, A., Greenleaf, W. B., Chen, X. S., 2005. The roles of the residues on the channel beta-hairpin and loop structures of simian virus 40 hexameric helicase. Proc Natl Acad Sci USA 102, 11248-53.

Shirazi, M., Mozafari, J., Rakhshandehroo, F., Shams-Bakhsh, M., 2014. Genetic diversity, host range, and distribution of tomato yellow leaf curl virus in Iran. Acta Virol 58, 128-36.

Shung, C. Y., Sunter, G., 2007. AL1-dependent repression of transcription enhances expression of Tomato golden mosaic virus AL2 and AL3. Virology 364, 112-22.

Singh, D. K., Islam, M. N., Choudhury, N. R., Karjee, S., Mukherjee, S. K., 2007. The 32 kDa subunit of replication protein A (RPA) participates in the DNA replication of Mung bean yellow mosaic India virus (MYMIV) by interacting with the viral Rep protein. Nucleic Acids Res 35, 755-70.

Singleton, M. R., Wigley, D. B., 2002. Modularity and specialization in superfamily 1 and 2 helicases. J Bacteriol 184, 1819-26.

Singleton, M. R., Sawaya, M. R., Ellenberger, T., Wigley, D. B., 2000. Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. Cell 101, 589-600.

Southern, E. M., 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98, 503-17.

Srivastava, K. M., Hallan, V., Raizada, R. K., Chandra, G., Singh, B. P., Sane, P. V., 1995. Molecular cloning of Indian tomato leaf curl virus genome following a simple method of concentrating the supercoiled replicative form of viral DNA. J Virol Methods 51, 297-304.

Stanley, J., Gay, M. R., 1983. Nucleotide sequence of cassava latent virus DNA. Nature 301, 260-262.

Stanley, J., Saunders, K., Pinner, M. S., Wong, S. M., 1997. Novel defective interfering DNAs associated with ageratum yellow vein geminivirus infection of Ageratum conyzoides. Virology 239, 87-96.

Stenger, D. C., Revington, G. N., Stevenson, M. C., Bisaro, D. M., 1991. Replicational release of geminivirus genomes from tandemly repeated copies: evidence for rolling-circle replication of a plant viral DNA. Proc Natl Acad Sci USA 88, 8029-33.

Stitt, B. L., Xu, Y., 1998. Sequential hydrolysis of ATP molecules bound in interacting catalytic sites of Escherichia coli transcription termination protein Rho. J Biol Chem 273, 26477-86.

Story, R. M., Weber, I. T., Steitz, T. A., 1992. The structure of the E. coli recA protein monomer and polymer. Nature 355, 318-25.

Sunter, G., Hartitz, M. D., Bisaro, D. M., 1993. Tomato golden mosaic virus leftward gene expression: autoregulation of geminivirus replication protein. Virology 195, 275-80.

Sunter, G., Hartitz, M. D., Hormuzdi, S. G., Brough, C. L., Bisaro, D. M., 1990. Genetic analysis of tomato golden mosaic virus: ORF AL2 is required for coat protein accumulation while ORF AL3 is necessary for efficient DNA replication. Virology 179, 69-77.

Suyal, G., Mukherjee, S. K., Choudhury, N. R., 2013. The host factor RAD51 is involved in mungbean yellow mosaic India virus (MYMIV) DNA replication. Arch Virol 158, 1931-41.

Tiwari, N., Singh, V. B., Sharma, P. K., Malathi, V. G., 2013. Tomato leaf curl Joydebpur virus: a monopartite begomovirus causing severe leaf curl in tomato in West Bengal. Arch Virol 158, 1-10.

Towbin, H., Staehelin, T., Gordon, J., 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci USA 76, 4350-4.

Trinks, D., Rajeswaran, R., Shivaprasad, P., Akbergenov, R., Oakeley, E. J., Veluthambi, K., Hohn, T., Pooggin, M. M., 2005. Suppression of RNA silencing by a geminivirus nuclear protein, AC2, correlates with transactivation of host genes. Virol J 79, 2517-2527.

Uchiyama, A., Shimada-Beltran, H., Levy, A., Zheng, J. Y., Javia, P. A., Lazarowitz, S. G., 2014. The Arabidopsis synaptotagmin SYTA regulates the cell-to-cell movement of diverse plant viruses. Front Plant Sci 5, 584.

Vanitharani, R., Chellappan, P., Pita, J. S., Fauquet, C. M., 2004. Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. Virol J 78, 9487-9498.

Varma, A., Malathi, V. G., 2003. Emerging geminivirus problems: A serious threat to crop production. Ann Appl Biol 142, 145-164.

Varsani, A., Shepherd, D. N., Dent, K., Monjane, A. L., Rybicki, E. P., Martin, D. P., 2009. A highly divergent South African geminivirus species illuminates the ancient evolutionary history of this family. Virol J 6, 36.

Varsani, A., Navas-Castillo, J., Moriones, E., Hernandez-Zepeda, C., Idris, A., Brown, J. K., Murilo Zerbini, F., Martin, D. P., 2014. Establishment of three new genera in the family Geminiviridae: Becurtovirus, Eragrovirus and Turncurtovirus. Arch Virol 159, 2193-203.

Vasudeva, R. S., Sam Raj, J. , 1948. A leaf curl disease of tomato. Phytopathology 38, 364-369.

Veaute, X., Delmas, S., Selva, M., Jeusset, J., Le Cam, E., Matic, I., Fabre, F., Petit, M. A., 2005. UvrD helicase, unlike Rep helicase, dismantles RecA nucleoprotein filaments in Escherichia coli. EMBO J 24, 180-9.

Voinnet, O., Pinto, Y. M., Baulcombe, D. C., 1999. Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. Proc Natl Acad Sci USA 96, 14147-14152.

Von Hippel, P. H., Delagoutte, E., 2001. A general model for nucleic acid helicases and their "coupling" within macromolecular machines. Cell 104, 177-90.

Wang, H., Hao, L., Shung, C.-Y., Sunter, G., Bisaro, D. M., 2003. Adenosine kinase is inactivated by geminivirus AL2 and L2 proteins. Plant Cell 15, 3020-3032.

Wartig, L., Kheyr-Pour, A., Noris, E., De Kouchkovsky, F., Jouanneau, F., Gronenborn, B., Jupin, I., 1997a. Genetic analysis of the monopartite tomato yellow leaf curl geminivirus: roles of V1, V2, and C2 ORFs in viral pathogenesis. Virology 228, 132-140.

Wartig, L., Kheyr-Pour, A., Noris, E., De Kouchkovsky, F., Jouanneau, F., Gronenborn, B., Jupin, I., 1997b. Genetic analysis of the monopartite tomato yellow leaf curl geminivirus: roles of V1, V2, and C2 ORFs in viral pathogenesis. Virology 228, 132-40.

Xia, H., Wang, P., Wang, G. C., Yang, J., Sun, X., Wu, W., Qiu, Y., Shu, T., Zhao, X., Yin, L., Qin, C. F., Hu, Y., Zhou, X., 2015. Human Enterovirus Nonstructural Protein 2CATPase Functions as Both an RNA Helicase and ATP-Independent RNA Chaperone. PLoS Pathog 11, e1005067.

Xie, Q., Suarez-Lopez, P., Gutierrez, C., 1995. Identification and analysis of a retinoblastoma binding motif in the replication protein of a plant DNA virus: requirement for efficient viral DNA replication. EMBO J 14, 4073-82.

Xie, Q., Sanz-Burgos, A. P., Guo, H., Garcia, J. A., Gutierrez, C., 1999. GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. Plant Mol Biol 39, 647-56.

Yadava, P., Suyal, G., Mukherjee, S. K., 2010. Begomovirus DNA replication and pathogenicity. Curr Sci (00113891) 98.

Yang, J. Y., Iwasaki, M., Machida, C., Machida, Y., Zhou, X., Chua, N. H., 2008. betaC1, the pathogenicity factor of TYLCCNV, interacts with AS1 to alter leaf development and suppress selective jasmonic acid responses. Genes Dev 22, 2564-77.

Yazdi, H. R., Heydarnejad, J., Massumi, H., 2008. Genome characterization and genetic diversity of beet curly top Iran virus: a geminivirus with a novel nonanucleotide. Virus Genes 36, 539-45.

Yoon-Robarts, M., Blouin, A. G., Bleker, S., Kleinschmidt, J. A., Aggarwal, A. K., Escalante, C. R., Linden, R. M., 2004. Residues within the B' motif are critical for DNA binding by the superfamily 3 helicase Rep40 of adeno-associated virus type 2. J Biol Chem 279, 50472-81.

Zhang, W., Olson, N. H., Baker, T. S., Faulkner, L., Agbandje-McKenna, M., Boulton, M. I., Davies, J. W., McKenna, R., 2001. Structure of the Maize streak virus geminate particle. Virology 279, 471-7.

Zhou, X., Xie, Y., Tao, X., Zhang, Z., Li, Z., Fauquet, C. M., 2003. Characterization of DNAbeta associated with begomoviruses in China and evidence for co-evolution with their cognate viral DNA-A. J Gen Virol 84, 237-47.

Zhou, Y., Rojas, M. R., Park, M. R., Seo, Y. S., Lucas, W. J., Gilbertson, R. L., 2011. Histone H3 interacts and colocalizes with the nuclear shuttle protein and the movement protein of a geminivirus. J Virol 85, 11821-32.

Zrachya, A., Glick, E., Levy, Y., Arazi, T., Citovsky, V., Gafni, Y., 2007. Suppressor of RNA silencing encoded by Tomato yellow leaf curl virus-Israel. Virology 358, 159-165.

ANNEXURE

APPENDIX-I

PREPARATION OF REAGENTS, BUFFERS AND MEDIA

Buffer/Reagent/Media	Method of Preparation

1. For agarose gel electrophoresis

i.	50X Tris acetate EDTA (TAE)	Tris base (242 g), 57.1 ml of Glacial acetic acid, 0.5 M EDTA was added and the volume was made up to 1000 ml with distilled H_2O. The pH was adjusted to 8.0.
ii.	6X DNA loading dye	0.25% (W/V) bromophenol blue, 0.25% (W/V) xylene cyanol FF, 30% (V/V) glycerol was dissolved in distilled H_2O and stored at 4 ºC.
iii.	DNA Molecular Marker	Two marker were used 1) Fermentas DNA marker-(ladder size in bp) 10000, 8000, 6000, 5000, 4000, 3500, 3000, 2500, 2000, 1500, 1000, 750, 500 and 250. 2) NEB DNA marker-(ladder size in bp) 10000, 8000, 6000, 5000, 4000, 3000, 2000, 1500, 1000 and 500.
iv.	Ethidium bromide (10mg/ml)	Ethidium bromide (1 g) was added to 100 ml of distilled H_2O, stirred at magnetic stirrer for several hours, transferred to a dark bottle and stored at room temperature.

2. For cloning

i.	Ampicillin	Stock solution (100 mg/ml) was made in double distilled H_2O, filter sterilized (through 0.22 μm filter) and aliquoted in 1.5 tubes and stored at -20ºC.
ii.	Kanamycin	Stock solution (50 mg/ml) was made in double distilled H_2O, filter sterilized (through 0.22 μm filter) and aliquoted in 1.5 tubes and stored at -20ºC.

iii.	Luria broth media	Bacto-tryptone (1 g), 0.5 g of yeast extract, 1 g of NaCl was added and dissolved in 95 ml of water. After adjusting the pH to 7.0, the solution was made upto 100 ml with with distilled H_2O and sterilized by auroclaving.
		For agrobacterium culture LB media was prepared and autoclaved. After autoclaving Rifampicin was added to the final concentration of 30 μg/ml. Filter sterilized glucose was added to the final concentration of 1%.
iv.	Luria agar media	Bacto-tryptone (1 g), 0.5 g of yeast extract, 1 g of NaCl was added and dissolved in 95 ml of water. After adjusting the pH to 7.0, the solution was made upto 100 ml with with distilled H_2O. Then 1.5 g of bacto-agar was added and sterilized by autoclaving.
v.	X-gal (5-bromo-4chloro-3-indolyl-b-D-galactopyranoside)	Stock solution (20 mg/ml) was prepared by dissolving X-gal in dimethyl foramide, and stored at -20ºC.
vi.	IPTG (Isopropyl thio-a-galactoside)	IPTG (2 g) was dissolved in distilled H_2O, and volume was adjusted to 10 ml then filter sterilized through 0.22 μm disposable filter and was stored at -20ºC.
vii.	100 mM $MgCl_2$	$MgCl_2.6H_2O$ (2.03 g) was dissolved in distilled H_2O, volume adjusted to 100 ml and sterilized by autoclaving.
viii.	100 mM $CaCl_2$	$CaCl_2.2H_2O$ (1.47 g) was dissolved in distilled H_2O, volume adjusted to 100 ml and sterilized by autoclaving.

3. For plasmid isolation

i.	Solution I (Resuspension buffer)	It was prepared by adjusting final concentration to 1M Tris HCl (25 mM), (pH 8.0)- 2.5 ml 20% Glucose (50 mM) -

4.5 ml 0.5 M EDTA (10 mM)- 2.0 ml and sterile distilled H_2O- Up to 100 ml.

ii. Solution II (Lysis buffer)	It was prepared freshly by mixing 10N NaOH (0.2 N)- 2.0 ml, 1% SDS- 5.0 ml, distilled H_2O- Up to 100 ml
iii. Solution III (pH 4.8) (Neutralization buffer)	Sodium acetate (40.81 g) was dissolved in minimum volume of distilled H_2O, and pH was adjusted to 4.8 with Glacial acetic acid, now volume adjusted to 100 ml with H_2O and sterilized by autoclaving.
iv. Phenol: Chloroform: Isoamylalcohol	It was prepared by mixing Tris-saturated Phenol (pH 7.5) 25 ml, Chloroform 24 ml, and Isoamylalcohol 1 ml.

4. For total DNA extraction

i. For Dellaporta method	Extraction buffer was prepared by adjusting final concentration to 100 mM Tris pH 8.0, 50 mM EDTA pH 8.0, 500 mM NaCl, 10 mM 2- mercaptoethanol.
ii. For CTAB method	Extraction buffer was prepared by adjusting final concentration to 100 mM TrisCl pH 8.0, 100 mM EDTA pH 8.0, 1.4 M NaCl, 2% CTAB (w/v). Solution was incubated this solution at 65ºC for 30-60 min (till the CTAB completely dissolves to make a homogenous solution), and 1% 2-mercaptoethanol was added just before use.

5. For southern hybridization

20X SSC	It was prepared by adding 175.3 g NaCl, 88.2 g Sodium citrate (pH adjusted to 7.0 by adding 14 N HCl) in 1 litre of H_2O.
Denaturation solution	It was prepared by adjusting final concentration of 1.5 M NaCl and 0.5 M NaOH, in H_2O.

Denhardt's solution(50X)	It was prepared by adding 1% (W/V) ficol, 1% (W/V) polyvinylpyrrolidone, 1% (W/V) BSA, and was dissolved in H_2O to prepare a 50X solution.
Depurination solution	It was prepared by adding 2.8 ml of HCl in 197.2 ml of sterile H_2O.
Neutralization solution	It was prepared by adding 1.5 M NaCl, 1.0 M Tris-HCI (pH 7.2).
Herring sperm	DNA (100 mg) was dissolved in 10 ml of sterile distilled H_2O and mixed by vortexing, stored at -20ºC.
Hybridization buffer	It was prepared by adding 0.25 M sodium phosphate buffer (pH 7.2), 7% (w/v) SDS, 1 mM EDTA.

6. For Western blotting

1X transfer buffer	Tris base (6 gm) and 3 gm glycine were dissolved in water. To this 4 ml 10% SDS and 100 ml of methanol was added. Volume of total mixture was made up to 1 lt by adding water.
20 X PBS	NaCl (160 gm), 4 gm KCl, 28.89 gm Disodiumhydrogen phosphate, and 4.8 gm potassiumdihydrogen phosphate were dissolved in 1 L water.
PBST	For 2 L PBST, 10 ml 20X PBS was diluted with 1890 ml of water and 100 ul of tween-20 was added.
Blocking buffer	Five gm of non-skimmed milk powder was (5% non- skimmed milk powder) added to 100 ml of 1X PBST.
Developing solution	50 µl of 1% DAB was added to 5 ml of PBS and 50 µl of 0.3% peroxide was added to this solution slowly and was mixed well before use.

7. For protein expression and Purification

Tris-Cl (1M): pH 6.8	In 800 ml of distilled water 121.1 gm of Tris base was dissolved. The pH was adjusted to 6.8 by adding concentrated HCl. The volume of the solution was then adjusted to 1L.
Tris-Cl (1.5 M): pH 8.8	In 800 ml of distilled water 181.6 gm of Tris base was dissolved. The pH was adjusted to 8.8 by adding concentrated HCl. The volume of the solution was then adjusted to 1 L.
Lysozyme (10mg/ml)	Lysozyme at a concentration of 10 mg/ml was dissolved in 10mM Tris-Cl (pH-8.0).
PMSF (100mM)	A stock solution was prepared by dissolving 17 mg of PMSF in 1 ml of isopropanol by vortexing and stored at -20ºC.
Acrylamide Solution (30%w/v)	

Acrylamide (DNA-sequencing grade)	29 gm
N, N'-methylenebisacrylamide	1 gm
H_2O	60 ml

Volume was adjusted to 100 ml with distilled water. Thereafter solution was filtered through a nitrocellulose filter. And stored in dark bottles at room temperature.

Ammonium Persulfate (10%w/v)	One gm of ammonium persulfate was dissolved in 10 ml of H_2O and stored at 4ºC
1X tris-glycine SDS buffer	Gel electrophoresis buffer was prepared with 250 mM glycine and 25 mM tris and pH was adjusted to 8.3. Then SDS was added to the final concentration of 0.1%.
6X SDS loading dye	For 20 ml of gel loading buffer 6 ml of 1 M tris was added to 10 ml water. 0.4 gm SDS and 0.02 gm bromophenol blue was added to it and was solubilized properly. Than 2 ml of 100% glycerol was added.
Dialysis buffer	50 mM Tris-HCl, 100 mM NaCl, 10% Glycerol, 1 mM PMSF, 5 mM β-Mercaptoethanol

Staining Solution (100 ml)

Methanol	45 ml
Acetic Acid	10 ml
Double distilled H_2O	45 ml
Commassie Brilliant Blue	0.25 g

Destaining Solution (100 ml)

Methanol	45 ml
Acetic Acid	10 ml
Double distilled H_2O	45 ml

Protein Extraction Buffers

	Lysis Buffer	Wash Buffer	Elution Buffer
Tris Cl pH 8.0	50 mM	50 mM	50 mM
NaCl	100 mM	150 mM	150 mM
PMSF	1 mM	1 mM	1 mM
$MgCl_2$	5 mM	5 mM	5 mM
Imidazole	10 mM	100 mM	500 mM
β- ME	5 mM	5 mM	5 mM

10 X SDS Running Buffer (pH 8.3)

Glycine	188 g
Tris	30 g
SDS	10 g
H_2O	1 litre (make up the volume after adjusting pH)

APPENDIX-II

PREPARATION OF COMMONLY USED STOCK SOLUTION

Calcium chloride (2.5 M)

$CaCl_2$.$6H_2O$ (11 g) was dissolved in a final volume of 20 ml of distilled H_2O. The solution was sterilized by passing it through a 0.22 µm filter. It was then stored in 1 ml aliquots at 4ºC.

Deoxyribonucleoside Triphosphates (dNTPs)

Each dNTP was dissolved in H_2O at an approximate concentration of 100 mM. 0.05 M Tris base and a micropipette was used to adjust the pH of each of the solutions to 7.0 (pH paper was used to check the pH). An aliquot of the neutralized dNTP was diluted appropriately and the optical density at the wavelengths as given below in the table, was recorded. The actual concentration of each dNTP was calculated. The solutions were diluted with H_2O to a final concentration of 50mM dNTP. Each was stored separately at -70ºC in small aliquots.

Base	Wavelength	Extinction Coefficient (E) ($M^{-1}cm^{-1}$)
A	259	1.54×10^4
G	253	1.37×10^4
C	271	9.10×10^3
T	267	9.60×10^3

EDTA (0.5M, pH 8.0)

Disodium EDTA-$2H_2O$ (186.1 g) was added to 800 ml of H_2O. It was stirred vigorously on a magnetic stirrer and pH was adjusted to 8.0 with NaOH (20 g of NaOH pellets). It was dispensed into aliquots and sterilized by autoclaving.

NaOH (10 N)

NaOH (400 g) pellets were added to 800 ml of H_2O, stirring continuously. The volume was adjusted to 1 litre with H_2O once the pellet got dissolved completely. The solution was stored in a plastic container at room temperature.

NaCl (Sodium Chloride, 5 M)

NaCl (292 g) was dissolved 800 ml of H_2O. The volume was adjusted to 1 litre with H_2O. It was dispensed into aliquots and then sterilized by autoclaving. The NaCl solution was stored at room temperature.

Potassium Acetate (5 M)

Potassium acetate (5 M)   - 60 ml

Glacial acetic acid   - 11.5 ml

H_2O - 28.5 ml

The resulting solution was 3 M with respect to potassium and 5 M with respect to acetate. The buffer was stored at room temperature.

SDS (20% w/v)

Electrophoresis-grade SDS (200 g) was dissolved in 900 ml of H_2O. It was heated to 68ºC and stirred with magnetic stirrer to assist dissolution. The volume was adjusted to 1 litre with H_2O and stored at room temperature.

Sodium Acetate (3 M, pH 5.2 and pH 7.0)

Sodium acetate $3H_2O$ (408.3 g) was dissolved into 800 ml of H_2O. The pH was adjusted to 5.2 with glacial acetic acid. It can also be adjusted to 7.0 with dilute acetic acid. The volume was adjusted to 1 litre with H_2O. It was later dispensed into aliquots and sterilized by autoclaving.

Tris HCl

Tris base (121.1 g) was dissolved in 800 ml of H_2O. The pH was adjusted to desired value by adding concentrated HCl, as mentioned below.

pH	HCl
7.4	70 ml
7.6	60 ml
8.0	42 ml

The solution was allowed to cool at room temperature and the pH was adjusted. Then, volume of solution was adjusted to 1 litre with H_2O. This was then aliquoted and sterilised by autoclaving.

X-gal Solution (2% w/v)

A stock solution was prepared by dissolving X-gal in di-methylformamide at a concentration of 20 mg/ml solution. A glass or polypropylene tube was used. The tube containing the solution was wrapped in aluminium foil to prevent damage by light and was stored at -20ºC. Sterilization of X- gal solution by filtration was not required.

Rifampicin (30 mg/ml)

Rifampicin (30 mg) was dissolved in 1 ml of methanol. It was mixed properly by vortexing and wrapped with aluminium foil and stored at -20°C.

Cefotaxime (200 mg/ml)

One gm of Cefotaxim was dissolved in 5 ml of sterile distilled water. The solution was stored at 4°C.

Streptomycin (1 gm/ml)

One gram of Streptomycin was dissolved in 1ml of sterile distilled water. The solution was stored at 4°C.
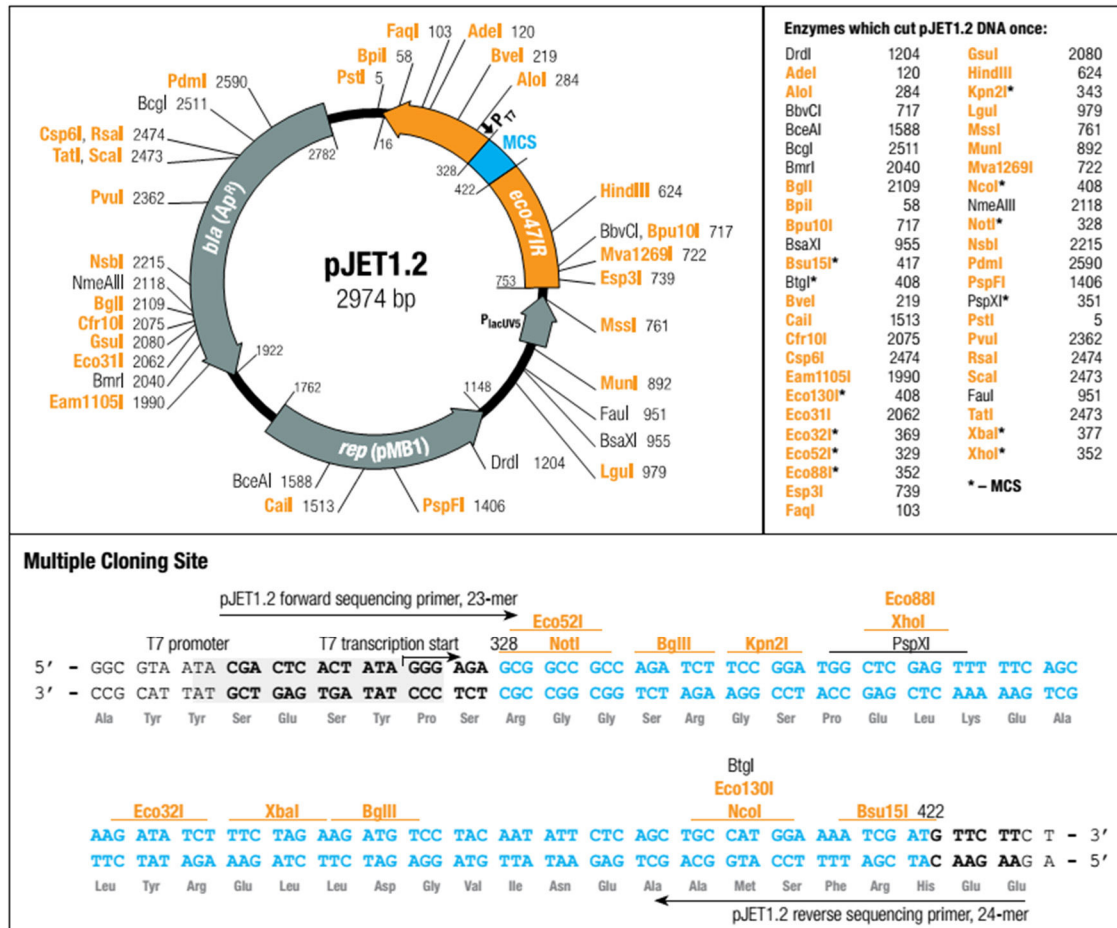
APPENDIX-III

VECTOR MAPS

Figure 1. Map of pJET 1.2 vector showing the position of multiple cloning sites (MCS). T7 promoter, replication origin and ampicillin resistance gene (*bla*). This vector contains a lethal restriction enzyme gene that is disrupted by ligation of a DNA insert into the MCS. As a result, only bacterial cells with recombinant plasmids are able to form colonies.
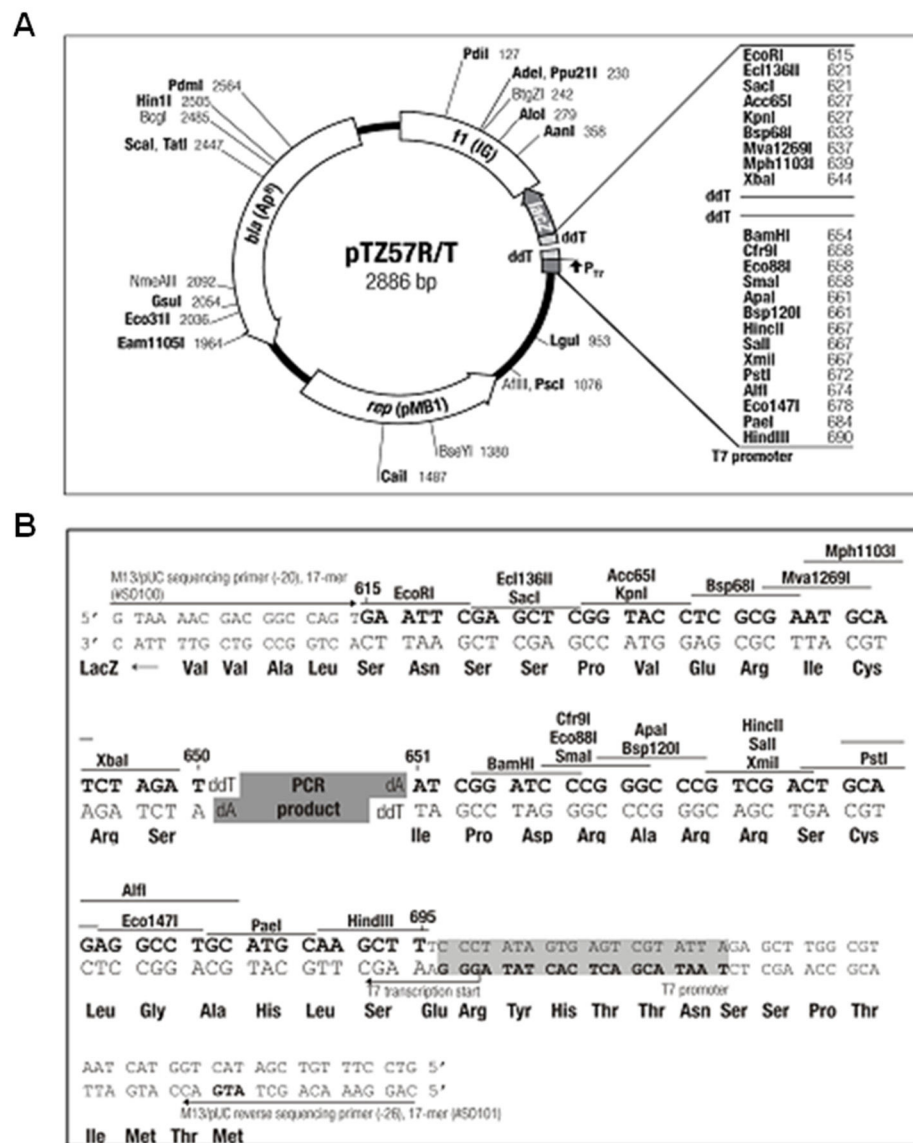
Figure 2. Map of pTZ57R/T vector showing unique restriction sites, LacZ gene, T7 promoter, replication origin and ampicillin resistance gene (A) and DNA sequence of MCS region.
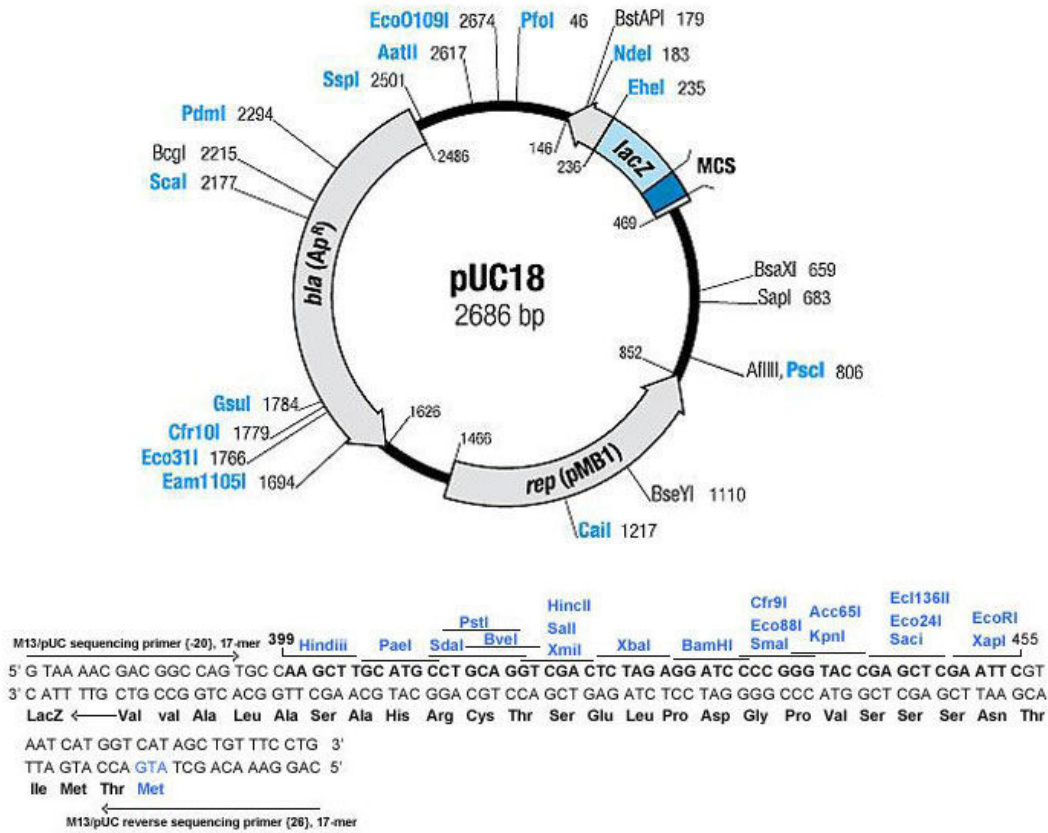
Figure 3. Map of pUC 18 vector showing the position of multiple cloning sites, LacZ promoter, replication origin and ampicillin resistance gene (bla). DNA sequence of the MCS is also indicated.
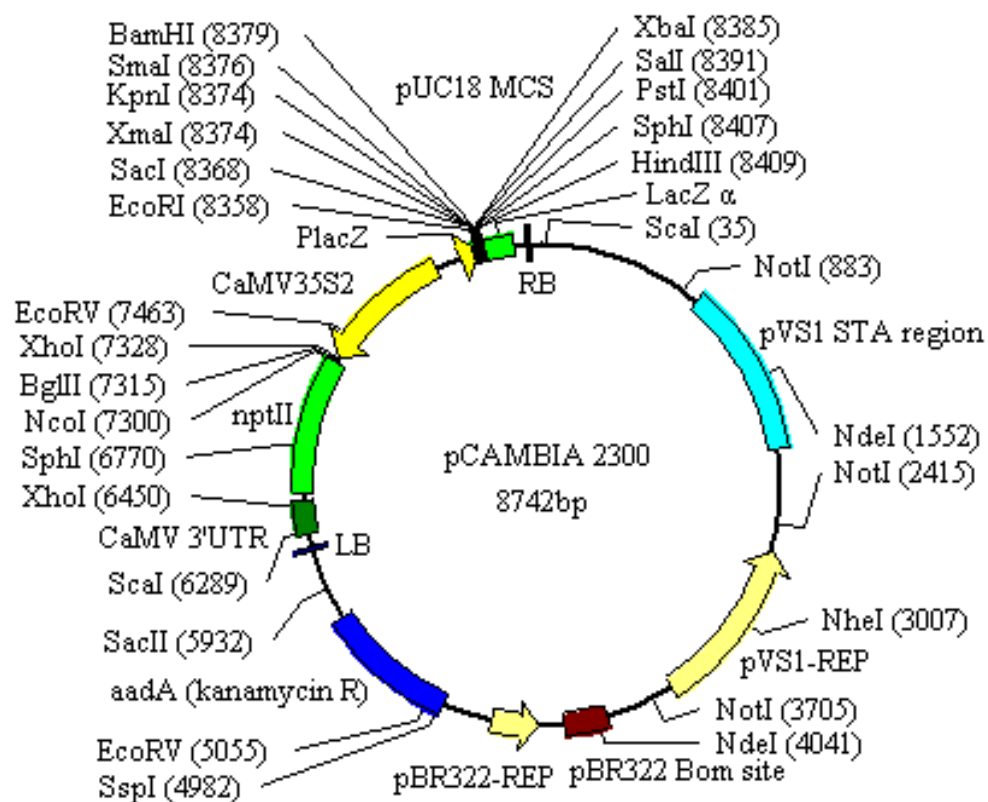
Figure 4. Schematic diagram of plant binary vector pCAMBIA 2300 showing position of MCS. NPTII gene, left and right borders and pBR322 replication origin site.
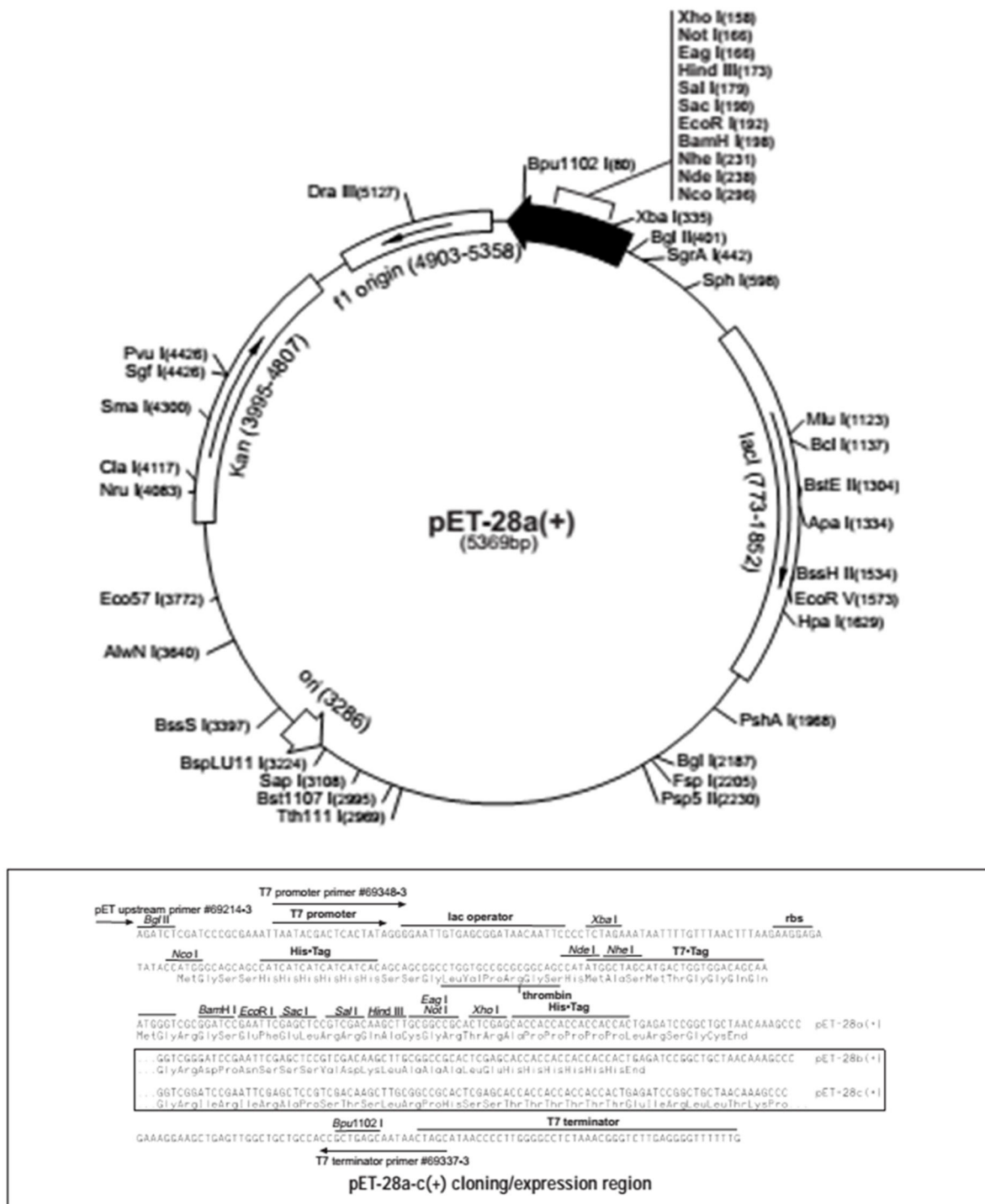
Figure 5. Map of pET 28 vector showing positions of multiple cloning site, T7 promoter, Kanamycin resistance gene and DNA sequence of cloning/expression region.
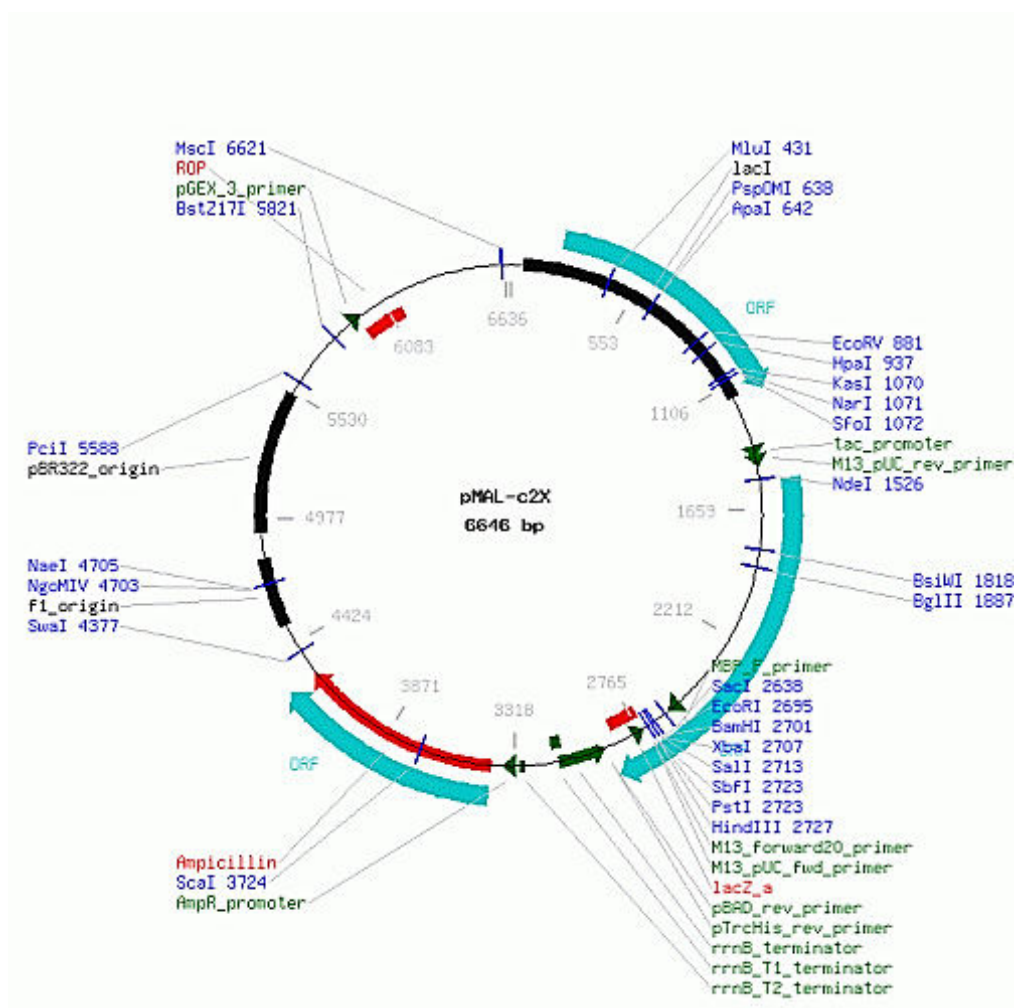
Figure 6. Map of pMAL-c2X vector showing position of multiple cloning sites,
ampicillin resistance gene and pBR322 origin of replication.

Publication

George B, **Ruhel R**, Mazumder M, Sharma VK, Jain SK, Gourinath S, Chakraborty S. 2014. Mutational analysis of the helicase domain of a replication initiator protein reveals critical roles of Lys 272 of the B' motif and Lys 289 of the β-hairpin loop in geminivirus replication. *Journal of General Virology*. 95(7):1591-602.

Mutational analysis of the helicase domain of a replication initiator protein reveals critical roles of Lys 272 of the B′ motif and Lys 289 of the β-hairpin loop in geminivirus replication

Biju George,[1,2] Rajrani Ruhel,[1] Mohit Mazumder,[1] Veerendra Kumar Sharma,[1] Swatantra Kumar Jain,[2] Samudrala Gourinath[1] and Supriya Chakraborty[1]

[1]School of Life Sciences, Jawaharlal Nehru University, New Delhi, India

[2]Department of Biotechnology, Jamia Hamdard University, New Delhi, India

Correspondence

Supriya Chakraborty

supriyachakrasls@yahoo.com

Replication initiator protein (Rep) is indispensable for rolling-circle replication of geminiviruses, a group of plant-infecting circular ssDNA viruses. However, the mechanism of DNA unwinding by circular ssDNA virus-encoded helicases is unknown. To understand geminivirus Rep function, we compared the sequence and secondary structure of Rep with those of bovine papillomavirus E1 and employed charged residue-to-alanine scanning mutagenesis to generate a set of single-substitution mutants in Walker A (K227), in Walker B (D261, 262), and within or adjacent to the B′ motif (K272, K286 and K289). All mutants were asymptomatic and viral accumulation could not be detected by Southern blotting in both tomato and *N. benthamiana* plants. Furthermore, the K272 and K289 mutants were deficient in DNA binding and unwinding. Biochemical studies and modelling data based on comparisons with the known structures of SF3 helicases suggest that the conserved lysine (K289) located in a predicted β-hairpin loop may interact with ssDNA, while lysine 272 in the B′ motif (K272) located on the outer surface of the protein is presumably involved in coupling ATP-induced conformational changes to DNA binding. To the best of our knowledge, this is the first time that the roles of the B′ motif and the adjacent β-hairpin loop in geminivirus replication have been elucidated.

INTRODUCTION

Geminiviruses cause devastating diseases in a wide range of crop plants worldwide (Stanley, 1983; Fauquet *et al.*, 2003). These viruses, which contain circular ssDNA, are either monopartite or bipartite (i.e. they possess single-component or two-component genomes, respectively). Tomato leaf curl Gujarat virus (ToLCGuV; genus *Begomovirus*, family *Geminiviridae*), is one of the most predominant monopartite begomoviruses causing severe losses to tomato production in the Indian subcontinent (Chakraborty *et al.*, 2003; Chakraborty, 2008).

The geminivirus genome (~2.8 kb) possesses a stem–loop secondary structural element and a direct repeat sequence that function as the origin of replication (ori) and the binding site for replication initiator protein (Rep), respectively. Binding of Rep to the origin leads to strand- and site-specific nicking of viral DNA in an ATP-independent manner. Rep remains covalently linked to the 5′ end of nicked DNA, while

the 3′-hydroxyl group is used for the synthesis of the nascent strand (Orozco *et al.*, 1997). Though the detailed three-dimensional structure of Rep is yet to be determined, the protein is known to possess modular functions. The N-terminal region of Rep possesses site-specific nicking, ligation and DNA binding activities (Fontes *et al.*, 1992; Orozco *et al.*, 1997; Chatterji *et al.*, 2000), while the C terminus (aa 120–361) functions autonomously as a 3′ to 5′ helicase (Choudhury *et al.*, 2006; Clérot & Bernardi, 2006). Nonetheless, mechanistic details of Rep-mediated DNA unwinding are currently unknown.

Comparative sequence alignments of the geminivirus Rep proteins have shown that they belong to the SF3 helicase family (Koonin, 1993). Helicases of this family possess three conserved signature motifs: Walker A [involved in ATP binding; GxxxxGK(T/S)], Walker B (involved in ATP hydrolysis; DxxD or xxxxDD), and motif C (a conserved asparagine residue which interacts with the gamma Pi of ATP and an 'apical' water molecule). The B′ [(K/R)x$_{3\text{-}4}$ G$_{X7\text{-}8}$K] motif, located between Walker B and motif C, has been identified in SF3 helicases only. The B′ motif has been

characterized in animal viruses such as simian virus 40 (SV40, large T antigen), bovine papillomavirus (BPV, E1) and adeno-associated virus (AAV, Rep) (Hickman & Dyda, 2005). Geminivirus Rep differs from other SF3 helicases, both in terms of its oligomerization properties and in that it lacks an arginine finger in the AAA+ domain (Clérot & Bernardi, 2006). Therefore, in an effort to understand the roles and contribution(s) of residues within and adjacent to the B′ motif in the geminivirus replication cycle, we have compared the sequence and secondary structure of ToLCGuV Rep with BPV E1, with which it shares a sequence identity of 16 %, and have constructed alanine mutants of K272, K286 and K289, which are located in and adjacent to the B′ motif. The effects of substituting K227 of Walker A, and D261 and D262 of Walker B, with alanines were also studied.

Based on the structural information provided by our modelling studies, we hypothesize that K289 of ToLCGuV Rep plays a role in ssDNA interactions and DNA unwinding. The position of K289 and its potential for binding ssDNA is illustrated by the model as well as demonstrated by DNA binding assays. Mutational studies of K272 also indicated its critical role in ssDNA binding. However, the modelled structure indicates that this residue is on the outer surface of the protein and therefore could not have direct interaction with ssDNA. ATP binding is prerequisite for the formation of an ssDNA–geminiviral Rep complex. However, aa residue K272 is also not involved in ATP binding. Therefore, we speculate that this residue couples conformational changes induced by ATP binding to ssDNA binding, which is required for translocation of the helicase along the DNA double helix. Roles of these residues in viral replication and pathogenesis are also elucidated. Our findings add further support to the structural and functional similarity of geminivirus Rep proteins with SF3 helicases.

RESULTS

Expression and purification of wt and substitution mutants of RepC

The N-terminal domain of Rep possesses a sequence non-specific DNA binding activity (Nash et al., 2011) that may interfere with the ssDNA binding activity required for a DNA unwinding assay. Therefore, we selected the C-terminal domain for our study. Comparison of the amino acid sequence and predicted secondary structure of ToLCGuV Rep with those of BPV E1 and SV40 T antigen (a prototypical member of the SF3 family) revealed the presence of conserved SF3 signature motifs (Fig. 1). In the present study, we investigated the role of residues in and adjacent to the B′ motif (K272, K286, K289), Walker A (K227) and Walker B (D261, D262) by substituting the corresponding residues with alanine. C-terminal fragments (RepC) of wt and mutant Rep proteins corresponding to aa 122–361 were expressed in E. coli at 18 °C with 0.2 mM IPTG, which yielded reasonably good quantities of soluble

protein required for the biochemical assays. Details of the purification are described in Methods and Fig. S1 (available in the online Supplementary Material). All of the purified proteins are shown in Fig. 2(a).

Effect of mutations on overall structure and multimerization of RepC

The effects of mutations on the overall structure of RepC were analysed using circular dichroism (CD; Keideiling, 1996). These studies revealed that RepC consists of about 50 % random coil, ~37 % alpha helix and ~13 % beta sheet. The CD spectra of the mutants were similar to that of wt RepC (Fig. 2b), suggesting that the mutations caused no overall changes in the conformation of RepC.

Oligomeric states of RepC were analysed from the sedimentation profiles either in the presence or absence of cofactors (ATPγS and ATPγS plus ssDNA). No appreciable change was observed in the oligomeric state of RepC in the presence of these cofactors (Fig. 2c). The molecular mass of RepC was estimated from these experiments to be ~660 kDa, indicating that RepC formed a complex oligomeric state consisting of ~24 subunits, and the oligomeric state of each mutant was found to be similar to that of wt protein (Fig. 2d). Sedimentation profiles were similar even at a high salt concentration (1 M NaCl) and at a low protein concentration (1.5 μM RepC) (Fig. 2d).

Effects of mutations on ATP hydrolysis and nucleotide (ATPγS) binding

ATPase activity of purified RepC was found to be protein concentration dependent (Fig. 3a). ATP hydrolysis activity of the Walker A mutant (K227A) was completely lost, as demonstrated in other geminiviruses (Choudhury et al., 2006; Clérot & Bernardi, 2006) (Fig. 3b, lane 5), whereas for other mutants it was comparable to wt RepC. Structural data revealed that the residue K272 corresponding to aa K391 of AAV Rep could be a potential candidate for interaction with nucleotides (James et al., 2003). Therefore, the nucleotide (ATPγS) binding affinities of wt RepC and of the K272A mutant were also studied. Additionally, K227A (a Walker A mutant known to interfere with ATP binding) was selected as a negative control. The ATPγS binding affinity of wt RepC and of the K272A mutant were (mean ± SE) 0.95 ± 0.07 μM and 1.27 ± 0.09 μM, respectively. This result suggests that K272 is not involved in ATP binding. As expected, no affinity of RepC K227A to the ligand ATPγS could be detected.

Effects of RepC mutations on ssDNA binding

The ability of RepC and its mutants to bind ssDNA was studied using an electrophoretic mobility shift assay (EMSA). The K286A, D262A and K227A mutants decreased the affinity of RepC (concentration, 0.75 μM) for the DNA construct by approximately 35, 30 and 90 %, respectively
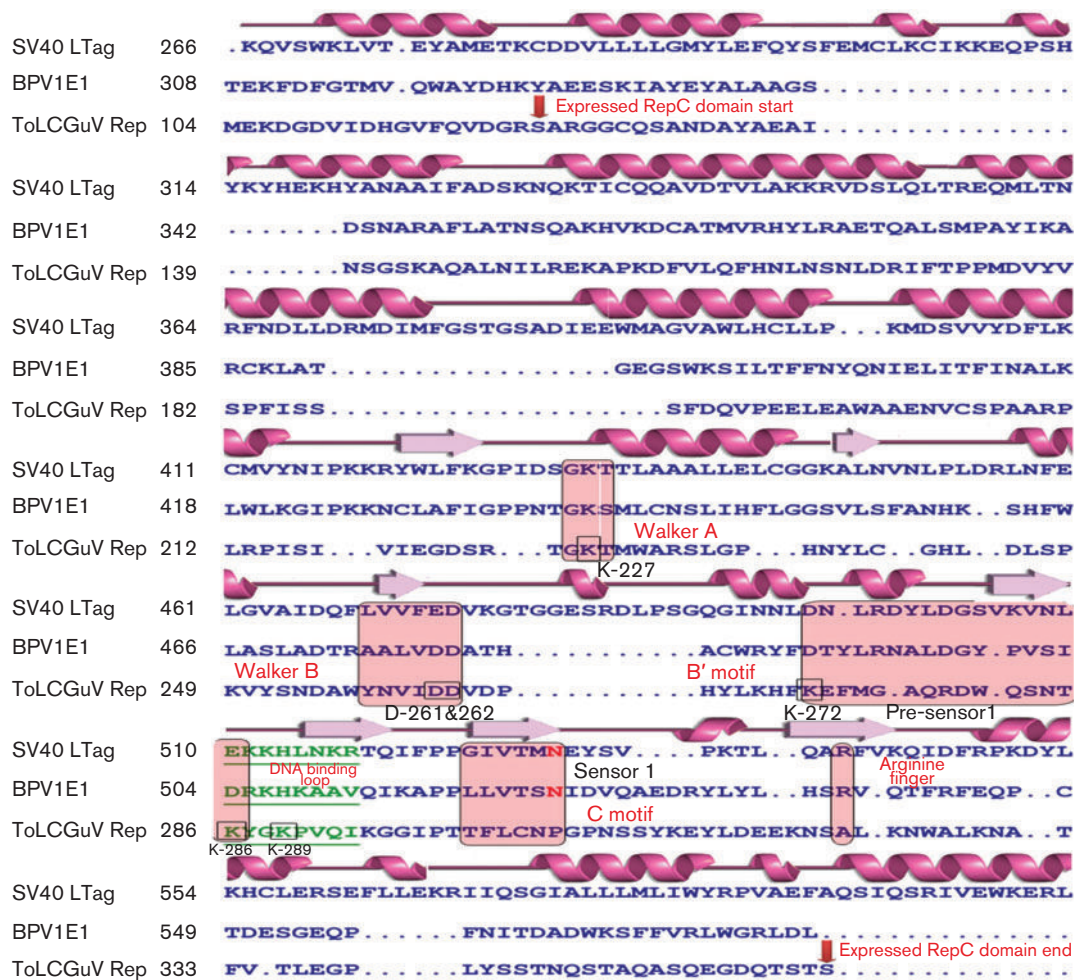
Fig. 1. Structure-based sequence alignment of the region spanning the signature motifs in replication initiator proteins from representative members of the SF3 family. Walker A, B, B′, C and arginine finger motifs are boxed in pink. The underlined green sequence, located between Walker B and C, represents the predicted β-hairpin loop. BPV E1: bovine papillomavirus type 1a E1 protein (PDB Code 2GXA); SV40 LTag: simian virus 40 large T antigen (PDB Code 1SVM). Relative positions of alpha helices, beta strands and loop regions are shown in pink.

(Fig. 4a, b). At similar concentration of RepC, complete loss of DNA binding activity was observed for K272A and K289A (Fig. 4a, b). Additionally, the binding of DNA to wt RepC, as well as to the K272A and K289A mutants, was also measured using a fluorescent spectroscopy assay. Wt RepC bound to a 17-mer oligonucleotide (Table S1) with a K_d value of 21.6 ± 3 nM, whereas the mutant K272A displayed a more than twofold reduction in ssDNA binding ($K_d = 54.9 \pm 7.8$ nM); the K289A mutant displayed drastically reduced ssDNA binding function (Fig. 4a, b) and could not be analysed by fluorescence spectroscopic assay.

Effects of mutations on DNA unwinding activity

The helicase activity of RepC was measured by its ability to displace a partial duplex helicase substrate. DNA unwinding by wt RepC was initially assayed in the presence of

various nucleotide triphosphates and at a protein concentration of 96 nM (Fig. 5a, b). GTP and ATP are equivalently able to act as cofactors in the reaction, unwinding approximately 30 % of helicase substrate, whereas CTP and TTP supported unwinding of ~20 % and 10 % of the helicase substrate, respectively. The small amount (~3 %) of displaced oligonucleotide observed in the presence of ADP or ATPγS were likely to be the unwound products, since they were absent in the EDTA-treated lane.

As predicted, mutants deficient at binding DNA could not unwind the helicase substrate (Fig. 5c). Mutation of K272, K289 or K227 to alanine completely abolished RepC helicase activity, while, at a concentration of 128 nM, the K286 and K262 mutants retained ~26 % and 37 % of substrate unwinding activity, respectively (Fig. 5c). At 128 nM concentration, wt RepC could unwind approximately 69 % of helicase substrate.
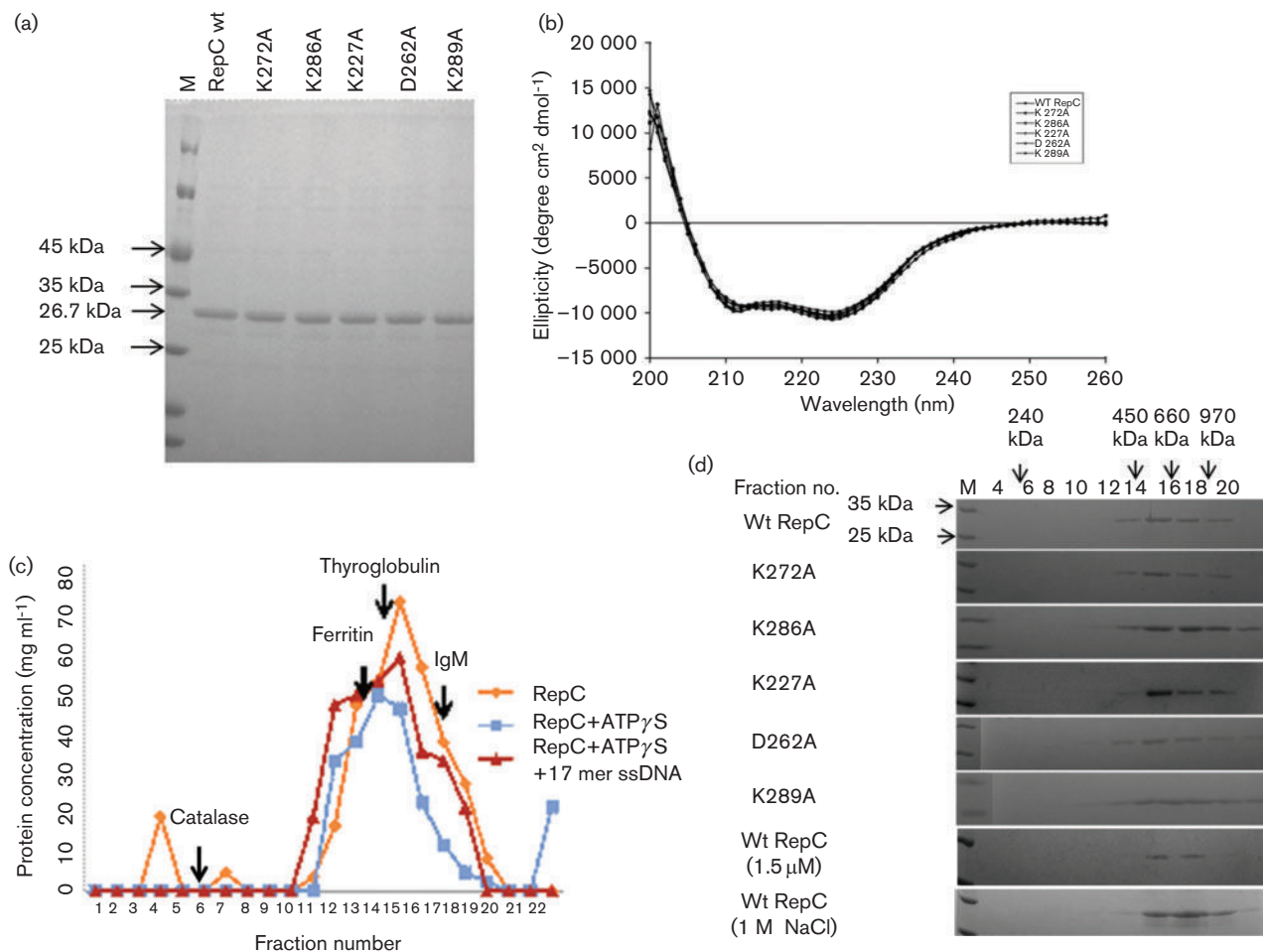
Fig. 2. Effect of mutations on the overall structure of ToLCGuV-RepC. (a) SDS-PAGE analysis of purified RepC proteins. Six micrograms of each protein were electrophoresed through a 10 % polyacrylamide gel. (b) CD spectra of the wt RepC together with all the mutants (0.1 mg ml^{-1} concentration) in 10 mM phosphate buffer (pH 8.0), 100 mM NaCl and 10 % glycerol. (c) Glycerol gradient sedimentation of wt RepC, and also of this protein in the presence of nucleotides or of nucleotides plus ssDNA. (d) Fractions collected from glycerol gradient sedimentation of wt and mutant RepC proteins resolved using 10 % SDS-PAGE. The downward pointing arrows indicate the positions of the peaks for the markers used in gradient centrifugation. Lane M corresponds to the molecular mass marker in kDa.

RepC mutants are unable to support viral replication *in vivo*

Since the N-terminal domain of Rep specifically interacts with the ori and is indispensable for replication initiation, infectious tandem repeat constructs of the DNA-A genome, harbouring either wt or mutated full-length Rep (RepF) in combination with the cognate DNA-B component of ToLCGuV, were used for agroinoculation to test their replication potential in a transient leaf disc assay. It is relevant to mention here that ToLCGuV is monopartite in nature and association of DNA-B increases its pathogenicity. A viral genome harbouring D261A, K227A and K289A mutants did not replicate in *N. benthamiana* leaf discs, while K272A, K286A and D262A mutants replicated at lower levels (~2 %, ~10 % and ~4 % respectively) as compared to the

wt virus (Fig. 6a). Presence of DNA-B was verified in all combinations of either wt or mutants where replication of DNA-A was not abolished (Fig. 6a, lower panel). *In vivo* complementation in *N. benthamiana* leaf discs indicated that all mutants could be rescued by concomitant expression of wt Rep, albeit at a much lower level in comparison to replication of DNA-A (wt Rep) and DNA-B (Fig. 6b).

Interestingly, *in planta* infectivity of wt and mutant viruses showed that none of the mutants could induce discernible symptoms (Fig. S2a) or have detectable levels of Rep by Southern blotting with a Rep-specific probe (Fig. S2b). ToLCGuV DNA-A could be detected from systemic asymptomatic leaves of plants inoculated with K272A, K286A and D261A mutants through Rep-specific PCR, while DNA-B was detected using a BV1 ORF probe in all mutant
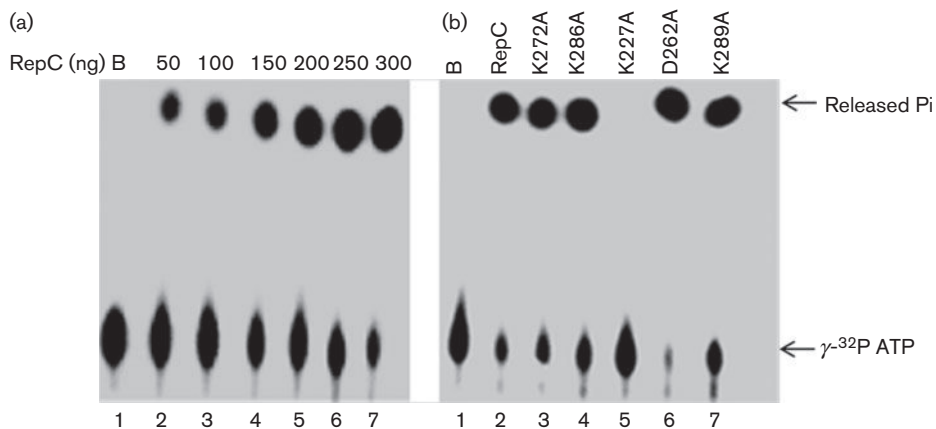
(a)

RepC (ng) B   50   100   150 200 250 300

← Released Pi

← *γ*-³²P ATP

1   2   3   4   5   6   7

(b)
B   RepC K272A K286A K227A D262A K289A

1   2   3   4   5   6   7

Fig. 3. ATPase activity of RepC. (a) Autoradiograph showing the ATPase activity of wt RepC. Concentration of the protein used is indicated at the top of the lane. (b) ATPase activities of RepC mutants. The respective mutation is indicated above each lane. Five hundred nanograms of each protein was used. The positions of γ-³²P-labelled ATP and released Pi are indicated. B indicates boiled protein.

(a)

Wt RepC   No protein   Without ATP*γ*S   K286A   D262A   No protein   K272A   No protein   Without ATP*γ*S   K227A   K289A   No protein

1 2 3 4   1 2 3 4   1 2 3 4   1 2 3 4   1 2 3 4   1 2 3 4

PDC →

Free 17-mer ssDNA →

(b)
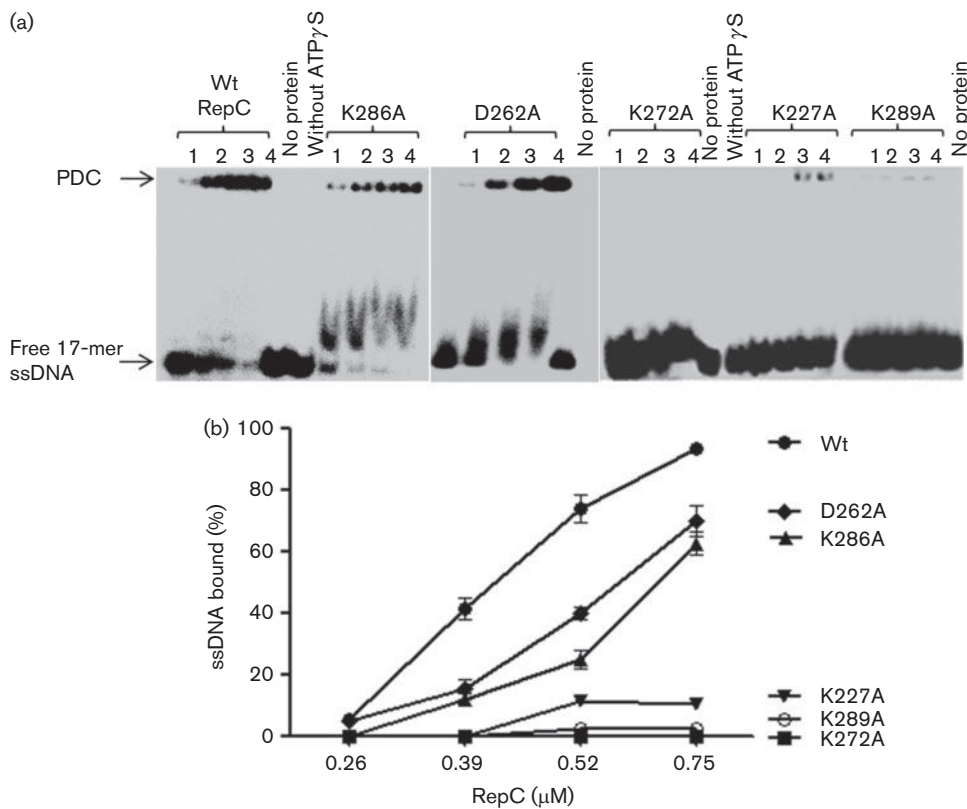
Fig. 4. DNA binding activities of RepC variants. (a) DNA binding activity by EMSA. Wt and mutant RepC proteins were used, each at concentrations of 0.26, 0.39, 0.52 and 0.75 µM, corresponding to lanes 1, 2, 3 and 4, respectively. The position of the protein–DNA complexes (PDCs) is indicated by an arrow. Lanes labelled without protein and without ATP*γ*S indicate negative controls. (b) Plot of the ssDNA binding activities as quantified from the autoradiograph shown in (a). Error bars represent SD of three independent experiments.

Fig. 5. dsDNA unwinding activity of wt and mutant RepC proteins. (a) DNA unwinding activity of wt RepC in the presence of various nucleotides. The reaction was carried out for 6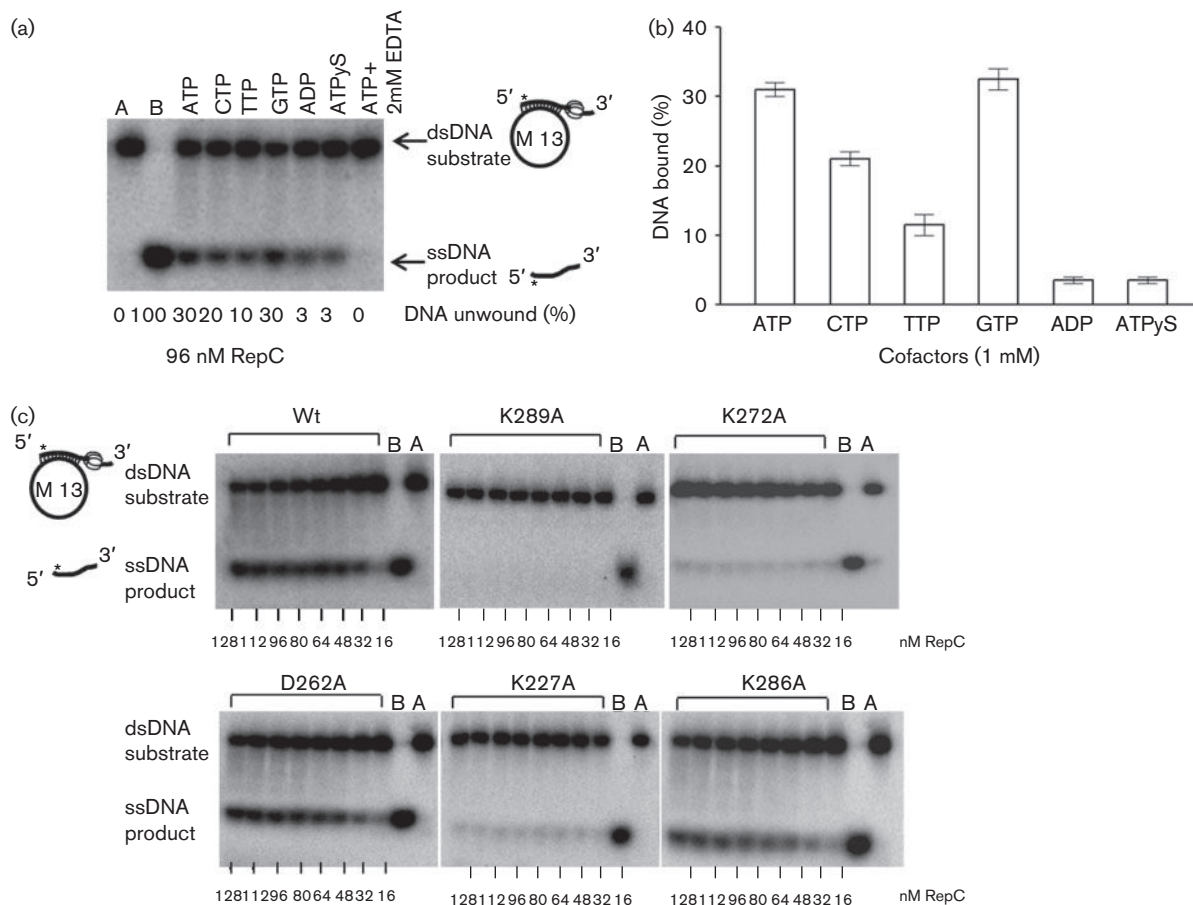0 min at 37 °C using 96 nM wt RepC with 1 mM of the indicated nucleotide. (b) Plot of the helicase activities quantified from the autoradiograph as shown in (a). (c) Representative gels of the unwinding reactions carried out by wt and mutant RepC as a function of protein concentration (indicated at the bottom of the lanes). The radiolabel is marked with an asterisk. Lanes A (annealed substrate) and B (boiled by heating to 100 °C for 5 min before loading onto the gel) represent the reaction mixture without protein.

combinations (Fig. S2c). Similar results were obtained from inoculated tomato plants (data not shown).

Comparative modelling of the helicase domain of ToLCGuV Rep

A full understanding of critical helicase domain residues, including those mutated in our experiments, would benefit from visualizing their environments in a three-dimensional structure. A crystal structure of ToLCGuV Rep, however, is unavailable. We therefore built a comparative predicted model of RepC using the experimentally determined structure of its close homologue BPV-E1 as a template (Fig. 7). Although the amino acid sequences of the helicase domains of these two proteins are only ~16 % identical, many features of the predicted model of the resulting hexamers justifies our biochemical data and is comparable with other hexameric helicases (Li *et al.*, 2003; Yoon-Robarts *et al.*, 2004; Enemark & Joshua-Tor, 2006). For example,

the model places the Walker A and Walker B motifs at the interface between subunits of the hexamer, suitable for ATP binding. In addition, according to this model, the *β* hairpin that precedes motif C/sensor 1 (the 'pre-sensor 1 hairpin') points into the central channel (as it does in SF3 helicases). The location of this loop, and the positive charge of two of its residues, K286 and K289, underscores the role of this loop in ssDNA binding in geminivirus Rep helicase. Interestingly, the B′ residue (K272) lies adjacent to the ATP binding pocket at the outer surface of the protein, and thus possibly has no role in direct interactions with DNA. Fig. 7 illustrates the positioning of the conserved residues studied here in the RepC hexamer.

DISCUSSION

Based on the sequence and secondary structure alignments with other helicases, ToLCGuV Rep appears to have a
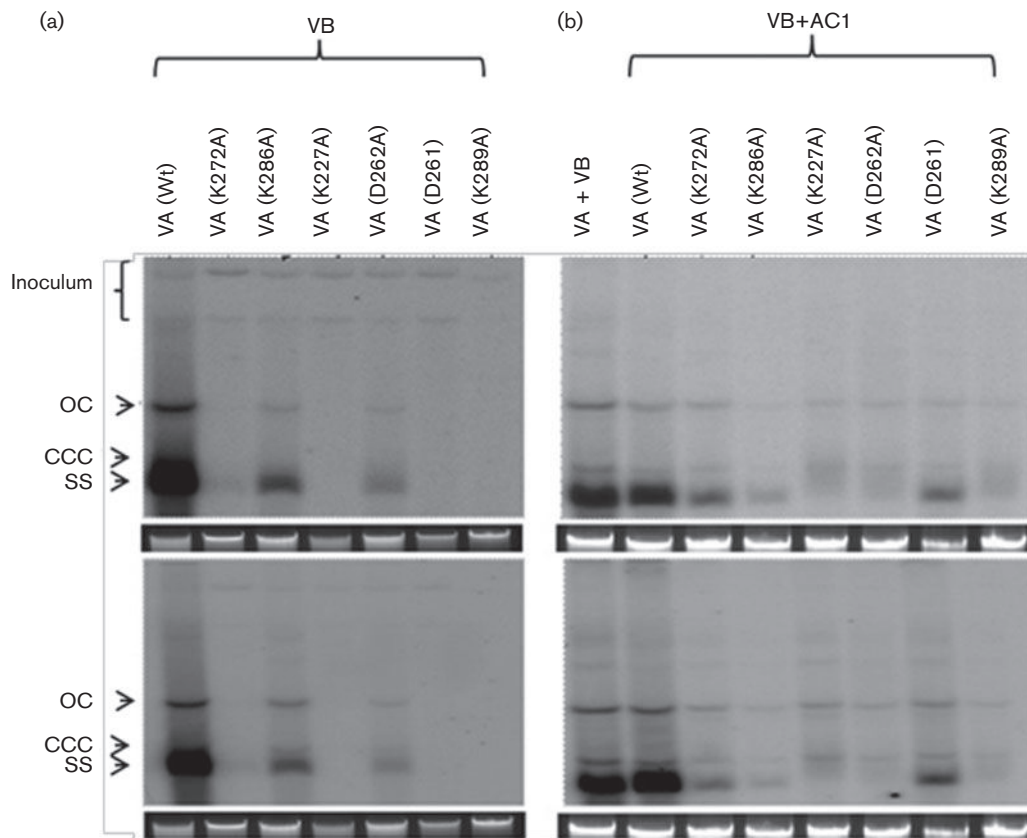
Fig. 6. Transient replication of wt and mutant ToLCGuV DNAs. (a) Southern blot showing DNA accumulation levels of either the wt or mutant ToLCGuV DNA-A component, together with infectious cognate DNA-B clone. OC, open circular; CCC, covalently closed circular and SS, single stranded, represent the different forms of replicated viral DNA. Inoculum represents the binary construct used for inoculation. (b) Infectious clones of wt DNA-A or its mutants with DNA-B were complemented with the wt Rep construct (Rep expressed under the control of the 35S promoter of pBinAR vector). Respective mutations in the Rep gene and loading controls are indicated. The same blot was stripped and probed with a BV1 probe (lower panels).

comparatively smaller AAA+ domain, which contains the SF3 signature motifs, and to lack an otherwise conserved arginine residue that forms the arginine finger (Clérot & Bernardi, 2006) (Fig. 1). Therefore, it is necessary to improve our understanding of the roles of SF3 signature motifs in geminivirus Rep helicases, such as that from ToLCGuV. In this study, we focused on characterizing the precise functions of structural elements present in the helicase domain of Rep in geminivirus replication. These studies included ATP binding, ATP hydrolysis, ssDNA binding, DNA unwinding and *in vivo* replication mediated by wt and mutant Rep proteins. Our data are summarized in Table 1. We then developed *in silico* a three-dimensional model of the helicase domain of Rep based on limited sequence homology with the BPV-E1 protein.

In order to determine the effect of mutations on the structural integrity of the Rep protein, the oligomeric state and the degree of secondary structure of wt and mutant proteins were estimated. In agreement with a previous report, where Mungbean yellow mosaic India virus Rep

residues 120–362, as well as full-length Rep, were shown to form oligomers of 24 subunits (Choudhury *et al.*, 2006), the sedimentation profile of ToLCGuV also indicated that an oligomeric species of RepC consisting of ~24 subunits was present (Fig. 2c, d). A broad distribution of protein in glycerol gradients indicates the possibility that multiple oligomeric species of the RepC protein may exist in solution. Populations of oligomers of various sizes, ranging from hexamers up to dodecamers or even higher order oligomers have been observed in the Rep122–359 protein of Tomato yellow leaf curl Sardinia virus (Clérot & Bernardi, 2006). In SF3 members, oligomerization is shown to be favoured in the presence of either ATP (San Martín *et al.*, 1997) or DNA (Fouts *et al.*, 1999; Maggin *et al.*, 2012; Zarate-Perez *et al.*, 2012). In AAV2 Rep, a short 23 amino acid residue linker is involved in DNA-dependent oligomerization and formation of higher order assemblies of the protein (Zarate-Perez *et al.*, 2012). However, monomeric or dimeric species of ToLCGuV RepC were never observed in our study nor in previous studies on oligomerization of other begomovirus Rep proteins (Choudhury *et al.*, 2006;
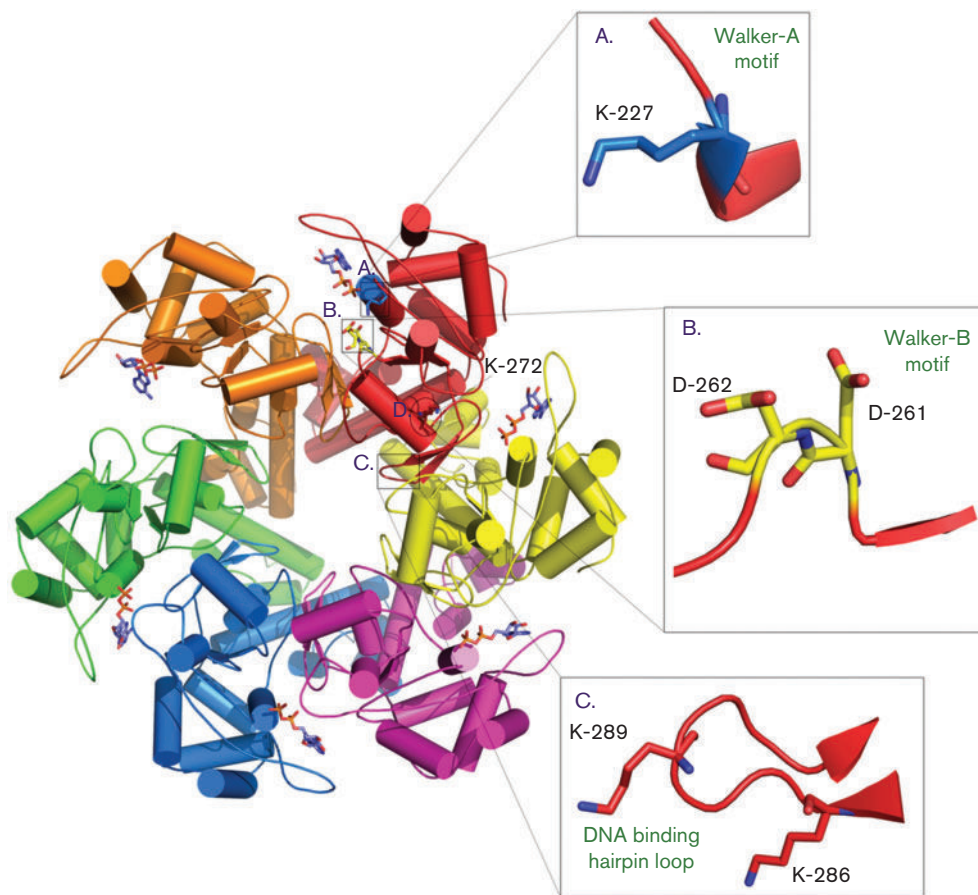
Fig. 7. Homology model of the ToLCGuV RepC (122–361) helicase based on the bovine papillomavirus E1 helicase domain as a template. Cartoon representation of the helicase hexamer where the six subunits are shown in different colours. The nucleotides (ADP) coupled to each subunit are depicted in sticks. Locations of residues relevant for this study are depicted in sticks in the magnified pictures (in boxes) from a single subunit.

Clérot & Bernardi, 2006). The oligomerization profile of ToLCGuV RepC was stable even at high salt or low protein concentrations, and only higher order oligomers of RepC were found even in the absence of nucleotide or ssDNA, indicating that the intrinsic structure of RepC allowed it to form stable higher order oligomers (Fig. 2c, d). In our

Table 1. Summary of biochemical activities of wt and mutant variants of ToLCGuV RepC

	ATP hydrolysis	ATPγS binding*	DNA binding	Helicase activity	Replication (leaf disc)	Replication (*in planta*)
Wt RepC	+ + +	+ + +	+ + +	+ + +	+ + +	+ + +
K272A	+ + +	+ + +	−	−	−	−
K286A	+ + +	ND	+ +	+	+ +	−
K227A	−	−	−	−	−	−
D262A	+ + +	ND	+ +	+ +	+	−
D261A	ND	ND	ND	ND	−	−
K289A	+ + +	ND	−	−	−	−

+ + +, + + and + indicate maximum, moderate and low catalytic activity, respectively, whereas '−' stands for no activity. 'ND' stands for experiment not carried out.
*For nucleotide binding, two residues located on the outer surface of the protein (based on modelled structure) were analysed.

studies of ToLCGuV, sedimentation profiles of all RepC mutants were shown to be similar to that of wt protein, indicating that these mutations do not affect the oligomerization state and hence the overall structure of RepC. This latter point is further supported by our observation that the wt protein and mutants also share nearly identical far-UV CD signatures.

We further studied the effects of mutations on various activities required for DNA unwinding. The K227A mutant of geminivirus Rep was previously shown to be ATPase deficient (Desbiez *et al.*, 1995; Choudhury *et al.*, 2006). Our results indicate that the defect in ATP hydrolysis caused by K227A is due to its inability to bind ATP rather than its hydrolysis. Note that all other mutants retain ATPase activity comparable to that of wt levels (Table 1). In addition to the importance of this positively charged residue, two negatively charged residues are found in the Walker B motif of most geminiviruses. D261 in geminivirus Rep is shown to be essential for ATP hydrolysis (Choudhury *et al.*, 2006). However, no role has been identified for the adjacent residue D262 in ATP hydrolysis despite it being conserved. Mutants D261A and D262A in ToLCGuV Rep were therefore studied here to explain these features. DNA replication studies in leaf discs indicate that D261 is indispensable for Rep function, but that D262A did not affect ATPase activity, which is consistent with its ability to allow replication of the viral genome. Our results and the previous report on a geminivirus Walker B residue mutant (D261K) (Choudhury *et al.*, 2006) indicate that at this position a single conserved aspartic acid residue is sufficient for ATP hydrolysis by Rep. Considerable variability in this motif among SF3 helicase members also supports this observation (Koonin, 1993) (Fig. S3). Additionally, it should be noted that we were unable to test ATP hydrolysis of D261A owing to lack of expression in a prokaryotic system.

In addition to being unable to bind ATP, the K227A mutant also lost the ability to bind ssDNA (Fig. 4), indicating that nucleotide binding is indispensable for formation of a geminivirus Rep–ssDNA complex. In addition, formation of an efficient RepC–ssDNA complex in the presence of ATPγS suggests that nucleotide binding rather than hydrolysis induces conformational changes in Rep, which, in turn, facilitate DNA binding. DNA binding and unwinding activities of K272A and K289A were completely abolished, whereas mutants with ssDNA binding ability retained the unwinding property. Our *in silico* RepC model suggests that K272A resides at the subunit interface near the ATP binding pocket. However, the K272A mutation does not alter its ATP binding. Hence, the defect in ssDNA binding and unwinding activity of this mutant is likely to be due to a lack of suitable conformation that mediates the interaction of the β-hairpin residue (K289) with the phosphate backbone of DNA. Note that K289 lies at the tip of the β-hairpin structure located in the central channel of the oligomeric protein, and may thus directly interact with ssDNA during geminivirus Rep-mediated unwinding. ssDNA binding by similar positively charged residues located in the β-hairpin

loop has been demonstrated in AAV2 Rep, SV40 T antigen and BPV E1 (Yoon-Robarts *et al.*, 2004; Shen *et al.*, 2005; Castella *et al.*, 2006). Besides the charged lysines, the hydrophobic residue Y287, which is equivalent in position to H513 of SV40 T antigen (Shen *et al.*, 2005), is also conserved in the predicted β-hairpin loop of gemini- and nanovirus Rep proteins. As suggested by Shen *et al.* (2005), it may be involved in hydrophobic stacking with the bases of DNA.

To further examine whether the hydrolysis energy of different nucleotides could be used for DNA unwinding, a helicase assay was performed in the presence of various nucleotides. Interestingly, RepC efficiently uses ATP and GTP as cofactors for DNA unwinding. Since Rep is indispensable for replication of geminiviruses in plants, we have analysed the effect of all mutants *in vivo*. The inability of all mutants to produce symptoms of infection in plants may be attributed to a sub-threshold level of viral genome replication at inoculation sites. To test the ability of viral mutants to replicate at the site of inoculation, leaf disc assays were performed. The results indicate that both K286A and D262A could replicate in *N. benthamiana*, albeit at a reduced extent in comparison to wt. Low levels of viral replication for these two mutants suggest additional roles for these residues, such as interactions with host proteins. Complete loss of replication in K227A, D261A and K289A mutants was observed (Fig. 6a). Consistent with the results of the biochemical analysis of these mutants, the above studies highlight the critical role of these residues for proper Rep function.

Aided by sequence alignment, biochemical experiments and a homology-based model, a possible role of nucleotide binding for geminivirus Rep function is proposed. Based on the position of positively charged residues located on predicted β-hairpin loops located in the central pore of the RepC model, their critical role in ssDNA binding, and the function of equivalent residues in other hexameric NTPases, it seems that K289 of Rep might play a role in the ability of ToLCGuV Rep protein to interact with ssDNA. In *E. coli* RecA, the binding of DNA to similar lysine-containing loops induces conformational changes in the catalytic residues located in the Walker B motif and motif C (Story *et al.*, 1992). In geminivirus Rep and other SF3 helicases, motif B′ is located similarly between Walker B and motif C (highlighted in Fig. S3).

It has been suggested that the B′ motif plays multiple roles during the helicase reaction (Yoon-Robarts *et al.*, 2004). Some residues are directly involved in DNA binding (Yoon-Robarts *et al.*, 2004; Shen *et al.*, 2005; Castella *et al.*, 2006), while others are involved in the interaction with ATP (Story *et al.*, 1992). Another set of residues within this motif then couples DNA binding to ATP catalysis (Singleton *et al.*, 2000). ATP binding and hydrolysis occurring at the interface of neighbouring subunits drive conformational changes that promote remodelling of the target substrates (Erzberger & Berger, 2006). Similarly, some residues located in the B′ motif have been postulated to move in response to ATP

binding and its hydrolysis, and might be involved in positioning DNA binding β-hairpin residues as proposed for SV40 large T antigen (Gai *et al.*, 2004). Based on the results of our experiments with the ToLCGuV RepC K272 mutant, we hypothesize that this residue is possibly involved in mediating ATP-induced allosteric changes in such a way as to allow DNA binding by amino acids exposed in the hairpin loops positioned in the central pore of the protein oligomers.

To the best of our knowledge, the present study provides information on the role of the geminivirus Rep B′ motif and β-hairpin loop for helicase function, similar to other SF3 helicases.

METHODS

Cloning, expression and purification of proteins. The full-length Rep (RepF) ORF from Tomato leaf curl Gujarat virus (ToLCGuV, GenBank accession no. AY190290) was cloned in the pET-28a (+) vector. The same clone was used as a template for site-directed mutagenesis to generate K227A, D261A, D262A, K272A, K286A and K289A mutants. A 723 bp C-terminal fragment (RepC) corresponding to amino acids 122 to 361 of Rep amplified from wt or mutant pET-RepF clones was ligated into pMAL-c2X vector at *Bam*HI and *Hind*III sites. All mutants were expressed and purified from *Escherichia coli* BL21 (DE3) at similar conditions as followed for wt RepC. Cells were grown at 37 °C to an OD_{600} of 0.6. Isopropyl β-D-thiogalactopyranoside (0.2 mM) was added, and incubated with shaking for a further 16 h at 18 °C. Pellets were resuspended in buffer A, containing 50 mM Tris/HCl (pH 8.0), 100 mM NaCl, 2 mM β-mercaptoethanol, 1 mM PMSF, 5 mM $MgCl_2$ and 10 % glycerol. MBP-tagged protein was purified on amylose resin (Qiagen) and was cleaved using Factor Xa (NEB) in buffer A supplemented with 2 mM $CaCl_2$. Protein was further purified by diethylaminoethanol (DEAE) chromatography (Sigma) and finally dialysed against 50 mM Tris/HCl (pH 8.0), 100 mM NaCl and 2 mM β-mercaptoethanol. All proteins were quantified by the Bradford method and Coomassie blue R-250 staining on SDS-PAGE. A list of primers and oligonucleotides used is provided in Table S1.

Construction of tandem repeat infectious clones. Specific point mutations in the C terminus of Rep were introduced by site-directed mutagenesis of the viral genome using a ToLCGuV DNA-A monomer cloned into pUC18 as template. All the clones were sequenced to ascertain that no mutation was introduced except the desired one. The tandem infectious clones of mutant ToLCGuV DNA-A were constructed as described for wt ToLCGuV DNA-A and DNA-B (Chakraborty *et al.*, 2003). Clones were confirmed by restriction digestion and by sequencing.

Secondary structure determination and glycerol gradient centrifugation. Circular dichroism (CD) spectra of RepC and all mutant proteins were obtained at 25 °C in a quartz cell with 1 mm light path using a Chirascan model of an Applied Photosystem CD spectrophotometer (Greenfield, 2006). A mean of 10 scans was taken and corrected for buffer contributions. The secondary structure composition was predicted using K2D software. For glycerol gradient centrifugation, about 22 μM of RepC or mutant protein was layered onto 15–40 % glycerol gradients in a buffer containing 25 mM Tris/HCl (pH 8.0), 250 mM NaCl and 1 mM DTT. Centrifugation was performed in a Beckman SW 55 Ti rotor at 45 000 r.p.m. min⁻¹ for 16 h at 4 °C. Fractions were collected, and analysed by SDS-PAGE. Standard markers, namely catalase (232 kDa), ferritin (450 kDa),

thyroglobulin (669 kDa) and IgM (980 kDa) were run in parallel gradients.

ATPase and helicase assay. ATPase assays were performed as described previously (Choudhury *et al.*, 2006). To study ATP hydrolysis, 0.2 μCi of [γ-^{32}P] ATP (6000 Ci mmol⁻¹) and indicated amounts of wt and mutant RepC were used. For helicase assays, the dsDNA substrate was obtained by annealing a ^{32}P-labelled 23 base-long oligonucleotide to circular M13mp18 ssDNA. Seventeen nucleotides of this oligonucleotide annealed to the M13 ssDNA, leaving a 6 nt 3′ overhang.

Helicase assays were performed at 37 °C in buffer containing 50 mM Tris/HCl (pH 8.0), 10 mM $MgCl_2$, 2.5 mM DTT, 0.1 mg ml⁻¹ BSA, 5 % glycerol, 2 mM ATP or indicated nucleotide and 10 nM of substrates with indicated amount of proteins, in a volume of 10 μl. After 30 min of incubation, the reaction was stopped by addition of (final concentrations) 8 % glycerol, 0.48 % SDS, 20 mM EDTA and 0.02 % bromophenol blue, and was subsequently analysed on native polyacrylamide gels. The gels were scanned and quantified using ImageQuant (Molecular Dynamics).

DNA binding assay. The single-stranded oligonucleotide substrate, M1317FW (Table S1) was labelled with [γ-^{32}P] ATP using T4 polynucleotide kinase. Electrophoretic mobility shift assay (EMSA) was performed at 4 °C as follows. In a total reaction volume of 12 μl, 0.26, 0.39, 0.52 and 0.75 μM of either wt or mutant protein was incubated with 10 nM radiolabelled DNA in a buffer consisting of 20 mM Tris/HCl pH 8.0, 2 mM DTT, 5 mM $MgCl_2$, 12 % glycerol, 25 mM KCl and 1 mM ATPγS. Reactions were incubated for 30 min at room temperature and loaded onto 3.4 % native polyacrylamide gels for analysis.

Fluorescence-based nucleotide and DNA binding. Fluorescence measurements were carried out at 25 °C using a Cary-Varian spectrofluorometer with a spectral bandpass of 5 nm and 10 nm for excitation and emission, respectively. Binding of the ATPγS with RepC or the DNA (5–186 nM) with ATPγS (2 mM) saturated RepC was measured by monitoring the intrinsic tryptophan fluorescence of either wt RepC or the mutants (1.75–1.85 μM) at 350 nm upon excitation at 295 nm. Measured fluorescence spectra were corrected for dilution and buffer background. All the binding data were analysed using a one-site saturation model (Nongkhlaw *et al.*, 2009).

Viral replication and complementation assay. Infectivity of the wt and mutant viral DNA was studied in three-week-old seedlings of *Nicotiana benthamiana* and tomato (*Solanum lycopersicum* cv. Punjab Chuhara) using standard agroinoculation methods (Chakraborty *et al.*, 2008). Infectious clones of wt DNA-A or mutants were co-inoculated with DNA-B to study replication *in planta* and in transient leaf disk assays (LDA). To study *in planta* viral replication, DNA from systemic leaves of *N. benthamiana* were isolated at 7, 14 and 28 days post-infection (p.i.) and hybridized with Rep probe (nt 1529–2614). For DNA-B detection, blots were stripped and reprobed with a ^{32}P-labelled ToLCGuV BV1 (nt 405–1208). PCR-based detection using ToLCGuV Rep and BV1 ORF specific primers was performed with annealing at 60 °C for 30 cycles. Southern blotting and LDA were performed as described previously (Kumari *et al.*, 2011). Infectious clones of wt DNA-A or its mutants with DNA-B in equal proportions were complemented with the wt Rep construct to study mutant complementation by the wt Rep gene in LDA.

Multiple sequence alignment and homology modelling. An NCBI BLAST search of the Protein Data Bank had shown little sequence similarity between ToLCGuV RepC and known structures. Therefore, conserved motifs were searched against the PROSITE (Sigrist *et al.*, 2002) and PRINTS (Attwood *et al.*, 2012) databases.

These results, combined with results from further analysis obtained from the PHYRE server (Kelley & Sternberg, 2009) predicted that the structure of ToLCGuV RepC was most similar to that of bovine papillomavirus E1 (BPV E1; 2GXA). Sequence alignments were generated using the CLUSTAL W package (Larkin *et al.*, 2007) and were subsequently improved based on structural alignment and motif conservation. Three-dimensional models were generated using the MODELLER package (Eswar *et al.*, 2006), using an alignment of residues 304–577 from the crystal structure of BPV E1 and residues 104–361 of ToLCGuV Rep. The best model was selected on the basis of it having the lowest DOPE score and highest GA341 assessment score, and was further validated by PROCHECK (Laskowski *et al.*, 1993). Furthermore, to obtain a hexameric state, the modelled structure was superimposed on the crystal structure of BPV E1 (2GXA). The ADP molecules were added at the active site of the ToLCGuV Rep model using the hexameric ADP bound crystal structure of the BPV E1 template. The structure was further relaxed to eliminate bad atomic contacts. The molecular minimization simulations were performed with the AMBER molecular dynamics package (Case *et al.*, 2006) using AMBER99SB force field and steepest descent algorithm to remove close van der Waals contacts, followed by conjugate gradient minimization until the energy was stable in sequential repetitions.

ACKNOWLEDGEMENTS

REFERENCES

Attwood, T. K., Coletta, A., Muirhead, G., Pavlopoulou, A., Philippou, P. B., Popov, I., Romá-Mateo, C., Theodosiou, A. & Mitchell, A. L. (2012). The PRINTS database: a fine-grained protein sequence annotation and analysis resource–its status in 2012. *Database (Oxford)* **2012**, bas019.

Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A. & other authors (2006). *Amber 9 User's Manual*. San Francisco: University of California.

Castella, S., Bingham, G. & Sanders, C. M. (2006). Common determinants in DNA melting and helicase-catalysed DNA unwinding by papillomavirus replication protein E1. *Nucleic Acids Res* **34**, 3008–3019.

Chakraborty, S. (2008). Tomato leaf curl viruses from India. In *Encyclopedia of Virology*, vol. 5, pp. 124–133. Oxford: Elsevier.

Chakraborty, S., Pandey, P. K., Banerjee, M. K., Kalloo, G. & Fauquet, C. M. (2003). Tomato leaf curl Gujarat virus, a new begomovirus species causing a severe leaf curl disease of tomato in Varanasi, India. *Phytopathology* **93**, 1485–1495.

Chakraborty, S., Vanitharani, R., Chattopadhyay, B. & Fauquet, C. M. (2008). Supervirulent pseudorecombination and asymmetric synergism between genomic components of two distinct species of begomovirus associated with severe tomato leaf curl disease in India. *J Gen Virol* **89**, 818–828.

Chatterji, A., Chatterji, U., Beachy, R. N. & Fauquet, C. M. (2000). Sequence parameters that determine specificity of binding of the replication-associated protein to its cognate site in two strains of tomato leaf curl virus-New Delhi. *Virology* **273**, 341–350.

Choudhury, N. R., Malik, P. S., Singh, D. K., Islam, M. N., Kaliappan, K. & Mukherjee, S. K. (2006). The oligomeric Rep protein of Mungbean yellow mosaic India virus (MYMIV) is a likely replicative helicase. *Nucleic Acids Res* **34**, 6362–6377.

Clérot, D. & Bernardi, F. (2006). DNA helicase activity is associated with the replication initiator protein rep of tomato yellow leaf curl geminivirus. *J Virol* **80**, 11322–11330.

Desbiez, C., David, C., Mettouchi, A., Laufs, J. & Gronenborn, B. (1995). Rep protein of tomato yellow leaf curl geminivirus has an ATPase activity required for viral DNA replication. *Proc Natl Acad Sci U S A* **92**, 5640–5644.

Enemark, E. J. & Joshua-Tor, L. (2006). Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* **442**, 270–275.

Erzberger, J. P. & Berger, J. M. (2006). Evolutionary relationships and structural mechanisms of AAA + proteins. *Annu Rev Biophys Biomol Struct* **35**, 93–114.

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U. & Sali, A. (2006). Comparative Protein Structure Modeling Using Modeller. *Curr Protoc Bioinformatics* **15**, 5.6.1–5.6.30.

Fauquet, C. M., Bisaro, D. M., Briddon, R. W., Brown, J. K., Harrison, B. D., Rybicki, E. P., Stenger, D. C. & Stanley, J. (2003). Revision of taxonomic criteria for species demarcation in the family *Geminiviridae*, and an updated list of begomovirus species. *Arch Virol* **148**, 405–421.

Fontes, E. P., Luckow, V. A. & Hanley-Bowdoin, L. (1992). A geminivirus replication protein is a sequence-specific DNA binding protein. *Plant Cell* **4**, 597–608.

Fouts, E. T., Yu, X., Egelman, E. H. & Botchan, M. R. (1999). Biochemical and electron microscopic image analysis of the hexameric E1 helicase. *J Biol Chem* **274**, 4447–4458.

Gai, D., Zhao, R., Li, D., Finkielstein, C. V. & Chen, X. S. (2004). Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. *Cell* **119**, 47–60.

Greenfield, N. J. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* **1**, 2876–2890.

Hickman, A. B. & Dyda, F. (2005). Binding and unwinding: SF3 viral helicases. *Curr Opin Struct Biol* **15**, 77–85.

James, J. A., Escalante, C. R., Yoon-Robarts, M., Edwards, T. A., Linden, R. M. & Aggarwal, A. K. (2003). Crystal structure of the SF3 helicase from adeno-associated virus type 2. *Structure* **11**, 1025–1035.

Keideiling, T. (1996). Vibrational circular dichroism application to conformational analysis of biomolecules. In *Circular Dichroism and the Conformational Analysis of Biomolecules*, pp. 555–598. Edited by G. D. Fasman. New York: Plenum Press.

Kelley, L. A. & Sternberg, M. J. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**, 363–371.

Koonin, E. V. (1993). A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. *Nucleic Acids Res* **21**, 2541–2547.

Kumari, P., Singh, A. K., Sharma, V. K., Chattopadhyay, B. & Chakraborty, S. (2011). A novel recombinant tomato-infecting begomovirus capable of transcomplementing heterologous DNA-B components. *Arch Virol* **156**, 769–783.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A. & other authors (2007). CLUSTAL W and CLUSTAL_X version 2.0. *Bioinformatics* **23**, 2947–2948.

Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* **26**, 283–291.

Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J. A., Fanning, E., Jochimiak, A., Szakonyi, G. & Chen, X. S. (2003). Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. *Nature* **423**, 512–518.

Maggin, J. E., James, J. A., Chappie, J. S., Dyda, F. & Hickman, A. B. (2012). The amino acid linker between the endonuclease and helicase domains of adeno-associated virus type 5 Rep plays a critical role in DNA-dependent oligomerization. *J Virol* **86**, 3337–3346.

Nash, T. E., Dallas, M. B., Reyes, M. I., Buhrman, G. K., Ascencio-Ibañez, J. T. & Hanley-Bowdoin, L. (2011). Functional analysis of a novel motif conserved across geminivirus Rep proteins. *J Virol* **85**, 1182–1192.

Nongkhlaw, M., Dutta, P., Hockensmith, J. W., Komath, S. S. & Muthuswami, R. (2009). Elucidating the mechanism of DNA-dependent ATP hydrolysis mediated by DNA-dependent ATPase A, a member of the SWI2/SNF2 protein family. *Nucleic Acids Res* **37**, 3332–3341.

Orozco, B. M., Miller, A. B., Settlage, S. B. & Hanley-Bowdoin, L. (1997). Functional domains of a geminivirus replication protein. *J Biol Chem* **272**, 9840–9846.

San Martín, M. C., Gruss, C. & Carazo, J. M. (1997). Six molecules of SV40 large T antigen assemble in a propeller-shaped particle around a channel. *J Mol Biol* **268**, 15–20.

Shen, J., Gai, D., Patrick, A., Greenleaf, W. B. & Chen, X. S. (2005). The roles of the residues on the channel beta-hairpin and loop structures of simian virus 40 hexameric helicase. *Proc Natl Acad Sci U S A* **102**, 11248–11253.

Sigrist, C. J., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., Bairoch, A. & Bucher, P. (2002). PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform* **3**, 265–274.

Singleton, M. R., Sawaya, M. R., Ellenberger, T. & Wigley, D. B. (2000). Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* **101**, 589–600.

Stanley, J. (1983). Infectivity of the cloned geminivirus genome requires sequences from both DNAs. *Nature* **305**, 643–645.

Story, R. M., Weber, I. T. & Steitz, T. A. (1992). The structure of the *E. coli recA* protein monomer and polymer. *Nature* **355**, 318–325.

Yoon-Robarts, M., Blouin, A. G., Bleker, S., Kleinschmidt, J. A., Aggarwal, A. K., Escalante, C. R. & Linden, R. M. (2004). Residues within the B′ motif are critical for DNA binding by the superfamily 3 helicase Rep40 of adeno-associated virus type 2. *J Biol Chem* **279**, 50472–50481.

Zarate-Perez, F., Bardelli, M., Burgner, J. W., II, Villamil-Jarauta, M., Das, K., Kekilli, D., Mansilla-Soto, J., Linden, R. M. & Escalante, C. R. (2012). The interdomain linker of AAV-2 Rep68 is an integral part of its oligomerization domain: role of a conserved SF3 helicase residue in oligomerization. *PLoS Pathog* **8**, e1002764.