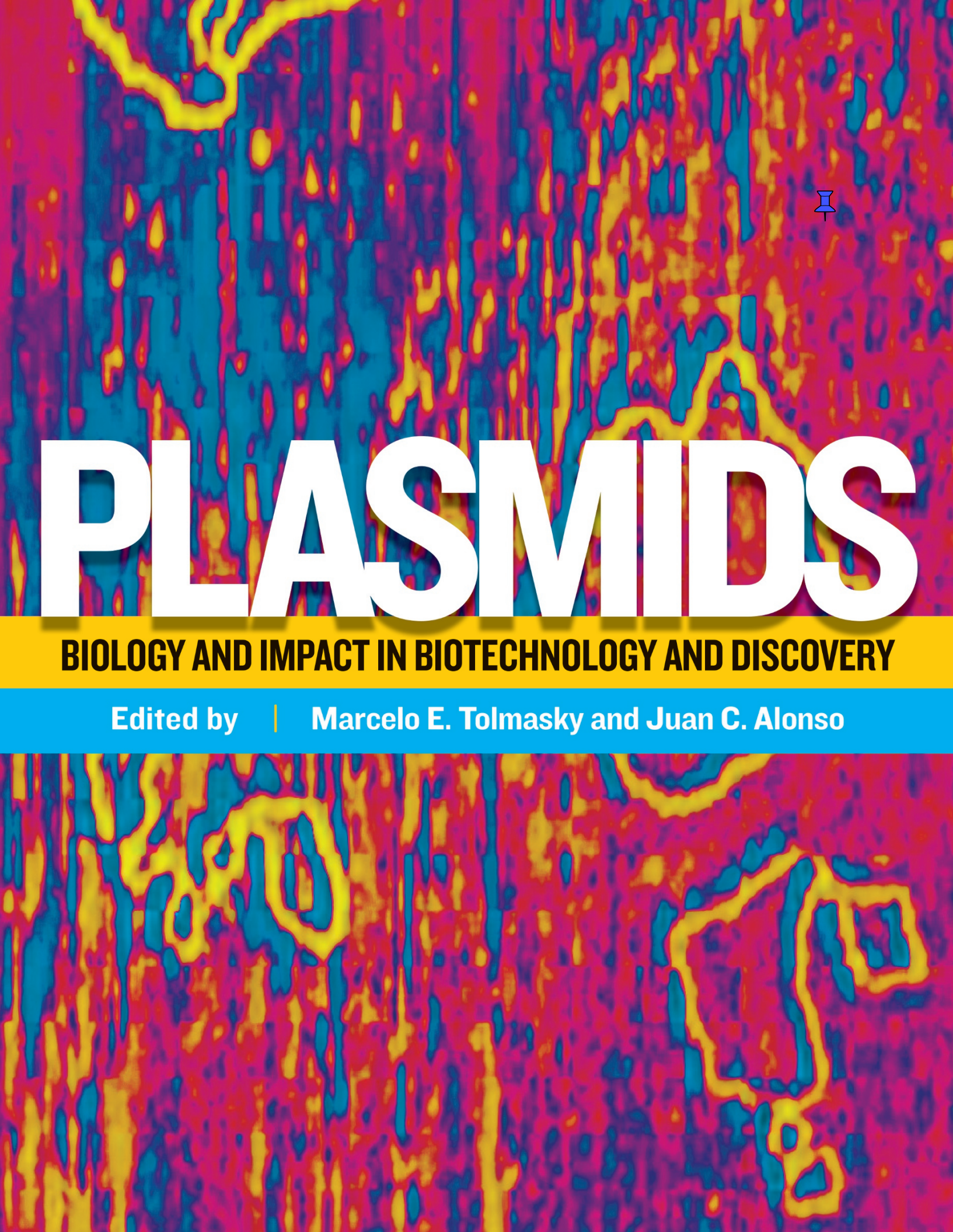
PLASMIDS
BIOLOGY AND IMPACT IN BIOTECHNOLOGY AND DISCOVERY
Edited by | Marcelo E. Tolmasky and Juan C. Alonso

PLASMIDS

BIOLOGY AND IMPACT IN BIOTECHNOLOGY AND DISCOVERY

PLASMIDS

BIOLOGY AND IMPACT IN BIOTECHNOLOGY AND DISCOVERY

Edited by |

Marcelo E. Tolmasky
Center for Applied Biotechnology Studies
Department of Biological Science
College of Natural Sciences and Mathematics
California State University, Fullerton

AND

Juan C. Alonso
Centro Nacional de Biotecnología, CSIC
Departamento de Biotecnología Microbiana,
Madrid, Spain

ASM Press, Washington, DC

Library of Congress Cataloging-in-Publication Data

Plasmids : biology and impact in biotechnology and discovery / edited by Marcelo Tolmasky, Department of Biological Sciences, California State University, Fullerton, and Juan Carlos Alonso, Centro Nacional de Biotecnología, Cantoblanco, Departamento de Biotecnología Microbiana, Madrid, Spain.
pages cm
Includes bibliographical references and index.
ISBN 978-1-55581-897-5 (hardcover : alk. paper) 1. Plasmids. I. Tolmasky, Marcelo, editor. II. Alonso, Juan Carlos (Biotechnologist), editor.
QR76.6.P563 2015
572.8'69—dc23
2015004787
eISBN: 978-1-55581-898-2 doi:10.1128/9781555818982

10 9 8 7 6 5 4 3 2 1

Address editorial correspondence to ASM Press, 1752 N St., N.W., Washington, DC 20036-2904, USA

Send orders to ASM Press, P.O. Box 605, Herndon, VA 20172, USA
Phone: 800-546-2416; 703-661-1593
Fax: 703-661-1501
E-mail: books@asmusa.org
Online: http://estore.asm.org

Cover images: on the front and back covers, the gold-colored ring structures in the background are atomic force microscopy (AFM) images showing one linear molecule and three covalently closed circular ones at different degrees of supercoiling (image provided by Sonia Trigueros); the spine has an electron microscopy (EM) image of a replicating plasmid (image produced by Jorge Crosa).

Cover design: Rings Leighton Design Group, Washington, DC

Dedicated to the memory of Jorge Crosa

Contents

Contributors

Vicki Adams
Department of Microbiology, Australian Research Council Centre of Excellence in Structural and Functional Microbial Genomics, Monash University, Clayton, Victoria 3800, Australia

Juan C. Alonso
Departamento de Biotecnología Microbiana, Centro Nacional de Biotecnología, CSIC, 28049 Madrid, Spain

Karsten Arends
Robert Koch-Institute, Nordufer 20, 13353 Berlin, Germany

Silvia Ayora
Departamento de Biotecnología Microbiana, Centro Nacional de Biotecnología, CNB-CSIC, 28049 Madrid, Spain

Chang-Ho Baek
Life Technologies, Carlsbad, CA 92008

Fernando Baquero
Ramón y Cajal University Hospital, IRYCIS, 28034 Madrid, Spain

Jamie C. Baxter
Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada

Ilaria Benedetti
Departamento de Biotecnología Microbiana, Centro Nacional de Biotecnología, CNB-CSIC, 3, Darwin Street, 28049 Madrid, Spain

LORENA BORDANABA-RUISECO
Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

SABINE BRANTL
AG Bakteriengenetik, Philosophenweg 12, Friedrich-Schiller-Universität Jena, D-07743 Jena, Germany

ALICIA BRAVO
Centro de Investigaciones Biológicas, CSIC, 28040 Madrid, Spain

KATARZYNA BURY
Department of Molecular and Cellular Biology, Intercollegiate Faculty of Biotechnology, University of Gdansk and Medical University of Gdansk, Gdansk, Poland

MANEL CAMPS
Department of Microbiology and Environmental Toxicology, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064

WAI TING CHAN
Centro de Investigaciones Biológicas, CSIC, 28040 Madrid, Spain

KENG-MING CHANG
Section of Molecular Genetics and Microbiology, Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, Texas 78712

JONATHAN CHESNUT
Life Technologies, Carlsbad, CA 92008

PETER J. CHRISTIE
Department of Microbiology and Molecular Genetics, University of Texas Medical School at Houston, Houston, TX 77005

KEVIN CLANCY
Life Technologies, Carlsbad, CA 92008

MIQUEL COLL
Institute for Research in Biomedicine (IRB-Barcelona), and Institut de Biologia Molecular de Barcelona, CSIC, Baldiri Reixac 10-12, 08028 Barcelona, Spain

LAURA C.C. COOK
Department of Medicinal Chemistry, University of Illinois, Chicago, IL 60607

TERESA M. COQUE
Department of Microbiology, Ramón y Cajal University Hospital, IRYCIS, 28034 Madrid, Spain

FRANÇOIS CORNET
CNRS, Laboratoire de Microbiologie et Génétique Moléculaires, F-31062 Toulouse, France

ESTELLE CROZAT
UPS, Laboratoire de Microbiologie et Génétique Moléculaires, Université de Toulouse, F-31062 Toulouse, France

Fernando De La Cruz
Instituto de Biomedicina y Biotecnología de Cantabria (IBBTEC), Universidad de Cantabria, CSIC, 39011 Santander, Spain

Gloria del Solar
Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

Eduardo Díaz
Department of Environmental Biology, Centro de Investigaciones Biológicas (CSIC), 28040 Madrid, Spain

Ramón Díaz-Orejas
Centro de Investigaciones Biológicas, CSIC, 28040 Madrid, Spain

Manuela Di Lorenzo
Department of Microbial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), 6708 PB Wageningen, The Netherlands

Gary M. Dunny
Department of Microbiology, University of Minnesota, Minneapolis, MN 55455

Manuel Espinosa
Consejo Superior de Investigaciones Científicas, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

Hsiu-Fang Fan
Department of Life Sciences and Institute of Genome Sciences, National Yang-Ming University, Taipei 112, Taiwan

Cris Fernández-López
Centro de Investigaciones Biológicas, CSIC, 28040 Madrid, Spain

Andrea T. Feßler
Institute of Farm Animal Genetics, Friedrich-Loeffler-Institut (FLI), Neustadt-Mariensee, Germany

Patrick Forterre
Institut Pasteur, 75015 Paris, France

Florian Fournes
UPS, Laboratoire de Microbiologie et Génétique Moléculaires, Université de Toulouse, F-31062 Toulouse, France

Barbara E. Funnell
Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada

José L. García
Department of Environmental Biology, Centro de Investigaciones Biológicas, CSIC, 28040 Madrid, Spain

M. Pilar Garcillán-Barcia
Instituto de Biomedicina y Biotecnología de Cantabria (IBBTEC), Universidad de Cantabria, CSIC, 39011 Santander, Spain

DANIELLE A. GARSIN
Department of Microbiology and Molecular Genetics, The University of Texas Health Science Center at Houston, Houston, Texas

MICHAEL S. GILMORE
Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Boston, MA 02114

NIKOLAUS GOESSWEINER-MOHR
Institute of Molecular Biosciences, University of Graz, 8010 Graz, Austria

JAY E. GORDON
Department of Microbiology and Molecular Genetics, University of Texas Medical School at Houston, Houston, TX 77005

ELISABETH GROHMANN
Faculty of Biology, Microbiology, Albert-Ludwigs-University Freiburg, 79104 Freiburg, Germany

BERNARD HALLET
Institut des Sciences de la Vie, UC Louvain, 4/5 L7.07.06 Place Croix du Sud, B-1348 Louvain-la-Neuve, Belgium

ANA MARÍA HERNÁNDEZ-ARRIAGA
Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, 28040 Madrid, Spain

ANGELES HUESO
Departamento de Biotecnología Microbiana, Centro Nacional de Biotecnología, CNB-CSIC, 3, Darwin Street, 28049 Madrid, Spain

N. PATRICK HIGGINS
Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL 35294

DAVID C. HOOPER
Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114

GEORGE A. JACOBY
Lahey Hospital and Medical Center, 41 Mall Road, Burlington, MA 01805

MAKKUNI JAYARAM
Department of Life Sciences and Institute of Genome Sciences, National Yang-Ming University, Taipei 112, Taiwan

SVEN JECHALKE
Julius Kühn-Institut, Federal Research Centre for Cultivated Plants (JKI), Institute for Epidemiology and Pathogen Diagnostics, Messeweg 11-12, 38104 Braunschweig, Germany

AASHIQ H. KACHROO
Section of Molecular Genetics and Microbiology, Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, Texas 78712

KRISTINA KADLEC
Institute of Farm Animal Genetics, Friedrich-Loeffler-Institut (FLI), Neustadt-Mariensee, Germany

Clarence I. Kado
Plant Pathology, University of California, Davis, Davis, CA 95616

Federico Katzen
Life Technologies, Carlsbad, CA 92008

Walter Keller
Institute of Molecular Biosciences, University of Graz, 8010 Graz, Austria

Ekaterina Kinnear
Mucosal Infection and Immunity Group, Section of Virology, Imperial College London, St Mary's Campus, London W2 1PG,United Kingdom

Igor Konieczny
Department of Molecular and Cellular Biology, Intercollegiate Faculty of Biotechnology, University of Gdansk and Medical University of Gdansk, Gdansk, Poland

Mart Krupovic
Institut Pasteur, 75015 Paris, France

Antonio Lagares
Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas, IBBM, Instituto de Biotecnología y Biología Molecular, CONICET, Universidad Nacional de La Plata, (1900) La Plata, Argentina

Val Fernández Lanza
Centro de Investigación en Red en Epidemiología y Salud Pública (CIBER-ESP), Melchor Fernández Almagro, 3-5, 28029 Madrid, Spain

Jihong Li
Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, 3550 Terrace Street, Pittsburgh, PA 15261

Joshua Lilly
Department of Microbiology and Environmental Toxicology, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064

David L. Lin
Department of Biological Science, College of Natural Sciences and Mathematics, Center for Applied Biotechnology Studies, California State University, Fullerton, 800 N. State College Blvd., Fullerton, CA 92831

Michael Liss
Life Technologies, Carlsbad, CA 92008

Yen-Ting Liu
Section of Molecular Genetics and Microbiology, Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, Texas 78712

Víctor de Lorenzo
Departamento de Biotecnología Microbiana, Centro Nacional de Biotecnología, CSIC, 3, Darwin Street, 28049 Madrid, Spain

Fabián Lorenzo-Díaz
Instituto Universitario de Enfermedades Tropicales y Salud Pública de Canarias, Universidad de La Laguna, 38071 Laguna, Spain

CHIEN-HUI MA
Section of Molecular Genetics and Microbiology, Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, Texas 78712

CRISTINA MACHÓN
Institute for Research in Biomedicine (IRB-Barcelona), and Institut de Biologia Molecular de Barcelona, CSIC, Baldiri Reixac 10-12, 08028 Barcelona, Spain

ALFONSO H. MAGADAN
Département de Biochimie, Microbiologie et Bio-Informatique, Groupe de Recherche en Écologie Buccale, Faculté des Sciences et de Génie, et de Médecine Dentaire, Félix d'Hérelle Reference Center for Bacterial Viruses, Université Laval, Quebec City, Quebec G1V 0A6, Canada

JOSÉ LUÍS MARTÍNEZ
Centro Nacional de Biotecnología, CNB, and Unidad de Resistencia a Antibióticos y Virulencia Bacteriana (HRYC-CSIC), Madrid, Spain

ESTEBAN MARTÍNEZ-GARCÍA
Departamento de Biotecnología Microbiana, Centro Nacional de Biotecnología, CNB-CSIC, 3, Darwin Street, 28049 Madrid, Spain

BRUCE A. MCCLANE
Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, 3550 Terrace Street, Pittsburgh, PA 15261

SYLVAIN MOINEAU
Groupe de Recherche en Écologie Buccale, Faculté de Médecine Dentaire, Université Laval, Quebec City, Quebec G1V 0A6, Canada

LÁZARO MOLINA
CIDERTA, Laboratorio de Investigación y Control Agroalimentario (LICAH), Parque Huelva Empresarial, 21007 Huelva, Spain

GABRIEL A. MONTEIRO
Department of Bioengineering, Centre for Biological and Chemical Engineering, IBB, Institute for Biotechnology and Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisboa, Portugal

ROBERT J. MOORE
Department of Microbiology, Australian Research Council Centre of Excellence in Structural and Functional Microbial Genomics, Monash University, Clayton, Victoria 3800, Australia

MARIANO PISTORIO
Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas, IBBM, Instituto de Biotecnología y Biología Molecular, CONICET, Universidad Nacional de La Plata, (1900) La Plata, Argentina

DUARTE MIGUEL F. PRAZERES
Department of Bioengineering, Instituto Superior Técnico, Centre for Biological and Chemical Engineering, IBB, Institute for Biotechnology and Bioengineering, Universidade de Lisboa, 1049-001 Lisboa, Portugal

Maria S. Ramirez
Department of Biological Science, Center for Applied Biotechnology Studies, College of Natural Sciences and Mathematics, California State University, Fullerton, 800 N. State College Blvd., Fullerton, CA 92831

Juan Luis Ramos
Environmental Protection Department, Profesor Albareda, Estación Experimental del Zaidin, CSIC, 18008 Granada, Spain

Nikolai V. Ravin
Center of Bioengineering, Russian Academy of Sciences, Prosp. 60-let Oktiabria, Bldg. 7-1, Moscow 117312, Russia

Kasie Raymann
Institut Pasteur, 75015 Paris, France

Julian I. Rood
Department of Microbiology, Australian Research Council Centre of Excellence in Structural and Functional Microbial Genomics, Monash University, Clayton, Victoria 3800, Australia

Philippe Rousseau
UPS, Laboratoire de Microbiologie et Génétique Moléculaires, Université de Toulouse, F-31062 Toulouse, France

Paul A. Rowley
Section of Molecular Genetics and Microbiology, Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, Texas 78712

José A. Ruiz-Masó
Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

Sofía Ruiz-Cruz
Centro de Investigaciones Biológicas, CSIC, 28040 Madrid, Spain

Julie E. Samson
Département de Biochimie, Microbiologie et Bio-Informatique, Groupe de Recherche en Écologie Buccale, Faculté des Sciences et de Génie, et de Médecine Dentaire, Félix d'Hérelle Reference Center for Bacterial Viruses, Université Laval, Quebec City, Quebec G1V 0A6, Canada

Juan Sanjuán
Departamento de Microbiología del Suelo y Sistemas Simbióticos. Estación Experimental del Zaidín, CSIC, Granada, Spain

Saumitra Sau
Section of Molecular Genetics and Microbiology, Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, Texas 78712

Stefan Schwarz
Institute of Farm Animal Genetics, Friedrich-Loeffler-Institut (FLI), Neustadt-Mariensee, Germany

Ana Segura
Environmental Protection Department, Profesor Albareda, Estación Experimental del Zaidin, CSIC, 1, 18008 Granada, Spain

JIANZHONG SHEN
Beijing Key Laboratory of Detection Technology for Animal-Derived Food Safety, College of Veterinary Medicine, China Agricultural University, Beijing 100193, P. R. China

KORNELIA SMALLA
Julius Kühn-Institut, Federal Research Centre for Cultivated Plants (JKI), Institute for Epidemiology and Pathogen Diagnostics, Messeweg 11-12, 38104 Braunschweig, Germany

NORA E. SOBERÓN
Departamento de Biotecnología Microbiana, Centro Nacional de Biotecnología, CNB-CSIC, 28049 Madrid, Spain

VIRTU SOLANO-COLLADO
Centro de Investigaciones Biológicas, CSIC, 28040 Madrid, Spain

NICOLAS SOLER
DynAMic, Université de Lorraine, UMR1128, INRA, Vandoeuvre-lès-Nancy, France

MICHIEL STORK
Process Development, Institute for Translational Vaccinology, 3720 AL Bilthoven, The Netherlands

JACOB STRAHILEVITZ
Hadassah-Hebrew University, Jerusalem 91120, Israel

ANA P. TEDIM
Department of Microbiology, Ramón y Cajal University Hospital, IRYCIS, 28034 Madrid, Spain

MARCELO E. TOLMASKY
Center for Applied Biotechnology Studies, Department of Biological Science, College of Natural Sciences and Mathematics, California State University, Fullerton, 800 N. State College Blvd., Fullerton, CA 92831

EVA M. TOPP
Department of Biological Sciences, University of Idaho, 875 Perimeter, MS 3051, Moscow, Idaho 83844-3051

MARÍA DE TORO
Instituto de Biomedicina y Biotecnología de Cantabria (IBBTEC), Universidad de Cantabria—CSIC, 39011 Santander, Spain

GERMAN M. TRAGLIA
Institute of Microbiology and Medical Parasitology, National Scientific and Technical Research Council (CONICET), University of Buenos Aires, Buenos Aires, Argentina

TUNG TRAN
Department of Biological Science, Center for Applied Biotechnology Studies, College of Natural Sciences and Mathematics, California State University, Fullerton, 800 N. State College Blvd., Fullerton, CA 92831

John S. Tregoning
Mucosal Infection and Immunity Group, Section of Virology, Imperial College London, St Mary's Campus, London, W2 1PG, United Kingdom

Francisco A. Uzal
California Animal Health and Food Safety Laboratory, San Bernardino Branch, School of Veterinary Medicine, University of California, Davis, San Bernardino, CA

Daria Van Tyne
Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, and Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02114

Andrea Volante
Departamento de Biotecnología Microbiana, Centro Nacional de Biotecnología, CNB-CSIC, 28049 Madrid, Spain

Alexander V. Vologodskii
Department of Chemistry, New York University, New York, NY 10003

Yang Wang
Beijing Key Laboratory of Detection Technology for Animal-Derived Food Safety, College of Veterinary Medicine, China Agricultural University, Beijing 100193, P. R. China

Aleksandra Wawrzycka
Department of Molecular and Cellular Biology, Intercollegiate Faculty of Biotechnology, University of Gdansk and Medical University of Gdansk, Gdansk, Poland

Katarzyna Wegrzyn
Department of Molecular and Cellular Biology, Intercollegiate Faculty of Biotechnology, University of Gdansk and Medical University of Gdansk, Gdansk Poland

Sarah Wendlandt
Institute of Farm Animal Genetics, Friedrich-Loeffler-Institut (FLI), Neustadt-Mariensee, Germany

Jessica A. Wisniewski
Australian Research Council, Centre of Excellence in Structural and Functional Microbial Genomics, Department of Microbiology, Monash University, Clayton, Victoria 3800, Australia

Cong-Ming Wu
Institute of Farm Animal Genetics, Friedrich-Loeffler-Institut (FLI), Neustadt-Mariensee, Germany

Preface

One of the biggest dreams of medicine from the 1940s, the complete defeat of infectious diseases caused by bacteria, was treated to a rough awakening with the rise and dissemination of antibiotic resistance, toxins, and pathogenicity functions. In the early 1960s it was found that this dissemination was usually associated with the acquisition of genes that were located in extrachromosomal elements analogous to those that Joshua Lederberg had called "plasmids" in 1952. The importance of the discovery led to intense research on plasmid biology, which in turn resulted in innumerable benefits to the development of science. The list of discoveries in the fields of cell and molecular biology is far too long to detail in this Preface. In addition, a monumental contribution of research on plasmids was instrumental in the development of molecular cloning and the biotechnology revolution that ensued. Their role in virulence and antibiotic resistance, together with the generalization of "omics" disciplines, has recently ignited a new wave of interest in plasmids. As models for understanding innumerable biological mechanisms of living cells, as tools for creating the most diverse therapies, and as invaluable helpers to understand the dissemination of microbial populations, plasmids continue to be at the center of research.

Marcelo E. Tolmasky
Juan C. Alonso

Introduction

I

Plasmids—Biology and Impact in Biotechnology and Discovery
Edited by Marcelo E. Tolmasky and Juan C. Alonso

doi:10.1128/microbiolspec.PLAS-0019-2013

Clarence I. Kado[1]

1 Historical Events That Spawned the Field of Plasmid Biology

INTRODUCTION

Extrachromosomal genetic elements, now widely known as plasmids, were recognized over 60 years ago. Historically, extrachromosomal genetic elements that transferred antibiotic resistance to recipient pathogenic bacteria were called R factors, and those that were conjugative were called T factors (1). Bacteria, particularly *Shigella* strains harboring R and T factors, were found in 1951 in Japan, then in Taiwan and Israel in 1960 (2), and in the United States and Europe in 1963 to 1968 (3). The F factor (for fertility) was the genetic element, also called the "sex factor," that was required for bacterial conjugation (4–8). The sex factor determined the ability of *Escherichia coli* strain K12 to conjugate and transfer genes to recipients.

All of these extrachromosomal elements that propagated either autonomously in the cytoplasm or as an integral part of the host chromosome were called episomes (9). To avoid unnecessary confusion in the usage of a number of terms related to extrachromosomal elements such as plasmagenes, conjugons, pangenes, plastogenes, choncriogenes, cytogenes, proviruses, etc., Lederberg (10; Fig. 1) coined the term *plasmid* to represent any extrachromosomal genetic entity. This term has been widely accepted and used with the understanding that these genetic elements are not organelles, individual genes, parasites (viruses), or symbionts (11). Henceforth, *plasmid(s)* became the conventional term used today.

Based on the established fact that plasmids can reside in *E. coli* and *Shigella* spp., a number of workers began searching for plasmids in other enteric bacteria as well as in pseudomonads and Gram-positive bacteria. By 1977, over 650 plasmids were listed and classified into 29 incompatibility groups (12). Recently, through DNA sequence comparisons of 527 plasmids, there has appeared to be a great deal of interchange of genes between plasmids due to horizontal gene transfer events (13). Incompatibility is determined when two plasmids introduced into a single cell can both replicate and be maintained stably. If the plasmids coexist (replicate and be maintained stably), they are considered compatible. If the plasmids cannot coexist stably, their replication systems are incompatible (14). The incompatible plasmids cannot share a common replication system. Thus, a plasmid classification system was developed that allowed researchers to make logical comparisons of their work on similar plasmids. The classification also provided a system that helped prevent instituting a different name or number for identical plasmids worked on by separate laboratories.

From these early studies, several basic areas of research on plasmids evolved. Researchers focused on

[1]Plant Pathology, University of California Davis, Davis, CA 95616.

Figure 1 Joshua Lederberg.
doi:10.1128/microbiolspec.PLAS-0019-2013.f1

(i) analyzing the physical structure and locating genes on plasmids; (ii) identifying the replication system and the mechanism of replication of plasmids, including how they partition; (iii) determining the conjugative machinery and the mechanism and regulation of plasmid transfer; (iv) dissecting the genetic traits conferred by plasmids, such as metabolic TOL plasmids, bacteriocin-producing Col plasmids, tumor-inducing Ti and virulence plasmids, heavy metal resistance pMOL plasmids, radiation resistant plasmids, etc; (v) restructuring plasmids for utilitarian use, e. g., gene vector development, reporter systems, genetic engineering of mammals and plants; and (vi) surveying the epidemiology and horizontal gene transfer events and reconstructing the evolution of plasmids.

BIRTH OF THE FIELD OF PLASMID BIOLOGY

The term *plasmid biology* was conceived in 1990 at the Fallen Leaf Lake Conference on Promiscuous Plasmids in Lake Tahoe, California. International conferences on plasmid biology were henceforth launched, being held in different countries including Germany, Canada, Spain, the United States, Austria, Mexico, the Czech Republic, Greece, Poland, and Argentina. An example of the proceedings of one of these conferences was published in 2007 (15). An Asian venue is yet to be selected. The International Society for Plasmid Biology was established in 2004 and remains an active internationally recognized professional society (www.ISPB.org).

EARLY STRUCTURAL STUDIES AND GENETIC MAPPING OF PLASMIDS

Knowledge gained from a novel method of separating closed circular DNA from linear DNA in HeLa cells using dye-buoyant CsCl density gradient centrifugation (16) made it possible to examine plasmid DNA derived from bacteria. Earlier studies used analytical centrifugation and density gradient centrifugation on an *E. coli* "episomal element" (F-*lac*) that was conjugatively transferred to *Serratia marcescens*. The 8% difference in guanine plus cytosine content between the episome of *E. coli* (50% GC) vs. *S. marcescens* DNA (58% GC) was sufficient to neatly separate the episome from chromosomal DNA and established the fact that the episome was indeed made of DNA (17). Further physical evidence led to the suggestion that bacteriophage φX174 DNA was circular (18). This was confirmed by electron microscopy by Kleinschmidt et al. (19). Kleinschmidt carefully prepared and used the Langmuir trough technique and examined over 1,000 electron-micrographs to obtain a perfect photograph (A. K. Kleinschmidt, personal communication, 1964). Moreover, phage PM2 DNA was observed by electron microscopy to be a closed circular double-stranded molecule (20). These findings prompted researchers to examine by electron microscopy bacterial extrachromosomal elements of their particular interest and confirmed that plasmids are indeed circular DNA molecules (although linear plasmids also exist).

RECOGNITION OF PLASMID REPLICATION AND PARTITIONING SYSTEMS

Replication of plasmids requires DNA synthesis proteins encoded by chromosomal genes of the hosting bacterial cell. Between one and eight proteins can be involved, depending on the plasmid (Table 1). DNA replication of plasmids is initiated by the binding of the initiator protein to specific binding sites at the replicative origin. Initiator binding promotes the localized unwinding of a discrete region from the DNA origin.

Table 1 Plasmid initiator proteins

Initiator	Replication mode	Plasmid	Molecular mass	References
RepA	Theta type	R1, R100	33 kDa	74, 75
RepA1	Theta type	EntP307	40 kDa	76
RepA	Theta type	pSC101	37.5 kDa	77, 78
RepC	Theta type	RSF1010	31 kDa	79
RepE	Theta type	F	29 kDa	80
TrfA	Theta type	RK2	33 kDa	81
π (*pir*)	Theta type	R6K	35 kDa	82
RepA	Rolling circle	pA1	5.6 kDa	83
RepB	Rolling circle	pLS1	24.2 kDa	84
RepC	Rolling circle	pT181	38 kDa	85
RepD	Rolling circle	pC221	38 kDa	86

A helicase is then directed to the exposed single-stranded DNA region followed by a prepriming complex to initiate DNA synthesis (21). Initiation of DNA replication by the initiator binding to the origin sequence(s) is a critical function in plasmid survival as an extracellular genetic element.

As part of the plasmid replication process, specific plasmid concentrations (copy number) occur as the host bacterial cell initiates cell division. Partitioning and stable segregation of the plasmid are initiated. Partition systems are categorically classified based on ATPase proteins (22). Type I is characterized by Walker box ATPases, while a subset, type Ia, occurs when the nucleotide-binding P-loop is preceded by an N-terminal regulatory domain, and in type Ib this is not the case. The mechanisms that contribute to the stable segregation of plasmids F, P1, R1, NR1, pSC101, and ColE1 have been reviewed (23). The locus responsible for partitioning of pSC101 was designated "*par*" (24). The *par* locus is able to rescue unstable pSC101-derived replicons in the *cis*, but not the *trans*, configuration. It is independent of copy number control, does not specify plasmid incompatibility, and is not associated directly with plasmid replication functions. From phylogenetic analysis of *par* loci from plasmids and bacterial chromosomes, two trans-acting proteins form a nucleoprotein complex at a *cis*-acting centromere-like site (22). One these proteins, identified as an ATPase, functions to tether plasmids and chromosomal origin regions to specific poles of the dividing cells. Therefore, the mitotic stability of plasmids depends on a centromere, a centromere-binding protein, and an ATPase. In the case of plasmid F, two genes, *sopA* and *sopB*, and a centromeric target site, *sopC*, function to ensure that both daughter cells receive a daughter plasmid during cell division. The products of *sopA* and *sopB* stabilize the plasmid bearing a centromere-like sequence in *sopC* (25). SopA hydrolyzes ATP by binding DNA (26). The centromere-like region contains a 43-bp sequence that is repeated 12 times in the same orientation (27), and each element contains a 7-bp inverted repeat targeted by SopB (28). Like *sopA*, *sopB*, and *sopC* of plasmid F, plasmid P1 has counterpart partition genes (*parA*, *parB*) and a target site (*pars*) (29).

Some plasmids such as ColE1 are partitioned randomly at cell division, and their inheritance is proportional to the number of plasmids present in the cell (30). High-copy-number plasmids usually do not require an active *par* system for stable maintenance because random distribution ensures plasmid segregation to the two daughter cells at the time of cell division, while larger, low-copy plasmids such as F, R100, and P1 possess genes that encode inhibitors of host cell growth. In the case of plasmid F, the *ccdA* and *ccdB* (for coupled cell division) genes encode an 8.7-kDa and an 11.7-kDa protein, respectively, the latter of which inhibits cell growth (31). This inhibitor functions in cells that have lost their plasmid due to errors in replication or cell division. The action of the inhibitor is prevented by the CcdA protein, which loses stability in the absence of the plasmid and therefore no longer functions to inhibit the action of the CcdB protein. Plasmid biologists have referred to this interesting mechanism of controlling plasmid copy number as a "killing" function that specifically kills cells lacking a plasmid (or postsegregational killing).

LANDMARKS LEADING TO PLASMID-MEDIATED CONJUGATIVE TRANSFER

The historical experiments on plating together two different triple auxotrophic mutants leading to prototrophic bacterial colonies that propagated indefinitely on minimal medium was the classical laboratory event that led Lederberg and Tatum (6, 7) to conclude that there was sex in bacteria (32). Examination of single cell isolates of these prototrophic strains showed that they were indeed heterozygotes. Hayes (4) showed the heterothallic nature of conjugation whereby recombination is mediated by the one-way transfer of genetic material from donor to recipient bacteria. Self-transmissible plasmids such as F, R1, R100, and R6K encode the capacity to promote conjugation. They all possess related transfer (*tra)* genes. Plasmid F (called sex factor)-mediated conjugation has received the most attention. *E. coli* harboring this sex factor produce a filamentous organelle called the F pilus (Fig. 2) that was needed for conjugation between sex factor-bearing

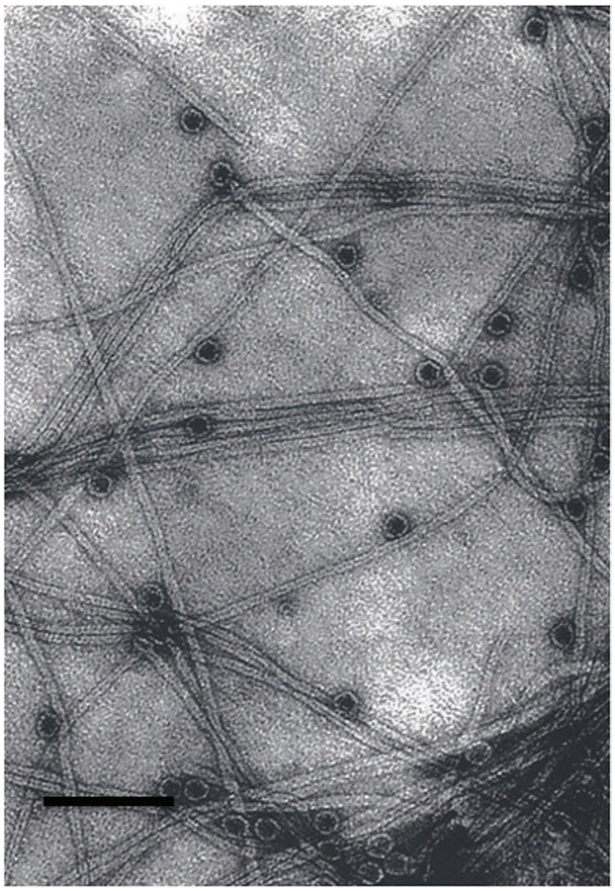

Figure 2 Purified F pili bearing spherical RNA MS2 phages. Electron micrograph courtesy of Professor Manabu Inuzuka, Fukui Medical University, Fukui, Japan. Bar = 2000 Å.
doi:10.1128/microbiolspec.PLAS-0019-2013.f2

donors (known as F^+ donors) and F^- recipients. Historically, the F pilus (or "sex pilus," coined by Harden and Meynell [33] and reviewed by Tomoeda et al. [34]) was suggested by Brinton (35) to serve as a conduit through which DNA passes. Somewhat similar to bacteriophage (T phage) tail retraction, the F pilus was proposed to retract and bring together conjugating cells into wall-to-wall contact (36, 37). Although the F pilus is needed for initial contact between F^+ and F^- cells, it is not necessary for DNA transfer after the contacts have stabilized (38). The formation of mating pairs involves a complex apparatus bridging the donor cell envelope that assembles the conjugative pilus. The pilus interacts with the recipient cell and apparently retracts by depolymerization into the donor cell, culminating in intimate wall-to-wall contact during mating-pair stabilization (38, 39).

This type of intimate contact, termed the conjugational junction, between stabilized mating pairs was examined by electron microscopy of thin sections of the junction (40). No specific substructure such as a plasma bridge was observed. Interestingly, the F pilus of *E. coli* was claimed to support stable DNA transfer in the absence of wall-to-wall contact between cells (41). In earlier work using micromanipulation, Ou and Anderson (42) showed DNA transfer in the absence of direct cell-to-cell contact. More recently, the F pilus was observed in real-time visualization to mediate DNA transfer at considerable cell-to-cell distances (43). Most (96%) of the transferred DNA integrated by recombination in the distal recipient cells.

Genetic and sequence analyses have provided further insights to the mechanism of plasmid DNA transfer. With conjugative plasmids, the genes required for mating pair formation and DNA transfer are located in one or two clusters identified as the transfer (*tra*) regions (44). The proteins involved in the unidirectional transfer of single-stranded DNA from donor to recipient are encoded by the *tra* operon of the F plasmid. These proteins form the relaxosome, which processes plasmid DNA at the origin of transfer (*oriT*). Sequence similarities were recognized between pilin-encoding genes of F-like plasmids (45). Studies of the promiscuous DNA transfer system encoded by the Ti (for tumor-inducing) plasmid of *Agrobacterium tumefaciens* revealed that the *virB* operon encodes a sex pilus involved in T-DNA transfer to plants (46). Moreover, the *virB* operon of the Ti plasmid exhibits close homologies to genes that are known to encode the pilin subunits and pilin assembly proteins of other conjugative plasmids such as F, R388, RP4, and even the *ptl* operon of *Bordetella pertussis* (46, 47). The components of these plasmid transfer apparatuses became classified as members of the type IV secretion family (48). The F plasmid transfer apparatus has homologs to VirB proteins encoded by the *virB* operon of the type IV secretion system (49). In fact, the VirB2 propilin protein is similar to the TraA propilin of F and is processed into their respective pilin subunit of a size (50–53) similar to the T-pilus (51, 54). Posttranslational processing also occurs with VirB1, a pilin-associated protein (55). Interestingly, the type IV DNA-protein transfer system of the Ti plasmid is highly promiscuous by promoting transfer between the domain *Bacteria* to members of the domain *Eukarya* (56).

Based on the intensive and excellent studies on plasmid DNA transfer systems of narrow and broad-host-range conjugative plasmids by a large number of excellent researchers past and present (reviewed in 57, 58), it appears that the transmission or transfer of plasmids is essential to their survival (see below).

FUNCTIONAL ATTRIBUTES REQUIRED FOR PLASMID PERSISTENCE AND SURVIVAL

Conjugative transfer of plasmids reflects an indispensable trait required for their ensured survival as selfish DNA molecules (56, 59). Traits such as conferring antibiotic resistance were first recognized as being plasmid-borne in *Shigella* and *Salmonella* spp. as described in the introduction above. Antibiotic resistance conferred by plasmid genes provided survival value to pathogens that would otherwise be killed by the antibiotic(s). This in turn offered survival and maintenance of the plasmid itself in the antibiotic-resistant pathogenic bacterial host. Likewise, metabolic/catabolic plasmids confer on host bacteria the ability to survive in harsh environments such as in sediments from industrial waste and from mining exudates of silver, copper, cadmium, tellurite, etc. Unusual environments such as sites containing an abundance of substrates such as aromatic hydrocarbons, toluene, xylene, pesticides, herbicides, and organic waste products all provided specialized niches for bacteria that live under the auspices of specialized enzymes that degrade or modify one or more of these compounds. These bacteria harbor plasmids that confer on their host cell the ability to metabolize, degrade, or modify substances that otherwise would be toxic or lethal to the host bacterial cell. The catabolic TOL plasmid pWWO, first described by Williams and Murray (60), is one of the best studied for its catabolic enzymes and genetic structure (61).

The selfishness of plasmids is exemplified by plasmids encoding bacteriocins that kill susceptible bacterial cells not harboring the same or like plasmids. The lethal action of these antibacterial proteins occurs through puncturing plasma membranes, degrading nucleic acids, or cleaving peptidoglycans. Examples of bacteriocins are colicin encoded by plasmid ColE1 (62), cloacin encoded by plasmid CloDF13 (63), and nisin F encoded by plasmid pF10 (64).

Of medical and veterinary relevance are plasmids that confer virulence traits on their bacterial hosts. Various pathogenic *E. coli* strains harbor plasmids that confer interesting virulence traits (65). Loss of the virulence-conferring plasmid results in the loss of its pathogenic trait unless the pathogenicity island transposes into the chromosome of the bacterial host. Another member of the *Enterobacteriaceae* are *Shigella* spp. All invasive *Shigella flexneri* strains, regardless of serotype, harbor a large virulence plasmid, pWR110 (66). Mutagenesis or curing of the plasmid results in the loss of pathogenicity. Plasmid-conferred virulence is not restricted to Gram-negative bacteria. Indeed, the pathogenicity of *Staphylococcus aureus* is highly dependent on its resident plasmid (67). The genes conferring the pathogenic trait and antibiotic resistance are highly conserved, and their spread among *S. aureus* strains is restrained. A number of plant pathogens also harbor virulence plasmids, a number of which encode secretion machinery for injection into their host plants (reviewed in Kado [68]).

RECONSTRUCTION OF PLASMIDS FOR BIOTECHNOLOGY AND BIOMEDICAL APPLICATIONS

The development of recombinant DNA techniques (69) has led to a multitude of possibilities of designing plasmid vector systems useful in fundamental research and industrial, agricultural, and medical applications. Early vector systems were based on ColE1 derivatives that were primarily restricted to *E. coli* owing to their replication machinery. The introduction of broad-host-range plasmids such as RK2 and RSF1010 made it possible to introduce recombinant DNA technologies into bacteria other than members of the *Enterobacteriaceae*. In recent times, a number of plasmid shuttle vector systems have become commercially available, too numerous to list in this paper. Plasmids constructed as vectors for various purposes are reviewed elsewhere (70). Some examples of useful vector systems are listed in Table 2. Vectors designed for pharmaceutical and genetic engineering of mammalian and plant cells have been recently reviewed (71–73).

CONCLUSION AND FUTURE OF PLASMID BIOLOGY

Plasmids have provided the basic foundation for recombinant DNA technologies. Significant insights are being gained from genome sequencing and reconstruction by computer modeling of prospective enzymes (proteins) encoded by sequenced plasmid genes. The commercially available kits for plasmid isolation, DNA amplification, sequencing, and a large number of purified enzymes have made earlier laborious procedures part of history. However, at the same time, there is the loss of insightful knowledge due to the absence of on-hand experiences for isolating nucleic acids and proteins and seeing exactly what they do in reconstruction experiments.

In-depth studies of how plasmids are maintained and dispersed, and how they acquire or lose encoded traits, and of why they persist in the natural and even in man-made environments all are important questions that remain in the field of plasmid biology. Plasmid biologists

Table 2 Examples of plasmid vector systems and their uses

Vector	Application	References
pBR322	General cloning, provided basis for ColE1 cloning vector derivatives	87
pUC	Multiple cloning sites, open reading frame DNA as *lacZ* fusions controlled by *lac* regulatory elements	88–90
pHG175	Multiple cloning sites, promoter probe for tetracycline resistance	91
pKUN9	A pUC9 derivative modified whereby both strands of a cloned DNA fragment can be obtained in a single-stranded form for expeditious sequencing	92
pUCD2335	Mini-T DNA vector bearing a high-copy vir region for genetic engineering of plants	93
pBIN19	Binary vector system for genetic engineering of plants	72
pUCD607	Luciferase reporter of real-time infection by bacteria in higher cells	94
pUCD800	Vector for positive selection of transposons and insertion elements via sucrose sensitivity conferred by the *sacB* gene that encodes levan sucraseLethal to enteric bacteria	95
pUCD2715	*Vibrio* luciferase vector for genetic engineering of plants to make them glow in the dark	96
pWS233	*sacRB* bearing vector bearing gentamicin and tetracycline resistance genes and Mob functions of RP4	97
pUCD4121	Vector that generates unmarked deletions in bacterial chromosomes; bears a *sacB* lethality and neomycin resistance gene	98
pGKA10CAT	A Bluescript pKS(+) derivative for functional analysis of enhancer domains of a transcriptional regulatory region	99
pXL1635	Derived from pRK290, contains RP4 *par* fragment and deleted *oriT* of RK2; for industrial use	100
pJQ200 & pJQ210	Suicide vectors bearing *sacB*, *ori* of pACYC184, and *oriT* and *mob* of RP4	101
pUCD5140	Light sensitivity-producing vector derived from pUCD2335 containing a *rbcS3A* promoter-*gus* fusion and CaMV35S promoter driving a phytochrome A gene of *Avena sativa*	102
pJAZZ	Linear vector for *E. coli* cloning, contains phage N15 *ori*, minimizes formation of nonrecombinants	103
pHP45Ω	A pBR322 derivative for insertional mutagenesis, bearing Ω, and streptomycin/spectinomycin resistance genes flanked by inverted repeats with transcription/translation termination signals and synthetic polylinkers	104

who, "outside of the box" (e.g., replication, partitioning, conjugation) have far-sighted visions of the future prospects of the field of plasmid biology will be the key contributors to the science.

Acknowledgment. *Conflicts of interest: I disclose no conflicts.*

Citation. Kado CI. 2014. Historical events that spawned the field of plasmid biology. Microbiol Spectrum **2**(5)**:**PLAS-0019-2013.

References

1. **Mitsuhashi S, Kameda M, Harada K, Suzuki M.** 1969. Formation of recombinants between non-transmissible drug-resistance determinants and transfer factors. *J Bacteriol* **97:**1520–1521.
2. **Nakaya R, Nakamura A, Murata Y.** 1960. Resistance transfer agents in *Shigella*. *Biochem Biophys Res Commun* **3:**654–659.
3. **Mitsuhashi S.** 1977. Epidemiology of R factors, p 25–43. *In* Mitsuhashi S (ed), *R Factor, Drug Resistance Plasmid*. University Park Press, Baltimore, MD.
4. **Hayes W.** 1952. Recombination in *Bact. coli* K12: unidirectional transfer of genetic material. *Nature (London)* **169:**118–119.
5. **Hayes W.** 1953. Observations on a transmissible agent determining sexual differentiation in *Bact. coli*. *J Gen Microbiol* **8:**72–88.
6. **Lederberg J, Tatum EL.** 1946. Novel genotypes in mixed cultures of biochemical mutants of bacteria. *Cold Spring Harbor Symp Quant Biol* **11:**113–114.
7. **Lederberg J, Tatum EL.** 1946. Gene recombination in *Escherichia coli*. *Nature (London)* **158:**558.
8. **Lederberg J, Cavalli LL, Lederberg EM.** 1952. Sex compatibility in *Escherichia coli*. *Genetics* **37:**720–730.
9. **Jacob F, Wollman EL.** 1958. Les épisomes, elements génétiques ajoutés. *C R Hebd. Seances Acad Sci* **247:** 154–156.
10. **Lederberg J.** 1952. Cell genetics and hereditary symbiosis. *Physiol Rev* **32:**403–430.
11. **Lederberg J.** 1998. Plasmid (1952–1997). *Plasmid* **39:** 1–9.
12. **Bukhari AI, Shapiro JA, Adhya SL (ed).** 1977. *DNA Insertion Elements, Plasmids, and Episomes*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
13. **Zhou Y, Call DR, Broschat SL.** 2012. Genetic relationships among 527 Gram-negative bacterial plasmids. *Plasmid* **68:**133–141.
14. **Novick RP.** 1987. Plasmid incompatibility. *Microbiol Rev* **51:**381–395.
15. **Kado CI, Helinski DR.** 2007. Proceedings of the international symposium on plasmid biology. *Plasmid* **57:** 182–243.
16. **Radloff R, Bauer W, Vinograd J.** 1967. A dye-buoyant-density method for the detection and isolation of closed

circular duplex DNA: the closed circular DNA in HeLa cells. *Proc Natl Acad Sci USA* **57:**1514–1521.

17. **Marmur J, Rownd R, Falkow S, Baron LS, Schildkraut C, Doty P.** 1961. The nature of intergeneric episomal infection. *Proc Natl Acad Sci USA* **47:**972–979.
18. **Fiers W, Sinsheimer RL.** 1962. The structure of the DNA of bacteriophage X174 III. Ultracentrifugal evidence for a ring structure. *J Mol Biol* **5:**424–434.
19. **Kleinschmidt AK, Burton A, Sinsheimer RL.** 1963. Electron microscopy of the replicative form of the DNA of the bacteriophage phi-X174. *Science* **142:**961.
20. **Espejo RT, Canelo ES, Sinsheimer RL.** 1969. DNA of bacteriophage PM2: a closed circular double-stranded molecule. *Proc Natl Acad Sci USA* **63:**1164–1168.
21. **Bramhill D, Kornberg A.** 1988. Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. *Cell* **52:** 743–755.
22. **Gerdes K, Møller-Jensen J, Bugge Jensen R.** 2000. Plasmid and chromosome partitioning: surprises from phylogeny. *Mol Microbiol* **37:**455–466.
23. **Nordström K, Austin SJ.** 1989. Mechanisms that contribute to the stable segregation of plasmids. *Annu Rev Genet* **23:**37–69.
24. **Meacock PA, Cohen SN.** 1980. Partitioning of bacterial plasmids during cell division: a *cis*-acting locus that accomplishes stable plasmid inheritance. *Cell* **20:**529–542.
25. **Ogura T, Hiraga S.** 1983. Partition mechanism of F plasmid: two plasmid gene-encoded products and a *cis*-acting region are involved in partition. *Cell* **32:** 351–360.
26. **Ah-Seng Y, Lopez F, Pasta F, Lane D, Bouet J-Y.** 2009. Dual role of DNA in regulating ATP hydrolysis by the SopA partition protein. *J Biol Chem* **284:**30067–30075.
27. **Mori H, Kondo A, Ohshima A, Ogura T, Hiraga S.** 1986. Structure and function of F plasmid genes essential for partitioning. *J Mol Biol* **192:**1–15.
28. **Hayakawa Y, Murotsu T, Matsubara K.** 1985. Mini-F protein that binds to a unique region for partition of mini-F plasmid DNA. *J Bacteriol* **163:**349–354.
29. **Abeles AL, Snyder KM, Chattoraj DK.** 1984. P1 plasmid replication: replicon structure. *J Mol Biol* **173:**307–324.
30. **Summers DK, Sherratt DJ.** 1984. Multimerization of high copy number plasmids causes instability: ColE1 encodes a determinant essential for plasmid monomerization and stability. *Cell* **36:**1097–1103.
31. **Jaffé A, Ogura T, Hiraga S.** 1985. Effects of the ccd function of the F plasmid on bacterial growth. *J Bacteriol* **163:**841–849.
32. **Lederberg J, Tatum EL.** 1953. Sex in bacteria: genetic studies, 1945–1952. *Science* **118:**169–175.
33. **Harden V, Meynell E.** 1972. Inhibition of gene transfer by antiserum and identification of serotypes of sex pili. *J Bacteriol* **109:**1067–1074.
34. **Tomoeda M, Inuzuka M, Date T.** 1975. Bacterial sex pili. *Prog Biophys Mol Biol* **30:**23–56.
35. **Brinton CC Jr.** 1965. The structure, function, synthesis and genetic control of bacterial pili and a molecular model for DNA and RNA transport in Gram negative bacteria. *Trans NY Acad Sci* **27:**1003–1054.
36. **Curtiss R.** 1969. Bacterial conjugation. *Annu Rev Microbiol* **23:**69–136.
37. **Marvin DA, Hohn B.** 1969. Filamentous bacterial viruses. *Bacteriol Rev* **33:**172–209.
38. **Achtman M, Morelli G, Schwuchow S.** 1978. Cell-cell interactions in conjugating *Escherichia coli*: role of F pili and fate of mating aggregates. *J Bacteriol* **135:** 1053–1061.
39. **Achtman M, Kennedy N, Skurray R.** 1977. Cell-cell interactions in conjugating *Escherichia coli*: role of *traT* protein in surface exclusion. *Proc Natl Acad Sci USA* **74:**5104–5108.
40. **Durrenberger MB, Villiger W, Bachi T.** 1991. Conjugational junctions: morphology of specific contacts in conjugating *Escherichia coli* bacteria. *J Struct Biol* **107:** 146–156.
41. **Harrington LC, Rogerson AC.** 1990. The F pilus of *Escherichia coli* appears to support stable DNA transfer in the absence of wall-to-wall contact between cells. *J Bacteriol* **172:**7263–7264.
42. **Ou JT, Anderson TF.** 1970. Role of pili in bacterial conjugation. *J Bacteriol* **102:**648–654.
43. **Babic A, Lindner AB, Vulic M, Stewart EJ, Radman M.** 2008. Direct visualization of horizontal gene transfer. *Science* **319:**1533–1536.
44. **Frost LS, Ippen-Ihler K, Skurray RA.** 1994. Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol Rev* **58:**162–210.
45. **Frost LS, Finlay BB, Opgenorth A, Paranchych W, Lee JS.** 1985. Characterization and sequence analysis of pilin from F-like plasmids. *J Bacteriol* **164:**1238–1247.
46. **Kado CI.** 1994. Promiscuous DNA transfer system of *Agrobacterium tumefaciens:* role of the *virB* operon in sex pilus assembly and synthesis. *Mol Microbiol* **12:** 17–22.
47. **Shirasu K, Kado CI.** 1993. The *virB* operon of the *Agrobacterium tumefaciens* virulence regulon has sequence similarities to B, C and D open reading frames downstream of the pertussis toxin-operon and to the DNA transfer-operons of broad-host-range conjugative plasmids. *Nucleic Acids Res* **21:**353–354.
48. **Cascales E, Christie PJ.** 2003. The versatile bacterial type IV secretion systems. *Nat Rev Microbiol* **1:**137–149.
49. **Lawley TD, Klimke WA, Gubbin MJ, Frost LS.** 2003. F factor conjugation is a true type IV secretion system. *FEMS Microbiol Lett* **224:**1–15.
50. **Jones AL, Lai EM, Shirasu K, Kado CI.** 1996. VirB2 is a processed pilin-like protein encoded by the *Agrobacterium* Ti plasmid. *J Bacteriol* **178:**5706–5711.
51. **Lai EM, Kado CI.** 1998. Processed VirB2 is the major subunit of the promiscuous pilus of *Agrobacterium tumefaciens*. *J Bacteriol* **180:**2711–2717.
52. **Lai EM, Eisenbrandt R, Kalkum M, Lanka E, Kado CI.** 2002. Biogenesis of T pili in *Agrobacterium tumefaciens*

requires precise VirB2 propilin cleavage and cyclization. *J Bacteriol* **184:**327–330.

53. **Shirasu K, Kado CI.** 1993. Membrane location of the Ti plasmid VirB proteins involved in the biosynthesis of a pilin-like conjugative structure on *Agrobacterium tumefaciens*. *FEMS Microbiol Lett* **111:**287–294.
54. **Lai EM, Kado CI.** 2000. The T-pilus of *Agrobacterium tumefaciens*. *Trends Microbiol* **8:**361–369.
55. **Zupan JR, Ward D, Zambryski P.** 1998. Assembly of the VirB transport complex for DNA transfer from *Agrobacterium tumefaciens* to plant cells. *Curr Opin Microbiol* **1:**649–655.
56. **Kado CI.** 2009. Horizontal gene transfer: sustaining pathogenicity and optimizing host-pathogen interactions. *Mol Plant Pathol* **10:**143–150.
57. **Novick RP.** 1969. Extrachromosomal inheritance in bacteria. *Bacteriol Rev* **33:**210–235.
58. **Phillips G, Funnel B.** 2004. *Plasmid Biology*. ASM Press, Washington, DC.
59. **Kado CI.** 1998. Origin and evolution of plasmids. *Antonie van Leeuwenhoek* **73:**117–126.
60. **Williams PA, Murray K.** 1974. Metabolism of benzoate and the methylbenzoates by *Pseudomonas putida* (arvilla) mt-2: evidence for the existence of a TOL plasmid. *J Bacteriol* **120:**416–423.
61. **Burlage RS, Hooper SW, Sayler GS.** 1989. The TOL (pWW0) catabolic plasmid. *Appl Environ Microbiol* **55:** 1323–1328.
62. **Staudenbauer WL.** 1978. Structure and replication of the colicin E1 plasmid. *Curr Top Microbiol Immunol* **83:**93–156.
63. **Van Tiel-Menkvled GJ, Rezee A, De Graaf FK.** 1979. Production and excretion of cloacin DF13 by *Escherichia coli* harboring plasmid CloDF13. *J Bacteriol* **140:** 415–423.
64. **De Kwaadsteniet M, ten Doeschate K, Dicks LMT.** 2007. Characterization of the structural gene encoding Nisin F, a new lantibiotic produced by a *Lactococcus lactis* subsp. *lactis* isolate from freshwater catfish (*Claria gariepinus*). *Appl Environ Microbiol* **74:**547–549.
65. **Johnson TJ, Nolan LK.** 2009. Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiol Molec Biol Rev* **73:**750–774.
66. **Sansonetti PJ, Kopecko DJ, Formal SB.** 1982. Involvement of a plasmid in the invasive ability of *Shigella flexneri*. *Infect Immun* **35:**852–860.
67. **McCarthy AJ, Lindsay JA.** 2012. The distribution of plasmids that carry virulence and resistance genes in *Staphylococcus aureus* is lineage associated. *BMC Microbiol* **12:**104. doi:10.1186/1471-2180-12-104.
68. **Kado CI.** 2010. *Plant Bacteriology*. APS Press, St. Paul, MN.
69. **Lobban P, Kaiser AD.** 1973. Enzymatic end-to-end joining of DNA molecules. *J Mol Biol* **79:**453–471.
70. **Rodriguez RL, Denhardt DT (ed).** 1988. *Vectors: A Survey of Molecular Cloning Vectors and Their Use*. Butterworths, London.
71. **Kaufman RJ.** 2000. Overview of vector design for mammalian gene expression. *Mol Biotechnol* **16:**151–160.
72. **Lee LY, Gelvin SB.** 2008. T-DNA binary vectors and systems. *Plant Physiol* **146:**325–332.
73. **Tolmachov OE.** 2011. Building mosaics of therapeutic plasmid gene vectors. *Curr Gene Ther* **11:**466–478.
74. **Masai H, Kaziro Y, Arai K.** 1983. Definition of *oriR*, the minimum DNA segment essential for initiation of R1 plasmid replication *in vitro*. *Proc Natl Acad Sci USA* **80:**6814–6818.
75. **Rosen J, Ryder T, Inokuchi H, Ohtsubo H, Ohtsubo E.** 1980. Genes and sites involved in replication and incompatibility of an R100 plasmid derivative based on nucleotide sequence analysis. *Mol Gen Genet* **179:** 527–537.
76. **Song H, Phillips SE, Parsons MR, Maas R.** 1996. Crystallization and preliminary crystallographic analysis of RepA1, a replication control protein of the RepFIC replicon of enterotoxin plasmid EntP307. *Proteins* **25:** 137–138.
77. **Churchward G, Linder P, Caro L.** 1983. The nucleotide sequence of replication and maintenance functions encoded by plasmid pSC101. *Nucleic Acids Res* **11:** 5645–5659.
78. **Vocke C, Bastia D.** 1983. DNA-protein interaction at the origin of DNA replication of the plasmid pSC101. *Cell* **35:**495–502.
79. **Scherzinger E, Haring V, Lurz R, Otto S.** 1991. Plasmid RSF1010 DNA replication *in vitro* promoted by purified RSF1010 RepA, RepB and RepC proteins. *Nucleic Acids Res* **19:**1203–1211.
80. **Komori H, Matsunaga F, Higuchi Y, Ishiai M, Wada C, Miki K.** 1999. Crystal structure of a prokaryotic replication initiator protein bound to DNA at 2.6 Å resolution. *EMBO J* **18:**4597–4607.
81. **Kongsuwan K, Josh P, Picault MJ, Wijffels G, Dalrymple B.** 2006. The plasmid RK2 replication initiator protein (TrfA) binds to the sliding clamp beta subunit of DNA derived from the amino-terminal portion of 33-kilodalton TrfA. *J Bacteriol* **188:**5501–5509.
82. **Germino J, Bastia D.** 1983. Interaction of the plasmid R6K-encoded replication initiator protein with its binding sites on DNA. *Cell* **34:**125–134.
83. **Vuicic M, Topisirovic L.** 1993. Molecular analysis of the rolling-circle replicating plasmid pA1 of *Lactobacillus plantarum* A112. *Appl Environ Microbiol* **59:**274–280.
84. **De la Campa AG, del Solar GH, Espinosa M.** 1990. Initiation of replication of plasmid pLS1: the initiator protein RepB acts on two distant DNA regions. *J Mol Biol* **213:**247–262.
85. **Koepsel RR, Murray RW, Rosenblum WD, Khan SA.** 1985. Purification of pT181-encoded RepC protein required for the initiation of plasmid replication. *J Biol Chem* **260:**8571–8577.
86. **Thomas CD, Baison DF, Shaw WV.** 1990. *In vitro* studies of the initiation of staphyloccal plasmid replication. Specificity of RepD for its origin (oriD) and characterization of the Rep-ori tyrosyl ester intermediate. *J Biol Chem* **265:**5519–5530.

87. **Balbás P, Soberón X, Merino E, Zurita M, Lomeli H, Valle F, Flores N, Bolivar F.** 1986. Plasmid vector pBR322 and its special-purpose derivatives: a review. *Gene* **50:**3–40.

88. **Messing J.** 1983. New M13 vectors for cloning, p 20–78. *In* Wu R, Grossman L, Moldave K (ed), *Methods in Enzymology*, Academic Press, Orlando, FL.

89. **Norrander J, Kempe T, Messing J.** 1983. Construction of improved M13 vectors using oligo-deoxynucleotide-directed mutagenesis. *Gene* **16:**101–106.

90. **Vieira J, Messing J.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19:** 259–268.

91. **Stewart GSAB, Lubinsky-Mink S, Jackson CG, Cassel A, Kuhn J.** 1986. pHG165: a pBR322 copy number derivative of pUC8 for cloning and expression. *Plasmid* **15:**172–181.

92. **Peeters BPH, Schoenmakers JGG, Konings RNH.** 1986. Plasmid pKUN9, a versatile vector for the selective packaging of both DNA strands into single-stranded DNA-containing phage-like particles. *Gene* **41:**39–46.

93. **Zyprian E, Kado CI.** 1990. *Agrobacterium*-mediated plant transformation by novel mini-T vectors in conjunction with a high-copy *vir* region helper plasmid. *Plant Mol Biol* **15:**245–256.

94. **Shaw JJ, Kado CI.** 1986. Development of a *Vibrio* bioluminescence gene-set to monitor phytopathogenic bacteria during the ongoing disease process in a non-disruptive manner. *Nat Biotechnol* **4:**560–564.

95. **Gay P, LeCoq D, Steinmetz M, Berkelman T, Kado CI.** 1985. Positive selection procedure for entrapment of insertion sequence elements in Gram-negative bacteria. *J Bacteriol* **164:**918–921.

96. **Okumura K, Chlumsky L, Baldwin TO, Kado CI.** 1992. Enhanced stable expression of a *Vibrio* luciferase under the control of the Ω-3 translational enhancer in transgenic plants. *World J Microbiol Biotechnol* **8:**638–644.

97. **Selbitschka W, Niemann S, Pühler A.** 1993. Construction of gene replacement vectors for Gram$^-$ bacteria using a genetically modified *sacRB* gene as a positive selection marker. *Appl Microbiol Biotechnol* **38:**615–618.

98. **Kamoun S, Tola E, Kamdar H, Kado CI.** 1992. Rapid generation of directed and unmarked deletions in *Xanthomonas*. *Mol Microbiol* **6:**809–816.

99. **Kumar G.** 1992. Two *cat* expression vectors for cloning and generation of 3′- and 5′-deletion mutants. *Gene* **110:**101–103.

100. **Crouzet J, Lévy-Schil S, Cauchois L, Cameron B.** 1992. Construction of a broad-host-range non-mobilizable stable vector carrying RP4 par-region. *Gene* **110:**105–108.

101. **Quandt J, Hynes MF.** 1993. Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria. *Gene* **127:**15–21.

102. **Kurata H, Furusaki S, Kado CI.** 1998. Light-enhanced target gene expression in tobacco BY-2 by the combination of overexpressed phytochrome and *rbcS3A* promoter. *Biotechnol Lett* **20:**463–468.

103. **Godiska R, Dhodda V, Gilbert V, Ravin N, Mead D.** 2007. Proceedings of the International Symposium on Plasmid Biology. *Plasmid* **57:**182–243.

104. **Prentki P, Krisch HM.** 1984. *In vitro* insertional mutagenesis with a selectable DNA fragment. *Gene* **29:** 303–313.

Plasmid Replication Systems and Their Control

II

Plasmids—Biology and Impact in Biotechnology and Discovery
Edited by Marcelo E. Tolmasky and Juan C. Alonso

doi:10.1128/microbiolspec.PLAS-0026-2014

Igor Konieczny,[1] Katarzyna Bury,[1] Aleksandra Wawrzycka,[1]
and Katarzyna Wegrzyn[1]

Iteron Plasmids

2

INTRODUCTION

Iteron plasmids are extrachromosomal genetic elements that can be found in all Gram-negative bacteria. Despite the fact that these plasmids bring antibiotic resistance to host bacterium, they can also bring other features, for example, genes for degradation of specific compounds or toxin production. Iteron plasmids possess characteristic directed repeats located within the origin of replication initiation that are called iterons. These plasmids became model systems for investigation of the molecular mechanisms for DNA replication initiation and for the analysis of mechanisms of control of plasmid copy number in bacterial cells. This research has provided our basic understanding of plasmid biology and the relationship between plasmid DNA and host cells. The control mechanisms utilized by iteron plasmids are based on the nucleoprotein complexes formed by the plasmid-encoded replication initiation protein (Rep). The Rep proteins interact with iterons, which initiates the process of plasmid DNA synthesis, but Rep proteins are also able to form complexes with iterons, which inhibits the replication initiation process. This inhibition is called "handcuffing." Also, Rep protein can interact with inverted repeated sequences, causing transcriptional auto-repression. Finally, various chaperone protein systems and proteases affect the Rep activity and, therefore, overall plasmid DNA metabolism.

STRUCTURE OF THE ORIGIN OF REPLICATION INITIATION

The origin region is one of the most important sequences within plasmid DNA; it ensures plasmid autonomous replication, independent of replication of the bacterial chromosome. As in other replicons, plasmid origins consist of characteristic motifs recognized by replication initiation proteins. In iteron-containing plasmids (Fig. 1), iterons that are directly repeated sequences play a crucial role during DNA replication initiation and are critical for plasmid copy number control (see also text below). They are quite short sequences, whose lengths vary from 17 bp in plasmid RK2 (1), 19 bp in plasmids F (2) and P1 (3), to 22 bp in R6K (4), pPS10 (5), and plasmids from the IncQ incompatibility group (6). Sometimes, such as in plasmid pXV2 from the IncW incompatibility group, direct repeats within the origin can vary in length. In pXV2 there are two 18-bp and two 19-bp repeats (7). The iteron number and spacing between iterons also can differ among iteron-containing plasmids. From the best-characterized plasmids the smallest iterons were identified in plasmid pSC101 (8), in which there are three iterons. In plasmids pPS10 (5) and F (2) there are four iterons; in RK2, 5 (1); and up to seven have been identified in R6K (4). In plasmids from the IncQ group there are three or four identical direct repeats, but sometimes

[1]Department of Molecular and Cellular Biology, Intercollegiate Faculty of Biotechnology of University of Gdansk and Medical University of Gdansk, Gdansk, Poland.

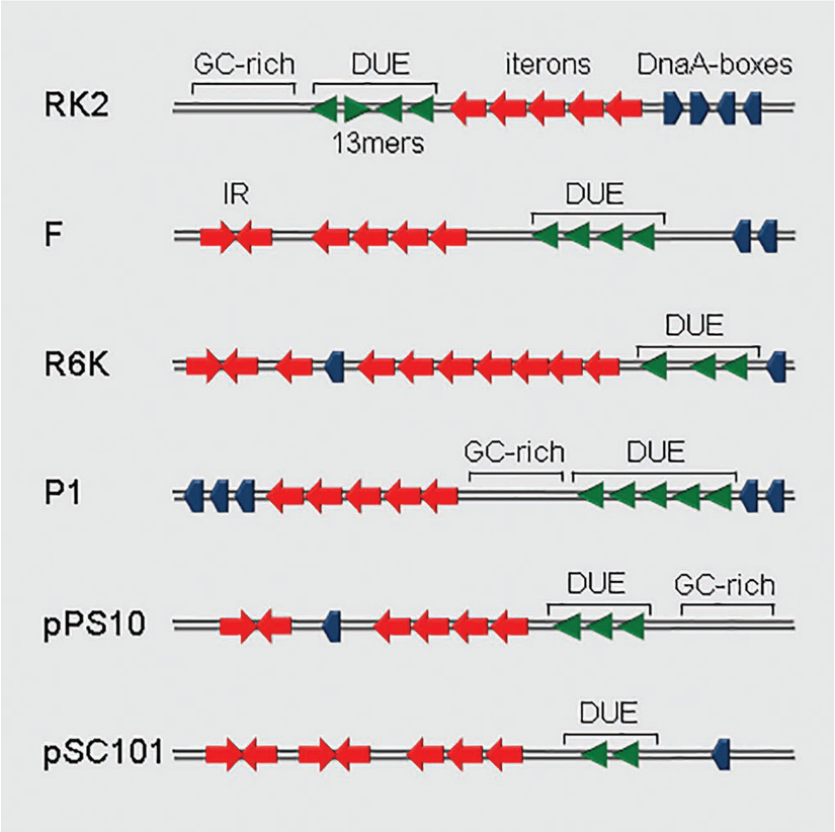

Figure 1 Scheme of the iteron-containing plasmid origin structure. The direct repeats—iterons—and inverted repeats (IR) are depicted as red arrows. The DUE region of each origin is marked, and repeated sequences within the region are depicted as green triangles. DnaA-box sequences are marked in blue. The region rich in guanidine and cytidine residues (GC-rich) is marked within the origins, if identified. The origins are not drawn to scale.
doi:10.1128/microbiolspec.PLAS-0026-2014.f1

the functional origin contains more iterons that are partly deleted, contain point mutations, or are incorrectly spaced (6). Plasmid R478 from the IncHI2 incompatibility group even contains several iterons that differ in length (eight 18-bp and nine 76-bp iterons) and are separated by a sequence of *rep* genes (9). Iterons are recognized by a plasmid-encoded Rep protein, and they are bound by a Rep monomeric form (10–13) in a cooperative manner (14, 15). Mutations within an iteron sequence can abolish the binding of Rep protein and, in consequence, plasmid replication. This was shown for plasmid R6K, in which changes in a sequence of iterons made impossible the binding of the π protein *in vitro* and replication of plasmids with mutated origins *in vivo* (16). Similarly, mutations within an iteron sequence in the origin of plasmid P1 reduced or completely prevented origin activity (17). Negative effects on plasmid replication are also exerted by changes in spacers between iterons. The importance of sequences adjacent to iterons was shown for plasmids P1 (17), RK2 (14), and pSC101 (18). Also, disturbances in the position of iterons in relation to other motifs present in the origin region, especially changes in proper helical phasing, have a negative influence on plasmid replication activity (19).

The binding of the plasmid initiator to double-stranded DNA (dsDNA) containing iterons results in local destabilization of the DNA duplex. Plasmid Rep protein is very often accompanied in its action by host initiator DnaA protein. DnaA protein binds a specific motif called DnaA-box, also localized within the plasmid origin. DnaA-boxes are 9-bp-long sequences with consensus sequences that are varied depending on the host bacteria (20). DnaA-boxes can be localized upstream from iterons (e.g., plasmids RK2 and pPS10), downstream from the region rich in adenine and thymine residues (AT-rich), where local destabilization of the duplex occurs (e.g., plasmids F and pSC101), or in

both these positions (e.g., plasmids R6K and P1) (21). In exceptional situations, such as in plasmid pXV2, DnaA-box overlaps with the first iteron (7). In some plasmid origins there is just one DnaA-box (e.g., plasmids pSC101, pPS10, and pXV2), and in others there are two (e.g., oriγ of plasmid R6K and plasmid F), four (e.g., plasmid RK2), or even five (e.g., plasmid P1) such motifs. Their length and sequence usually correspond to the consensus sequence of DnaA-boxes present in the origin of the *Escherichia coli* chromosome (*oriC*). If there are some deviations from consensus, they usually do not exceed point mutations. Examples include DnaA-boxes from plasmids P1 and RK2, which contain one or two mismatches. The position of DnaA-boxes is as important as their sequence. Insertions of more or less than a helical turn between DnaA-boxes and iterons within the plasmid RK2 origin resulted in inactivation of the origin's replication activity (19). The binding of DnaA protein to DnaA-boxes in the origin of broad-host-range iteron-containing plasmids can vary in different host bacteria. For instance, in the plasmid RK2 origin, DnaA-boxes 3 and 4 should be present when replication takes place in *E. coli* and *Pseudomonas putida* cells. However, they can be missed during plasmid replication in *Pseudomonas aeruginosa* (22). For *E. coli* chromosome *oriC* it was shown that beside DnaA-boxes, DnaA protein bound with ATP can interact with ATP-DnaA-boxes localized within AT-rich repeats (23). However, in iteron-containing plasmid origin regions, motifs for ATP-DnaA binding, similar to those observed in *oriC*, have not been identified to date.

The third motif, in addition to iterons and DnaA-boxes, that can be distinguished within the iteron-containing plasmid origin is the AT-rich region. This is the sequence, usually located near iterons, where local destabilization of the double-stranded helix occurs during the process of replication initiation. This region is therefore considered a DNA unwinding element (DUE) where single-stranded DNA (ssDNA) is created. Although the thermodynamic stability of the AT-rich region can differ in different origins, usually it has much lower free energy (ΔG) than the overall profile of adjacent sequences (21). In the AT-rich region, it is possible to discern short repeated sequences, usually oriented directly. The exception can be the origin of plasmid RK2, where one of the repeated sequences is inverted in relation to the other ones (24). Repeated sequences within the AT-rich region are located tandemly one after the other (e.g., origin of plasmids RK2 [24] and pSC101 [25]), or they are separated with spacers of different length (e.g., 7-, 1-, and 6-bp spacers between AT-rich repeats of plasmid F [25] and 29- and 9-bp spacers in oriγ of plasmid R6K [26]). The length of those repeated sequences can preserve 13 nucleotides (13-mers), as are present in the AT-rich region of *E. coli oriC* (e.g., plasmid RK2 [1] and pSC101 [25]). But more often they are shorter, such as in plasmids R6K (10 nucleotides [26, 27]), F (8 nucleotides [25]), and P1 (7 nucleotides [28, 29]). Also, the number of repeats can be different, and there can be two repeats in the AT-rich region of plasmid pSC101, four in plasmids RK2 and F, and up to five in plasmid P1 (21).

Although the consensus sequences for AT-rich repeats in different origins are difficult to identify, the consensus can be established for particular plasmid origins. The presence of all repeats within the AT-rich region, as well as their sequence, is very important for the proper replication activity of the origin. Even point mutations within these sequences can completely abolish plasmid replication (17, 30, 31). Also, substitution of one AT-rich repeat in a plasmid origin into a repeat from a bacterial chromosome origin results in a lack of replication activity *in vitro* and a decrease of activity *in vivo* (30). Although the presence and sequence of AT-rich repeats is critical for plasmid replication, the exact role of these motifs is still ambiguous.

The presence of binding sites for replication initiation proteins, iterons, and DnaA-boxes, as well as the region where duplex opening occurs, is very important for the replication initiation process. However, these motifs are not the only ones that can be distinguished within the origin of iteron-containing plasmids. In some plasmid origins the binding site for integration host factor (IHF) can be identified. Such a situation occurs, for instance, in the plasmid P1 origin, where the IHF binding site is located downstream from the cluster of three DnaA-boxes (32). The binding of IHF protein results in the bending of the DNA molecule; however, not only the bend but also its proper phasing for the downstream DNA is required for the activity of the origin (32). Insertions of less than a helical turn between IHF binding sites and DnaA-boxes in the P1 origin had a negative effect on origin activity. The IHF binding sites were also present in plasmids pSC101 (33) and R6K (34). This motif was identified as well in the plasmid RK2 origin, but the IHF deficiency in *E. coli* seemed not to alter plasmid replication efficiency or plasmid copy number control (35).

Other motifs that can be identified within some plasmid origins but are not directly involved during the replication process are sites, GATC motifs recognized by Dam methylotransferase. They are usually overlapped AT-rich repeated sequences (e.g., in plasmids P1 [36] and pSC101 [31]) or are located adjacent to

these repeats (e.g., in plasmid P1 [36]). The methylated GATC sequence becomes hemimethylated during replication and in this form is recognized by the SeqA protein (37, 38), which sequestrates newly synthesized DNA (39). SeqA negatively regulates DNA replication by blocking the GATC sites and preventing replication proteins from binding. Apart from the GATC motif, a region rich in guanidine and cytidine residues (GC-rich) can be identified in some plasmids' origins (e.g., plasmids RK2, P1, pPS10, and IncQ). Its exact role is unknown, and in plasmid RK2 it can be deleted without any effects on origin activity (19). In plasmid P1, in which a GT-rich sequence plays the role of a spacer between iterons and AT-rich repeats, the sequence of this region can vary considerably, but its length must be preserved (36).

In a few plasmids identification of motifs other than those described here was reported. For example, in plasmids F and R1 the binding site for the IciA protein was detected (25). The IciA protein, which binds the site located in the AT-rich region of plasmid origins, probably, like in *E. coli oriC* (40), inhibits the unwinding process at the AT-rich region. In the origin of plasmid R6K, binding sites for other regulatory proteins, Fis (factor for inversion stimulation) were found (41). It was shown that plasmid replication depends on the Fis protein when the gene for the copy-up mutant of the π protein and the penicillin resistance gene were present on plasmid DNA (41).

It could be concluded that for the proper activity of the origin of iteron-containing plasmids, not only the presence and the sequence of essential motifs, such iterons, DnaA-boxes, and AT-rich repeats, is important. The appropriate location of these motifs in relation to each other also has a great impact on replication activity. In particular, changes in proper helical phasing have a negative influence on plasmid replication.

Rep PROTEIN STRUCTURE

Although many plasmid Rep proteins have been identified, the crystallographic data are limited to a few replicons. This is due to a high instability of the Rep proteins, so understanding the initiators' role in the structural context is a challenge. Plasmid replication initiators such as the RepA initiator of plasmid pPS10, RepE of plasmid F, and the π protein of plasmid R6K are best characterized in terms of structure. The RepA initiator of pPS10 was the first Rep protein whose structure was predicted to consist of two winged helix (WH) domains (42). These findings have been confirmed by the crystal structure of the monomer of a homologous RepE initiator of plasmid F, bound to iteron DNA (43). The other crystal structure of a plasmid Rep protein was determined for the monomeric form of the π initiator protein of plasmid R6K as a complex with a single copy of its cognate DNA-binding site (iteron) (44). The crystal structures of both RepE and π proteins are depicted in Fig. 2. Although the crystal structures of RepE and π proteins shed new light on the Rep monomers' interaction with DNA, the molecular nature of Rep activation remained unknown until the crystal structure of the dimeric N-terminal domain of the plasmid pPS10 initiator (dRepA) was resolved (45). Nonetheless, the crystallographic data obtained for plasmid Rep proteins are limited to the WH domain description. Rep proteins are composed of two WH domains—N-terminal WH1 and C-terminal WH2—that are responsible for interaction with DNA (42) (Fig. 2). The WH2 domain contains a putative helix-turn-helix motif, which is the main determinant of Rep binding to both the iteron sequences and the inverted repeats (partially homologous to the iteron sequence), which was shown for the RepE initiation protein of the mini-F plasmid and RepA of plasmid pPS10 (46, 47). A formation of nucleoprotein complex by Rep protein results in the bending of the DNA molecule. Iteron interaction with the WH1 and WH2 domains of the Rep monomer, or interaction of inverted repeats with both WH2 domains of Rep dimer, induce DNA bending (42, 48). In Rep monomers, the WH2 domain binds to the 3′-half of the iteron, while the WH1 domain changes structure and contacts the 5′-iteron end, through both the phosphodiester backbone and the minor grove (42).

In contrast to initiation proteins of replicons F, R6K, and pPS10, the crystal structures of the TrfA protein of RK2 as well as P1 RepA have not been determined. The structure prediction using fold-recognition homology modeling was carried out in both cases. The N-terminal part of TrfA does not show a unique three-dimensional structure with the absence of stabilizing factors; it seems to be disordered in solution as opposed to the C-terminal part of the protein, which is expected as two copies of WH domains. Helices of both WH structures interact with major grooves of the DNA phosphate backbone (49). A series of mutations located within the WH1WH2 domains have been found to affect the TrfA-DNA interaction (50, 51). The structure predicted for P1 RepA, similar to TrfA, contains WH domains. By means of fold-recognition programs, it was shown that despite the lack of sequence similarity, RepA shares structural homology with plasmid F RepE. The model predicted that RepA binds one half of the binding site through interactions with the N-terminal

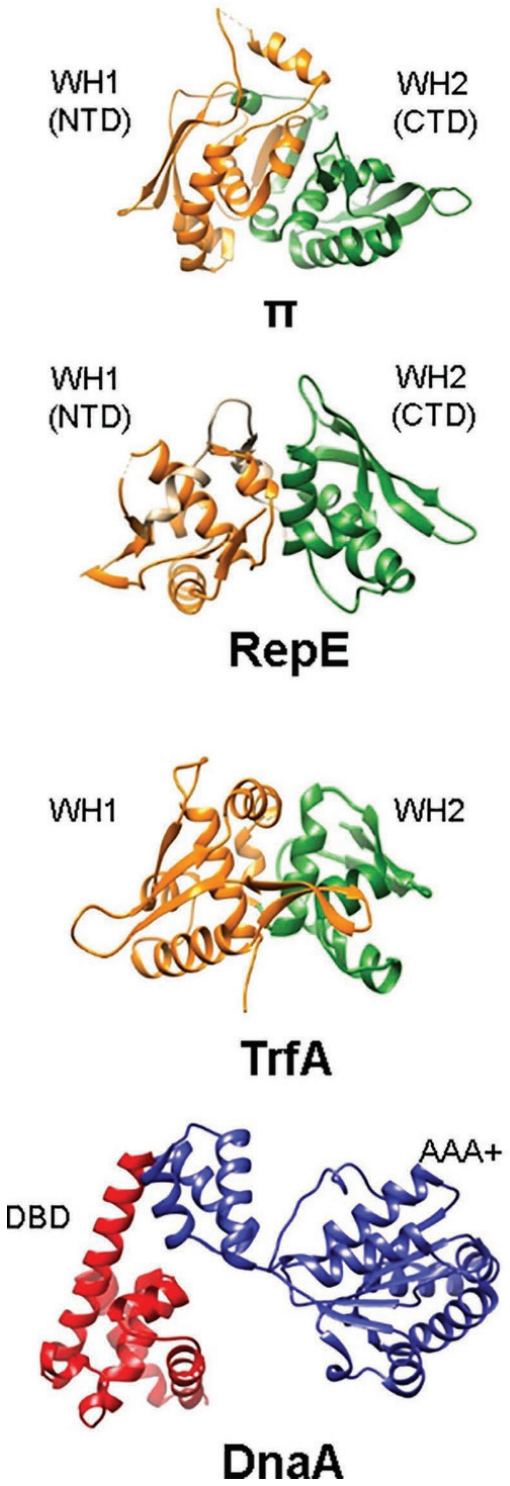

Figure 2 Structure of replication initiators. DnaA of *A. aeolicus*, RepE54 from *E. coli* mini-F plasmid, π from R6K, and the C-terminal part of the TrfA protein (190-382 aa) of plasmid RK2 are depicted. Structure of the DnaA, RepE54, and π are derived from crystallographic data (PDB entry 1L8Q, 1REP, and 2NRA, respectively). The TrfA model was developed based on homology modeling. The AAA+ domain is colored in blue, the DNA binding domain (DBD) is shown in red, and Winged-Helix domains (WH1 and WH2) are colored in yellow and green, respectively. References and detailed information for crystallographic data of the DnaA, RepE54, π, and TrfA model are given in the text.
doi:10.1128/microbiolspec.PLAS-0026-2014.f2

DNA binding domain (WH1) and the second half through interactions with the C-terminal domain (WH2) (52). Interestingly, the residues involved in Rep-DNA interactions located outside the WH domains have been determined with the use of RK2 initiator TrfA mutants (50, 51). These results assume the existence of an additional DNA binding motif, apart from WH1WH2 domains.

Like in plasmid-encoded Rep proteins, WH domains responsible for DNA binding were found in *Archaea* and *Eukaryota* initiators. However, the AAA+ domain (ATPases associated with various cellular activities) commonly present in *Archaea* and *Eukaryota* initiation proteins was not found in plasmid Reps (53) (Fig. 2). Thus, with regard to the DNA binding mechanism, the plasmid Rep proteins are similar to eukaryotic replication initiators. The results of biochemical and spectroscopic experiments revealed functional similarities between pPS10 RepA and archaeal/eukaryal initiators (53). The crystal structure determined for the archaeal initiator Cdc6 confirmed these findings (54). Interestingly, it was reported that similar to the mammalian proteins PrP and α-synuclein, the WH1 domain of the pPS10 RepA can assemble into amyloid fibers upon binding to DNA *in vitro* and in *E. coli* cells (55–58). It opens a direct means to untangle the general pathway (s) for protein amyloidosis in a host with reduced genome and proteome (59).

Plasmid Rep proteins exist in cells mostly as dimers (12, 60). The dissociation of dimers by the action of chaperones or interaction with iteron-containing DNA (see also text below) results in conformational changes in the Rep structure (61). A compact arrangement of the two WH domains, competent for binding to the inversely repeated sequences, becomes a more elongated form, which is suited for iteron binding (42). These conformational changes consist of a significant increase of the overall β-sheet at the expense of the α-helical one (61). The situation is different for the Rep dimers that interact with inversely repeated sequences. Binding of Rep dimers to the inverted repeats does not result in dissociation to monomeric forms or change in the dimers' conformation (61). Although only the monomeric

form of Rep proteins is replication-active, dimers of Rep can bind to an inversely repeated sequence localized close to the promoter region of the *rep* gene, which results in transcription auto-repression (see text below). This was shown for the RepA initiator of the pSC101 (62) F RepE initiation protein (43) and the π initiator of the plasmid R6K (63). In the dimeric form of Rep, the WH2 domain binds to inverted repeats via the major groove, whereas the WH1 domain acts as the dimerization interface (61). Dimerization of pPS10 RepA is determined by interactions between β-sheets of the monomers that are originated due to a conformational change in the protein that involves a leucine zipper (LZ)-like motif (42). The LZ-like motif, present in several eukaryotic regulatory proteins (64), has also been found in the WH1 domain of RepA of pSC101(65), RepE of F (43), and π of R6K (66). The dimerization interface is also localized in the WH1 of the model predicted for plasmid RK2 TrfA replication initiation protein (Fig. 2). Similar to the proposal for pPS10 RepA (42), this interface is located on an extended antiparallel β-sheet forming two hairpins (49).

Besides the indirect effect of the LZ motif in Rep protein dimerization, the LZ-like motif was characterized as responsible for Rep interaction with host replication factors. The mutations, described either in pPS10 or in the *E. coli* chromosome, have revealed evidence of a WH1-mediated interaction between RepA and the chromosomal initiator DnaA (67). Nonetheless, protein-protein interaction of Reps are not restricted to the LZ-like region. The best evidence for this statement is a TrfA initiator of plasmid RK2 existing in two replicationally active forms of different molecular mass. The smaller, 33-kDa protein, TrfA-33, is the result of an independent in-frame translational start in the open reading frame used for the larger, 44-kDa protein, TrfA-44 (68–70). The mutation at the N-terminal end of the *trfA* gene (resulting in the availability of the TrfA-33 version only) changes the host range of plasmid RK2, but the binding of DNA remains unaffected. These results demonstrate that the N-terminal end of TrfA is involved in interaction with host replication factors (71). With the use of the evolution experiment, IncP1 plasmids were shown to specialize to a novel host due to the single mutations reported at the N-terminal region of replication initiation protein TrfA (72, 73). In *P. aeruginosa* the TrfA-44 residues between 20 and 30 are responsible for DnaB recruiting (71), and in *E. coli* TrfA-33 interacts *in vitro* with DnaB helicase (74). It also acts with the *E. coli* Hda regulator, which inactivates DnaA and this way prevents overinitiation of RK2 (75). In addition, the specific motif characteristic of proteins interacting with the β clamp of *E. coli* DNA polymerase III was reported in TrfA and TrfA/RepA orthologues from plasmids related to RK2 and pMLb (76), but the relevance of this interaction needs to be elucidated.

The replication of iteron-containing plasmids requires the plasmid-encoded replication initiator, but the host-encoded initiation protein is also involved. The chromosomal initiator, *E. coli* DnaA, is composed of four functional domains (77–79). Crystallographic data obtained for the DnaA conserved core domains III/IV of the thermophilic bacterium *Aquifex aeolicus* revealed that, in contrast to plasmid initiators, this protein is composed of the AAA+ and DBD (DNA binding domain) domains (79) (Fig 2). These domains are involved in DnaA oligomerization and DNA binding/remodeling functions, which are the critical aspects of origin processing. It is crucial for the interaction with ssDNA DUE at chromosomal replication origins and formation of filament structure (80–82). Since plasmid Rep does not possess an AAA+ domain is responsible for nucleotide binding, it could be considered that WH domains, responsible for the binding of iterons within the dsDNA origin, can also bind ssDNA arising after dsDNA melting.

MECHANISM OF ITERON PLASMID DNA REPLICATION INITIATION

Origin Recognition

Models presenting steps of DNA replication initiation of iteron-containing plasmid and bacterial chromosomes are presented in Fig. 3. The first step of replication initiation at the plasmid origin is the formation of an initial complex facilitated by the specific interaction of Rep proteins with iterons. It has been demonstrated that replication initiation of iteron plasmids usually requires cooperative interaction of Rep monomers with iterons. pPS10 RepA as well as RK2 TrfA initiators cooperatively bind iterons at the plasmid replication origins (14, 15). Although the pPS10 RepA dimers and monomers both interact with iterons, only monomers initiate DNA replication. It is noteworthy that the existence of an early transient complex between a dimeric pPS10 RepA and an iteron half has been reported, and based on this, a model for iteron-induced dimeric pPS10 RepA dissociation and conformational activation has been proposed (61). Also, the TrfA protein functionally interacts with plasmid RK2 iterons as a monomer (12). Similar to pPS10 and RK2, the origin of the narrow host range plasmid P1 is recognized by the

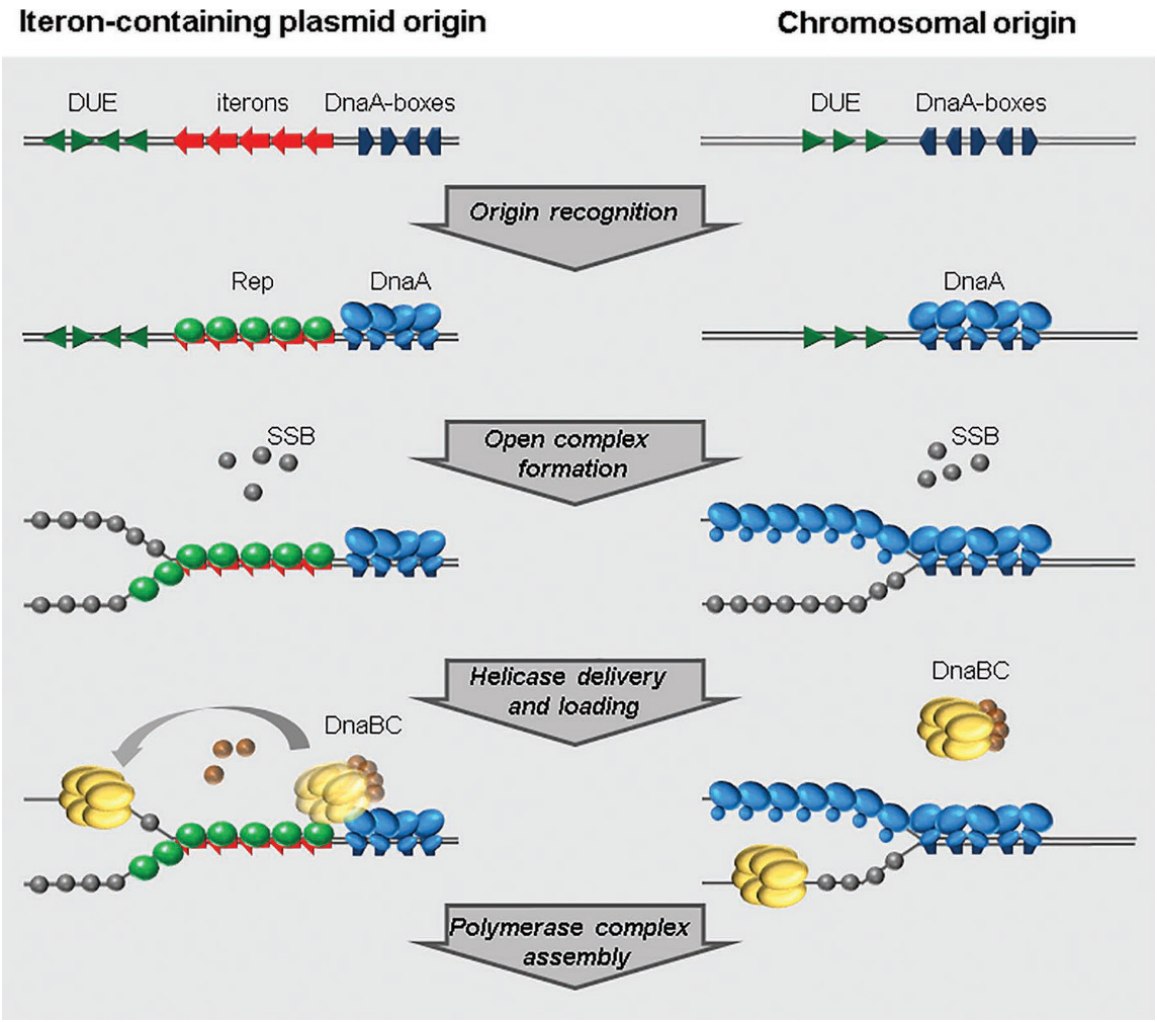

Figure 3 Model of replication initiation: comparison of the processes occurring on the iteron-containing plasmid origin with the replication initiation of bacterial chromosomes. The iteron-containing plasmid origin is recognized by the plasmid-encoded initiator (Rep), which binds cooperatively to the iterons. The interaction of Rep with iterons results in the formation of an open complex and destabilization of the DNA unwinding element (DUE), which creates ssDNA. In RK2, pPS10, F, R6K, P1, and pSC101 the formation of the open complex requires cooperation of the plasmid Rep and host DnaA proteins, while at the chromosomal origin the DnaA protein is sufficient for this process. During the chromosomal origin opening DnaA forms filament on the ssDNA. Helicase delivery and loading requires interaction with the replication initiators; in addition, in *E. coli* the DnaB helicase delivery at the chromosomal *oriC,* as well as at the plasmid RK2 *oriV,* requires the DnaC accessory protein. During the RK2 replication initiation in *E. coli* the host-encoded DnaBC helicase complex is delivered to the DnaA-box sequence through interaction with DnaA, and subsequently the plasmid initiator TrfA translocates the helicase to the opened plasmid origin. The interactions between *E. coli* DnaB and the R6K π protein, F RepE, and pSC101 RepA have also been established as essential for helicase complex formation at the plasmids' origins. The helicase unwinds the DNA double helix, and after a short RNA fragment is synthesized by a primase, a polymerase complex is assembled. Single-stranded DNA binding protein (SSB) is required for replication initiation of both chromosomal and iteron-containing plasmid DNA. The HU/IHF proteins' contribution in DNA replication initiation was omitted in the scheme. For a detailed description see the text.
doi:10.1128/microbiolspec.PLAS-0026-2014.f3

monomer of the P1 initiation protein RepA (83, 84). The interaction between the Rep protein and iterons has also been shown for RepE of plasmid F (85), RepA of plasmid pSC101 (86), and the π initiation protein of *E. coli* plasmid R6K (87). The narrow host range plasmid R6K contains three origins of replication, α, β, and γ, but only two elements, the *ori* and *pir* gene product π proteins, are required for a minimal replicon. The binding of seven iterons by the π initiator has been demonstrated as required for proper *ori* activity (88–91). The π initiator efficiently binds to *ori* iterons but not to the *ori* or to the *ori* iterons (92).

Origin Opening

It was determined that Rep plasmid interaction with iterons generates a localized strand destabilization of DUE, leading to an open complex formation at the origin of plasmid replication. Although the involvement of the plasmid initiator is essential, the host-encoded DnaA and histone-like proteins are also required for plasmid origin opening. It was demonstrated for pPS10 that mutations within the DnaA-box sequence affect the replication *in vivo* (5). DnaA is mainly needed for the enhancement or stabilization of the Rep plasmid-induced open complex formation and histone-like protein (HU and/or IHF) interaction with the DNA-enhanced DNA-bending process. It was determined with KMnO4 assay that TrfA interaction with iterons generates a localized strand destabilization, and *E. coli* DnaA protein enhanced the TrfA-induced open complex (24). It was shown that this reaction occurs only in the presence of the *E. coli* HU protein (24). Similar to RK2 initiator TrfA, the binding of the RepE initiator of plasmid F to iterons induces a localized opening in the origin region, with the assistance of HU (93). The addition of DnaA increases the opening of the F plasmid origin (93) and is also required for the pSC101 origin (94) and R6K *ori* (88, 89). The open complex formation by pSC101 RepA monomers in cooperation with host DnaA also requires the presence of the IHF protein (33, 95). The open complex at the R6K *ori* is formed as a result of cooperative π monomers binding to the iterons and host DnaA interaction with its cognate binding sites (15). KMnO4 footprinting has shown that, in contrast to the RK2 initiator TrfA and F RepE, the P1 RepA alone is not sufficient for *oriR* opening, but in the presence of DnaA, the addition of RepA increased the KMnO4 reactivity of the origin (96). The replication initiation of plasmid RK2 might occur in a DnaA-dependent or DnaA-independent way, depending on the host bacterium. In *E. coli* RK2 efficiently replicates and is maintained in the presence of TrfA and a host DnaA protein, while in *Pseudomonas* the longer form (44 kDa) of the replication initiator is required and DnaA is indispensable (97, 98). In *Caulobacter crescentus* both DnaA-dependent and DnaA-independent models of RK2 plasmid replication initiation are possible (99). Interestingly, the structure of DnaA protein itself might influence the host range of plasmids. Narrow-host-range plasmid pPS10 usually replicates only in the phytopathogen *Pseudomonas savastanoi* cells, due to the ability to bind DnaA-box in the pPS10 origin only by DnaA protein from this bacterium. It has been demonstrated that both the mutation in the LZ motif of pPS10 RepA and mutations in the sequence of *E. coli* DnaA promote the efficient establishment of plasmid pPS10 in the *E. coli* host (67, 100). These results suggest that mutations in plasmid and bacterial initiators that result in expanding the host range of the plasmid probably favor efficient and functional interactions between those proteins. Although the chromosomal initiator, DnaA protein, alone is insufficient for the efficient formation of an open complex at the origin of plasmids F, RK2, pSC101, and R6K (13, 24, 90, 93, 101), it has been shown to be both sufficient and indispensable in opening the AT-rich region at the origin of the bacterial chromosome (see Fig. 3). DnaA interaction with DnaA-box sequences localized within the origin of chromosomal replication (*oriC*) results in destabilization of the DUE, leading to open complex formation. The histone-like proteins HU and IHF stimulate the assembly of the open complex at *oriC* (102–104). This nucleoprotein structure formation requires ATP due to *E. coli* DnaA ATP-dependent conformational changes that promote the formation of the DnaA filament on ssDNA of DUE that is essential for the opening of the replication origin (81, 82, 105). The formation of an open complex at the plasmid origin, in contrast to *E. coli* chromosomal replication, is an ATP-independent process (24, 90, 93, 96, 106, 107), but the presence of ATP or its nonhydrolyzable analogue (ATPγS) promotes the extension of the open region (24). It is not known if plasmid Rep proteins can interact with the ssDNA and form filament structures to promote origin opening, like the DnaA replication initiator does.

Helicase Delivery and Loading

The origin opening generates ssDNA, which is a key element for replication complex assembly at the replication origin. The first step in the assembly of the replication complex is delivering helicase at the replication origin and loading it on ssDNA. While plasmids belonging to the IncP incompatibility group extensively

use the replication proteins from the host cell for their own DNA synthesis, they utilize different host-specific mechanisms for helicase delivery and loading (71, 108, 109). Both *in vivo* (97, 98) and *in vitro* (108, 109) analysis with the use of purified proteins from *E. coli* and *Pseudomonas* sp. revealed different host-dependent requirements for RK2 replication initiation. In *E. coli* the DnaB helicase complex with DnaC is initially recruited by DnaA protein interaction (110). The DnaA bound at DnaA-boxes located at the plasmid origin recruits host helicase (111). Then, as a result of translocation into the AT-rich region of the plasmid origin and interaction with the 33-kDa version of the plasmid replication initiator, the helicase is activated for the unwinding of the plasmid dsDNA template. The mechanism of helicase recruitment and loading during the RK2 plasmid replication in *P. aeruginosa* is DnaA-independent and relies on the 44-kDa TrfA protein, while in *P. putida* cells two variants of TrfA protein can be utilized (108, 109). The helicase complex formation during RK2 replication in *C. crescentus* cells might proceed through two different modes: DnaA-independent employing TrfA-44 and DnaA-dependent relying on the shorter version of the replication initiator (99). *In vitro* activity of *C. crescentus* DnaB helicase on the RK2 DNA template was observed in the presence of TrfA-44, and *C. crescentus* DnaA was not required for this process. *In vivo* the mini-RK2 plasmid encoding only TrfA-33 was as stably maintained as those encoding TrfA-44 or both. In contrast, TrfA-33 in cooperation with *C. crescentus* DnaA *in vitro* was unable to activate *C. crescentus* DnaB. The homologue of the *E. coli* DnaC protein needed for proper helicase loading into the open complex might be required for *C. crescentus* DnaB helicase activation. To date, no data about this kind of protein either in *Pseudomonas* or in *Caulobacter* cells have been reported, and its identification requires further investigation (99).

Rep-Helicase Interaction

Similar to the RK2 plasmid initiator TrfA, the interactions between other iteron-containing plasmid Rep proteins and host-encoded helicases have also been reported. *E. coli* DnaB interacts with plasmid replication initiators as was shown for the R6K π protein (112) plasmid F RepE (113) and pSC101 RepA (114). These interactions have been established as essential for helicase complex formation at the mentioned plasmid origins. A DnaB mutant, which does not interact with pSC101 RepA, was unable to activate the replication initiation at the pSC101 origin. Nonetheless, this mutant was able to support *E. coli* chromosomal replication (114). The R6K π protein and pSC101 RepA have also been shown to form complexes with *E. coli* DnaA (90, 101). Similar to R6K and pSC101, the helicase complex formation at the origins of pPS10 and P1 replicons, in addition to the plasmid-encoded initiator, depends on host DnaA protein and requires other host-encoded factors such as DnaC and HU/IHF (67, 115, 116).

The lack of ability for stable complex formation between the plasmid Rep protein and a host helicase might be one of the reasons for plasmid host range restrictions as was shown for *E. coli* plasmid F. The helicase complex at the F origin composed of the replication proteins from the nonnative hosts (*P. aeruginosa* and *P. putida*) might be formed in the presence of F initiator RepE. However, the interactions between RepE and DnaB of *P. aeruginosa* and *P. putida* were unstable, contrary to RepE interaction with *E. coli* DnaB helicase (113).

Polymerase Complex Assembly

Synthesis of iteron-containing plasmid DNA depends on the initial activity of a plasmid replication initiator and utilization of host replication machinery. Because plasmids do not encode their own polymerases, the host bacterium polymerase is utilized for the plasmid DNA replication. The mechanism of the events leading to the formation of the polymerase complex at the plasmid origin of replication still needs to be elucidated. Even though the DNA replication of plasmids RK2 (111), R6K (117), and F (118) has been reconstituted *in vitro* with purified proteins, and specific requirements for this reaction have been identified, the molecular mechanism for the assembly of the polymerase complex at plasmid origins is still not known. The *in vitro* analysis showed that in addition to the plasmid Rep protein, the *E. coli* proteins DnaA, HU, DnaB helicase, DnaC, SSB, DnaG primase, DNA gyrase, and Pol III holoenzyme are required for plasmid DNA synthesis. Interestingly, the specific motif (QL[S/D]LF) determining interaction with the β clamp subunit of Pol III has been identified in plasmid Rep proteins (119), though the relevance of the interaction between the β clamp and Rep proteins has not been determined. The loading of the β clamp is a composite reaction involving clamp opening and then positioning around the DNA with the use of the γ-complex (reviewed in reference 120). β clamp interaction with primed DNA is the first of subsequent events leading to polymerase complex assembly at the chromosomal origin of replication (121). Although the direct involvement of a replication initiation protein in the process of polymerase recruitment has not been reported to date, the plasmid Rep

protein interaction with specific Pol III holoenzyme subunits might determine the mechanism for an efficient recruitment of host-encoded replication machinery to the plasmid origin.

CONTROL MECHANISMS OF REPLICATION IN ITERON-CONTAINING PLASMIDS

The iteron-containing plasmid replicons have evolved a number of strategies to ensure their hereditary stability and maintenance at the specific copy number. These plasmids occur in a low-copy number per bacterial cell, so their maintenance requires tight regulation of replication. The main elements involved in the regulation of these plasmid replications are iterons.

Control by Handcuffing

"Handcuffing" is a mechanism of replication inhibition observed in iteron-containing plasmids. The handcuff structure formation is based on the ability of the initiator protein to couple two *ori* regions located on separate plasmid molecules. The *ori* coupling occurs via binding of the Rep protein to iterons. This pairing of iterons is believed to cause steric hindrance to their function that prevents a new round of replication initiation (Fig. 4) (122) by inhibiting origin melting (123). It is considered that handcuffing is a major mechanism that controls the plasmid copy number.

There are three alternative models of Rep-mediated handcuffing. The first one assumes that the handcuff structures are created by the action of Rep dimers, which can bridge two DNA particles. This model was proposed for the replication protein of plasmid R6K (124, 125). Here, the major role of π dimers in the creation of R6K handcuff complexes was detected by electron microscopy (124) and ligation enhancement assays (66, 126, 127). Both of these techniques enable detection of handcuff structures in reaction to the dimeric form of the π protein. In the ligation assay, the monomeric variant of Rep was less efficient in forming ligated products (125). In contrast, the mutant of the π initiator, which binds iterons exclusively as a dimer (13), handcuffed DNA more efficiently than the wild type of the π protein. To summarize, the π dimers have a greater affinity to participate in handcuff structure creation than π monomers. The indirect evidence supporting this model is the fact of handcuffing being counteracted by molecular chaperones (DnaK-J/GrpE triad), which mediate the dissociation of dimers to monomers (123, 128).

The handcuff structure creation in the second model assumes the participation of Rep monomers in the creation of such structures by direct interactions between two arrays of Rep monomers bound to iterons in two plasmid molecules (56). This model is based on the fact that monomers of Rep initiators have a higher affinity for the iteron repeats than the dimeric forms (42, 124). Moreover, it has been reported for plasmid pPS10 that the dimeric Rep mutant is unable to create handcuff structures (56), and iterons of this plasmid play an active role in displacing the equilibrium between Rep dimers and monomers (61).

The third model of handcuff structure is a combination of the other two models. In this model, two monomers bound to the iterons of two separate plasmid molecules, are bridged by the dimer of the Rep protein. Such a model was proposed for handcuffing of plasmids RK2 (129) and F (123). The evidence for this model was obtained in a purified *in vitro* replication system (123). The handcuffing was found to be most proficient only when monomeric and dimeric forms of Rep protein were present simultaneously. Models involving participation of Rep protein dimers are also supported by the fact that handcuffing-defective mutants (Rep monomers of RK2 and R6K) were found to have abnormally high copy numbers (130). Therefore, it can be concluded that the handcuffing has a substantial role in iteron-mediated plasmid copy number control.

If the role of the handcuff is to block the origin and inhibit the replication, then there must be a mechanism that acts in an opposite way and "uncuffs" the coupled origin structures, which enables the reinitiation of plasmid replication. However, the mechanism of handcuff reversal is still unclear. There are results suggesting the participation of the chaperones in handcuff structure disruption (128), showing that the efficiency of handcuffing decreases in the presence of chaperones. Those results indicate that an increasing ratio of monomers over dimers is predominantly responsible for handcuffing reversal. It has also been discovered that the efficiency of handcuff structure creation increases with increasing Rep-bound iteron concentration and decreases when the reaction mixture is diluted. However, the dilution did not decrease Rep binding to the iterons (128).

Control by Auto-Repression

A high concentration of Rep protein initiator may result in more frequent, uncontrolled initiation replication events. To prevent this, the control mechanism that limits the amount of Rep initiator in the cell has to exist. Transcriptional auto-repression is a well-known mechanism for maintaining levels of gene

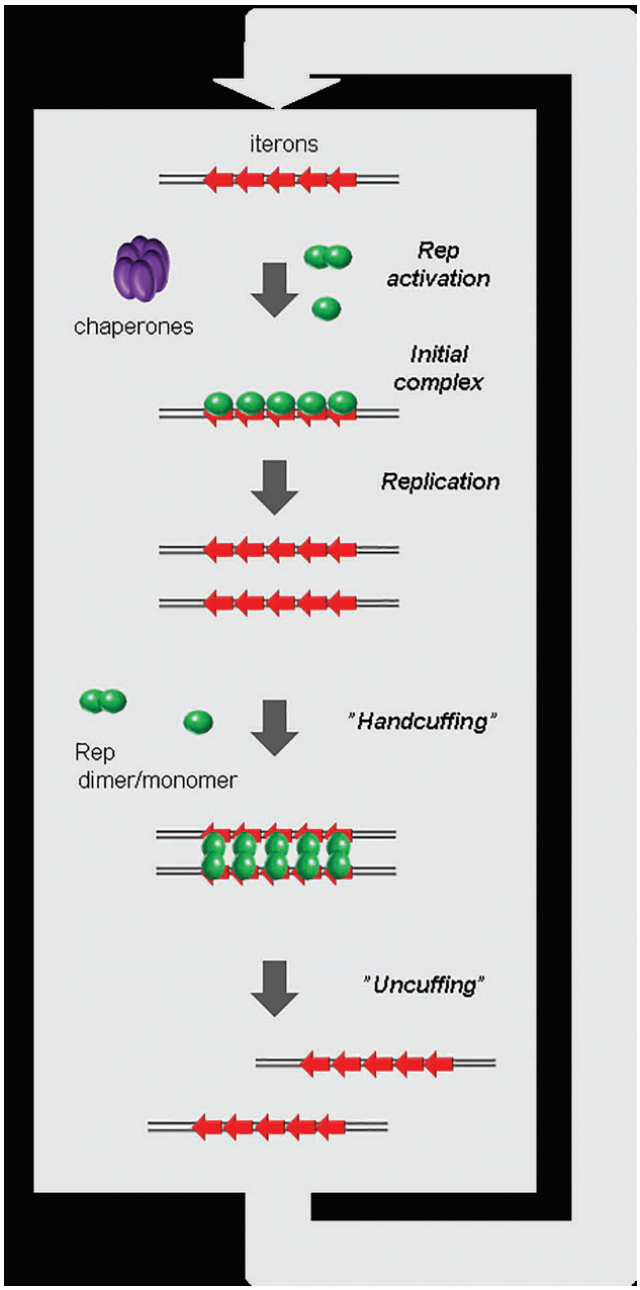

Figure 4 Regulation of iteron-containing plasmid replication initiation by the iterons. Rep protein activation occurs by the action of chaperones that convert the Rep dimer to the active monomeric form. Monomers bind to the iteron sequences and perform the initial complex that leads to replication of DNA. Rep protein may also act as a negative regulator of DNA replication by creating "handcuff" structures. Rep proteins couple origins of two separate plasmid particles in a process termed "handcuffing." In the literature suggestions of chaperone proteins' participation in the "uncuffing" process can be found, but the mechanism of the handcuff structures' reversal is still unclear. For details see the text.
doi:10.1128/microbiolspec.PLAS-0026-2014.f4

product within narrow limits (131, 132). In many plasmid systems (F, R6K, pPS10, and pSC101), autorepression is mediated by binding of the Rep dimer to inverted repeats located adjacent to the origin region (Fig. 5) (11, 46, 65, 133). A sequence of inverted repeats overlaps with the *rep* gene promoter. This kind of regulation mechanism inhibits transcription initiation starting from the *rep* gene promoter, and this effect is promoter-specific (63). The affinity of the Rep dimers is higher for inverted than direct repeats, so the Reps must have specific, dimeric conformation for binding to these sites (133). Symmetrical motifs in the Rep dimer recognize the symmetry of inverted repeats (134). The mechanism of auto-regulation appears to be one of steric hindrance. When the promoter site is occupied by Rep protein, the RNA polymerase cannot displace it from the binding site. However, it has been shown that the initiator proteins can displace RNA polymerase from the promoter, and the addition of the RNA polymerase before the Rep protein does not prevent binding of Rep protein to its binding site (63, 135, 136). This inhibition of RNA polymerase binding resembles typical repressor-polymerase competition and, in this model, the Rep dimer acts as a repressor. An explanation for this auto-regulation mechanism is a higher affinity of the initiator protein for DNA sequence than that of RNA polymerase for the same sequence (136).

Activation and Proteolysis of Rep

As mentioned above, Rep proteins exist in monomer-dimer equilibrium, but only the monomeric form of the proteins can bind specifically to the iterons (12). Saturation of iterons in the replication origin by Rep monomers allows replication initiation. To create such a complex, conformational activation of Rep proteins is required. Dissociation of the Rep dimers into monomers simultaneously changes the conformation of the proteins and makes them competent for the iteron binding. The dissociation may be spontaneous and could occur just by dilution to low/sub-micromolar concentration. This phenomenon has been found for P1 and pSC101 plasmids (62, 83, 137). However, those monomeric forms of Rep proteins require the chaperones for refolding into the active form and for DNA binding (62, 137). The conversion of a dimer to an active monomer can also be mediated by dissociation induced by interaction with iteron-containing DNA. It has been shown that micromolar amounts of DNA, which contain a single iteron, actively induce *in vitro* the dissociation dimers into both monomers and conformational changes (61, 138).

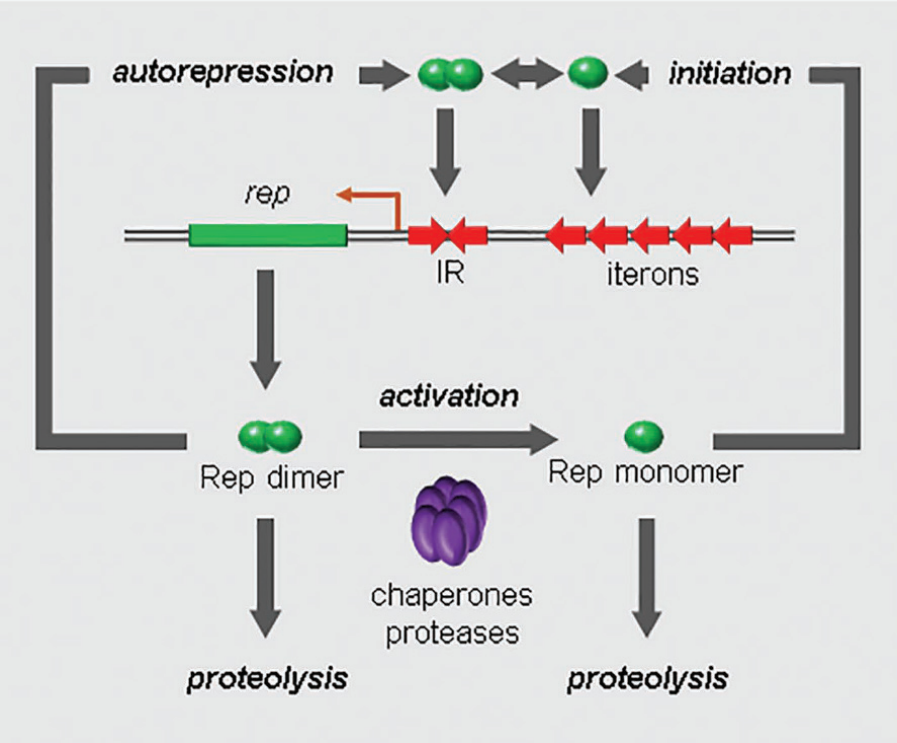

Figure 5 Regulation of iteron-containing plasmid replication initiation by the auto-repression mechanism. Binding of Rep dimers to inverted repeats inhibits the initiation of transcription starting from the *rep* gene promoter. This phenomenon is called auto-repression. An active, monomeric form of Rep protein arises as a result of the action of chaperones. It binds to the iteron sequences that lead to the initiation of DNA replication. Proteases are another factor that may influence the replication process. They limit the amount of both dimer and monomer forms of the Rep protein. For details see the text.
doi:10.1128/microbiolspec.PLAS-0026-2014.f5

The monomeric form of Rep may also arise by the action of molecular chaperons, which actively convert dimers to monomers (139–142). *In vitro* techniques demonstrated that both the ClpX chaperone (139) and the ClpB/DnaK/DnaJ/GrpE system (140) activate the plasmid RK2 replication initiation protein TrfA by converting inactive dimers to an active monomer form. It has been also shown that DnaK/DnaJ/GrpE heat shock proteins are required for the activation of Rep initiators of F, R6K, and P1 plasmids (10, 60, 83, 138, 142, 143). Monomerization of the P1 plasmid initiator may also occur by the action of the ClpA protein, which alone functions as a molecular chaperone (144, 145).

The proteases are other factors affecting iteron-containing plasmid metabolism. They may influence the replication process by proteolysis of the replication initiator. In *E. coli*, four cytosolic proteases have been identified to date: ClpXP, ClpAP, ClpYQ, and Lon (146). The proteases limit the half-life of Rep initiator proteins, which is important for replication initiation. It has been shown that initiator proteins of bacteriophages lambda and Mu and of plasmid RK2 are proteolyzed by *E. coli* ClpXP protease (49, 147, 148) and that ClpAP protease degrades the Rep initiator of plasmid P1 (144). Additionally, it has been described for the TrfA initiator of plasmid RK2 that DNA is a factor that stimulates TrfA proteolysis by ClpAP and Lon proteases (149). Moreover, the Lon protease degrades the TrfA protein only in the nucleoprotein complex, while ClpAP-dependent degradation of TrfA is substantially stimulated in the presence of iteron-containing plasmid DNA (149). This specific stimulation of proteolysis could be important in terms of understanding nucleoprotein complex stability. It may also have an effect on the iteron-containing plasmid copy number, by interaction with the nucleoprotein complex handcuff structure or the other complexes of Rep protein with iteron-containing plasmid DNA.

CONCLUSIONS

All the described mechanisms that affect plasmid metabolism are intended to control the plasmid replication frequency and thereby to control the plasmid copy number. The iteron-containing plasmids, as described above, predominantly use the limitation of Rep protein concentration to control initiation of replication. The limited amount of initiator is achieved by the autorepression mechanism. This kind of replication regulation was initially proposed to be the sole mechanism of replication control, but subsequent experiments showed the marginal effect of surplus initiator. This proved that this mechanism is insufficient (150–152). In iteron-containing plasmids, origin inactivation by handcuffing is an essential mechanism for effective replication regulation. It assumes that the iteron concentration, rather than the level of Rep expression, determines the rate of replication. Another critical parameter that influences the replication initiation is the dimer/monomer ratio of the Rep initiator. The efficient control of plasmid replication initiation requires a combination of all the above-mentioned regulatory mechanisms. Furthermore, it has been reported that all these mechanisms need to work in concert and no single mechanism alone is able to regulate plasmid replication effectively (130). Therefore, it seems to be clear why there are multiple modes of control and that all these modes appear to be cooperative rather than mutually exclusive, which explains why they have been conserved.

Acknowledgments. *This work was supported by the Polish National Science Center (grant number 2012/04/A/NZ1/00048]) and the Foundation for Polish Science (TEAM/2009-3/5). Conflicts of interest: We disclose no conflicts.*

Citation. Konieczny I, Bury K, Wawrzycka A, Wegrzyn K. 2014. Iteron plasmids. Microbiol Spectrum **2**(6)**:**PLAS-0026-2014.

References

1. **Stalker DM, Thomas CM, Helinski DR.** 1981. Nucleotide sequence of the region of the origin of replication of the broad host range plasmid RK2. *Mol Gen Genet* **181:**8–12.
2. **Murotsu T, Matsubara K, Sugisaki H, Takanami M.** 1981. Nine unique repeating sequences in a region essential for replication and incompatibility of the mini-F plasmid. *Gene* **15:**257–271.
3. **Abeles AL, Snyder KM, Chattoraj DK.** 1984. P1 plasmid replication: replicon structure. *J Mol Biol* **173:** 307–324.
4. **Filutowicz M, McEachern MJ, Mukhopadhyay P, Greener A, Yang SL, Helinski DR.** 1987. DNA and protein interactions in the regulation of plasmid replication. *J Cell Sci Suppl* **7:**15–31.
5. **Nieto C, Giraldo R, Fernandez-Tresguerres E, Diaz R.** 1992. Genetic and functional analysis of the basic replicon of pPS10, a plasmid specific for *Pseudomonas* isolated from *Pseudomonas syringae* patovar *savastanoi*. *J Mol Biol* **223:**415–426.
6. **Loftie-Eaton W, Rawlings DE.** 2012. Diversity, biology and evolution of IncQ-family plasmids. *Plasmid* **67:** 15–34.
7. **Wu LT, Tseng YH.** 2000. Characterization of the IncW cryptic plasmid pXV2 from *Xanthomonas campestris* pv. *vesicatoria*. *Plasmid* **44:**163–172.
8. **Churchward G, Linder P, Caro L.** 1984. Replication functions encoded by the plasmid pSC101. *Adv Exp Med Biol* **179:**209–214.
9. **Page DT, Whelan KF, Colleran E.** 2001. Characterization of two autoreplicative regions of the IncHI2 plasmid R478: RepHI2A and RepHI1A((R478)). *Microbiology* **147:**1591–1598.
10. **Wickner S, Hoskins J, McKenney K.** 1991. Monomerization of RepA dimers by heat shock proteins activates binding to DNA replication origin. *Proc Natl Acad Sci USA* **88:**7903–7907.
11. **Ishiai M, Wada C, Kawasaki Y, Yura T.** 1994. Replication initiator protein RepE of mini-F plasmid: functional differentiation between monomers (initiator) and dimers (autogenous repressor). *Proc Natl Acad Sci USA* **91:**3839–3843.
12. **Toukdarian AE, Helinski DR, Perri S.** 1996. The plasmid RK2 initiation protein binds to the origin of replication as a monomer. *J Biol Chem* **271**(12)**:**7072–7078.
13. **Kruger R, Konieczny I, Filutowicz M.** 2001. Monomer/dimer ratios of replication protein modulate the DNA strand-opening in a replication origin. *J Mol Biol* **306:** 945–955.
14. **Perri S, Helinski DR.** 1993. DNA sequence requirements for interaction of the RK2 replication initiation protein with plasmid origin repeats. *J Biol Chem* **268:** 3662–3669.
15. **Bowers LM, Kruger R, Filutowicz M.** 2007. Mechanism of origin activation by monomers of R6K-encoded pi protein. *J Mol Biol* **368:**928–938.
16. **McEachern MJ, Filutowicz M, Helinski DR.** 1985. Mutations in direct repeat sequences and in a conserved sequence adjacent to the repeats result in a defective replication origin in plasmid R6K. *Proc Natl Acad Sci USA* **82:**1480–1484.
17. **Brendler TG, Abeles AL, Reaves LD, Austin SJ.** 1997. The iteron bases and spacers of the P1 replication origin contain information that specifies the formation of a complex structure involved in initiation. *Mol Microbiol* **23:**559–567.
18. **Ohkubo S, Yamaguchi K.** 1997. A suppressor of mutations in the region adjacent to iterons of pSC101 ori. *J Bacteriol* **179:**2089–2091.
19. **Doran KS, Konieczny I, Helinski DR.** 1998. Replication origin of the broad host range plasmid RK2. Positioning of various motifs is critical for initiation of replication. *J Biol Chem* **273:**8447–8453.

20. **Messer W.** 2002. The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. *FEMS Microbiol Rev* **26:**355–374.
21. **Rajewska M, Wegrzyn K, Konieczny I.** 2012. AT-rich region and repeated sequences: the essential elements of replication origins of bacterial replicons. *FEMS Microbiol Rev* **36:**408–434.
22. **Doran KS, Helinski DR, Konieczny I.** 1999. Host-dependent requirement for specific DnaA boxes for plasmid RK2 replication. *Mol Microbiol* **33:**490–498.
23. **Speck C, Messer W.** 2001. Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA. *EMBO J* **20:**1469–1476.
24. **Konieczny I, Doran KS, Helinski DR, Blasina A.** 1997. Role of TrfA and DnaA proteins in origin opening during initiation of DNA replication of the broad host range plasmid RK2. *J Biol Chem* **272:**20173–20178.
25. **Wei T, Bernander R.** 1996. Interaction of the IciA protein with AT-rich regions in plasmid replication origins. *Nucleic Acids Res* **24:**1865–1872.
26. **Bramhill D, Kornberg A.** 1988. A model for initiation at origins of DNA replication. *Cell* **54:**915–918.
27. **Stalker DM, Kolter R, Helinski DR.** 1979. Nucleotide sequence of the region of an origin of replication of the antibiotic resistance plasmid R6K. *Proc Natl Acad Sci USA* **76:**1150–1154.
28. **Chattoraj DK, Snyder KM, Abeles AL.** 1985. P1 plasmid replication: multiple functions of RepA protein at the origin. *Proc Natl Acad Sci USA* **82:**2588–2592.
29. **Abeles AL, Reaves LD, Austin SJ.** 1990. A single DnaA box is sufficient for initiation from the P1 plasmid origin. *J Bacteriol* **172:**4386–4391.
30. **Kowalczyk L, Rajewska M, Konieczny I.** 2005. Positioning and the specific sequence of each 13-mer motif are critical for activity of the plasmid RK2 replication origin. *Mol Microbiol* **57:**1439–1449.
31. **Rajewska M, Kowalczyk L, Konopa G, Konieczny I.** 2008. Specific mutations within the AT-rich region of a plasmid replication origin affect either origin opening or helicase loading. *Proc Natl Acad Sci USA* **105:**11134–11139.
32. **Fekete RA, Venkova-Canova T, Park K, Chattoraj DK.** 2006. IHF-dependent activation of P1 plasmid origin by dnaA. *Mol Microbiol* **62:**1739–1751.
33. **Stenzel TT, Patel P, Bastia D.** 1987. The integration host factor of *Escherichia coli* binds to bent DNA at the origin of replication of the plasmid pSC101. *Cell* **49:**709–717.
34. **Filutowicz M, Appelt K.** 1988. The integration host factor of *Escherichia coli* binds to multiple sites at plasmid R6K gamma origin and is essential for replication. *Nucleic Acids Res* **16:**3829–3843.
35. **Shah DS, Cross MA, Porter D, Thomas CM.** 1995. Dissection of the core and auxiliary sequences in the vegetative replication origin of promiscuous plasmid RK2. *J Mol Biol* **254:**608–622.
36. **Brendler T, Abeles A, Austin S.** 1991. Critical sequences in the core of the P1 plasmid replication origin. *J Bacteriol* **173:**3935–3942.
37. **Brendler T, Abeles A, Austin S.** 1995. A protein that binds to the P1 origin core and the oriC 13mer region in a methylation-specific fashion is the product of the host seqA gene. *EMBO J* **14:**4083–4089.
38. **Slater S, Wold S, Lu M, Boye E, Skarstad K, Kleckner N.** 1995. *E. coli* SeqA protein binds oriC in two different methyl-modulated reactions appropriate to its roles in DNA replication initiation and origin sequestration. *Cell* **82:**927–936.
39. **Lu M, Campbell JL, Boye E, Kleckner N.** 1994. SeqA: a negative modulator of replication initiation in *E. coli*. *Cell* **77:**413–426.
40. **Hwang DS, Kornberg A.** 1990. A novel protein binds a key origin sequence to block replication of an *E. coli* minichromosome. *Cell* **63:**325–331.
41. **Wu F, Wu J, Ehley J, Filutowicz M.** 1996. Preponderance of Fis-binding sites in the R6K gamma origin and the curious effect of the penicillin resistance marker on replication of this origin in the absence of Fis. *J Bacteriol* **178:**4965–4974.
42. **Giraldo R, Andreu JM, Diaz-Orejas R.** 1998. Protein domains and conformational changes in the activation of RepA, a DNA replication initiator. *EMBO J* **17:**4511–4526.
43. **Komori H, Matsunaga F, Higuchi Y, Ishiai M, Wada C, Miki K.** 1999. Crystal structure of a prokaryotic replication initiator protein bound to DNA at 2.6 A resolution. *EMBO J* **18:**4597–4607.
44. **Swan MK, Bastia D, Davies C.** 2006. Crystal structure of pi initiator protein-iteron complex of plasmid R6K: implications for initiation of plasmid DNA replication. *Proc Natl Acad Sci USA* **103:**18481–18486.
45. **Giraldo R, Fernandez-Tornero C, Evans PR, Diaz-Orejas R, Romero A.** 2003. A conformational switch between transcriptional repression and replication initiation in the RepA dimerization domain. *Nat Struct Biol* **10:**565–571.
46. **Garcia de Viedma D, Serrano-Lopez A, Diaz-Orejas R.** 1995. Specific binding of the replication protein of plasmid pPS10 to direct and inverted repeats is mediated by an HTH motif. *Nucleic Acids Res* **23:**5048–5054.
47. **Matsunaga F, Kawasaki Y, Ishiai M, Nishikawa K, Yura T, Wada C.** 1995. DNA-binding domain of the RepE initiator protein of mini-F plasmid: involvement of the carboxyl-terminal region. *J Bacteriol* **177:**1994–2001.
48. **Diaz-Lopez T, Davila-Fajardo C, Blaesing F, Lillo MP, Giraldo R.** 2006. Early events in the binding of the pPS10 replication protein RepA to single iteron and operator DNA sequences. *J Mol Biol* **364:**909–920.
49. **Pierechod M, Nowak A, Saari A, Purta E, Bujnicki JM, Konieczny I.** 2009. Conformation of a plasmid replication initiator protein affects its proteolysis by ClpXP system. *Protein Sci* **18:**637–649.
50. **Lin J, Helinski DR.** 1992. Analysis of mutations in trfA, the replication initiation gene of the broad-host-range plasmid RK2. *J Bacteriol* **174:**4110–4119.
51. **Cereghino JL, Helinski DR, Toukdarian AE.** 1994. Isolation and characterization of DNA-binding mutants

of a plasmid replication initiation protein utilizing an *in vivo* binding assay. *Plasmid* **31:**89–99.

52. **Sharma S, Sathyanarayana BK, Bird JG, Hoskins JR, Lee B, Wickner S.** 2004. Plasmid P1 RepA is homologous to the F plasmid RepE class of initiators. *J Biol Chem* **279:**6027–6034.
53. **Giraldo R.** 2003. Common domains in the initiators of DNA replication in Bacteria, Archaea and Eukarya: combined structural, functional and phylogenetic perspectives. *FEMS Microbiol Rev* **26:**533–554.
54. **Liu J, Smith CL, DeRyckere D, DeAngelis K, Martin GS, Berger JM.** 2000. Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control. *Mol Cell* **6:**637–648.
55. **Giraldo R.** 2007. Defined DNA sequences promote the assembly of a bacterial protein into distinct amyloid nanostructures. *Proc Natl Acad Sci USA* **104:**17388–17393.
56. **Gasset-Rosa F, Mate MJ, Davila-Fajardo C, Bravo J, Giraldo R.** 2008. Binding of sulphonated indigo derivatives to RepA-WH1 inhibits DNA-induced protein amyloidogenesis. *Nucleic Acids Res* **36:**2249–2256.
57. **Fernandez-Tresguerres ME, de la Espina SM, Gasset-Rosa F, Giraldo R.** 2010. A DNA-promoted amyloid proteinopathy in *Escherichia coli*. *Mol Microbiol* **77:**1456–1469.
58. **Giraldo R, Moreno-Diaz de la Espina S, Fernandez-Tresguerres ME, Gasset-Rosa F.** 2011. RepA-WH1 prionoid: a synthetic amyloid proteinopathy in a minimalist host. *Prion* **5:**60–64.
59. **Giraldo R.** Amyloid assemblies: protein legos at a crossroads in bottom-up synthetic biology. *Chembiochem* **11:**2347–2357.
60. **Kawasaki Y, Wada C, Yura T.** 1990. Roles of *Escherichia coli* heat shock proteins DnaK, DnaJ and GrpE in mini-F plasmid replication. *Mol Gen Genet* **220:**277–282.
61. **Diaz-Lopez T, Lages-Gonzalo M, Serrano-Lopez A, Alfonso C, Rivas G, Diaz-Orejas R, Giraldo R.** 2003. Structural changes in RepA, a plasmid replication initiator, upon binding to origin DNA. *J Biol Chem* **278:**18606–18616.
62. **Ingmer H, Fong EL, Cohen SN.** 1995. Monomer-dimer equilibrium of the pSC101 RepA protein. *J Mol Biol* **250:**309–314.
63. **Filutowicz M, Davis G, Greener A, Helinski DR.** 1985. Autorepressor properties of the pi-initiation protein encoded by plasmid R6K. *Nucleic Acids Res* **13:**103–114.
64. **Landschulz WH, Johnson PF, McKnight SL.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* **240:**1759–1764.
65. **Ingmer H, Cohen SN.** 1993. Excess intracellular concentration of the pSC101 RepA protein interferes with both plasmid DNA replication and partitioning. *J Bacteriol* **175:**7834–7841.
66. **Miron A, Mukherjee S, Bastia D.** 1992. Activation of distant replication origins *in vivo* by DNA looping as revealed by a novel mutant form of an initiator protein defective in cooperativity at a distance. *EMBO J* **11:**1205–1216.
67. **Fernandez-Tresguerres ME, Martin M, Garcia de Viedma D, Giraldo R, Diaz-Orejas R.** 1995. Host growth temperature and a conservative amino acid substitution in the replication protein of pPS10 influence plasmid host range. *J Bacteriol* **177:**4377–4384.
68. **Shingler V, Thomas CM.** 1984. Analysis of the trfA region of broad host-range plasmid RK2 by transposon mutagenesis and identification of polypeptide products. *J Mol Biol* **175:**229–249.
69. **Kornacki JA, West AH, Firshein W.** 1984. Proteins encoded by the trans-acting replication and maintenance regions of broad host range plasmid RK2. *Plasmid* **11:**48–57.
70. **Toukdarian AE, Helinski DR, Perri S.** 1996. The plasmid RK2 initiation protein binds to the origin of replication as a monomer. *J Biol Chem* **271:**7072–7078.
71. **Zhong Z, Helinski D, Toukdarian A.** 2003. A specific region in the N terminus of a replication initiation protein of plasmid RK2 is required for recruitment of *Pseudomonas aeruginosa* DnaB helicase to the plasmid origin. *J Biol Chem* **278:**45305–45310.
72. **Hughes JM, Lohman BK, Deckert GE, Nichols EP, Settles M, Abdo Z, Top EM.** 2012. The role of clonal interference in the evolutionary dynamics of plasmid-host adaptation. *MBio* **3**(4)**:**e00077-e00012.
73. **Sota M, Yano H, Hughes JM, Daughdrill GW, Abdo Z, Forney LJ, Top EM.** 2010. Shifts in the host range of a promiscuous plasmid through parallel evolution of its replication initiation protein. *ISME J* **4:**1568–1580.
74. **Pacek M, Konopa G, Konieczny I.** 2001. DnaA box sequences as the site for helicase delivery during plasmid RK2 replication initiation in *Escherichia coli*. *J Biol Chem* **276:**23639–23644.
75. **Kim PD, Banack T, Lerman DM, Tracy JC, Camara JE, Crooke E, Oliver D, Firshein W.** 2003. Identification of a novel membrane-associated gene product that suppresses toxicity of a TrfA peptide from plasmid RK2 and its relationship to the DnaA host initiation protein. *J Bacteriol* **185:**1817–1824.
76. **Kongsuwan K, Josh P, Picault MJ, Wijffels G, Dalrymple B.** 2006. The plasmid RK2 replication initiator protein (TrfA) binds to the sliding clamp beta subunit of DNA polymerase III: implication for the toxicity of a peptide derived from the amino-terminal portion of 33-kilodalton TrfA. *J Bacteriol* **188:**5501–5509.
77. **Sutton MD, Kaguni JM.** 1997. The *Escherichia coli* dnaA gene: four functional domains. *J Mol Biol* **274:**546–561.
78. **Messer W, Blaesing F, Majka J, Nardmann J, Schaper S, Schmidt A, Seitz H, Speck C, Tungler D, Wegrzyn G, Weigel C, Welzeck M, Zakrzewska-Czerwinska J.** 1999. Functional domains of DnaA proteins. *Biochimie* **81:**819–825.
79. **Erzberger JP, Pirruccello MM, Berger JM.** 2002. The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation. *EMBO J* **21:**4763–4773.

80. **Erzberger JP, Mott ML, Berger JM.** 2006. Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nat Struct Mol Biol* **13:**676–683.
81. **Duderstadt KE, Chuang K, Berger JM.** 2011. DNA stretching by bacterial initiators promotes replication origin opening. *Nature* **478:**209–213.
82. **Ozaki S, Kawakami H, Nakamura K, Fujikawa N, Kagawa W, Park SY, Yokoyama S, Kurumizaka H, Katayama T.** 2008. A common mechanism for the ATP-DnaA-dependent formation of open complexes at the replication origin. *J Biol Chem* **283:**8351–8362.
83. **DasGupta S, Mukhopadhyay G, Papp PP, Lewis MS, Chattoraj DK.** 1993. Activation of DNA binding by the monomeric form of the P1 replication initiator RepA by heat shock proteins DnaJ and DnaK. *J Mol Biol* **232:** 23–34.
84. **Wickner S, Hoskins J, McKenney K.** 1991. Function of DnaJ and DnaK as chaperones in origin-specific DNA binding by RepA. *Nature* **350:**165–167.
85. **Kawasaki Y, Wada C, Yura T.** 1992. Binding of RepE initiator protein to mini-F DNA origin (ori2). Enhancing effects of repE mutations and DnaJ heat shock protein. *J Biol Chem* **267:**11520–11524.
86. **Manen D, Upegui-Gonzalez LC, Caro L.** 1992. Monomers and dimers of the RepA protein in plasmid pSC101 replication: domains in RepA. *Proc Natl Acad Sci USA* **89:**8923–8927.
87. **Germino J, Bastia D.** 1983. Interaction of the plasmid R6K-encoded replication initiator protein with its binding sites on DNA. *Cell* **34:**125–134.
88. **Kelley WL, Bastia D.** 1992. Activation *in vivo* of the minimal replication origin beta of plasmid R6K requires a small target sequence essential for DNA looping. *New Biol* **4:**569–580.
89. **Kelley WL, Patel I, Bastia D.** 1992. Structural and functional analysis of a replication enhancer: separation of the enhancer activity from origin function by mutational dissection of the replication origin gamma of plasmid R6K. *Proc Natl Acad Sci USA* **89:**5078–5082.
90. **Lu YB, Datta HJ, Bastia D.** 1998. Mechanistic studies of initiator-initiator interaction and replication initiation. *EMBO J* **17:**5192–5200.
91. **Kruger R, Filutowicz M.** 2000. Dimers of pi protein bind the A+T-rich region of the R6K gamma origin near the leading-strand synthesis start sites: regulatory implications. *J Bacteriol* **182:**2461–2467.
92. **Mukherjee S, Erickson H, Bastia D.** 1988. Enhancer-origin interaction in plasmid R6K involves a DNA loop mediated by initiator protein. *Cell* **52:**375–383.
93. **Kawasaki Y, Matsunaga F, Kano Y, Yura T, Wada C.** 1996. The localized melting of mini-F origin by the combined action of the mini-F initiator protein (RepE) and HU and DnaA of *Escherichia coli*. *Mol Gen Genet* **253:**42–49.
94. **Hasunuma K, Sekiguchi M.** 1977. Replication of plasmid pSC101 in *Escherichia coli* K12: requirement for dnaA function. *Mol Gen Genet* **154:**225–230.
95. **Gamas P, Burger AC, Churchward G, Caro L, Galas D, Chandler M.** 1986. Replication of pSC101: effects of mutations in the *E. coli* DNA binding protein IHF. *Mol Gen Genet* **204:**85–89.
96. **Mukhopadhyay G, Carr KM, Kaguni JM, Chattoraj DK.** 1993. Open-complex formation by the host initiator, DnaA, at the origin of P1 plasmid replication. *EMBO J* **12:**4547–4554.
97. **Durland RH, Helinski DR.** 1987. The sequence encoding the 43-kilodalton trfA protein is required for efficient replication or maintenance of minimal RK2 replicons in *Pseudomonas aeruginosa*. *Plasmid* **18:**164–169.
98. **Fang FC, Helinski DR.** 1991. Broad-host-range properties of plasmid RK2: importance of overlapping genes encoding the plasmid replication initiation protein TrfA. *J Bacteriol* **173:**5861–5868.
99. **Wegrzyn K, Witosinska M, Schweiger P, Bury K, Jenal U, Konieczny I.** 2013. RK2 plasmid dynamics in *Caulobacter crescentus* cells: two modes of DNA replication initiation. *Microbiology* **159:**1010–1022.
100. **Maestro B, Sanz JM, Faelen M, Couturier M, Diaz-Orejas R, Fernandez-Tresguerres E.** 2002. Modulation of pPS10 host range by DnaA. *Mol Microbiol* **46:** 223–234.
101. **Sharma R, Kachroo A, Bastia D.** 2001. Mechanistic aspects of DnaA-RepA interaction as revealed by yeast forward and reverse two-hybrid analysis. *EMBO J* **20:** 4577–4587.
102. **Skarstad K, Baker TA, Kornberg A.** 1990. Strand separation required for initiation of replication at the chromosomal origin of *E. coli* is facilitated by a distant RNA-DNA hybrid. *EMBO J* **9:**2341–2348.
103. **Ryan VT, Grimwade JE, Nievera CJ, Leonard AC.** 2002. IHF and HU stimulate assembly of pre-replication complexes at *Escherichia coli* oriC by two different mechanisms. *Mol Microbiol* **46:**113–124.
104. **Hwang DS, Kornberg A.** 1992. Opening of the replication origin of *Escherichia coli* by DnaA protein with protein HU or IHF. *J Biol Chem* **267:**23083–23086.
105. **Duderstadt KE, Mott ML, Crisona NJ, Chuang K, Yang H, Berger JM.** 2010. Origin remodeling and opening in bacteria rely on distinct assembly states of the DnaA initiator. *J Biol Chem* **285:**28229–28239.
106. **Schnos M, Zahn K, Inman RB, Blattner FR.** 1988. Initiation protein induced helix destabilization at the lambda origin: a prepriming step in DNA replication. *Cell* **52:**385–395.
107. **Park K, Mukhopadhyay S, Chattoraj DK.** 1998. Requirements for and regulation of origin opening of plasmid P1. *J Biol Chem* **273:**24906–24911.
108. **Jiang Y, Pacek M, Helinski DR, Konieczny I, Toukdarian A.** 2003. A multifunctional plasmid-encoded replication initiation protein both recruits and positions an active helicase at the replication origin. *Proc Natl Acad Sci USA* **100:**8692–8697.
109. **Caspi R, Pacek M, Consiglieri G, Helinski DR, Toukdarian A, Konieczny I.** 2001. A broad host range replicon with different requirements for replication initiation in three bacterial species. *EMBO J* **20:**3262–3271.

110. **Davey MJ, Fang L, McInerney P, Georgescu RE, O'Donnell M.** 2002. The DnaC helicase loader is a dual ATP/ADP switch protein. *EMBO J* **21:**3148–3159.
111. **Konieczny I, Helinski DR.** 1997. Helicase delivery and activation by DnaA and TrfA proteins during the initiation of replication of the broad host range plasmid RK2. *J Biol Chem* **272:**33312–33318.
112. **Ratnakar PV, Mohanty BK, Lobert M, Bastia D.** 1996. The replication initiator protein pi of the plasmid R6K specifically interacts with the host-encoded helicase DnaB. *Proc Natl Acad Sci USA* **93:**5522–5526.
113. **Zhong Z, Helinski D, Toukdarian A.** 2005. Plasmid host-range: restrictions to F replication in *Pseudomonas*. *Plasmid* **54:**48–56.
114. **Datta HJ, Khatri GS, Bastia D.** 1999. Mechanism of recruitment of DnaB helicase to the replication origin of the plasmid pSC101. *Proc Natl Acad Sci USA* **96:** 73–78.
115. **Giraldo R, Diaz R.** 1992. Differential binding of wild-type and a mutant RepA protein to oriR sequence suggests a model for the initiation of plasmid R1 replication. *J Mol Biol* **228:**787–802.
116. **Wickner SH, Chattoraj DK.** 1987. Replication of mini-P1 plasmid DNA *in vitro* requires two initiation proteins, encoded by the repA gene of phage P1 and the dnaA gene of *Escherichia coli*. *Proc Natl Acad Sci USA* **84:**3668–3672.
117. **Abhyankar MM, Zzaman S, Bastia D.** 2003. Reconstitution of R6K DNA replication *in vitro* using 22 purified proteins. *J Biol Chem* **278:**45476–45484.
118. **Zzaman S, Abhyankar MM, Bastia D.** 2004. Reconstitution of F factor DNA replication *in vitro* with purified proteins. *J Biol Chem* **279:**17404–17410.
119. **Dalrymple BP, Kongsuwan K, Wijffels G.** 2007. Identification of putative DnaN-binding motifs in plasmid replication initiation proteins. *Plasmid* **57:**82–88.
120. **Bloom LB.** 2009. Loading clamps for DNA replication and repair. *DNA Repair (Amst)* **8:**570–578.
121. **Simonetta KR, Kazmirski SL, Goedken ER, Cantor AJ, Kelch BA, McNally R, Seyedin SN, Makino DL, O'Donnell M, Kuriyan J.** 2009. The mechanism of ATP-dependent primer-template recognition by a clamp loader complex. *Cell* **137:**659–671.
122. **Park K, Han E, Paulsson J, Chattoraj DK.** 2001. Origin pairing ('handcuffing') as a mode of negative control of P1 plasmid copy number. *EMBO J* **20:**7323–7332.
123. **Zzaman S, Bastia D.** 2005. Oligomeric initiator protein-mediated DNA looping negatively regulates plasmid replication *in vitro* by preventing origin melting. *Mol Cell* **20:**833–843.
124. **Urh M, Wu J, Forest K, Inman RB, Filutowicz M.** 1998. Assemblies of replication initiator protein on symmetric and asymmetric DNA sequences depend on multiple protein oligomerization surfaces. *J Mol Biol* **283:**619–631.
125. **Kunnimalaiyaan S, Inman RB, Rakowski SA, Filutowicz M.** 2005. Role of pi dimers in coupling ("handcuffing") of plasmid R6K's gamma ori iterons. *J Bacteriol* **187:** 3779–3785.
126. **McEachern MJ, Bott MA, Tooker PA, Helinski DR.** 1989. Negative control of plasmid R6K replication: possible role of intermolecular coupling of replication origins. *Proc Natl Acad Sci USA* **86:**7942–7946.
127. **Miron A, Patel I, Bastia D.** 1994. Multiple pathways of copy control of gamma replicon of R6K: mechanisms both dependent on and independent of cooperativity of interaction of tau protein with DNA affect the copy number. *Proc Natl Acad Sci USA* **91:**6438–6442.
128. **Das N, Chattoraj DK.** 2004. Origin pairing ('handcuffing') and unpairing in the control of P1 plasmid replication. *Mol Microbiol* **54:**836–849.
129. **Toukdarian AE, Helinski DR.** 1998. TrfA dimers play a role in copy-number control of RK2 replication. *Gene* **223:**205–211.
130. **Paulsson J, Chattoraj DK.** 2006. Origin inactivation in bacterial DNA replication control. *Mol Microbiol* **61:** 9–15.
131. **Becskei A, Serrano L.** 2000. Engineering stability in gene networks by autoregulation. *Nature* **405:**590–593.
132. **Simpson ML, Cox CD, Sayler GS.** 2003. Frequency domain analysis of noise in autoregulated gene circuits. *Proc Natl Acad Sci USA* **100:**4551–4556.
133. **York D, Filutowicz M.** 1993. Autoregulation-deficient mutant of the plasmid R6K-encoded pi protein distinguishes between palindromic and nonpalindromic binding sites. *J Biol Chem* **268:**21854–21861.
134. **Pabo CO, Sauer RT.** 1992. Transcription factors: structural families and principles of DNA recognition. *Annu Rev Biochem* **61:**1053–1095.
135. **Vocke C, Bastia D.** 1985. The replication initiator protein of plasmid pSC101 is a transcriptional repressor of its own cistron. *Proc Natl Acad Sci USA* **82:**2252–2256.
136. **Kelley W, Bastia D.** 1985. Replication initiator protein of plasmid R6K autoregulates its own synthesis at the transcriptional step. *Proc Natl Acad Sci USA* **82:** 2574–2578.
137. **Chattoraj DK, Ghirlando R, Park K, Dibbens JA, Lewis MS.** 1996. Dissociation kinetics of RepA dimers: implications for mechanisms of activation of DNA binding by chaperones. *Genes Cells* **1:**189–199.
138. **Nakamura A, Wada C, Miki K.** 2007. Structural basis for regulation of bifunctional roles in replication initiator protein. *Proc Natl Acad Sci USA* **104:**18484–18489.
139. **Konieczny I, Helinski DR.** 1997. The replication initiation protein of the broad-host-range plasmid RK2 is activated by the ClpX chaperone. *Proc Natl Acad Sci USA* **94:**14378–14382.
140. **Konieczny I, Liberek K.** 2002. Cooperative action of *Escherichia coli* ClpB protein and DnaK chaperone in the activation of a replication initiation protein. *J Biol Chem* **277:**18483–18488.
141. **Kruklitis R, Welty DJ, Nakai H.** 1996. ClpX protein of *Escherichia coli* activates bacteriophage Mu transposase in the strand transfer complex for initiation of Mu DNA synthesis. *EMBO J* **15:**935–944.

142. Zzaman S, Reddy JM, Bastia D. 2004. The DnaK-DnaJ-GrpE chaperone system activates inert wild type pi initiator protein of R6K into a form active in replication initiation. *J Biol Chem* **279:**50886–50894.

143. Sozhamannan S, Chattoraj DK. 1993. Heat shock proteins DnaJ, DnaK, and GrpE stimulate P1 plasmid replication by promoting initiator binding to the origin. *J Bacteriol* **175:**3546–3555.

144. Wickner S, Gottesman S, Skowyra D, Hoskins J, McKenney K, Maurizi MR. 1994. A molecular chaperone, ClpA, functions like DnaK and DnaJ. *Proc Natl Acad Sci USA* **91:**12218–12222.

145. Pak M, Wickner S. 1997. Mechanism of protein remodeling by ClpA chaperone. *Proc Natl Acad Sci USA* **94:** 4901–4906.

146. Dougan DA, Mogk A, Bukau B. 2002. Protein folding and degradation in bacteria: to degrade or not to degrade? That is the question. *Cell Mol Life Sci* **59:**1607–1616.

147. Levchenko I, Luo L, Baker TA. 1995. Disassembly of the Mu transposase tetramer by the ClpX chaperone. *Genes Dev* **9:**2399–2408.

148. Wojtkowiak D, Georgopoulos C, Zylicz M. 1993. Isolation and characterization of ClpX, a new ATP-dependent specificity component of the Clp protease of *Escherichia coli*. *J Biol Chem* **268:**22609–22617.

149. Kubik S, Wegrzyn K, Pierechod M, Konieczny I. 2012. Opposing effects of DNA on proteolysis of a replication initiator. *Nucleic Acids Res* **40:**1148–1159.

150. Tsutsui H, Fujiyama A, Murotsu T, Matsubara K. 1983. Role of nine repeating sequences of the mini-F genome for expression of F-specific incompatibility phenotype and copy number control. *J Bacteriol* **155:**337–344.

151. Pal SK, Chattoraj DK. 1988. P1 plasmid replication: initiator sequestration is inadequate to explain control by initiator-binding sites. *J Bacteriol* **170:**3554–3560.

152. Durland RH, Helinski DR. 1990. Replication of the broad-host-range plasmid RK2: direct measurement of intracellular concentrations of the essential TrfA replication proteins and their effect on plasmid copy number. *J Bacteriol* **172:**3849–3858.

Plasmids—Biology and Impact in Biotechnology and Discovery
Edited by Marcelo E. Tolmasky and Juan C. Alonso

doi:10.1128/microbiolspec.PLAS-0029-2014

Joshua Lilly[1]
Manel Camps[1]

3 Mechanisms of Theta Plasmid Replication

INTRODUCTION

Plasmids have been used as convenient models for the study of molecular mechanisms of replication and DNA repair due to their small size, dispensability to the host, and easy manipulation. In addition, plasmids are key facilitators for the evolution and dissemination of drug resistance and for the evolution of complex interactions with animal or plant hosts. Understanding plasmid replication and maintenance therefore has significant practical implications for the clinic and for bioremediation.

Circular plasmids use a variety of replication strategies depending on the mechanism of initiation of DNA replication and depending on whether leading- and lagging-strand synthesis are coupled or uncoupled. This article focuses on replication of circular plasmids whose lagging strand is synthesized discontinuously, a mechanism known as theta replication because replication intermediates have the shape of the Greek letter θ (theta). Our discussion will focus on replication initiation, which informs different biological properties of plasmids (size, host range, plasmid copy number, etc.), and on how initiation is regulated in these plasmids. To highlight unique aspects of theta plasmid replication, this mode of replication will also be compared with another mode of circular plasmid replication, strand-displacement.

REPLICATION INITIATION

General Structure of Plasmid Origins of Replication

Replication initiation depends on a section of sequence known as the plasmid origin of replication (*ori*). *Basic replicon* refers to the minimal sequence that supports replication, preserving the regulatory circuitry. *Minimal replicon* refers to the minimal portion of sequence supporting plasmid replication even though replication may not be properly regulated, as seen in alterations in plasmid copy number or in the compatibility properties of the plasmids. Finally, there is an even narrower definition of *ori*, which refers to the portion of sequence that is targeted by replication initiation factors *in trans* to initiate replication. In this article we will use the term *origin of replication*, or *ori*, to refer to the *cis-ori*, and *replicon* to refer to basic or minimal replicons.

Rep proteins are plasmid-encoded initiators of replication, although some theta plasmids rely exclusively on host initiation factors for replication. Rep recognition

[1]Department of Microbiology and Environmental Toxicology, University of California Santa Cruz, Santa Cruz, CA 95064.

sites typically consist of direct repeats or *iterons*, whose specific sequence and spacing are important for initiator recognition. Spacing is critically relevant so that the distance matches the helical periodicity of the DNA double helix, allowing recognition of specific DNA sequences (1). Iterons are intrinsically bent, and iteron curvature is enhanced by Rep binding.

Rep proteins are essential and rate-limiting for plasmid replication initiation. Controlled expression of two Rep proteins (π of R6K and RepA of ColE2) can produce a wide range of plasmid copy numbers per cell (between 1 and 250 copies), providing a convenient system for gene dosage optimization of recombinant proteins (2).

Plasmid replicons have a modular structure. Replicons often have motifs that are recognized by plasmid-encoded Reps, A+T-rich areas, G+C-rich areas, methylation sites, and binding sites for host initiation and/or remodeling factors. *Rep* loci, when present, are typically upstream of the plasmid *ori*, immediately adjacent or in close proximity to it.

Replication Initiation: Duplex Melting and Replisome Assembly

Depending on the replicon, duplex melting can be either dependent on transcription or mediated by plasmid-encoded *trans*-acting proteins (Reps). Rep binding of *ori* iterons generally leads to the formation of a nucleoprotein complex that opens up the DNA duplex at the A+T-rich segment.

Opening of the DNA duplex is necessary for replisome assembly, which in theta-type plasmids can be DnaA-dependent or PriA-dependent. DnaA-dependent assembly closely resembles replication initiation at *oriC*, the site initiating chromosomal replication. By contrast, PriA-dependent assembly parallels replication restart following replication fork arrest, which depends on D-loop formation, with the extra DNA strand supplied by homologous recombination (3–5).

In theta-type plasmids, Rep-mediated duplex melting leads to loading of DnaB on the replication fork, often with DnaA assistance. In plasmids that instead rely on transcription for duplex melting, the transcript itself can be processed and becomes the primer for extension. Continuous extension of this primer initiates leading-strand synthesis, facilitating the formation of a displacement loop, or D-loop, as the nascent single-stranded DNA (ssDNA) strand separates the two strands of the DNA duplex and hybridizes with one of them. In this case, PriA (initiator of primosome assembly) can be recruited to the forked structure of the D-loop; alternatively, PriA can be recruited to a hairpin structure that forms when the double-stranded DNA opens (6). PriA promotes both the unwinding of the lagging-strand arm and assembly of two additional proteins (PriB and DnaT) to load DnaB onto the lagging strand template. Thus, in this case loading of DnaB is independent of DnaA.

After loading of DnaB, both DnaA-dependent and -independent modes of replication converge. In both cases, replisome assembly involves the following additional players: SSB (single-stranded binding protein), DnaB (helicase), DnaC (loading factor), the DnaG (primase), and the DNA polymerase III (Pol III) holoenzyme. SSB is recruited to exposed areas of ssDNA, stabilizing them. DnaB is loaded onto the replication fork in the form of a complex with DnaC and recruits DnaG (the primase), which distributively synthesizes RNA primers for lagging-strand synthesis (7). Replisome assembly is completed by loading of the Pol III holoenzyme (8). This holoenzyme contains a core (with α, a catalytic, and ε, a 3′→5′, catalytic subunit), a β_2 processivity factor, and a DnaX complex ATPase that loads β_2 onto DNA and recruits the Pol III core to the newly loaded β_2 (9). DnaB helicase activity is stimulated through its interaction with Pol III and modulated through its interaction with DnaG, facilitating the coordination of leading-strand synthesis with that of lagging-strand synthesis during slow primer synthesis on the lagging strand (10).

Unlike Gram-negative bacteria, which have a single replicative polymerase (Pol III), Gram-positive bacteria have two replicative polymerases: PolC and DnaE. PolC is a processive polymerase responsible for leading-strand synthesis, while DnaE extends DnaG-synthesized primers before handoff to PolC at the lagging strand (11, 12).

In theta plasmids, lagging-strand synthesis is discontinuous and coordinated with leading-strand synthesis. The replicase extends a free 3′-OH of an RNA primer, which can be generated by DnaG primase (in Gram-negative bacteria), by the concerted action of DnaE and DnaG primase (in Gram-positive bacteria), or by alternative plasmid-encoded primases. Discontinuous lagging-strand synthesis involves repeated priming and elongation of Okazaki fragments and is comparable in plasmids and chromosomes, although Okazaki fragments were found to be smaller in a ColE1-like plasmid, approximately one-third the length of Okazaki fragments in the chromosome (13).

DNA polymerase I (Pol I) contributes to plasmid replication in several ways. In ColE1 and ColE1-like plasmids, Pol I can extend a primer to initiate leading-strand synthesis and open the DNA duplex; this process

can expose a hairpin structure in the lagging strand, known as a single-strand initiation (*ssi*) site or primosome assembly (*pas*) site, and/or generate a D-loop. Both hairpins and forked structures recruit PriA, which is the first step in the replisome initiation complex. Following replisome assembly, Pol I plays a critical role in discontinuous lagging-strand synthesis, removing RNA primers through its 5′→3′ exonuclease activity and filling in the remaining gap through its polymerase activity (14). In addition, two lines of evidence suggest that Pol I can functionally replace Pol III in *Escherichia coli*: (i) Pol I is essential for *polC* (Poll III-minus) strain viability, showing that both polymerases are functionally redundant (15). (ii) Mutations generated through error-prone Pol I replication of a ColE1-like plasmid *in vivo* strongly suggest that Pol I replicates both plasmid strands with similar frequency beyond the point where the switch to Pol III is expected, again suggesting that Pol I can be redundant with the Pol III replisome (16).

THETA PLASMID REPLICATION

Three modes of replication can be distinguished for circular plasmid replication: theta, strand-displacement, and rolling circle. This review focuses on theta. This mode of replication is similar to chromosomal replication in that the leading and lagging strands are replicated coordinately, with discontinuous lagging-strand synthesis. No DNA breaks are required for this mode of replication. Coordinated replication of both strands leads to the formation of bubbles in the early stages of replication, seen as the Greek letter θ under electron microscopy. Four classes of theta-type plasmids can be distinguished based on their mode of replication initiation, although the last two categories show hybrid features of the first two and will be discussed together (see theta replication section in Table 1).

Class A Theta Replication

Class A theta plasmids include R1, RK2, R6K, pSC101, pPS10, F, and P. All these plasmids depend on Rep proteins for replication initiation: RepA for R1, pSC101, pPS10, and P1; Trf1 for RK1; and π for R6K. Note that the name of these Reps is incidental, so sharing a name is not an indication of related structure or mode of action. Rep proteins bind direct repeats (iterons) in the plasmid *ori*. In class A, these iterons are rarely identical, although they frequently conform to a consensus motif. In plasmid P1, RepA monomers contact each iteron through two consecutive turns of the helix, leading to in-phase bending of the DNA, which wraps around RepA (17). Similarly, in R6K plasmids, π binding of its cognate iterons bends the DNA and generates a wrapped nucleoprotein structure (18).

There are two prominent exceptions to the presence of multiple iterons in class A theta plasmid *ori*s: (i) Plasmid R1, which features two partial palindromic sequences instead of iterons; however, similar to other plasmids of this class, R1 palindromic sequences are recognized by RepA. (ii) The R6K plasmid, which has three *ori*s, only one of which has multiple iterons: γ (with seven iterons), a second origin (α) with a single

TABLE 1 Comparison of the three basic modes of plasmid replication initiation in circular plasmids

	Leading-strand synthesis		Lagging-strand synthesis		
Type of replication	Plasmid initiation factors	Host factors	Coupling with leading strand	Plasmid factors	Host factors
Theta class A	Rep (duplex melting)	DnaA-replisome	Yes	No	Replisome
Theta class B	None	RNAP Pol I RNase H PriA-replisome	Yes	No	Replisome
Theta class C	Rep (duplex melting, primase)	Replisome	Yes	No	Replisome
Theta class D	Rep (duplex melting, RNA processing?)	RNAP PriA-replisome	Yes	No	Replisome
Strand-displacement	Rep A (helicase) Rep B (primase) Rep C (initiator)	Replisome (recruited by RepA)	No (simultaneous)	Rep A (helicase) Rep B (primase) Rep C (initiator)	None

iteron, and a third origin (β) with only half an iteron. It appears that the γ *ori* is an establishment origin, allowing replication initiation immediately following mobilization, when levels of π protein are low, whereas α and β *ori*s would be maintenance origins in cells inheriting the plasmid by vertical transmission (19). In any case, γ ori acts as an enhancer, favoring the long-range activation of α and β oris by transfer of π. Thus, α and β *ori*s are still dependent on the multiple iterons present in *ori* γ.

Rep binding of a cognate sequence in the plasmid *ori* mediates the earliest step in replication initiation: duplex DNA melting. A Rep-DnaA interaction is frequently involved, although the importance of this interaction varies between individual *ori*s. In plasmid pSC101, RepA serves to stabilize DnaA binding to distant *dnaA* boxes, leading to strand melting (20). Plasmid P1's *ori* has two sets of tandem *dnaA* boxes at each end; DnaA binding loops up the DNA, leading to preferential loading of DnaB to one of the strands (21). By contrast, RK2's TrfA was shown to mediate open complex formation and DnaB helicase loading in the absence of *dnaA* boxes, although the presence of DnaA protein was still required (22).

As mentioned above, the double strand melts in response to iteron binding by Rep protein. Melting occurs at an AT-rich region. Similar to chromosomal *oriC*, AT-rich segments of sequence frequently have sites for host factors playing an architectural role such as histone-like protein, integration host factor, and factor for inversion stimulation. These host factors help with DNA melting and with the structural organization of the initiation complex (1, 23, 24).

Class B Theta Replication

Class B theta plasmids include ColE1 and ColE1-like plasmids, which are frequently used for recombinant gene expression. Unlike class A, class B plasmids rely exclusively on host factors for both double-strand melting and primer synthesis. The DNA duplex is opened in this case by transcription of a long (~600 bp) preprimer called RNA II, which is transcribed from a constitutive promoter P2. Constitutive expression from this promoter is enhanced by a 9-bp motif *5′*-AAGATCTTC, which is located immediately upstream of the -35 box (25). The *3′* end of the preprimer RNA forms a stable hybrid with the *5′* end of the lagging-strand DNA template of *ori*. This stable RNA-DNA hybridization (R-loop formation) is facilitated by the pairing of a stretch of G-rich sequence on the transcript with a C-rich stretch on the lagging-strand DNA template and by a hairpin structure located between the G- and C-rich stretches (26). Following R-loop formation, the RNA preprimer is processed by RNase H (which recognizes the AAAAA motif in RNAII), producing a free 3′-OH end. Extension of this RNA primer by Pol I initiates leading-strand synthesis. The point where the RNA primer is extended (known as RNA/DNA switch) is considered the replication start point (reviewed in references 27–29).

As mentioned above, the nascent leading strand separates the two strands of the DNA duplex and can hybridize with the leading-strand template, forming a D-loop. PriA is recruited to the forked structure of the D-loop; alternatively, PriA can be recruited to hairpin structures forming on the lagging-strand template when the duplex opens. Indeed, *priA* strains do not support ColE1 plasmid replication, and hypomorphic mutations in *priA priB* result in a reduced ColE1 plasmid copy number (30–32).

When the Pol III holoenzyme is loaded (27, 28) this polymerase continues leading-strand synthesis and initiates lagging-strand synthesis. Pol III replication of the lagging strand toward the RNA II sequence is arrested 17 bp upstream of the DNA/RNA switch, at a site known at *terH*, ensuring unidirectional replication (33). Lagging-strand replication by Pol III appears to end a few hundred nucleotides upstream of the *terH* site (33), leaving a gap that is filled by Pol I (16).

The only step that is essential in this process of replication initiation is R-loop formation; deficits in RNase H and/or Pol I do not prevent initiation, although they have a substantial impact on the efficiency of replication initiation. In the absence of RNase H, unprocessed transcripts can still be extended with some frequency, and in the absence of Pol I, the Pol III replisome can still be loaded on an R-loop formed by the transcript and lagging-strand template (28).

R-loop formation can happen as a result of local supercoiling in the trail of the advancing RNA polymerase during transcription and is highly deleterious because R-loops block transcription and the elongation step during translation (34). Therefore, cells have mechanisms to suppress unscheduled R-loop formation. The most important ones are relaxation of the DNA template by type I topoisomerase activity, RNA degradation by RNase H, RecG dissociation of R-loops by branch migration, factor-dependent transcriptional termination, and coupling transcription to translation (reviewed in reference 35). Accordingly, titration of R-loop-suppressing factors through uncoupling transcription from translation (by starvation, temperature shift, or chloramphenicol treatment) results in increased ColE1 plasmid copy number (36), whereas

RecG overexpression dramatically suppresses replication initiation (37). However, loss of topoisomerase I and RNase H activity do not increase plasmid copy number despite inducing increased R-loop formation because these activities are also required for plasmid replication initiation (particularly RNase H).

Hybrid Classes of Theta Replication (Classes C and D)

Classes C and D have specialized priming mechanisms combined with elements of class A and class B replication. Like class A plasmids, class C and D plasmids have Rep proteins, located immediately upstream of *ori*. Like class B plasmids, however, both initiate leading-strand synthesis by Pol I extension of a free 3′-OH. Class C and D plasmids both have termination signals in the 3′ direction of lagging-strand synthesis, making replication of these plasmids unidirectional.

Class C and D theta plasmid replication is based on the evolution of more efficient ways to prime replication initiation. The evolution of plasmid-specific primases exploits the specificity provided by Rep interaction with *ori* to minimize the size of the *cis-ori* sequence. Such specificity is not possible when multiple primers are needed, as in the case of lagging-strand synthesis in the chromosome. Also, the evolution of specialized priming mechanisms broadens the host range of these plasmids by reducing dependence on host factors (38).

Class C includes ColE2 and ColE3 plasmids. The *ori*s for these two plasmids are the smallest described so far (32 bp for ColE2 and 33 bp for ColE3); these two *ori*s differ only at two positions, one of which determines plasmid specificity (39). ColE2 and ColE3 *oris* have two iterons and show two discrete functional subregions: one specializing in stable binding of the Rep protein (region I) and the other specializing in initiation of DNA replication (region III), with an area of overlap in between (region II) (40). Unlike class A initiator Rep proteins, the Rep protein in class C plasmids has primase activity, synthesizing a unique primer RNA (ppApGpA) that is extended by Pol I at a fixed site in the origin region (41). Class C replication is unidirectional, as the 3′ end of the lagging-strand DNA fragment was mapped to a specific site at the end of the *ori* region. The Rep protein may stay bound to the *ori* after initiation of replication, blocking progression of the replisome synthesizing the lagging strand (42).

Class D includes large, low-copy streptococcal plasmids that replicate in a broad range of Gram-positive bacteria. Examples include pAMβ1 from *Enterococcus faecalis*, pIP501 from *Streptococcus agalactiae*, and pSM19035 from *Streptococcus pyogenes*. In these plasmids, replication shares some features with class B theta replication, specifically a requirement for transcription across the *ori* sequence, Pol I extension and PriA-dependent replisome assembly (43). In this case, the transcript is generated from a promoter controlling expression of rep, which is immediately upstream of the *ori* (43). The replication process has been studied in detail for pAMβ1, although the Rep proteins (RepE for pAMβ1, RepF for pIP501, and RepS for pSM19035) are 97% identical for all three plasmids, and the three plasmids share a replisome structure, suggesting that they share mechanisms for replication initiation and termination. Replication depends on transcription through the origin. Rep binds specifically and rapidly to a unique site immediately upstream of the replication initiation site. This binding denatures an AT-rich sequence immediately downstream of the binding site to form an open complex (44). Compared to class A, this open complex is atypical on several counts: (i) the cognate sequence does not have multiple iterons, (ii) binding does not induce strong bending of the origin, and (iii) melting does not require additional host factors. In addition to opening of the double strand, RepE appears to have an active role in primer processing, as melting increases RepE binding and RepE can cleave transcripts from the repE operon in close proximity to the RNA/DNA switch (45).

Class D replisome assembly is PriA-dependent. A primosome assembly signal can be found 150 nucleotides (nt) downstream from the *ori* on the lagging-strand template. There is a site for replication arrest induced by Topb, a plasmid-encoded topoisomerase related to topo III, 190 nt downstream for the *ori* (46). A second replication arrest site can be found 230 nt downstream from the plasmid *ori*; in this case arrest is caused by collision with a site-specific resolvase, Resb, which is a plasmid-borne gene responsible for plasmid segregation stability (47). The presence of two independent checkpoints for Pol I progression in pAMβ1 is intriguing; this may be a mechanism that ensures Pol I availability for chromosomal replication and/or that facilitates recruitment of PriA, as PriA is known to be recruited to sites of replication fork arrest. In any case the two replication blocks appear to be largely redundant, as Topb is dispensable for pAMBβ1 replication (46).

COMPARISON OF THE THETA AND STRAND-DISPLACEMENT MODES OF PLASMID REPLICATION

Plasmids that replicate using the strand-displacement mode of replication include *E. coli* incompatibility

group Q (IncQ) plasmids of γ-proteobacteria such as RSF1010. Strand-displacement replication depends on a specialized primase: RepB. In this case, the function of replication initiator function is provided by a different Rep (RepC). Similar to initiator Rep proteins in class A theta plasmids, Rep C binds cognate iteron sequences, bending the DNA and melting duplex DNA at an adjacent A+T-rich region. An additional plasmid-encoded protein (a helicase, RepA) helps melt the DNA, recruit Pol III, and support continuous replication of one strand. This single-stranded replication produces a daughter ssDNA strand, which separates the two strands of the DNA duplex and allows hybridization with one of them, creating a D-loop (hence the name of this mode of replication).

A model for strand-displacement replication is presented in Fig. 1. After RepC-induced melting of the duplex, RepA monomers assemble around the exposed ssDNA and catalyze bidirectional unwinding of the DNA. This exposes the two different *ssi* sites, which are adjacent and are both palindromic, resulting in inverted repeats on the two DNA strands. When these two sites are exposed in single-stranded configuration, base-pair complementarity favors the formation of two hairpins, one for each strand, (Fig. 1, panel II) (48). Hairpin formation is assisted by a slowdown in RepA progression at a G+C-rich region (reviewed in reference 49). The base of each hairpin contains the start point for DNA synthesis, which is recognized by Rep B, and primer synthesis ensues (50, 51). The Pol III holoenzyme extends off of the synthesized primer (Fig. 1, panels III to V). Initiation can occur at either site independently and is continuous. As replication progresses, facilitated by the RepA helicase, a theta-type intermediate forms (Fig. 1, panels III and IV). Ligation of the two daughter strands produces two double-stranded circles (Fig. 1, panel VI).

Unlike theta-type replication, strand-displacement replication initiation is independent of host factors. This autonomous replication initiation gives these plasmids a very broad range of operation (52). As mentioned above, strand-displacement replication initiation has some similarities to class C theta plasmid replication (with a specialized, plasmid-encoded primase) and similarities to class A theta plasmid replication (with a Rep initiator involved in melting the duplex), but strand displacement presents three major differences relative to theta plasmid replication: (i) no involvement of DnaBC, as RepA is loaded on *ssi* sites exposed in the ssDNA configuration, recruiting the replicase; (ii) priming is carried out by RepB, functionally replacing the host primase DnaG; and (iii) Pol III replicates each strand continuously, initiating at two single-stranded motifs located on opposite strands (*ssiA* and *ssiB*). Note that continuous replication includes the lagging strand, which in this case does not involve synthesis of Okazaki primers (53).

REGULATION OF REPLICATION INITIATION

The frequency of replication initiation is regulated by negative feedback loop mechanisms. These regulatory mechanisms allow for rapid expansion when plasmids colonize a new permissive cell (establishment phase) and later tune the frequency of replication so that, on average, there is one replicative event per plasmid copy number per cell cycle (steady state phase), minimizing fluctuations in copy number (54).

Types of Feedback Regulatory Mechanisms

Plasmid copy number regulation needs mechanisms to monitor the plasmid copy number through a "sensor" and mechanisms to modulate replication initiation in response to feedback through an "effector" (55). The sensor mechanism depends on molecules whose concentration in the cytoplasm is proportional to plasmid copy number. In theta plasmids, inhibition of replication occurs at the initiation step and depends on three types of mechanisms: (i) antisense RNAs that hybridize to a complementary region of an essential RNA (countertranscribed RNAs, or ctRNAs) – dual mechanisms involving ctRNA and an additional protein repressor also occur –; (ii) Rep binding of iterons located in the Rep promoter, suppressing transcription; and (iii) steric hindrance between plasmids by interaction between Rep initiator proteins bound to different plasmids, which "handcuffs" them. Note that in all three cases sensor and effector functions are performed by the same molecule.

Countertranscribed RNA Inhibition

These feedback mechanisms share the following elements: two promoters in opposite orientations, one directing the synthesis of an RNA essential for replication and the other directing the synthesis of an inhibitor ctRNA. The ctRNA is complementary to a region near the 5′ end of the essential RNA, is typically strongly expressed, and has a short half-life, whereas its target RNA is expressed at constitutive but low levels. Examples of targets include maturation of a primer required for replication initiation (ColE1 plasmids), inhibition of repA translation (R1), and premature termination of translation of a rep mRNA (class D

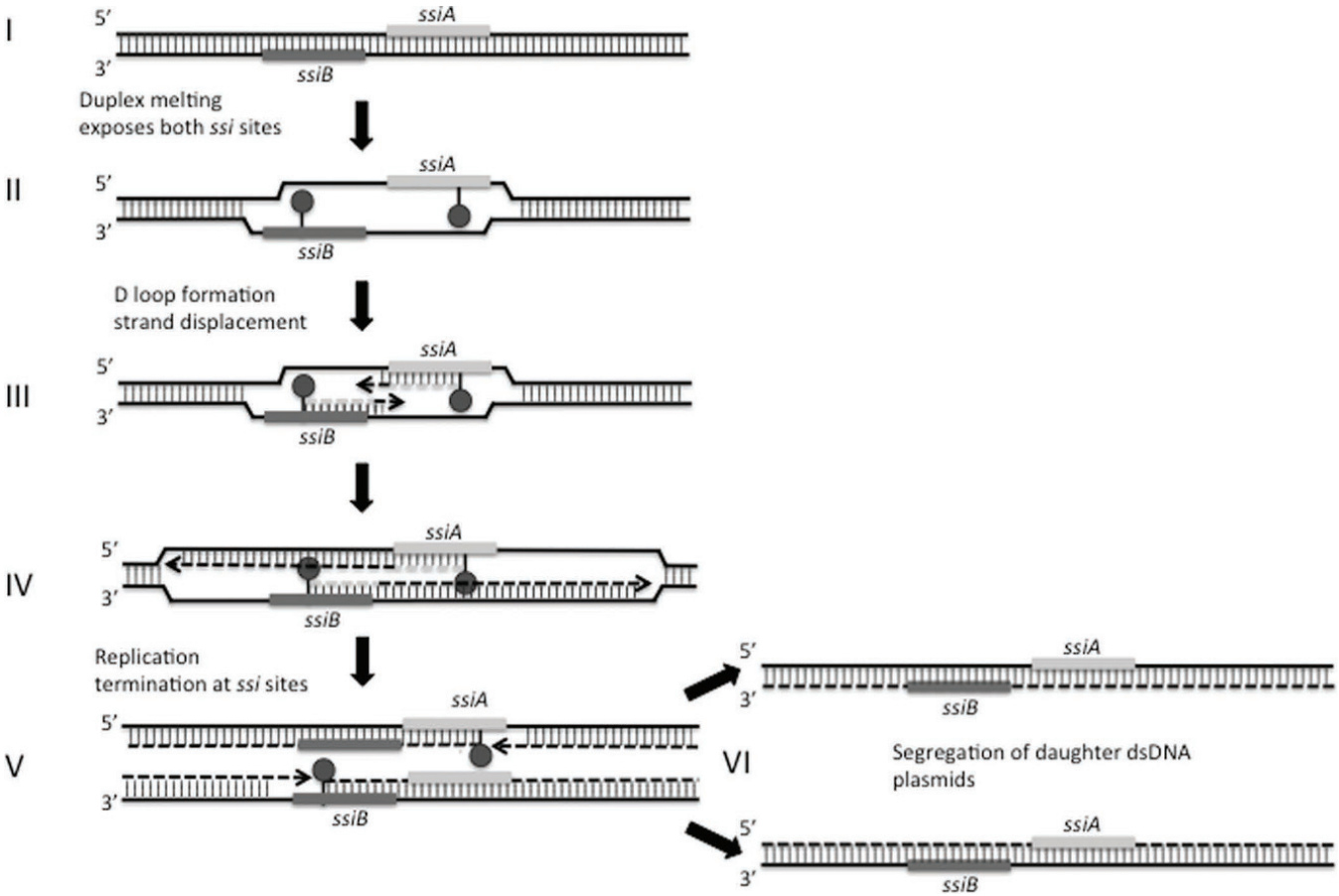

Figure 1 Model of plasmid replication by the strand-displacement mechanism. (I) Parental DNA duplex (solid black lines) depicting the two single-stranded replication initiation sites, *ssiA* (light gray box) and *ssiB* (dark gray box). Vertical lines show hybridization between DNA strands. (II) The DNA duplex is melted through binding of RepC (possibly in concert with the RepA helicase), allowing the two *ssi* sites to form hairpins (ball and stick). (III) The base of the hairpin is recognized by RepB′, which initiates the synthesis of an RNA primer (light gray dashed line). Extension of the free 3′-OH of the primer by Pol III (assisted by the RepA helicase) is shown as dashed black arrows. Two D-loops are formed, one for each direction of synthesis, as parental strands are displaced and dissociate from each other, leaving ssDNA intermediates. This is shown as areas where one of the strands has no hydrogen bonding. (IV) Synthesis continues in both directions, extending the area of D-loop formation. (V) Elongation is completed and termination of replication occurs on both strands at the *ssi* sites in which replication began. At this point, the *ssi* sites on the newly synthesized daughter strands are restored. (VI) Segregation: the two daughter strands are ligated, resulting in two DNA duplexes, each containing a parental strand (solid black line) and daughter strand (dashed black line). doi:10.1128/microbiolspec.PLAS-0029-2014.f1

plasmids). Antisense RNA regulation of plasmid replication has been extensively reviewed elsewhere (55–57).

RNAI (ColE1, ColE2) and CopA (R1) ctRNA molecules are highly structured. Given that the target preprimer and ctRNA sequence are complementary, higher-order structures for both RNAs are mirror images of each other. The first contact between sense and antisense RNAs occurs by pairing between complementary sequences at the loop portion of stem-loops, a rate-limiting step known as the "kissing complex formation" (58). Point mutations at the loop portion of stem-loops are frequently tolerated, as mutations in the template DNA introduce complementary changes in sense and ctRNA at the same time, preserving base-pairing. These mutations modulate the affinity of sense RNA–ctRNA interaction, with A-U pairs generally decreasing affinity relative to G-C pairs (for ColE1 plasmids reviewed in reference 27).

Several ctRNAs (ColE1 and ColE2 RNAI and R1 copA) have a short half-life due to the presence of an RNase E cleavage site, which consists of the U-rich sequence and a hairpin structure at the 3′ end. Conditional expression of a hyperactive variant of RNase E has been used for controlled overproduction of ColE1 plasmid DNA (59). RNase E cleavage produces monophosphorylated decay intermediates lacking short portions of the 5′ end. In the case of ColE1 and ColE2, these pRNAI cleavage intermediates are polyadenylated by PAPI, facilitating exonucleotidic digestion by PNPase (60, 61). Deletion of *pcnB*, the gene encoding PAPI, leads to increased cytoplasmic levels of pRNAI cleavage intermediates and to a 5- to 10-fold (ColE1) and 2-fold (ColE2) decrease in plasmid copy number (61, 62). RNase III has also been reported to degrade ColE1 RNAI upstream of RNase E (63). In ColE2, the differential stability between RNAI and its target rep-mRNA is partially due to differential exonuclease recruitment by RNase E (64).

Single Mechanisms Involving ctRNA Inhibition

In ColE1, the ctRNA (RNAI) is transcribed from P1, a promoter located 108 bp downstream from the sense promoter P2. Both preprimer and ctRNA form three stem-loops (SL1-3); the loop portion consists of six to seven unpaired residues. These residues are critical, as their pairing with their complementary counterparts initiates hybridization. Next, the 5′ end of RNAI (antitail) nucleates the hybridization between the two RNAs to form a duplex.

Hybridization between the preprimer and ctRNA leads to conformational changes in the preprimer, blocking R-loop formation further downstream, a phenomenon known as "action at a distance" (reviewed in references 27 and 65). This conformational change is mediated by the interaction of a sequence domain (β) in the preprimer with another sequence domain further downstream (γ), making the preprimer incompetent for R-loop formation. In addition to being short-lived, ColE1 RNA I has a short window of action, because as soon as RNAII is transcribed past position 200 downstream of the RNA/DNA switch, hybridization of the β domain with another sequence domain (α) forms a new loop (SL4), which makes RNAII refractory to RNAI inhibition.

SL1 to SL3 bear a structural resemblance to the cloverleaf structure of tRNAs and even have homology to the anticodon loops of 11 tRNAs (66). Competitive hybridization between tRNA and RNAI or RNAII appears to interfere with RNAI/RNAII hybrid formation (66). In addition, uncharged $tRNA^{ala}$ cleaves RNAI both *in vitro* and *in vivo* (67), and there is evidence suggesting that the 3′-CAA terminus of uncharged tRNAs hybridizes stably with RNAI (68). This functional cross-talk between RNAI and tRNAs may contribute to plasmid copy number deregulation associated with amino-acid starvation in *relA* strains used for recombinant gene expression; one of the key factors is the limiting yield of large-scale recombinant expression (69). Cross-talk between ctRNA and tRNAs may also explain the conservation of the 5′-UUGGCG-3′ sequence at the loop region of many of the antisense RNAs and their targets involved in regulation of replication, suggesting that this sequence is under common and strong selective pressure (70).

In ColE2 plasmids the ctRNA is also known as RNAI and has a complex secondary structure. In this case, RNAI is complementary to the 5′ end of rep mRNA containing an untranslated sequence. Given that the 5′ end portion of RNAI does not cover the initiation codon of Rep or its immediate vicinity, inhibition in this case appears to be caused by structural disruption of secondary or tertiary structures required for translation (70).

Dual Mechanisms Involving ctRNA

These mechanisms are plasmid copy number regulatory systems that include two elements: a ctRNA and a transcriptional repressor protein. In these systems, Rep expression is controlled by a strong, repressor-regulated promoter so that there is a high rate of Rep transcription when the repressor does not operate. The two best-studied examples are the R1 plasmid, where the ctRNA is CopA and the repressor is CopB, and pIP501, where the ctRNA is RNAIII and the repressor is CopR. These dual mechanisms may represent an advantage during the establishment phase, particularly for mobilizable plasmids such as class D plasmids.

In R1, repA can be transcribed from an upstream promoter P1 or from an alternative promoter further downstream, P2. Expression of repA is translationally coupled to that of tap, a small leader peptide. CopA inhibits repA expression by inhibiting translation of tap. The second element is a transcriptional repressor of P2, CopB. CopB expression is under the control of P1 but not P2. When levels of CopB are high, tap+repA are transcribed as polycystron copB-tap-RepA RNA from P1 (as P2 is silenced by CopB), but when they are low, the P2 promoter becomes derepressed and tap+repA can also be expressed from that alternative promoter, leading to a transient increase in tap+repA expression (71).

Class D plasmids have a cop-ctRNA-rep modular structure. In this case the two regulatory elements are RNAIII and a Cop protein. RNAIII is transcribed in the

opposite orientation relative to its target DNA (5′ end of rep) from promoter pIII, whereas pI and pII control CopR and Rep expression, respectively, in the sense orientation. In pIP501 plasmids, RNAIII hybridization to its complementary sequence induces folding of RNA into a transcriptional terminator structure that prevents transcription of repR. This mechanism only operates on nascent (<260-nt-long) RNAs, as longer rep transcripts form an alternative secondary structure that is refractory to repR-induced transcriptional attenuation (72). CopR (whose levels reflect plasmid copy number in the cell) inhibits the sense promoter pII. A decreased plasmid copy number leads to pII derepression, resulting in increased RepR expression. In addition, induction of pII (*repR*) transcription results in a substantial decrease in pIII transcription because pIII is supercoiling-sensitive. In pAMBβ1, CopF (the equivalent of CopR), in addition to suppressing RepF transcription, decreases primer formation since CopF transcription generates the primer for replication initiation (see class D in the "Hybrid Classes of Theta Replication" section above).

Transcriptional Regulation by Rep Binding

In some class A theta plasmids, a different mechanism of regulation involves inhibiting Rep transcription by Rep itself. In these plasmids, iterons are located in the promoter of the Rep operon, outside the plasmid *ori*. Rep binding to these cognate sequences inhibits Rep expression and thus acts as an autoregulatory mechanism.

Rep binding of two alternative binding sites (Rep promoter and plasmid *ori*) involves changes in the conformation and oligomerization status of the Rep protein. These changes have been studied in detail in the RepA protein of pPS10 (73). This protein has two winged-helix domains (WH1 and WH2). When Rep A is in dimeric form, it acts as a transcriptional repressor, with the WH1 domain functioning as a dimerization interface. Low concentrations of RepA favor dissociation of Rep dimers into monomers, which are the only form that is active as an initiator. Monomerization involves conversion of the dimerization domain into a second origin-binding sequence and remodeling of the WH1 sequence to bind the opposite iteron end (73). In some cases, monomerization can be assisted by chaperones or by the allosteric effect of binding iterons at the *ori* (74–77).

Steric Hindrance

A different feedback mechanism, known as steric hindrance or handcuffing, was initially proposed for P1 and R6K plasmids (78, 79) but could operate in more iteron-containing plasmids. According to this model, as the number of plasmids in the cell increases, Rep molecules bound to iterons of one origin begin to interact with similar complexes generated in other origins. This pairing (known as handcuffing) produces plasmid pairs linked through Rep-Rep interactions, causing a steric hindrance to both origins that interferes with origin melting (80). Rep molecules are paired through zipping-up DNA-bound RepA monomers (78). A difference between this model and the autoregulation model is that the rate of replication depends on iteron concentration, not Rep expression level. Both mechanisms of autoregulation could be working together for initiators that are limiting (81).

CONCLUDING REMARKS

Plasmids contribute to the adaptation of bacterial hosts to an ever-changing environment through mobilization and amplification of selected genes. Different circular plasmids show differences in duplex melting, leading-strand priming, and lagging-strand synthesis. Learning more about the diversity of the replication mechanisms present in plasmids can help us understand the mechanisms that cells have available to replicate and repair their DNA. Organellar replication and restoration of replication after replication fork arrest are two examples of processes that occur in cells that are mechanistically closely related to plasmid replication. Also, learning more about these mechanisms will improve our understanding of plasmid biology, as mechanisms of replication limit plasmid size, host range, and mobilization capacity. Finally, maintaining a stable plasmid copy number is critical for the host, as loss of the plasmid entails losing the adaptive functions carried in the plasmid sequence, and runaway plasmid replication is lethal. Thus, mechanisms of plasmid replication regulation represent potential targets for antimicrobial intervention.

Acknowledgments. *We thank Saleem Khan and Douglas Rawlings for critical reading of the manuscript. This work was partially supported by a National Cancer Institute award (K08CA116429), by a National Institute of Environmental Health Sciences grant (R01ES019625) award, and by an NSF Advances in Biological Informatics Innovation grant (ABI1262435). Conflicts of interest: We disclose no conflicts.*

Citation. Lilly J, Camps M. 2015. Mechanisms of theta plasmid replication. Microbiol Spectrum **3**(1):PLAS-0029-2014.

References

1. **Giraldo R, Fernandez-Tresguerres ME.** 2004. Twenty years of the pPS10 replicon: insights on the molecular mechanism for the activation of DNA replication in iteron-containing bacterial plasmids. *Plasmid* **52:** 69–83.

2. **Kittleson JT, Cheung S, Anderson JC.** 2011. Rapid optimization of gene dosage in *E. coli* using DIAL strains. *J Biol Eng* **5:**10.
3. **Gregg AV, McGlynn P, Jaktaji RP, Lloyd RG.** 2002. Direct rescue of stalled DNA replication forks via the combined action of PriA and RecG helicase activities. *Mol Cell* **9:**241–251.
4. **Kogoma T.** 1997. Stable DNA replication: interplay between DNA replication, homologous recombination, and transcription. *Microbiol Mol Biol Rev* **61:**212–238.
5. **Sandler SJ, Marians KJ.** 2000. Role of PriA in replication fork reactivation in *Escherichia coli*. *J Bacteriol* **182:** 9–13.
6. **Masai H, Arai K.** 1996. DnaA- and PriA-dependent primosomes: two distinct replication complexes for replication of *Escherichia coli* chromosome. *Front Biosci* **1:** d48–d58.
7. **Wu CA, Zechner EL, Marians KJ.** 1992. Coordinated leading- and lagging-strand synthesis at the *Escherichia coli* DNA replication fork. I. Multiple effectors act to modulate Okazaki fragment size. *J Biol Chem* **267:** 4030–4044.
8. **Johnson A, O'Donnell M.** 2005. Cellular DNA replicases: components and dynamics at the replication fork. *Annu Rev Biochem* **74:**283–315.
9. **McHenry CS.** 2011. DNA replicases from a bacterial perspective. *Annu Rev Biochem* **80:**403–436.
10. **Tanner NA, Hamdan SM, Jergic S, Loscha KV, Schaeffer PM, Dixon NE, van Oijen AM.** 2008. Single-molecule studies of fork dynamics in *Escherichia coli* DNA replication. *Nat Struct Mol Biol* **15:**998.
11. **Lenhart JS, Schroeder JW, Walsh BW, Simmons LA.** 2012. DNA repair and genome maintenance in *Bacillus subtilis*. *Microbiol Mol Biol Rev* **76:**530–564.
12. **Rannou O, Le Chatelier E, Larson MA, Nouri H, Dalmais B, Laughton C, Janniere L, Soultanas P.** 2013. Functional interplay of DnaE polymerase, DnaG primase and DnaC helicase within a ternary complex, and primase to polymerase hand-off during lagging strand DNA replication in *Bacillus subtilis*. *Nucleic Acids Res* **41:**5303–5320.
13. **Allen JM, Simcha DM, Ericson NG, Alexander DL, Marquette JT, Van Biber BP, Troll CJ, Karchin R, Bielas JH, Loeb LA, Camps M.** 2011. Roles of DNA polymerase I in leading and lagging-strand replication defined by a high-resolution mutation footprint of ColE1 plasmid replication. *Nucleic Acids Res* **39:**7020–7033.
14. **Patel PH, Suzuki M, Adman E, Shinkai A, Loeb LA.** 2001. Prokaryotic DNA polymerase I: evolution, structure, and "base flipping" mechanism for nucleotide selection. *J Mol Biol* **308:**823–837.
15. **Maki H, Bryan SK, Horiuchi T, Moses RE.** 1989. Suppression of dnaE nonsense mutations by pcbA1. *J Bacteriol* **171:**3139–3143.
16. **Troll C, Yoder J, Alexander D, Hernandez J, Loh Y, Camps M.** 2014. The mutagenic footprint of low-fidelity Pol I ColE1 plasmid replication in *E. coli* reveals an extensive interplay between Pol I and Pol III. *Curr Genet* **60:**123–134.
17. **Mukhopadhyay G, Chattoraj DK.** 1993. Conformation of the origin of P1 plasmid replication. Initiator protein induced wrapping and intrinsic unstacking. *J Mol Biol* **231:**19–28.
18. **Urh M, Wu J, Wu J, Forest K, Inman RB, Filutowicz M.** 1998. Assemblies of replication initiator protein on symmetric and asymmetric DNA sequences depend on multiple protein oligomerization surfaces. *J Mol Biol* **283:** 619–631.
19. **Filutowicz M, Dellis S, Levchenko I, Urh M, Wu F, York D.** 1994. Regulation of replication of an iteron-containing DNA molecule. *Prog Nucleic Acid Res Mol Biol* **48:**239–273.
20. **Stenzel TT, MacAllister T, Bastia D.** 1991. Cooperativity at a distance promoted by the combined action of two replication initiator proteins and a DNA bending protein at the replication origin of pSC101. *Genes Dev* **5:** 1453–1463.
21. **Park K, Chattoraj DK.** 2001. DnaA boxes in the P1 plasmid origin: the effect of their position on the directionality of replication and plasmid copy number. *J Mol Biol* **310:**69–81.
22. **Doran KS, Helinski DR, Konieczny I.** 1999. Host-dependent requirement for specific DnaA boxes for plasmid RK2 replication. *Mol Microbiol* **33:**490–498.
23. **del Solar G, Giraldo R, Ruiz-Echevarria MJ, Espinosa M, Diaz-Orejas R.** 1998. Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62:**434–464.
24. **Rakowski SA, Filutowicz M.** 2013. Plasmid R6K replication control. *Plasmid* **69:**231–242.
25. **Wu YC, Liu ST.** 2010. A sequence that affects the copy number and stability of pSW200 and ColE1. *J Bacteriol* **192:**3654–3660.
26. **Masukata H, Tomizawa J.** 1986. Control of primer formation for ColE1 plasmid replication: conformational change of the primer transcript. *Cell* **44:**125–136.
27. **Camps M.** 2010. Modulation of ColE1-like plasmid replication for recombinant gene expression. *Recent Pat DNA Gene Seq* **4:**58–73.
28. **Cesareni G, Helmer-Citterich M, Castagnoli L.** 1991. Control of ColE1 plasmid replication by antisense RNA. *Trends Genet* **7:**230–235.
29. **Wang Z, Yuan Z, Hengge UR.** 2004. Processing of plasmid DNA with ColE1-like replication origin. *Plasmid* **51:** 149–161.
30. **Lee EH, Kornberg A.** 1991. Replication deficiencies in priA mutants of *Escherichia coli* lacking the primosomal replication n' protein. *Proc Natl Acad Sci USA* **88:** 3029–3032.
31. **Sandler SJ, Samra HS, Clark AJ.** 1996. Differential suppression of priA2::kan phenotypes in *Escherichia coli* K-12 by mutations in priA, lexA, and dnaC. *Genetics* **143:**5–13.
32. **Jaktaji RP, Lloyd RG.** 2003. PriA supports two distinct pathways for replication restart in UV-irradiated *Escherichia coli* cells. *Mol Microbiol* **47:**1091–1100.
33. **Nakasu S, Tomizawa J.** 1992. Structure of the ColE1 DNA molecule before segregation to daughter molecules. *Proc Natl Acad Sci USA* **89:**10139–10143.

34. **Drolet M.** 2006. Growth inhibition mediated by excess negative supercoiling: the interplay between transcription elongation, R-loop formation and DNA topology. *Mol Microbiol* **59:**723–730.
35. **Drolet M, Broccoli S, Rallu F, Hraiky C, Fortin C, Masse E, Baaklini I.** 2003. The problem of hypernegative supercoiling and R-loop formation in transcription. *Front Biosci* **8:**d210–d221.
36. **Gowrishankar J, Harinarayanan R.** 2004. Why is transcription coupled to translation in bacteria? *Mol Microbiol* **54:**598–603.
37. **Fukuoh A, Iwasaki H, Ishioka K, Shinagawa H.** 1997. ATP-dependent resolution of R-loops at the ColE1 replication origin by *Escherichia coli* RecG protein, a Holliday junction-specific helicase. *EMBO J* **16:**203–209.
38. **del Solar G, Alonso JC, Espinosa M, Diaz-Orejas R.** 1996. Broad-host-range plasmid replication: an open question. *Mol Microbiol* **21:**661–666.
39. **Yasueda H, Horii T, Itoh T.** 1989. Structural and functional organization of ColE2 and ColE3 replicons. *Mol Gen Genet* **215:**209–216.
40. **Aoki K, Shinohara M, Itoh T.** 2007. Distinct functions of the two specificity determinants in replication initiation of plasmids ColE2-P9 and ColE3-CA38. *J Bacteriol* **189:** 2392–2400.
41. **Takechi S, Matsui H, Itoh T.** 1995. Primer RNA synthesis by plasmid-specified Rep protein for initiation of ColE2 DNA replication. *EMBO J* **14:**5141–5147.
42. **Takechi S, Itoh T.** 1995. Initiation of unidirectional ColE2 DNA replication by a unique priming mechanism. *Nucleic Acids Res* **23:**4196–4201.
43. **Bruand C, Ehrlich SD.** 1998. Transcription-driven DNA replication of plasmid pAMbeta1 in *Bacillus subtilis*. *Mol Microbiol* **30:**135–145.
44. **Le Chatelier E, Janniere L, Ehrlich SD, Canceill D.** 2001. The RepE initiator is a double-stranded and single-stranded DNA-binding protein that forms an atypical open complex at the onset of replication of plasmid pAMbeta 1 from Gram-positive bacteria. *J Biol Chem* **276:**10234–10246.
45. **Brantl S.** 2014. Antisense-RNA mediated control of plasmid replication: pIP501 revisited. *Plasmid*. [Epub ahead of print.] doi:10.1016/j.plasmid.2014.07.004.
46. **Bidnenko V, Ehrlich SD, Janniere L.** 1998. *In vivo* relations between pAMbeta1-encoded type I topoisomerase and plasmid replication. *Mol Microbiol* **28:**1005–1016.
47. **Janniere L, Bidnenko V, McGovern S, Ehrlich SD, Petit MA.** 1997. Replication terminus for DNA polymerase I during initiation of pAM beta 1 replication: role of the plasmid-encoded resolution system. *Mol Microbiol* **23:** 525–535.
48. **Sakai H, Komano T.** 1996. DNA replication of IncQ broad-host-range plasmids in Gram-negative bacteria. *Biosci Biotechnol Biochem* **60:**377–382.
49. **Loftie-Eaton W, Rawlings DE.** 2012. Diversity, biology and evolution of IncQ-family plasmids. *Plasmid* **67:**15–34.
50. **Honda Y, Akioka T, Takebe S, Tanaka K, Miao D, Higashi A, Nakamura T, Taguchi Y, Sakai H, Komano T.** 1993. Mutational analysis of the specific priming signal essential for DNA replication of the broad host-range plasmid RSF1010. *FEBS Lett* **324:**67–70.
51. **Miao DM, Honda Y, Tanaka K, Higashi A, Nakamura T, Taguchi Y, Sakai H, Komano T, Bagdasarian M.** 1993. A base-paired hairpin structure essential for the functional priming signal for DNA replication of the broad host range plasmid RSF1010. *Nucleic Acids Res* **21:**4900–4903.
52. **Rawlings DE, Tietze E.** 2001. Comparative biology of IncQ and IncQ-like plasmids. *Microbiol Mol Biol Rev* **65:**481–496.
53. **Tanaka K, Kino K, Taguchi Y, Miao DM, Honda Y, Sakai H, Komano T, Bagdasarian M.** 1994. Functional difference between the two oppositely oriented priming signals essential for the initiation of the broad host-range plasmid RSF1010 DNA replication. *Nucleic Acids Res* **22:**767–772.
54. **Nordstrom K, Wagner EG.** 1994. Kinetic aspects of control of plasmid replication by antisense RNA. *Trends Biochem Sci* **19:**294–300.
55. **del Solar G, Espinosa M.** 2000. Plasmid copy number control: an ever-growing story. *Mol Microbiol* **37:**492–500.
56. **Wagner EG, Altuvia S, Romby P.** 2002. Antisense RNAs in bacteria and their genetic elements. *Adv Genet* **46:** 361–398.
57. **Brantl S.** 2014. Plasmid replication control by antisense RNAs. *In* Tolmasky ME, Alonso JC (ed), *Plasmids: Biology and Impact in Biotechnology and Discovery*. ASM Press, Washington, DC. In press.
58. **Hjalt TA, Wagner EG.** 1995. Bulged-out nucleotides in an antisense RNA are required for rapid target RNA binding *in vitro* and inhibition *in vivo*. *Nucleic Acids Res* **23:**580–587.
59. **Go H, Lee K.** 2011. A genetic system for RNase E variant-controlled overproduction of ColE1-type plasmid DNA. *J Biotechnol* **152:**171–175.
60. **Nishio SY, Itoh T.** 2008. Replication initiator protein mRNA of ColE2 plasmid and its antisense regulator RNA are under the control of different degradation pathways. *Plasmid* **59:**102–110.
61. **Xu FF, Gaggero C, Cohen SN.** 2002. Polyadenylation can regulate ColE1 type plasmid copy number independently of any effect on RNAI decay by decreasing the interaction of antisense RNAI with its RNAII target. *Plasmid* **48:**49–58.
62. **Nishio SY, Itoh T.** 2008. The effects of RNA degradation enzymes on antisense RNAI controlling ColE2 plasmid copy number. *Plasmid* **60:**174–180.
63. **Binnie U, Wong K, McAteer S, Masters M.** 1999. Absence of RNASE III alters the pathway by which RNAI, the antisense inhibitor of ColE1 replication, decays. *Microbiology* **145**(Pt 11):3089–3100.
64. **Nishio SY, Itoh T.** 2009. Arginine-rich RNA binding domain and protein scaffold domain of RNase E are important for degradation of RNAI but not for that of the Rep mRNA of the ColE2 plasmid. *Plasmid* **62:**83–87.
65. **Polisky B.** 1988. ColE1 replication control circuitry: sense from antisense. *Cell* **55:**929–932.

66. **Yavachev L, Ivanov I.** 1988. What does the homology between *E. coli* tRNAs and RNAs controlling ColE1 plasmid replication mean? *J Theor Biol* **131:**235–241.

67. **Wang Z, Yuan Z, Xiang L, Shao J, Wegrzyn G.** 2006. tRNA-dependent cleavage of the ColE1 plasmid-encoded RNA I. *Microbiology* **152:**3467–3476.

68. **Wang Z, Le G, Shi Y, Wegrzyn G, Wrobel B.** 2002. A model for regulation of ColE1-like plasmid replication by uncharged tRNAs in amino acid-starved *Escherichia coli* cells. *Plasmid* **47:**69–78.

69. **Grabherr R, Nilsson E, Striedner G, Bayer K.** 2002. Stabilizing plasmid copy number to improve recombinant protein production. *Biotechnol Bioeng* **77:**142–147.

70. **Hiraga S, Sugiyama T, Itoh T.** 1994. Comparative analysis of the replicon regions of eleven ColE2-related plasmids. *J Bacteriol* **176:**7233–7243.

71. **Malmgren C, Engdahl HM, Romby P, Wagner EG.** 1996. An antisense/target RNA duplex or a strong intramolecular RNA structure *5′* of a translation initiation signal blocks ribosome binding: the case of plasmid R1. *RNA* **2:**1022–1032.

72. **Brantl S, Birch-Hirschfeld E, Behnke D.** 1993. RepR protein expression on plasmid pIP501 is controlled by an antisense RNA-mediated transcription attenuation mechanism. *J Bacteriol* **175:**4052–4061.

73. **Giraldo R, Fernandez-Tornero C, Evans PR, Diaz-Orejas R, Romero A.** 2003. A conformational switch between transcriptional repression and replication initiation in the RepA dimerization domain. *Nat Struct Biol* **10:**565–571.

74. **DasGupta S, Mukhopadhyay G, Papp PP, Lewis MS, Chattoraj DK.** 1993. Activation of DNA binding by the monomeric form of the P1 replication initiator RepA by heat shock proteins DnaJ and DnaK. *J Mol Biol* **232:**23–34.

75. **Diaz-Lopez T, Lages-Gonzalo M, Serrano-Lopez A, Alfonso C, Rivas G, Diaz-Orejas R, Giraldo R.** 2003. Structural changes in RepA, a plasmid replication initiator, upon binding to origin DNA. *J Biol Chem* **278:**18606–18616.

76. **Kawasaki Y, Wada C, Yura T.** 1990. Roles of *Escherichia coli* heat shock proteins DnaK, DnaJ and GrpE in mini-F plasmid replication. *Mol Gen Genet* **220:**277–282.

77. **Wickner S, Skowyra D, Hoskins J, McKenney K.** 1992. DnaJ, DnaK, and GrpE heat shock proteins are required in oriP1 DNA replication solely at the RepA monomerization step. *Proc Natl Acad Sci USA* **89:**10345–10349.

78. **Gasset-Rosa F, Diaz-Lopez T, Lurz R, Prieto A, Fernandez-Tresguerres ME, Giraldo R.** 2008. Negative regulation of pPS10 plasmid replication: origin pairing by zipping-up DNA-bound RepA monomers. *Mol Microbiol* **68:**560–572.

79. **Park K, Han E, Paulsson J, Chattoraj DK.** 2001. Origin pairing ('handcuffing') as a mode of negative control of P1 plasmid copy number. *EMBO J* **20:**7323–7332.

80. **Zzaman S, Bastia D.** 2005. Oligomeric initiator protein-mediated DNA looping negatively regulates plasmid replication *in vitro* by preventing origin melting. *Mol Cell* **20:**833–843.

81. **Das N, Valjavec-Gratian M, Basuray AN, Fekete RA, Papp PP, Paulsson J, Chattoraj DK.** 2005. Multiple homeostatic mechanisms in the control of P1 plasmid replication. *Proc Natl Acad Sci USA* **102:**2856–2861.

Plasmids—Biology and Impact in Biotechnology and Discovery
Edited by Marcelo E. Tolmasky and Juan C. Alonso

doi:10.1128/microbiolspec.PLAS-0035-2014

José A. Ruiz-Masó,[1] Cristina Machón,[2,3] Lorena Bordanaba-Ruiseco,[1] Manuel Espinosa,[1] Miquel Coll,[2,3] and Gloria del Solar[1]

4 Plasmid Rolling-Circle Replication

GENERAL ASPECTS OF PLASMID ROLLING-CIRCLE REPLICATION

The main features that characterize rolling-circle replication (RCR) (see Fig. 1A) derive from its singular initiation mechanism, which relies on the sequence-specific cleavage, at the nick site of the double-strand origin (*dso*), of one of the parental DNA strands by an initiator Rep protein. This cleavage generates a 3′-OH end that allows the host DNA polymerases to initiate the leading strand replication. Therefore, the RCR initiation circumvents the synthesis of a primer RNA that is required in all other modes of replication of circular double-stranded DNA (dsDNA). Elongation of the leading strand takes place as the parental double helix is unwound by a host DNA helicase and the cleaved nontemplate strand is covered with the single-stranded DNA binding protein. Since the nascent DNA is covalently attached to the parental DNA, termination of a round of leading-strand replication implies a new cleavage event at the reconstituted nick site. This reaction is assumed to be catalyzed by the same Rep molecule that carried out the initiation cleavage and remained bound to the 5′ end of the parental strand while traveling along with the replication fork. A trans-esterification then occurs that joins this 5′ end to the 3′ end generated in the termination cleavage, releasing the displaced parental strand as a circular single-stranded DNA (ssDNA). This replicative intermediate serves as the template for the synthesis of the lagging strand, which depends solely on host-encoded enzymes and is initiated from a highly structured region of the ssDNA, termed the single-strand origin (*sso*).

Thus, the entire process of asymmetric RCR yields, in two separate steps (this is what asymmetric refers to), two circular dsDNAs containing either the newly synthesized leading or lagging strand and the complementary parental template strand. The DNA ligase and gyrase of the host cell next convert the new daughter DNA molecules in supercoiled forms indistinguishable from the rest of the plasmid pool. Generation of the ssDNA replicative intermediates is the hallmark of RCR, and detection of intracellular strand-specific plasmid ssDNA provides valuable clues about whether a given plasmid replicates by the rolling-circle mechanism (1, 2).

The basic catalytic mechanism operating in initiation and termination of RCR, i.e., the cleavage and rejoining

[1]Centro de Investigaciones Biológicas (CSIC), Ramiro de Maeztu 9, 28040 Madrid, Spain; [2]Institute for Research in Biomedicine (IRB-Barcelona), Baldiri Reixac 10-12, 08028 Barcelona, Spain; [3]Institut de Biologia Molecular de Barcelona (CSIC), Baldiri Reixac 10-12, 08028 Barcelona, Spain.

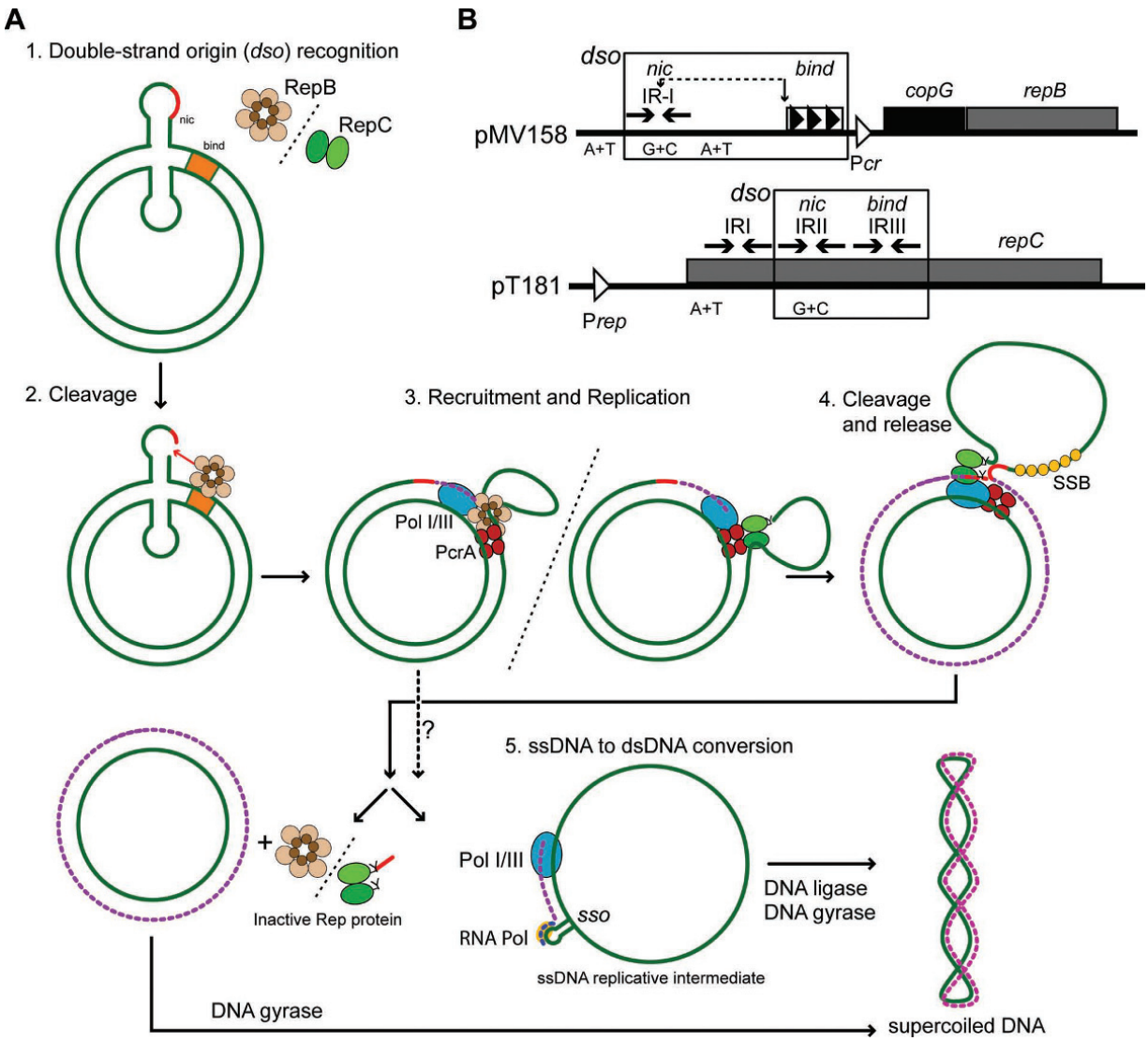

Figure 1 (A) A model for plasmid RCR based on pMV158 and pT181 replicons. Detailed information about the RCR process is given in the text. In the pMV158 replication model, a possible mechanism is shown in which, upon assembly and cleavage at the nick site, the hexameric ring of RepB encircles one of the plasmid strands within the central channel. As discussed in the text, the strand enclosure may confer high processivity to the replisome complex. The RepB-mediated mechanism that, at the termination step, yields the dsDNA replication product and the ssDNA intermediate, as well as the mechanism of RepB inactivation, remain undisclosed (dotted arrow with ? symbol). (B) Scheme of the *dso*s and of the adjacent regions of the pMV158 and pT181 RCR plasmids. The symbols used are as follows: direct repeats in the replication region are indicated by solid boxed arrows; the inverted arrows represent the two arms of the inverted repeat elements; promoters are indicated by open arrowheads. The AT- and GC-rich sequences (A+T and G+C, respectively) are also indicated. The dotted line above the pMV158 map indicates that the direct repeats of the *bind* locus are separated by 84 bp from the nick site. SSB, single-stranded DNA binding protein. doi:10.1128/microbiolspec.PLAS-0035-2014.f1

of ssDNA using an active-site Tyr that forms a transient 5′-phosphotyrosine bond with the cleaved DNA, is involved in a range of processes that take place in mobile genetic elements in all three domains of life. The enzymes that exhibit this catalytic mechanism are mainly included in the widespread His-bulky hydrophobic residue-His (HUH) endonuclease superfamily and have key roles in the replication of plasmids, bacteriophages, and plant and animal viruses; in plasmid conjugative transfer; and in transposition (3). RCR was discovered in ssDNA coliphage ΦX174 some 45 years ago (4–6). The pioneer characterization of gene A protein made the initiator of ΦX174 RCR the first member of the HUH endonuclease superfamily (7–10).

Plasmid RCR was first evidenced for the *Staphylococcus aureus* plasmid pT181 based on the characterization of the origin-specific nicking-closing activity of the purified pT181-encoded RepC protein (11). Shortly afterward, several other small plasmids from staphylococci, bacilli, streptococci, and streptomyces were also found to replicate by the RCR mechanism (12–14), which led to the assumption that most, if not all, small multicopy plasmids in Gram-positive bacteria use RCR. However, this premise proved inaccurate, as some small plasmids isolated from Gram-positive organisms were later reported to replicate by the theta mode (1). Moreover, although RCR plasmids are particularly abundant in Gram-positive bacteria, they have also been identified in various Gram-negative organisms, in archaea, and in mitochondria of the higher plant *Chenopodium album* (1, 2, 15).

Natural RCR plasmids range in size from as low as the 846 bp of the *Thermotoga* plasmids pRQ7, pMC24, and pRKU1 (16–18) to the almost 30 kb of pCG4 from *Corynebacterium glutamicum* (19). The nearly identical plasmids pRQ7, pMC24, and pRKU1 are the smallest found so far and consist of only the basic replicon, i.e., the backbone regions involved in replication and copy-number control. The basic replicon of RCR plasmids should include an essential module containing the *dso* and the genes that encode the initiator Rep protein and the replication control element(s), as well as at least one host-recognized *sso*, which, although not strictly essential, provides efficient synthesis of the lagging strand and hence is present in all natural RCR plasmids (Fig. 2). Homology in the essential module of the basic replicon has been the criterion used to classify RCR plasmids into replicon families (see below).

Apart from the basic replicon, some larger RCR plasmids contain additional backbone genes and elements that contribute to their maintenance or help them transfer between host cells (Fig. 2). Of special relevance, because of its frequent presence in RCR plasmids, is the MOB module, which is involved in the conjugative mobilization of the plasmid and consists of the transfer origin (*oriT*) and the *mob* gene(s) that encode the relaxase protein and, in some cases, auxiliary proteins (20). The apparent lack of active partition systems in RCR plasmids is consistent with the medium copy number (10 to 30 per chromosome equivalent) that they exhibit in their natural hosts. This feature ensures the stable inheritance of RCR plasmids by only random segregation to the daughter cells, providing that the replication control system efficiently corrects fluctuations of the plasmid copy number in single cells and that the plasmid molecules are maintained as individual copies. In this sense, the presence of homologs to components of toxin-antitoxin (TA) systems in some RCR plasmids is intriguing (21). It is noteworthy that whereas the TA systems were first proposed to play a role in plasmid stability through postsegregational killing of plasmid-free cells, the more recent competition hypothesis postulates that acquisition of these modules allows plasmids to exclude competing TA-free plasmids (22–24).

Some RCR plasmids also carry accessory genes that encode functions that can benefit the host cell under special conditions, thus reflecting the adaptation of the bacteria to their environment (Fig. 2). Antibiotic resistance determinants are among the most frequent traits encoded by RCR plasmids isolated from a variety of bacteria (25). Other accessory genes have been found to be relatively abundant in RCR plasmids from a given host. This is the case of small heat shock protein (*shsp*) genes carried by *Streptococcus thermophilus* plasmids belonging to the pC194 replicon family (26, 27). The presence of *shsp*-containing plasmids has been reported to increase cell survival at the high temperatures reached during different stages of fermentation in the dairy industry (27). Another striking example is the presence, in some *Bacillus thuringiensis* plasmids, of open reading frames encoding collagen-like proteins that are thought to play a role in aggregation formation or in adherence to other cells or substrates (28).

RCR plasmids are considered to contain promiscuous replicons, as many of them have been shown to replicate in species, genera, or even phyla other than those from which they were isolated (25). The simplicity of the RCR initiation, with only the plasmid-encoded Rep protein participating in recognition of the origin and priming of the leading strand synthesis, may underlie the usual promiscuity of these plasmids. The broadness of the host range of RCR plasmids would depend on the balanced expression of their essential

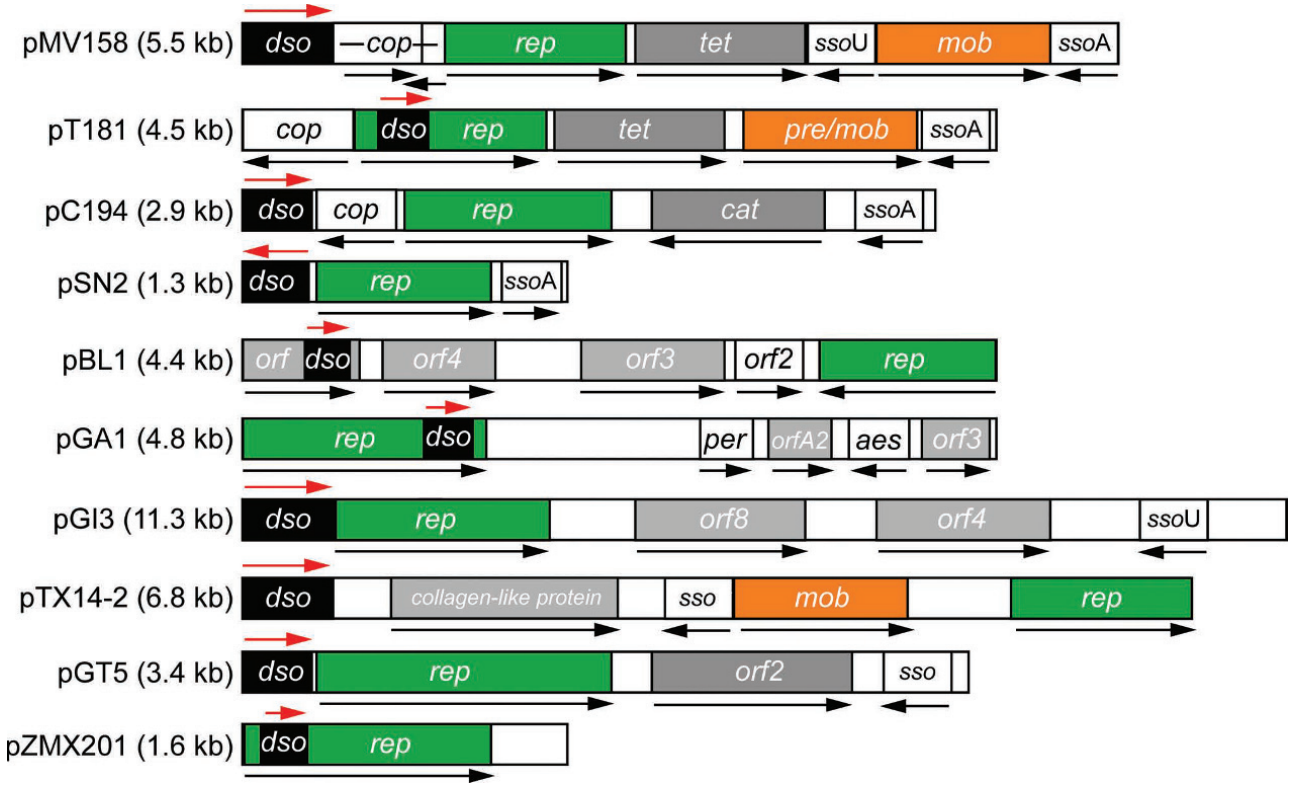

Figure 2 Functional organization of the RCR plasmids. Plasmids representative of the different families are shown. The arrows point to the direction of transcription (black) or the direction of replication (red) from the *dso* (leading strand) and *sso* (lagging strand). Inside the boxes, *rep* is the replication gene; *cop* represents the copy number control gene(s); *dso* is the double-strand origin of replication; *sso* is the single-strand origin of replication; *cat* and *tet* are chloramphenicol- and tetracycline-resistant genes, respectively; *pre/mob* represents the conjugative mobilization gene; *orf* indicates an open reading frame with unknown homology. The positions of the copy number control genes *per* and *aes* of pGA1, and of the *collagen-like protein* gene of pTX14-2 are also indicated.
doi:10.1128/microbiolspec.PLAS-0035-2014.f2

genes involved in initiation and control of replication as well as on the formation of a functional Rep-host helicase complex that can extensively unwind the plasmid DNA in a variety of bacteria (29–31). The broad host range of RCR plasmids is best exemplified by the pMV158-family prototype, which was initially isolated from *Streptococcus agalactiae* and subsequently transferred to a variety of *Firmicutes* (several *Streptococcus* and *Bacillus* species, *Listeria, S. aureus, Lactococcus lactis, Enterococcus faecalis, Clostridium*), *Actinobacteria* (*C. glutamicum, Brevibacterium*), and the γ-proteobacterium *Escherichia coli*. Moreover, the fact that members of each replicon family have been isolated from a variety of bacteria suggests the promiscuity of the ancestors from which these plasmids derive. In turn, plasmid adaptation to a new host can lead to the narrowing of the host range of the adapted plasmid. This seems to be the case for two *Mycoplasma mycoides* plasmids of the pMV158-replicon family, namely pADB201 and pKMK1, whose *rep* genes contain at least one UGA codon, which encodes tryptophan in this bacterium but is a stop codon in other bacteria, so that the host range of these plasmids is restricted to *Mycoplasma* species (32).

Due to their general smallness, high copy number, and promiscuity, RCR plasmids appear to be well suited for the construction of vectors for gene cloning and expression, provided a functional *sso* is present to minimize the generation of the recombinogenic ssDNA intermediates, which can lead to structural and segregational plasmid instability (33–37). Nevertheless, it has been reported that cloning of heterologous DNA in RCR plasmid vectors can result in the generation of linear high-molecular-weight (HMW) plasmid multimers

in relative amounts that correlate positively with the size of the DNA insert (38, 39). The formation of HMW by RCR plasmids has also been implicated in both structural (40) and segregational (41) instability. The generation of HMW plasmid DNA was at first related to a replication defect, as plasmids lacking *sso* were prone to accumulate HMW DNA (42). Accumulation of HMW plasmid DNA was enhanced in the absence of the ExoV enzyme (RecBCD in Gram-negative or AddAB in Gram-positive bacteria) (41). Despite the potential instability problems, vectors based on RCR plasmids have been developed and successfully used in pneumococci, enterococci, lactococi, and corynebacteria (43–45), for which genetic and biotechnological tools are scarce and hence welcome. It is worth mentioning that most of the nonintegrative plasmid vectors available in *Streptococcus pneumoniae* are based on pMV158 and that inducible expression vector pLS1ROM and recombinant pLS1ROM-GFP (containing the *gfp* gene, encoding the *Aequorea victoria* green fluorescent protein, cloned under control of the maltose-inducible P_M promoter) have proved to be structurally and segregationally stable in pneumococcus, even under induction conditions (45). Similarly, most of the autonomously replicating vectors for the industrial microorganism *C. glutamicum* are based on plasmids pBL1, pCG1, and pGA1 from *C. glutamicum* or on the broad-host-range plasmid pNG2 from *Corynebacterium diphtheriae*, all of them replicating by the rolling circle mode (46). These RCR plasmid vectors were found to be stably maintained in *C. glutamicum* cells grown under nonselective conditions (47).

An aspect of recognized relevance when pursuing the biotechnological use of plasmid vectors is the metabolic cost that carriage of these extrachromosomal elements imposes on the host, since a significant burden can lead to the overgrowth of the culture by plasmid-free cells even though plasmid inheritance is quite stable. Little information is available on the burden caused by RCR plasmids, as this subject has only been analyzed for the pMV158 replicon. Small (4.4 kb), medium-copy-number (~20 copies per chromosome equivalent) pMV158 derivatives that are stably inherited in pneumococcus and harbor an *sso* element efficiently recognized in this host slightly burden the *S. pneumoniae* cells, causing a 7 to 8% increase in the bacterial doubling time (48). Nevertheless, fitness impairment of pneumococcal cells harboring pMV158 derivatives has not been found to negatively affect the segregational stability of pLS1ROM and pLS1ROM-GFP (45).

This chapter aims to provide an updated review of the major findings in the study of the RCR plasmids and to highlight the pending questions and challenges for the detailed understanding of this kind of plasmid replication. Most of these issues have been dealt with in previous reviews on this subject (1, 25, 32, 49, 50).

Apart from the above-referenced asymmetric RCR, which is initiated by the Rep-mediated cleavage of one parental strand, a different, recombination-dependent replication mechanism that also leads to σ-shaped circular intermediates consisting of a circular DNA attached to a growing linear DNA has been reported to play an essential role during the replication cycle of many dsDNA viruses. Single origin-dependent replication of bacterial genomes and of many dsDNA viruses with circular genomes proceeds by the θ (circle to circle) mechanism. The trade-off between different DNA transactions could lead to the stall or collapse of the replication machinery, so that origin-independent remodeling and assembly of a new replisome at the stalled fork is required to restart the replication process. In dsDNA viruses (e.g., bacteriophage lambda, SPP1, etc.), replication restart becomes dependent on recombination proteins with a switch from the origin-mediated θ type to a σ type recombination-dependent replication. The replication shift from θ to σ generates the concatemeric viral DNA substrate needed to produce mature viral particles. This RCR-like σ mode has been reviewed by Lo Piano et al. (51) and will not be addressed here.

THE DOUBLE-STRAND ORIGIN

Replication of the leading strand of RCR plasmids initiates and proceeds in a unidirectional manner from their *dso*, a plasmid DNA region highly specific for its cognate initiator protein that contains the sequences involved in the initiation and termination of the leading strand. The *dso*, along with the *rep* gene and the control elements, is part of an essential module that harbors the functions for plasmid replication. Based on the homologies found in this essential module, up to 17 RCR plasmid families have been defined. Only three of these plasmid families have been studied in depth, their prototypes being the staphylococcal plasmids pT181/pC221 (2 and references therein; 52) and pC194/pUB110 (53) and the streptococcal plasmid pMV158 (1). The following plasmid families have also been studied although less thoroughly: the staphylococcal plasmid pSN2 family (54), the pBL1 and pCG1 plasmid families from *C. glutamicum* (55 and references therein), the pSTK-1 and pTX14-2 plasmid families from *B. thuringiensis* (28 and references therein), and the pGRB1 (56) and pGT5 (57) plasmid families from archaea.

The *dso*s of RCR plasmids can be found located upstream of the *rep* gene (pC194, pMV158, and pSN2 families), embedded within the 5′ portion (pT181 family) or the 3′ portion (pCG1 family) of the sequence coding their respective Rep proteins, or even downstream from the *rep* gene stop codon (pTX14-2 family). The *dso* can be physically and functionally divided into two regions, namely *bind*, which contains the specific binding sequence for the initiator protein, and *nic*, where Rep specifically cleaves the DNA at the nick site. The two loci can be either adjacent to each other (pT181 and pC194 families) or separated by a spacer region of up to 100 bp (pMV158 family) (Fig. 1B). The *dso*s of plasmids of the same family are characterized by a high degree of conservation in the *nic* region and by the presence of a less well-conserved *bind* region. In fact, Rep proteins encoded by different plasmids of the same family can perform *in vitro* the nicking-closing reaction on the *dso*s of all the plasmids belonging to the same family, but there is little or no cross-interaction with the *bind* region, which is indicative of the replicon-specificity of the *bind* locus. Interestingly, the pT181-encoded RepC initiator has been shown to drive *in vitro* replication of plasmid pC221, although this was greatly reduced if a competing pT181-*dso* was present (58). In spite of such *in vitro* recognition and extensive homologies of the Rep proteins and the *dso*s of pT181 and pC221, there is no cross-reactivity between the Rep proteins and the *dso*s of these plasmids *in vivo*, unless the Rep proteins are overproduced (59).

In the case of the pMV158 family, the DNA sequence of the *bind* locus was reported to consist of two or three direct repeats (DRs), whose lengths ranged from 5 to 21 bp (35), separated from the nick sequence by an intervening sequence of variable length (Fig. 1B). The *dso* of pJB01, a member of the pE194 subfamily, contains as the Rep-binding site three 7-bp nontandem DRs located 77 bp downstream from the nick site (60). Interestingly, the existence of distant DRs has not been elucidated in some plasmids of this subfamily (unpublished observation). The role of the different regions of the pMV158-*dso* in the interaction with the plasmid-encoded RepB initiator protein has been addressed in a systematic study (35, 61–64). RepB binds with high affinity to the *bind* locus, which is made up of three 11-bp tandem DRs located 84 bp downstream from the nick site. These repeats do not constitute an incompatibility determinant toward pMV158 and seem to be essential for plasmid *in vivo* replication but not for *in vitro* relaxation of supercoiled DNA mediated by RepB. A second RepB binding site is located in a region around the nick site, within the *nic* locus. Characterization of the relative affinity of RepB for the *bind* and *nic* loci revealed that the three DRs of the *bind* locus constitute the primary binding site, whereas the weaker binding of RepB to the *nic* locus could be involved in recognition of the nick site during initiation of replication (64). In plasmids of the pT181 and pC194 families, the DNA sequences of the *bind* (IRIII) and *nic* (IRII) loci are located in contiguous inverted repeats (IR) (Fig. 1B). In pT181, both the spacing and the phasing of IRII to IRIII are crucial for origin functionality (65). In addition, the proximal arm and the central part of the IRIII are important for sequence-specific recognition (65). A similar picture is found in plasmids of the pC194 family.

A typical feature of the *nic* regions is the presence of secondary structures such as hairpins and cruciform. The Rep nick sequence is generally located on an unpaired region within these hairpins, as exemplified by IRII of pT181 and IR-I of pMV158, which accounts for the requirement of plasmid DNA supercoiling to render the cleavage sequence a suitable ssDNA substrate for replication (66–68). The presence of secondary structures is likely to be involved in efficient recruitment and utilization of the initiator protein. Additionally, binding of the initiator protein to the *nic* locus could promote the melting of the substrate nick sequence. This seems to be the case in pMV158, where the extrusion frequency of the cruciform involving IR-I is very low at the growth temperature of the plasmid host (37°C) (69). *In vitro* footprinting experiments performed with supercoiled pMV158 DNA showed that binding of RepB to the *nic* locus promotes the extrusion of the IR-I cruciform, which in turn indicates that initiation of replication would take place only when specific binding of RepB occurs (64). Genetic analysis of the pC194 *dso* pointed to the existence of a hairpin located downstream of the nick site (70) that was shown to be important for replication of the plasmid (71). In contrast, RepU, the initiator protein of pUB110, does not require the presence of hairpins for efficient recognition of the *oriU*. Hairpin II, located downstream from the nick site, seems to be dispensable for initiation of replication of pUB110, although its absence provokes the accumulation of multimers, which is indicative of the involvement of this structure in termination of replication (72).

Out of the three plasmid family prototypes that have been studied in more detail (pT181, pC194, and pMV158), available information regarding the characteristics of the *dso* is limited to a few plasmids of different families. In the *dso* sequences of pJV1, pIJ101, and pSN22, three plasmids belonging to the same subfamily inside the