Estudios estructurales de relaxasas involucradas en el proceso de conjugación bacteriana

Silvia Russi

Barcelona, 2009

5. CONCLUSIONES

5. Conclusiones

5.1. La relaxasa TrwC del plásmido conjugativo R388

El análisis de la unión del complejo TrwC-ADN25 a diferentes metales divalentes (Zn^{2+}, Cu^{2+}, Ni^{2+}, Mn^{2+}, Mg^{2+}, Ca^{2+}) por difracción anómala de rayos X, junto con los ensayos bioquímicos, permitió establecer que el metal fisiológicamente empleado por la proteína TrwC en la actividad catalítica es el Zn^{2+}.

El metal desarrolla un papel clave en el mecanismo enzimático, en el que no solo interviene en la coordinación de la tríada de histidinas localizada en el sitio activo de la proteína, sino que además su unión al grupo fosfato lo polariza, y debilita el enlace fosfo-diéster a cortar.

En los complejos TrwC-ADN25-M^{2+} el lazo α_1-β_1, donde se localiza la segunda tirosina catalítica Tyr26, pudo ser completamente trazado. Si bien la posición de este residuo es bastante lejana del sitio activo, no puede descartarse su participación en el segundo corte de la cadena de ADN. Es posible que ocurran arreglos conformacionales, o incluso que la proteína actúe como dímero, para permitir que el corte se realice. La información estructural no permite establecer sin ambigüedad el rol exacto que juega este residuo en el mecanismo catalítico.

La estructura del complejo TrwC-ADN27 en la que está presente el enlace fosfo-diéster a cortar permitió la localización del mismo respecto al sitio activo, ayudando a comprender el papel del metal en la reacción de corte de la cadena de ADN.

En base a la información estructural obtenida para los complejos TrwC-ADN25-M^{2+} y el complejo TrwC-ADN27, es posible proponer un mecanismo a nivel molecular, del primer paso de la reacción de corte de la cadena de ADN. La Tyr18, activada por el Asp85, realiza el ataque nucleofílico al grupo fosfato, siendo ayudada por el ión metálico, que como se ha indicado, lo polariza favoreciendo el ataque.

La posición de grupos sulfato en la estructura TrwC-ADN25, que mimetizan la

posición de las bases de la cadena de ADN, junto con lo observado en la estructura del complejo TrwC-ADN27, sugieren la existencia de dos posibles rutas de salida para la cadena de ADN cuya bifurcación se produce cerca del sitio de unión al metal. Una hipótesis plausible es que la proteína utiliza ambos caminos en su mecanismo.

5.2. La relaxasa MobM del plásmido movilizable pMV158

Se logró determinar la estructura del dominio N-terminal de la relaxasa MobM en complejo con ADN, encontrándose en ella los elementos estructurales conservados de la familia de las relaxasas: el núcleo central formado por cinco hebras beta flanqueadas por ambos lados por hélices alfa.

Las interacciones de la cadena de ADN con la proteína son múltiples, dada la extensa superficie de interacción que involucra (un 30% de la superficie total accesible de la proteína). Entre ellas, destacan el reconocimiento de la proteína a la cadena de ADN mediante el lazo α2-β3. Dos giros betas consecutivos de dicho lazo orientan dos residuos arginina (Arg71 y Arg74) hacia el estrecho surco menor de la doble cadena de ADN, penetrando en él y estableciendo enlaces de hidrógeno que permiten el reconocimiento específico del ADN.

La arquitectura del sitio activo es semejante a la observada en la estructura de los complejos TrwC-ADN, con la característica tríada de histidinas encargada de coordinar el ión metálico. Las His126 e His133 adoptan conformaciones diferentes que en la estructura de la TrwC producto de la ausencia del metal en el sitio activo.

La identidad del residuo nucleofílico es ambigua. Los resultados sugieren que en la MobM la Tyr122 podría cumplir este papel catalítico, aunque se encuentra más alejada del grupo fosfato a cortar que en el caso de las relaxasas de las bacterias Gram-negativas TrwC o TraI. La tirosina Tyr44, por su parte, se localiza en una posición homóloga a la Tyr26 de la TrwC y podría cumplir un papel similar al de esta. La His22 podría ser también un residuo esencial.

El extremo 3’ de la cadena de ADN no entra al sitio activo y en su lugar se coloca el lazo de la horquilla de una cadena de ADN vecina. La posición relativa de dicha cadena parece estar mimetizando la conformación IR-2 en la que el sitio de corte estaría en el lazo de la horquilla del ADN.

Si bien no puede establecerse con certeza la razón por la cual el extremo 3’ de la cadena de ADN no entra al sitio activo, la comparación de la estructura obtenida con la de la relaxasa TrwC sugiere que el constructo de MobM con el que se trabajó carece de los elementos de secuencia secundaria indispensables para guiar al ADN hacia el sitio de corte, indicando que las hélices o “dedos” en la parte C-terminal del dominio serían imprescindibles para el correcto posicionamiento del substrato.

6. BIBLIOGRAFÍA

6. Bibliografía

Atmakuri, K., Cascales, E., and Christie, P.J. (2004). Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion. Mol Microbiol *54*, 1199-1211.

Backert, S., and Meyer, T.F. (2006). Type IV secretion systems and their effectors in bacterial pathogenesis. Curr Opin Microbiol *9*, 207-217.

Boer, D.R., Ruíz-Masó, J.A., López-Blanco, J.R., Blanco, A.G., Vives-Llàcer, M., Chacón, P., Uson, I., Gomis-Rüth, F.X., Espinosa, M., Llorca, O.*, et al.* (2009). Plasmid replication initiator RepB forms a hexamer reminiscent of ring helicases and has mobile nuclease domains. EMBO J *28*, 1666-1678.

Boer, R., Russi, S., Guasch, A., Lucas, M., Blanco, A.G., Perez-Luque, R., Coll, M., and de la Cruz, F. (2006). Unveiling the molecular mechanism of a conjugative relaxase: The structure of TrwC complexed with a 27-mer DNA comprising the recognition hairpin and the cleavage site. J Mol Biol *358*, 857-869.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem *72*, 248-254.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S.*, et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr *54 (Pt 5)*, 905-921.

Burdett, V. (1980). Identification of tetracycline-resistant R-plasmids in Streptococcus agalactiae (group B). Antimicrob Agents Chemother *18*, 753-760.

Burns, D.L. (2003). Type IV transporters of pathogenic bacteria. Curr Opin Microbiol *6*, 29-34.

Byrd, D.R., and Matson, S.W. (1997). Nicking by transesterification: the reaction catalysed by a relaxase. Mol Microbiol *25*, 1011-1022.

Cabezon, E., and de la Cruz, F. (2006). TrwB: an F(1)-ATPase-like molecular motor involved in DNA transport during bacterial conjugation. Res Microbiol *157*, 299-305.

Cabezon, E., Sastre, J.I., and de la Cruz, F. (1997). Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation. Mol Gen Genet *254*, 400-406.

Campos-Olivas, R., Louis, J.M., Clerot, D., Gronenborn, B., and Gronenborn, A.M. (2002). The structure of a replication initiator unites diverse aspects of nucleic acid metabolism. Proc Natl Acad Sci U S A *99*, 10310-10315.

Cascales, E., and Christie, P.J. (2003). The versatile bacterial type IV secretion systems. Nat Rev Microbiol *1*, 137-149.

Collaborative Computational Project, N. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr *50*, 760-763.

Cowtan, K. (2002). Generic representation and evaluation of properties as a function of position in reciprocal space. J Appl Cryst *35*, 655-663.

Chen, I., Christie, P.J., and Dubnau, D. (2005). The ins and outs of DNA transfer in bacteria. Science *310*, 1456-1460.

Christie, P.J. (2001). Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. Mol Microbiol *40*, 294-305.

Christie, P.J. (2004). Type IV secretion: the Agrobacterium VirB/D4 and related conjugation systems. Biochim Biophys Acta *1694*, 219-234.

Christie, P.J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S., and Cascales, E. (2005). Biogenesis, architecture, and function of bacterial type IV secretion systems.

Annu Rev Microbiol *59*, 451-485.

Christie, P.J., and Cascales, E. (2005). Structural and dynamic properties of bacterial type IV secretion systems (review). Mol Membr Biol *22*, 51-61.

Christie, P.J., and Vogel, J.P. (2000). Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. Trends Microbiol *8*, 354-360.

Datta, S., Larkin, C., and Schildbach, J.F. (2003). Structural insights into single-stranded DNA binding and cleavage by F factor TraI. Structure *11*, 1369-1379.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S.*, et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res *35*, W375-383.

de Antonio, C., Farias, M.E., García de Lacoba, M., and Espinosa, M. (2004). Features of the Plasmid pMV158-encoded MobM, a Protein Involved in its Mobilization. . J Mol Biol *335*, 733-743.

de la Cruz, F., and Davies, J. (2000). Horizontal gene transfer and the origin of species: lessons from bacteria. Trends Microbiol *8*, 128-133.

Ding, Z., Atmakuri, K., and Christie, P.J. (2003). The outs and ins of bacterial type IV secretion substrates. Trends Microbiol *11*, 527-535.

Draper, O., Cesar, C.E., Machon, C., de la Cruz, F., and Llosa, M. (2005). Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. Proc Natl Acad Sci U S A *102*, 16385-16390.

Dreiseikelmann, B. (1994). Translocation of DNA across bacterial membranes. Microbiol Rev *58*, 293-316.

Durrenberger, M.B., Villiger, W., and Bachi, T. (1991). Conjugational junctions:

morphology of specific contacts in conjugating Escherichia coli bacteria. J Struct Biol *107*, 146-156.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

Enemark, E.J., Stenlund, A., and Joshua-Tor, L. (2002). Crystal structures of two intermediates in the assembly of the papillomavirus replication initiation complex. Embo J *21*, 1487-1496.

Evans, P.R. (1993). Data reduction. In Proceedings of CCP4 Study Weekend. Data Collection & Processing, 114-122.

Farias, M.E., Grohmann, E., and Espinosa, M. (1999). Expression of the mobM gene of the streptococcal plasmid pMV158 in Lactococcus lactis subsp. lactis. FEMS Microbiol Lett *176*, 403-410.

Firth, N., Ippen-Ihler, K., and Skurray, R.A. (1996). Structure and function of the F factor and mechanism of conjugation. In Escherichia Coli and Salmonella: Cellular and Molecular Biology, N. FC, ed. (American Sociaty of Microbiology), pp. 2377-2401.

Francia, M.V., Varsaki, A., Garcillan-Barcia, M.P., Latorre, A., Drainas, C., and de la Cruz, F. (2004). A classification scheme for mobilization regions of bacterial plasmids. FEMS Microbiol Rev *28*, 79-100.

Furuya, N., and Komano, T. (2000). Initiation and termination of DNA transfer during conjugation of IncI1 plasmid R64: roles of two sets of inverted repeat sequences within oriT in termination of R64 transfer. J Bacteriol *182*, 3191-3196.

Garcillan-Barcia, M.P., Jurado, P., Gonzalez-Perez, B., Moncalian, G., Fernandez, L.A., and de la Cruz, F. (2007). Conjugative transfer can be inhibited by blocking relaxase activity within recipient cells with intrabodies. Molecular microbiology *63*, 404-416.

Gomis-Ruth, F.X., Sola, M., Acebo, P., Parraga, A., Guasch, A., Eritja, R., Gonzalez,

A., Espinosa, M., del Solar, G., and Coll, M. (1998). The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. Embo J *17*, 7404-7415.

Gomis-Ruth, F.X., Sola, M., de la Cruz, F., and Coll, M. (2004). Coupling factors in macromolecular type-IV secretion machineries. Curr Pharm Des *10*, 1551-1565.

Grandoso, G., Avila, P., Cayon, A., Hernando, M.A., Llosa, M., and de la Cruz, F. (2000). Two active-site tyrosyl residues of protein TrwC act sequentially at the origin of transfer during plasmid R388 conjugation. J Mol Biol *295*, 1163-1172.

Grohmann, E., Muth, G., and Espinosa, M. (2003). Conjugative plasmid transfer in gram-positive bacteria. Microbiol Mol Biol Rev *67*, 277-301, table of contents.

Guasch, A., Lucas, M., Moncalian, G., Cabezas, M., Perez-Luque, R., Gomis-Ruth, F.X., de la Cruz, F., and Coll, M. (2003). Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. Nat Struct Biol *10*, 1002-1010.

Guzmán, L.M., and Espinosa, M. (1997). The mobilization protein MobM, of the streptoccocal plasmid pMV158, specifically cleaves supercoiled DNA at the plasmid oriT. J Mol Biol *266*, 688-702.

Hickman, A.B., Ronning, D.R., Kotin, R.M., and Dyda, F. (2002). Structural unity among viral origin binding proteins: crystal structure of the nuclease domain of adeno-associated virus Rep. Mol Cell *10*, 327-337.

Howard, M.T., Nelson, W.C., and Matson, S.W. (1995). Stepwise assembly of a relaxosome at the F plasmid origin of transfer. J Biol Chem *270*, 28381-28386.

Kingsman, A., and Willetts, N. (1978). The requirements for conjugal DNA synthesis in the donor strain during flac transfer. J Mol Biol *122*, 287-300.

Kramer, M.G., Khan, S.A., and Espinosa, M. (1997). Plasmid rolling circle replication: identification of the RNA polymerase-directed primer RNA and requirement for DNA

polymerase I for lagging strand synthesis. Embo J *16*, 5784-5795.

Lanka, E., and Wilkins, B.M. (1995). DNA processing reactions in bacterial conjugation. Annu Rev Biochem *64*, 141-169.

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

Lederberg, J., and Tatum, E. (1946). Gene recombination in E. coli. Nature *158*, 558.

Leslie, A.G.W. (1992). Recent changes to the MOSFLM package for processing film and image plate data Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography *26*.

Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. Science *252*, 1162-1164.

Luscombe, N.M., Laskowski, R.A., and Thornton, J.M. (1997). NUCPLOT: a program to generate schematic diagrams of protein-nucleic acid interactions. Nucleic Acids Res *25*, 4940-4945.

Llosa, M., and de la Cruz, F. (2005). Bacterial conjugation: a potential tool for genomic engineering. Res Microbiol *156*, 1-6.

Llosa, M., Gomis-Ruth, F.X., Coll, M., and de la Cruz Fd, F. (2002). Bacterial conjugation: a two-step mechanism for DNA transport. Mol Microbiol *45*, 1-8.

Llosa, M., Grandoso, G., Hernando, M.A., and de la Cruz, F. (1996). Functional domains in protein TrwC of plasmid R388: dissected DNA strand transferase and DNA helicase activities reconstitute protein function. J Mol Biol *264*, 56-67.

Llosa, M., and O'Callaghan, D. (2004). Euroconference on the Biology of Type IV Secretion Processes: bacterial gates into the outer world. Mol Microbiol *53*, 1-8.

Marsden, R.L., McGuffin, L.J., and Jones, D.T. (2002). Rapid protein domain assignment from amino acid sequence using predicted secondary structure. Protein Sci *11*, 2814-2824.

Matthews, B.W. (1968). Solvent content of protein crystals. J Mol Biol *33*, 491-497.

Meinke, G., Bullock, P.A., and Bohm, A. (2006). Crystal structure of the simian virus 40 large T-antigen origin-binding domain. J Virol *80*, 4304-4312.

Moncalian, G., Cabezon, E., Alkorta, I., Valle, M., Moro, F., Valpuesta, J.M., Goni, F.M., and de La Cruz, F. (1999a). Characterization of ATP and DNA binding activities of TrwB, the coupling protein essential in plasmid R388 conjugation. J Biol Chem *274*, 36117-36124.

Moncalian, G., and de la Cruz, F. (2004). DNA binding properties of protein TrwA, a possible structural variant of the Arc repressor superfamily. Biochim Biophys Acta *1701*, 15-23.

Moncalian, G., Grandoso, G., Llosa, M., and de la Cruz, F. (1997). oriT-processing and regulatory roles of TrwA protein in plasmid R388 conjugation. J Mol Biol *270*, 188-200.

Moncalian, G., Valle, M., Valpuesta, J.M., and de la Cruz, F. (1999b). IHF protein inhibits cleavage but not assembly of plasmid R388 relaxosomes. Mol Microbiol *31*, 1643-1652.

Moscoso, M., Eritja, R., and Espinosa, M. (1997). Initiation of replication of plasmid pMV158: mechanisms of DNA strand-transfer reactions mediated by the initiator RepB protein. J Mol Biol *268*, 840-856.

Murray, J., and Garman, E. (2002). Investigation of possible free-radical scavengers and metrics for radiation damage in protein cryocrystallography. J Synchrotron Radiat *9*, 347-354.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr *53*, 240-255.

Murzin, A.G., Brenner, S.E., Hubbard, T., and Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol *247*, 536-540.

Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000). Lateral gene transfer and the nature of bacterial innovation. Nature *405*, 299-304.

Priebe, S.D., and Lacks, S.A. (1989). Region of the streptococcal plasmid pMV158 required for conjugative mobilization. J Bacteriol *171*, 4778-4784.

Ramachandran, G.N., and Sasisekharan, V. (1968). Conformation of polypeptides and proteins. Adv Protein Chem *23*, 283-438.

Read, R.J. (2001). Pushing the boundaries of molecular replacement with maximum likelihood. Acta Crystallogr D Biol Crystallogr *57*, 1373-1382.

Rice, P.A., Yang, S., Mizuuchi, K., and Nash, H.A. (1996). Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. Cell *87*, 1295-1306.

Rossmann, M.G., and Blow, D.M. (1962). The detection of sub-units within the crystallographic asymmetric unit. Acta Crystallogr *15*, 24-31.

Roussel, A., and Cambilleau, C. (1989). Turbo-Frodo. In Silicon Graphics geometry partners directory (Mountain View, CA, Silicon Graphics), pp. 77-79.

Santini, J.M., and Stanisich, V.A. (1998). Both the fipA gene of pKM101 and the pifC gene of F inhibit conjugal transfer of RP1 by an effect on traG. J Bacteriol *180*, 4093-4101.

Scott, J.R., and Churchward, G.G. (1995). Conjugative transposition. Annu Rev Microbiol *49*, 367-397.

Schneider, T.R., and Sheldrick, G.M. (2002). Substructure solution with SHELXD. Acta Crystallogr D Biol Crystallogr *58*, 1772-1779.

Schroder, G., and Lanka, E. (2005). The mating pair formation system of conjugative plasmids-A versatile secretion machinery for transfer of proteins and DNA. Plasmid *54*, 1-25.

Seubert, A., Hiestand, R., de la Cruz, F., and Dehio, C. (2003). A bacterial conjugation machinery recruited for pathogenesis. Mol Microbiol *49*, 1253-1266.

Sexton, J.A., and Vogel, J.P. (2002). Type IVB secretion by intracellular pathogens. Traffic *3*, 178-185.

Sheldrick, G. (2002). Macromolecular phasing with SHELXE. Z Krist *217*, 644-650.

Silverman, P.M. (1997). Towards a structural biology of bacterial conjugation. Mol Microbiol *23*, 423-429.

Szpirer, C.Y., Faelen, M., and Couturier, M. (2001). Mobilization function of the pBHR1 plasmid, a derivative of the broad-host-range plasmid pBBR1. J Bacteriol *183*, 2101-2110.

Terwilliger, T. (2004). SOLVE and RESOLVE: automated structure solution, density modification and model building. J Synchrotron Radiat *11*, 49-52.

Ton-That, H., and Schneewind, O. (2003). Assembly of pili on the surface of Corynebacterium diphtheriae. Mol Microbiol *50*, 1429-1438.

Uson, I., and Sheldrick, G.M. (1999). Advances in direct methods for protein crystallography. Curr Opin Struct Biol *9*, 643-648.

Vagin, A., and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. J Appl Cryst *30*, 1022-1025.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. Embo J *1*, 945-951.

Waters, V.L. (2001). Conjugation between bacterial and mammalian cells. Nature genetics *29*, 375-376.

Weik, M., Ravelli, R.B., Kryger, G., McSweeney, S., Raves, M.L., Harel, M., Gros, P., Silman, I., Kroon, J., and Sussman, J.L. (2000). Specific chemical and structural damage to proteins produced by synchrotron radiation. Proc Natl Acad Sci U S A *97*, 623-628.

Wilkins, B., and Lanka, E. (1993). DNA processing and replication during plasmid transfer between Gram-negative bacteria (New York, Plenum Press).

Ziegelin, G., Furste, J.P., and Lanka, E. (1989). TraJ protein of plasmid RP4 binds to a 19-base pair invert sequence repetition within the transfer origin. J Biol Chem *264*, 11989-11994.

7. PUBLICACIONES

7. Publicaciones

Los resultados derivados de este proyecto de tesis están incluidos en las siguientes publicaciones.

Artículos publicados:

Unveiling the Molecular Mechanism of a Conjugative Relaxase: The Structure of TrwC Complexed with a 27-mer DNA Comprising the Recognition Hairpin and the Cleavage Site, R. Boer[*], S. Russi[*], A. Guasch, M. Lucas, A.G. Blanco, R. Pérez-Luque, M. Coll1 & F. de la Cruz. J. Mol. Biol., 2006, Vol. 358, 857-869. [*]Ambos autores contribuyeron por igual al trabajo

Capítulos de libros:

*Molecular machinery for DNA translocation in bacterial conjugation***,** S. Russi, R. Boer & M. Coll, 2008, PLASMIDS: current research and future trends, edited by G. Lipps (Caister Academic Press).

Artículos en preparación:

Crystal structure of the N-terminal relaxase domain of the mobilization protein MobM, from plasmid pMV158, in complex with its cognate DNA. Russi *et al*., 2009.

doi:10.1016/j.jmb.2006.02.018 J. Mol. Biol. (2006) 358, 857–869

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Unveiling the Molecular Mechanism of a Conjugative Relaxase: The Structure of TrwC Complexed with a 27-mer DNA Comprising the Recognition Hairpin and the Cleavage Site

Roeland Boer[1]†, Silvia Russi[1]†, Alicia Guasch[1,3], María Lucas[2] Alexandre G. Blanco[1], Rosa Pérez-Luque[1], Miquel Coll[1]* and Fernando de la Cruz[2]

[1]*Institut de Biologia Molecular de Barcelona (CSIC) and Institut de Recerca Biomèdica Parc Científic de Barcelona Josep Samitier 1-5, 08028 Barcelona, Spain*

[2]*Departamento de Biología Molecular (Laboratorio asociado al CIB, CSIC), Universidad de Cantabria, Herrera Oria, s/n 39011 Santander, Spain*

[3]*Plataforma Automatitzada de Cristal·lografia (PCB-CSIC) Parc Cientific de Barcelona Josep Samitier 1-5, 08028 Barcelona, Spain*

**Corresponding author*

TrwC is a DNA strand transferase that catalyzes the initial and final stages of conjugative DNA transfer. We have solved the crystal structure of the N-terminal relaxase domain of TrwC in complex with a 27 base-long DNA oligonucleotide that contains both the recognition hairpin and the scissile phosphate. In addition, a series of ternary structures of protein–DNA complexes with different divalent cations at the active site have been solved. Systematic anomalous difference analysis allowed us to determine unambiguously the nature of the metal bound. Zn^{2+}, Ni^{2+} and Cu^{2+} were found to bind the histidine-triad metal binding site. Comparison of the structures of the different complexes suggests two pathways for the DNA to exit the active pocket. They are probably used at different steps of the conjugative DNA-processing reaction. The structural information allows us to propose (i) an enzyme mechanism where the scissile phosphate is polarized by the metal ion facilitating the nucleophilic attack of the catalytic tyrosine, and (ii) a probable sequence of events during conjugative DNA processing that explains the biological function of the relaxase.

Keywords: X-ray structure; relaxase; bacterial conjugation; DNA transfer; DNA–protein complex

Introduction

Bacterial conjugation is a process by which a DNA molecule is transferred from a donor to a recipient bacterium. The mechanism is efficient, and allows bacteria to acquire new adaptive traits such as antibiotic resistance. Thus, it is considered an important mechanism in bacterial evolution.[1,2] The underlying biochemical process can be operationally divided into two steps, DNA processing and DNA transport, each of which is carried out by a specific set of proteins encoded by the *tra* genes of a given conjugative plasmid. Conjugative DNA processing starts by cleavage of a specific phosphodiester bond (the *nic* site) in the donor supercoiled DNA by a plasmid-specific relaxase. The resulting nucleoprotein complex, the relaxosome, contacts the transport site, where a multi-protein DNA transport apparatus effects the transfer process of the cleaved DNA strand to the recipient cell. Presumably, the relaxase religates the transferred DNA strand, and finally host proteins replicate both single strands in donor and recipient bacteria to regenerate the double-stranded conjugative plasmid.[3,4]

Most conjugative systems contain phylogenetically related relaxases, according to their amino acid sequences.[5] Remarkably conspicuous is a histidine triad that has been intimately involved in the catalytic mechanism. A few selected relaxases have been analyzed at a biochemical level. The purified proteins can specifically cleave oligonucleotides containing their respective *nic* site sequences so that the 5′ phosphoryl end of the cleaved product becomes covalently bound to the hydroxyl group of a specific tyrosyl residue of the protein. Relaxases can then transfer the bound

† R.B. and S.R. contributed equally to this work.

E-mail address of the corresponding author: mcoll@ibmb.csic.es

DNA to an appropriate acceptor oligonucleotide by a second DNA strand-transfer reaction, so that a hybrid oligonucleotide is released from the enzyme.[6] The P-family relaxase TraI of plasmid RP4 is one of the best analyzed with respect to the biochemical details of these reactions. By using oligonucleotides bound to a solid support, Pansegrau & Lanka[7] showed that the TraI oligonucleotide adduct was incapable of carrying out the strand transfer reaction when challenged with a second oligonucleotide containing *nic*. The interpretation was that a second TraI molecule is required to complete the reaction, presumably by providing a new Tyr to catalyze the second strand transfer reaction.

The F-family relaxase TrwC, from the IncW plasmid R388, is a large protein composed of an N-terminal domain with DNA-relaxase activity and a C-terminal domain with DNA helicase activity.[8] It can cleave a supercoiled plasmid DNA containing *oriT in vitro* in the absence of accessory proteins.[9] TrwC contains two active-site tyrosyl residues (instead of the single one in P-family relaxases) that play different functional roles in the DNA processing reactions. Thus, it was proposed that conjugative DNA processing of plasmid R388 occurs by a variant of the flip-flop mechanism used in ϕX-174 replication[10] in which the two active tyrosine residues catalyze the initiation and termination steps in conjugative DNA replication.[11]

The atomic structures of two F-family relaxases, R388 TrwC[12] and F plasmid TraI,[13] have been reported. The TrwC structure shows a complex of the protein with the *nic* DNA up to the cleavage site. This complex was shown both as a binary DNA–protein complex and as a ternary complex with a Zn^{2+} in the three-histidine pocket. Although these structures provided valuable information with respect to the DNA and metal binding sites, they did not reveal the identity of the metal ion and did not provide a complete view of the reaction mechanism catalyzed by relaxases. In particular, both the exit path for the DNA and the position of the scissile phosphate with respect to the catalytic residues were unclear because the DNA used for the DNA–TrwC complex[12] did not include any residue downstream of the cleavage site nor the scissile phosphate (A26, according to our naming). Recently, the structure of TraI complexed with a 10 bp oligonucleotide, including the scissile bond but not the full recognition hairpin, has been reported.[14] Here we present additional structural information of TrwC with the description of a relaxase ternary complex with a 27-mer DNA oligonucleotide encompassing both the recognition hairpin and the cleavage site. Also, we present a detailed structural analysis of the binding of different metal ions. Finally, we provide a structure where the loop α1-β1 that includes the second catalytic residue Tyr26 is fully traced, in contrast to previous structures where it was disordered. As a result we clarify crucial aspects of the mechanism for the successive DNA-strand transfer reactions catalyzed by relaxases.

Results

Crystal structure of TrwC^{Y18F}–DNA27

For this complex a Tyr18Phe TrwC mutant relaxase was used, impairing cleavage of the DNA at the scissile O3′(T25)–P(A26) bond. In this crystal, the space group and crystal packing arrangement is the same as the one observed for the 1OSB structure.[12] The overall structures of the protein and the DNA in the DNA27 complex (Figure 1(a)) are similar to those found in the metal-free DNA25 complex structures determined previously.[12] Briefly, the DNA folds in a hairpin structure, with a double-helical stem that includes bases G3 to G7 which are paired, in a Watson–Crick manner, to bases C16 to G12, respectively. Bases G8 and A9, at the tip of the hairpin, continue the stack but are not paired, while A10 is folded back into the minor groove, establishing H-bonds with G8. A11 is swung out. Further downstream of the hairpin a U-turn occurs from base T21 to C24, prior to the DNA entry into the active site (Figure 1(b)). At the turn, bases T21, G22 and T23 form a hydrophobic cage that traps the side-chain of the first residue of the protein, Met1. When comparing the DNA25 and DNA27 structures in the respective complexes with TrwC, some differences appear at the 5′ end, where the first two residues were disordered in the DNA27 complex and, thus, not modelled. More significant differences from a mechanistic viewpoint occur at the 3′ end of the DNA. Base T25 has changed its orientation (when compared to 1OSB) to form a non-complementary base-pairing with G22, with a three-centered (bifurcated) H-bond between the O2(T25) and N2(G22) and N1(G22) and another H-bond between N3(T25) and O6(G22) (Figure 1(b)). The resulting conformation of residue T25 orients the 3′ oxygen atom and the phosphate group of the next residue (A26) towards the metal-binding site which is, in this structure, occupied by a water molecule (Figure 2(a)). This water molecule established H-bonds with His150, His161, His163 and one of the oxygen atoms of the scissile phosphate A26, most probably mimicking the metal ion coordination. No SO_4^{2-} was found close to that metal site, contrary to what occurred in the DNA25 metal-free complex, since that position is now occupied by phosphate A26. Interestingly, the second sulphate is still present, signalling the alternative path for the DNA exit (see below).

Residual density appeared at the 3′ end of T25 and adenosine A26 could be placed and refined in one of the molecules of the asymmetric unit (in the other molecules the density was weak and residue A26 was not built in). Adenine A26 stacks with one face on base T25 (Figure 1), while the backbone continues in a, more or less, straight trajectory. The other face of A26 is stacked on Met5. No additional

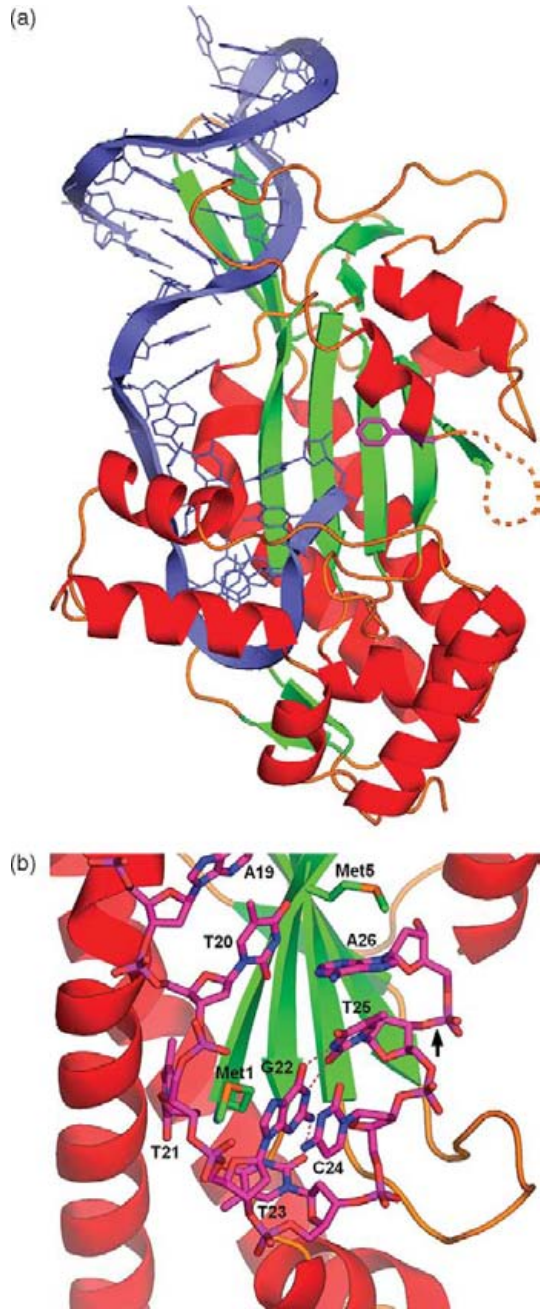

Figure 1. (a) Ribbon representation of the TrwC^{Y18F}–DNA27 complex. α-Helices are shown in red, β-strands in green and the DNA in blue. The side-chain of Phe18 is included in magenta, the location of the disordered loop Glu20–Asp30 in the sequence of TrwC is indicated by the broken line. (b) Ribbon representation of the TrwC^{Y18F}–DNA27 complex showing the U-turn of the DNA. The hydrogen bonds at the G22:T25 mismatch pairing and between G22 and T23 are shown as dots. The interactions of Met1 at the hydrophobic cage formed by T21, G22 and T23 are shown, as well as the T25/A26/Met5 stack. The black arrow indicates the scissile bond.

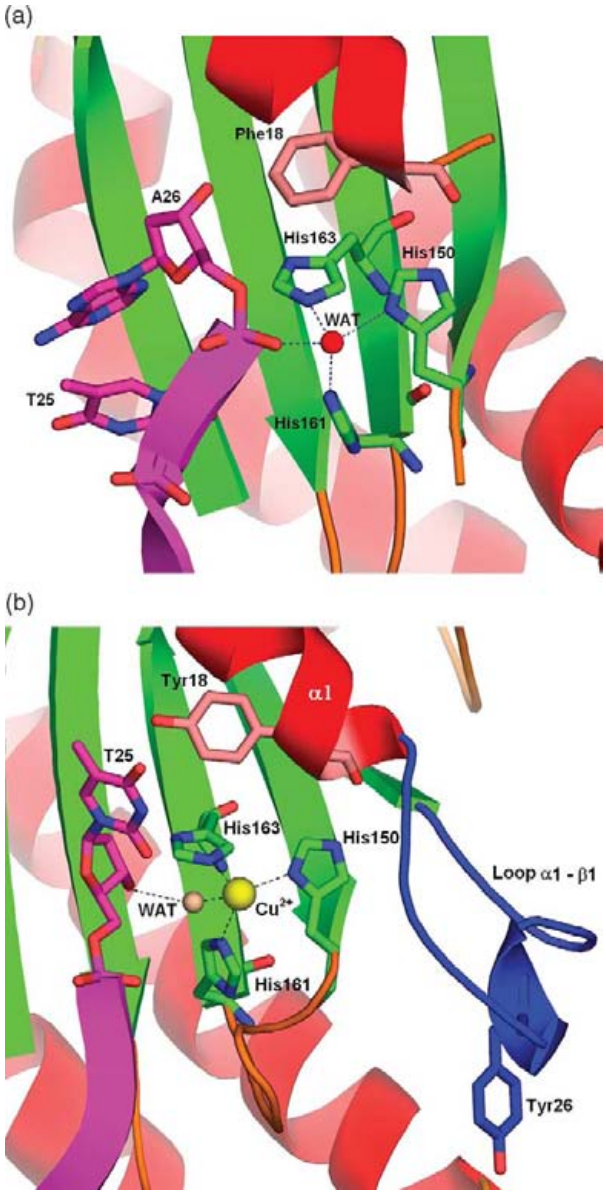

Figure 2. (a) View of the active site of the TrwC^{Y18F}–DNA27 complex showing the mutated Tyr18Phe residue location. A water molecule occupies the metal site, which is tetrahedraly H-bonded by the histidine triad and the O3′ atom of the scissile phosphate A26. The adenine base of residue A26 is stacked on T25. (b) View of the active site of the TrwC–Cu^{2+}–DNA25 structure showing the coordination of the Cu^{2+}, the catalytic residues Tyr18 and Tyr26, and the flexible loop α1–β1. The metal ion is tetra-coordinated by the histidine triad and a water molecule.

density for the last residue of the DNA sequence, T27, appeared during refinement, and thus it was not modelled. These observations, in combination with increasing *B* factors at residues T25 and A26 suggest a high degree of flexibility for the 3′ end of the DNA. The location and conformation of DNA residues T25 and A26 implies that the DNA backbone follows a parallel path to the hairpin stem (upwards in Figure 1).

A substantial displacement of protein residue Phe18 was observed, when compared to the position of Tyr18 in the native protein structure in the DNA25 complex. Tyr18 has been observed in

slightly different positions in different crystals (see below and Figure 5(a)), but the Tyr18Phe mutation further frees the side-chain, since the hydrogen bond of the Tyr18 hydroxyl group to Asp85 is not possible anymore.

Crystal structures with different metal ions

Relaxases, such as TrwC, require the addition of a divalent metal ion for oligonucleotide cleavage activity. The reaction is practically instantaneous, and is expressed as a percentage cleavage at equilibrium. We tested a series of divalent ions for cleavage of an oligonucleotide containing the *nic* site. Without externally added divalent ion, erratic figures were obtained, probably due to the presence of trace amounts of metal ions in buffers and/or protein or DNA preparations. Thus, 10 μM EDTA was added in all experiments shown in Figure 3. Under these conditions, and without added metal ions, no cleavage product was detected. Cleavage activity was detected when adding any of the following metal ions: Mg^{2+}, Mn^{2+}, Ca^{2+}, Ni^{2+}, Zn^{2+} or Cu^{2+} (see Figure 3). Similar equilibrium states could be achieved irrespective of the metal used, suggesting an ample flexibility in the nature of the metal required. However, while Mn^{2+}, Ni^{2+}, Zn^{2+} and Cu^{2+} cleaved at concentrations as low as 10 μM, concentrations of at least 100 μM were required to see high yields of cleavage with Mg^{2+} or Ca^{2+}. In some cases, such as Cu^{2+} and Zn^{2+} the percentage cleavage dropped at high metal concentrations due to macroscopic protein precipitation. It is also noteworthy that significant cleavage was obtained in all cases at 1 μM metal concentration, even in the presence of 10 μM EDTA. This result underscores the notion that trace amounts of metal are sufficient to catalyze some oligonucleotide cleavage.

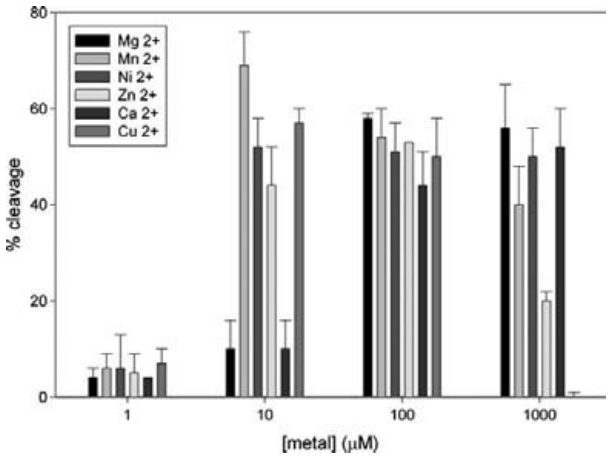

Figure 3. TrwC cleavage activity in the presence of different divalent metal ions. Cleavage experiments were performed by 30 min incubation of 0.5 μM oligonucleotide R388(12+18) with 1 μM TrwC–N293 at 37 °C in 10 mM Tris (pH 7.6), 100 mM NaCl, 10 μM EDTA and 1–1000 μM of either $MgCl_2$, $MnCl_2$, $NiCl_2$, $ZnCl_2$, $CuCl_2$ or $CaCl_2$. The percent cleavage is represented as a function of metal ion concentration.

To assess the structural effects of metal binding, we decided to obtain complexes of protein–DNA with different metals. Figure 4 shows the refined structures and anomalous difference maps for each of the metal-soaked crystals of the TrwC–DNA25 complex. Figure 4(a) corresponds to crystals soaked with Zn^{2+} and Mg^{2+}. With data collected at the peak of the Zn edge, strong difference density appears (yellow) at the metal site. However, the anomalous difference map for the same crystal calculated with data collected below the Zn edge (purple) shows no density at the metal site. This result demonstrates that Zn^{2+} is the metal ion bound to the three active site histidine residues.

Figure 4(b) corresponds to crystals soaked with Zn^{2+} and Mn^{2+}. With data collected at the peak of the Zn edge, strong anomalous difference density appears (yellow) at the metal site. However, when the anomalous map is calculated with data collected at the Mn absorption edge (Figure 4(c)) no density is found at the metal site. These results demonstrate that, again, Zn^{2+}, and not Mn^{2+}, is the metal ion bound at the TrwC active site.

Figure 4(d) corresponds to the anomalous map of crystals soaked only with Mn^{2+} and collected at the energy corresponding to the Mn absorption edge, as determined from a $MnCl_2$ solution placed in the crystal-mounting loop and irradiated at the beamline. No density is observed at the metal site and thus it can be concluded that the Mn^{2+} is not bound. No evidence of Mn^{2+} binding was found either from the X-ray fluorescence absorption scans on any of the Mn^{2+}-soaked crystals. The absence of metal ion at the active site is further confirmed by the orientation of His150 (Figure 4(d)), which is not appropriate for metal coordination. The orientation of this histidine residue is similar to the one observed in the metal-free 1OMH structure.[12]

Figure 4(e) and (f) shows the anomalous maps with data collected at the Ni and Cu absorption edges from crystals soaked in Ni^{2+} and Cu^{2+}, respectively. They clearly show the presence of these two metals at the metal binding site, with triple histidine coordination.

Other soakings with calcium and magnesium were also performed. As in the case of Mn^{2+}, the X-ray data (not shown) showed that binding of Ca^{2+} did not take place. To rule out Mg^{2+} binding was more difficult, it is not possible to collect anomalous data for this metal. The electron density maps showed the presence of an atom located at the metal site that could be firstly assigned as Mg^{2+} or an oxygen atom of a water molecule, since from the electron density height they are nearly indistinguishable at the resolution of the data described here. However, the coordination geometry of this atom at the active site in the structure of crystals soaked in Mg^{2+} does not correspond to the characteristic octahedral coordination around a magnesium atom.

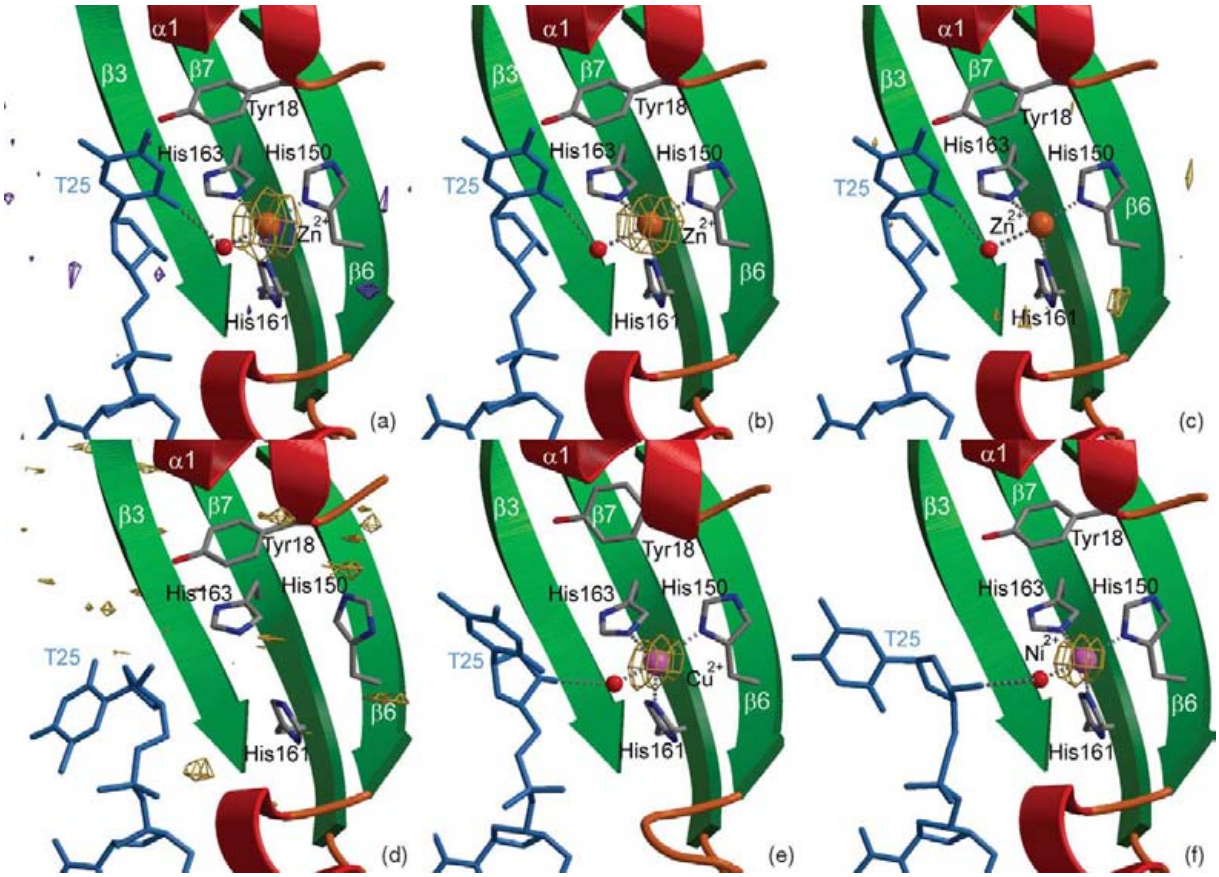

Figure 4. Bijvoet maps of the active site of TrwC, using anomalous differences of data measured at the edge of the appropriate element. (a) Crystal soaked in a solution containing both Zn^{2+} and Mg^{2+}. In yellow is shown the anomalous difference map calculated with data measured at the peak of the Zn absorption edge (data set 25ZNpkMG), contoured at the 7σ level. The contours in purple are from a dataset measured at an energy just below the absorption edge (data set 25ZNrmMG) using the same crystal and contoured at noise level (2.8σ). (b) Anomalous difference map calculated with data measured at the Zn edge (data set 25ZNpkMN) from TrwC–DNA25 crystals soaked in a solution containing both Zn^{2+} and Mn^{2+}, contoured at 7σ. (c) Anomalous difference map calculated using data measured at the Mn absorption edge (data set 25ZNMNpk) on the same crystal used for (b), contoured at noise level (2.8σ). (d) Anomalous difference map calculated with data measured on the Mn edge (data set 25MN) from a TrwC–DNA25 crystal soaked in $MnCl_2$ only, contoured at noise level (2.8σ). (e) Anomalous map calculated with data set measured at the Cu absorption edge (data set 25CU), from a TrwC–DNA25 crystal soaked in a solution containing Cu^{2+}, contoured at the 7σ level. (f) Anomalous map calculated using data measured at the Ni absorption edge (data set 25NI), from a Ni^{2+}-soaked TrwC–DNA25 crystal, contoured at 7σ level.

Metal binding caused some rearrangements in the conformation of the protein. For example, in the metal-free structure the loop α1–β1, encompassing protein residues 20 to 30, had extremely weak and fragmented electron density, an indication of its flexibility, and could not be traced. In contrast, clear density appeared in the Ni^{2+} and Zn^{2+}-bound structures and residues 20 to 26 could be built. In the case of the Cu^{2+}-bound structure, the whole loop was traced (Figure 2(b)). The conformation of this loop was similar in the structures with Ni^{2+} and Cu^{2+}, but slightly different in the Zn^{2+}-bound structure, at residues 21 to 24.

Also, a significant effect of the metal binding on the conformation of the catalytic Tyr18 is observed. In comparison to the metal-free 1OMH structure,[12] displacements of 2.39 Å in the Zn^{2+}-bound structure, 2.17 Å in the Ni^{2+}-bound structure and 0.68 Å in the Cu^{2+}-bound structure are found (Figure 5(a)) for the OH(Tyr18) atom, indicating that this side-chain is able to swing about its χ_1 torsion angle. In addition, nucleotide T25 shows a high degree of flexibility. The corresponding electron density is weak and dispersed, which makes it difficult to correctly place T25 in the metal-bound TrwC structures. However, careful inspection suggests that this residue orients predominantly the 3′ oxygen towards the metal-binding pocket in the Cu^{2+} and Zn^{2+}-bound structures and towards the opposite direction in the Ni^{2+}-bound structure (Figure 5(b)).

Discussion

The atomic structure of TrwC relaxase in complex with the 25-mer *nic* DNA[12] unveiled how the relaxase recognises the extruded inverted repeat hairpin and provided a detailed view of the architecture of the active site. However, since the DNA in the complex did not extend past the cleavage site but ended up

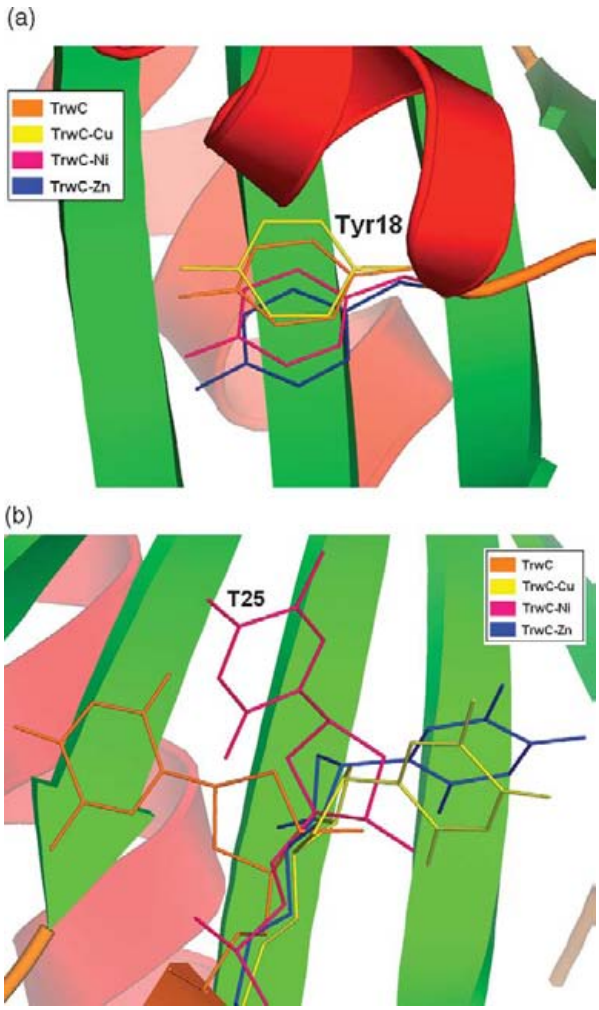

Figure 5. Different conformations of (a) Tyr18 and (b) T25, observed in the different structures of TrwC–DNA25 complexes solved.

before the scissile phosphate, it left some questions unanswered. For example, the precise position of the scissile phosphate was unknown and, thus, the kind of interactions that it could establish with the protein residues and the metal ion. The pathway of the DNA 3′ to the cleavage site was also unclear, although two sulphate ions, at adequate distances, seemed to indicate where the succeeding DNA phosphate groups after the scissile bond should be located. The precise catalytic mechanism and, in particular the role of the metal ion, was also a matter to decipher. Finally, the loop α1–β2 could not be fully modelled because of the weak electron density in that area. This loop includes the second catalytic Tyr26 and thus crucial structural information necessary to understand how the protein functions was missing. This work addresses both questions.

DNA conformation and its recognition by TrwC

The structure of the TrwC^{Y18F}–DNA27 complex confirms our previous findings[12] with respect to the DNA folding as a hairpin and its recognition by the relaxase by means of a β-sheet entering the major groove and a loop segment, flanked by two turns, intruding the opposite minor groove (Figure 1(a)). After the hairpin, a stretch (G17–T21) of extended single-stranded DNA follows, until a peculiar U-turn sharply inverts the direction of the DNA backbone. The present structure confirms the 180° turn, where Met1 is nestled, and further shows that the DNA continues antiparallel to the extended G17–T21 stretch (Figure 1(b)) until A26. It also shows a mismatch pairing G22:T25, an interaction that was not observed in the complexes with shorter DNA, where T25 is found to be very mobile (Figure 5(b)). The strength of this mismatch G22:T25 interaction must be marginal, since it does not hold when the sugar-phosphate backbone ends at residue 25, i.e. after cleavage. The functional significance of this fact is not clear. It is interesting, though, that both guanine and thymine are conserved at the equivalent position of G22 and T25 in the different *nic* sequences of the F-family of conjugative plasmids (see Figure 4b of Guasch *et al.*[12]). In the TraI–DNA10 complex,[14] the equivalent thymine and guanine are also H-bonded. In any case, this interaction must help to orient the DNA to the active site entrance and to place scissile phosphate A26 at close distance from the histidine-triad metal binding site.

Catalytic mechanism and role of the metal ion

One of the non-bridging oxygen atoms of the scissile phosphate is in the appropriate orientation to be the fourth ligand of the metal ion (Figure 2). No metal ion is present in the TrwC^{Y18F}–DNA27 active site despite efforts to co-crystallize the TrwC–DNA27 complex with metal ions and soaking attempts of derived crystals. The co-crystallization experiment did not produce crystals and soaking damaged the crystals as evidenced by lack of diffraction. However, a water molecule binds at the metal site and mimics the metal position and interactions. Indeed, when comparing the active sites of the metal-bound TrwC–DNA25 (e.g. the Cu-bound structure shown in Figure 2(b)) and TrwC^{Y18F}–DNA27 (Figure 2(a)) we can see that the tetrahedral coordination is similar and that the water molecule that is the fourth ligand in the TrwC–DNA25 structure is replaced by the phosphate oxygen in the TrwC^{Y18F}–DNA27 structure. Superimposition of both structures shows that this oxygen atom of the scissile phosphate would be located at a coordination distance (∼2.2 Å) from the metal ion. Moreover, in the TraI–DNA10 complex[14] two oxygen atoms of the scissile phosphate group are coordinated with the metal ion. These observations and the fact that a metal ion is necessary for the cleavage reaction allow us to propose a catalytic mechanism for the first transesterification reaction (steps 1–2 in Figure 7(b), below) where the scissile phosphate is polarized by the metal ion promoting the nucleophilic attack of the hydroxyl group of Tyr18 (Figure 6).

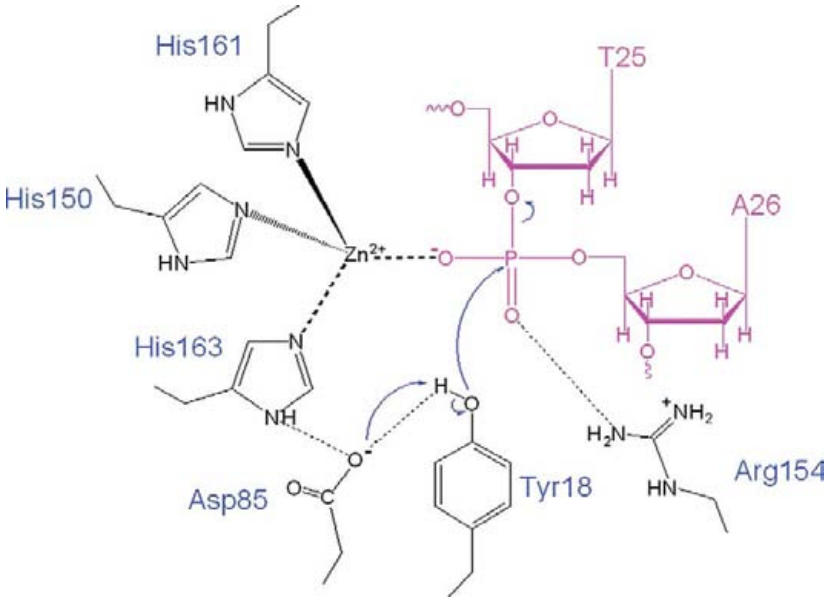

Figure 6. Schematic diagram illustrating the proposed catalytic mechanism for the first DNA cleavage reaction (see Discussion).

From the different structures we see that Tyr18 has the capacity to reorient its side-chain and, thus, it would be able to approach its reactive hydroxyl oxygen atom to the phosphate for the formation of a penta-coordinate intermediate. The metal ion would also stabilize this intermediate, which would be further stabilized by Arg154. In the $TrwC^{Y18F}$–DNA27 structure Arg154 is H-bonded to the non-bridging phosphate oxygen atom that does not interact with the metal. Although difficult to assess, the movement of the side-chain of Tyr18 may be correlated to the increased order in the loop region upstream in the sequence, as is the case in the metal-bound structures of TrwC–DNA25 (see below). The concomitant changes in this loop region are of interest because it contains residue Tyr26, a residue proposed to participate in the second cleavage reaction catalyzed by TrwC.

How Tyr18 gets activated for nucleophilic attack is an open question, but the proximity of Asp85, at hydrogen bond distance from the Tyr18 hydroxyl group, and the fact that its replacement with Ala impairs the reaction led us to propose[12] that this residue acts as a general base, abstracting a proton from the tyrosine hydroxyl group. The present data do not provide any evidence against that hypothesis: all structures presented herein (except for the Tyr18Phe mutant) show an H-bond between Asp85 and the hydroxyl group of Tyr18, even though the side-chain of the tyrosine suffers slight displacements. On the other hand, from the solved structures, there is no clear candidate for a general acid, a residue that should provide the proton to oxygen O3′(T25) to resume the cleaving reaction.

The nature of the metal ion

Another open question was what divalent metal ion was physiologically used by TrwC. *In vitro* reaction assays show that, in the absence of any metal, TrwC is unable to cleave the DNA. On the other hand, the addition of either Mg^{2+}, Mn^{2+}, Ca^{2+}, Zn^{2+}, Ni^{2+}, Cu^{2+} or Co^{2+} results in significant DNA cleavage. It should be noted, however, that a much higher concentration was needed for Mg or Ca-catalyzed cleavage, when compared to Zn^{2+}, Ni^{2+} and Cu^{2+}. This is in agreement with the fact that the affinity for Mg^{2+} is four orders of magnitude lower than for Zn^{2+} (M.L. & F. de la C., unpublished results). Our results with the soaking experiments confirm that Zn^{2+} binds readily to the active site, the histidine triad providing an appropriate metal binding site for this kind of ion. Cu^{2+} and Ni^{2+} display similar characteristics. On the other hand we did not find any evidence for Mg^{2+} or Mn^{2+} binding, although our cleavage reaction assays with oligonucleotides (see Results) indicate that both metal ions activate the enzyme (Mg^{2+} requiring a higher concentration than Mn^{2+}). A survey of all Mg^{2+} protein structures presently available does not show any case where this cation is triple-coordinated by histidine residues, except for the relaxase TraI structure where the electron density at the metal site was assigned to a Mg^{2+}.[4,13] In fact, all other Mg^{2+}-containing proteins in the PDB coordinate the metal ion with at least one carboxylate group. Furthermore, the most common Zn^{2+}-chelating sphere in proteins (when this cation is catalytic) is a tetrahedral arrangement of three histidine residues and a water molecule.[15] Such arrangement is observed in all our metal-bound structures. On the contrary, there is no indication in the TrwC active site architecture for the typical octahedral Mg^{2+} or Mn^{2+} coordination sphere.[16] It has to be noted, though, that all metal-binding data we present correspond to experiments carried out with crystals containing the DNA25 oligonucleotide, thus not including the scissile phosphate group near the metal site. Unfortunately, all soaking experiments

with the DNA27 oligonucleotide resulted in crystals that either lacked any metal at the active site (short time soaking) or did not diffract. Therefore, we can not rule out the binding of Mg^{2+} or Mn^{2+} to the TrwC active site, since the presence of phosphate A26 could provide a better environment for these metals to bind.

Two exit paths for the DNA

One important feature of the TrwC^{Y18F}–DNA27 complex is the fact that the DNA downstream of the cleavage site does not follow one of the paths that could be inferred from our previous work with the metal-free TrwC–DNA25 complex (PDB code 1OMH).[12] From that structure, it was apparent that the deep active site crevice continued after the metal binding site. Two sulphate ions, one close to the metal binding site and another 7 Å further away (labelled S_O and S_B in Figure 7(a)) suggested favorable locations for DNA phosphate groups after residue 25. Our prediction has been proven correct for the first position, since in the present TrwC^{Y18F}–DNA27 structure the scissile phosphate (A26) occupies the first sulphate S_O position. But after that the DNA does not follow the crevice. Instead, it continues "upwards" (Figure 1), anti-parallel to the single-stranded DNA stretch that extends between the hairpin and the U-turn. Since the crevice is almost closed over the active site by residues Arg14 and Arg154, the DNA seems to go, in the TrwC^{Y18F}–DNA27 structure, towards a *cul-de-sac*. However, closer inspection indicates that an exit path for the DNA is possible, as shown by model building. Indeed, a third sulphate ion was observed in that path, bound to Arg14 and Arg154 (sulphate S_A and path A in Figure 7(a)). In agreement, in TraI–DNA10 complex,[14] the DNA follows a similar path A. Still, we may speculate that an alternative functional path exists (path B) following the active site crevice as suggested by sulphate S_B. This sulphate ion is present in both the TrwC–DNA25[12] structure and in the TrwC^{Y18F}–DNA27 structure reported here, close to conserved residues Glu151 (an Asp in some relaxases of the family) and Thr152 (Figure 7(a)).

An attractive hypothesis is that both paths are used by the enzyme. In this hypothesis, the initiation and termination reactions will occur as schematized in Figure 7(b). We propose that, for the initiation reaction, the DNA follows path A. This conformation might be the preferred one in the constrained binding to supercoiled DNA, where only Tyr18 can cleave the DNA.[11] After the first cleavage, Tyr18 forms a covalent complex with the DNA 3′ to the cleavage site. The DNA moiety containing the 3′-OH leaves the active site and is extended by rolling-circle replication. Meanwhile, the relaxase is shot to the recipient cell, where it tracks on the incoming DNA by means of its 5′→3′ DNA helicase activity. When the reformed *nic* site DNA enters the recipient cell, it is recognized again and bound by the relaxase following path B (Figure 7(b), blue + green strand). It has to be so because path A is still occupied by the 3′ moiety of the DNA covalently attached to Tyr18. Besides, it is known that TrwC recognizes only specific bases 5′ to the cleavage site.[12] Since Tyr18 can swing (Figure 5(a)), it might well get slightly away from the active site to leave room for the new DNA (Figure 7(b), blue + green strand).

Role of the second catalytic tyrosine

The second cleavage reaction could be catalyzed by Tyr26, located at the tip of the α1–β2 loop as shown in Figure 2(b). After second cleavage, a new 3′-OH appears at the protein active site and the religation reaction with the 5′ end phosphate could take place by reversal of the initial cleavage reaction, thus re-circularizing the transferred DNA strand, as shown in Figure 7(b). We know from the different structures that a 3′-free T25 can rotate by an angle of 180° (Figure 5(b)). This ability could facilitate the adjustment of the 3′-OH position for attacking the phospho-tyrosine linkage.

In the different structures loop α1–β2 appears either fully disordered from residue 20 to 30 and not modelled (TrwC–DNA25, TrwC^{Y18F}–DNA27), partially ordered and partially modelled (TrwC–Ni^{2+}–DNA25) or fully modelled (TrwC–Cu^{2+}–DNA25). Still, in this later case, the temperature factors are rather high indicating a considerable flexibility. Where Tyr26 could be built in (Figure 2(b)) it appears to be at 17 Å distance from the metal site, its hydroxyl group pointing out in the opposite direction and making a hydrogen bond with the main-chain of Ala96, a residue of helix α4. A long movement, with the loop α1–β2 folding back towards the active site, must take place if the Tyr26 has to come in close proximity to the scissile phosphate. We cannot exclude, though, that this second cleavage is performed by a second TrwC molecule. In favor of this hypothesis is the fact that Tyr26 is located at the surface of the protein. However, in order to reach the active site of another relaxase molecule, the loop α1–β2 must further extend and stick out of the protein surface because the active site of the protein is located in a rather deep crevice. An example of such *trans*-attack mechanism is IS608 transposase. IS608 transposase belongs to the family of rolling-circle transposases, which contain catalytic Tyr residues and histidine triad motifs in their active sites. They are thus related to the superfamily of rolling-circle replication initiation proteins and conjugative relaxases.[17] The atomic structure of IS608 transposase[18] shows that the IS608 active Tyr is located in an α-helix that reaches the scissile bond of the target DNA located at the histidine triad catalytic site of the adjacent protein monomer.

Provided one relaxase molecule has been shot to the recipient, as proposed in the shoot and pump conjugation model[3] and recently confirmed experimentally,[19] there are two possible locations for termination reaction, the donor or the recipient

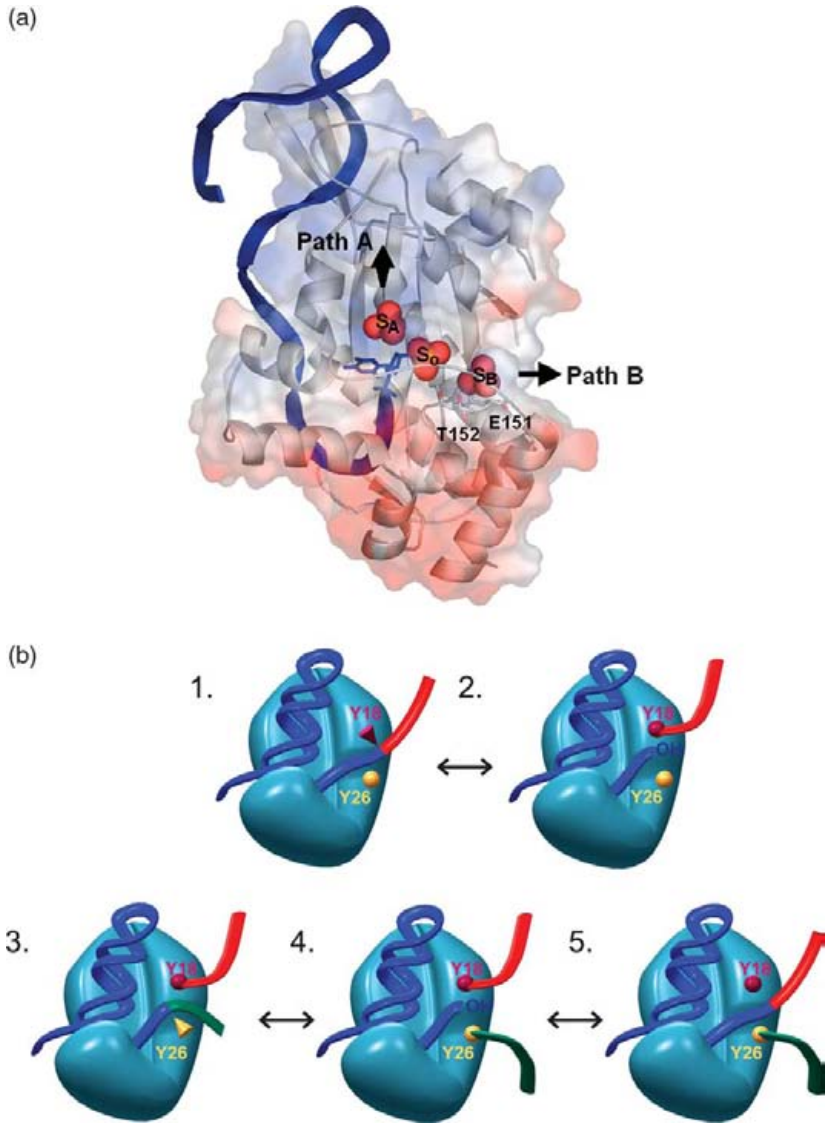

Figure 7. (a) Electrostatic potential surface representation of the TrwC^{Y18F}–DNA27 complex. The 27-mer DNA oligonucleotide is shown as a blue ribbon. Three sulphate ions, labelled S_A, S_O and S_B, are present in the active site crevice. Two putative exit paths for the DNA (A and B) are indicated with black arrows. (b) Schematic representation of the molecular mechanism of the TrwC relaxase.

cell. If termination occurs in the donor, it would have to be catalyzed by a second relaxase molecule, since the first has been transported to the recipient cell already. The second molecule can cleave the trailing *nic* site, and when the resulting 3′-OH reaches the recipient, it is recognized by the transported relaxase which can catalyze the reverse of the initial cleavage reaction or ligation reaction, resulting in circularization of the transferred DNA. Alternatively, if second cleavage occurs in the recipient, two possibilities remain. Either a second relaxase molecule produces the second cleavage or it is performed by the same molecule using a second catalytic residue. In the first two scenarios no second catalytic Tyr is needed, since re-circularization will occur by reversal of the initial cleavage reaction, as proposed by Byrd & Matson.[20] Although the two-Tyr flip-flop mechanism is appealing, and it is backed by biochemical data,[11] with the present data it is not possible to rule out any of the three alternative scenarios.

Materials and Methods

Protein preparation

The N-terminal relaxase domain of TrwC (residues 1 to 293), and mutant Tyr18Phe (named TrwC and TrwC^{Y18F}, respectively, here) were expressed and purified as follows.[21] *Escherichia coli* strain C43-DE3[22] containing plasmid pSU1588 was grown in a 2 l micro-DCU fermentation system (B. Biotech International). Plasmid pSU1588 carries the *trwC* fragment corresponding to the protein residues 1 to 293 cloned between the NdeI and BamHI sites of vector pET3a. Overexpression was induced with 0.5 mM IPTG for 2 h. Cells were harvested by centrifugation and resuspended in 50 mM Tris–HCl (pH 7.6), 10 mM EDTA and 10% (v/v) sucrose, and stored at −70 °C. Frozen cells were thawed and incubated with 10 mM benzamidine, 1 mM phenylmethylsulfonyl fluoride and 0.5 mg/ml lysozyme for 45 min at 0 °C. An equal volume of a solution containing 50 mM Tris–HCl (pH 7.6), 400 mM NaCl, 0.1 mM EDTA, 1 mM DTT and 0.5% (v/v) Triton-X-100 was added and the mixture incubated for

another 5 min on ice. The lysate was centrifuged at 145,000*g* for 30 min at 4 °C. Supernatants were applied to a P11-phosphocellulose column equilibrated in buffer A (50 mM Tris–HCl (pH 7.6), 200 mM NaCl, 0.1 mM EDTA) and eluted at 600 mM NaCl. Fractions containing N293–TrwC were pooled, diluted to 200 mM NaCl, loaded to a MonoS column (Amersham), and eluted with a linear NaCl gradient (200 mM–600 mM NaCl) in buffer A. Finally, a gel filtration step was carried out in a Superdex75HR column (Amersham) equilibrated in 50 mM Tris–HCl (pH 7.6), 500 mM NaCl, 0.1 mM EDTA. TrwC–N293 eluted as a monomer of about 40 kDa. Its concentration was estimated by UV absorbance at 280 nm using an extinction coefficient of 31,500 M^{-1} cm^{-1}, which was calculated as described by Edelhoch *et al*.[23] No loss of activity was observed after six months storage at −70 °C.

Purification of N293–TrwC labelled with seleno-methionine (SeMet) was performed using B834(DE3) strain and New Minimal Medium supplemented with SeMet.[24] The purification was performed according to the described procedure. For subsequent crystallization trials, both protein variants were concentrated to 2–6 mg/ml and stored at −20 °C in 50 mM Hepes/NaOH (pH 7.6), 450 mM NaCl.

Oligonucleotide cleavage reactions

Reaction mixtures contained 0.5 μM fluorescein-labeled oligonucleotide R388(12+18) (TGCGTATTGTCT/ATAGCCCAGATTTAAGGA), 1 μM TrwC–N293 in cleavage buffer (10 mM Tris (pH 7.6), 100 mM NaCl, 10 μM EDTA) plus or minus a variety of divalent metal salts ($MgCl_2$, $MnCl_2$, $NiCl_2$, $ZnCl_2$, $CuCl_2$ and $CaCl_2$). Metal ion concentrations ranged from 1 μM to 10 mM. After incubation for 30 min at 37 °C, samples were treated with 0.6 mg/ml proteinase K and 0.05% (w/v) SDS for 20 min at 37 °C. Reaction products were separated and quantified by capillary electrophoresis, using the CE Oligonucleotide Analysis Kit (BioRad) in the capillary system BioFocus®2000 (BioRad). The capillary used was a BioCAP Oligonucleotide Analysis Capillary (30 cm×75 μm i.d.×375 μm o.d.). Samples were introduced in the capillary by pressure (200 psi/s). Electrophoresis was carried out at 12 kV at 40 °C. A laser-induced detector was employed for fluorescence detection. Peak information (migration time, peak area and height) was obtained using the CE Integrator Software (BioRad).

Crystallization and structure determination of TrwC^{Y18F}–DNA27 complex

In order to prepare the protein–DNA complex, a solution 0.23 mM SeMet TrwC^{Y18F}, 450 mM NaCl, 50 mM Hepes (pH 7.6) was mixed with a solution of 0.23 mM of annealed DNA oligonucleotide, encompassing the *nic* sequence 25 nucleotides upstream and two nucleotides downstream from the cleavage site (DNA27=5′-GCGCACCGAAAGGTGCGTATTGTCT-3′), dissolved in the same buffer and salt. The resulting TrwC^{Y18F}–DNA27 complex was purified by size-exclusion chromatography and concentrated to 6.9 mg/ml of protein. Crystals were obtained by vapor diffusion, at 4 °C, from hanging drops prepared by mixing 2 μl of the concentrated TrwC solution and 1 μl of the precipitant solution containing 32% (w/v) PEG MME 2000, 0.3 M ammonium sulphate, 0.1 M sodium acetate (pH 4.6). Crystals appeared after seven days and continued to grow to a maximum size after 27 days. A complete dataset was measured at beamline ID14-4, at ESRF (Grenoble). Data collection is summarized in Table 1. The TrwC–DNA25 (see below for definition of DNA25) metal-free structure previously published[12] (PDB code 1OSB), modified to correctly represent the Tyr18Phe mutation, was used as a starting model for molecular replacement. Refinement was carried out with REFMAC5.2, part of the CCP4 program suite.[25] Initial refinement indicated a change in the position of the base T25 of the DNA strand. Subsequently, this residue and water molecules in and around the active site were removed, after which refinement was restarted. The DNA residue T25 could be correctly placed upon inspection of the $2F_o-F_c$ and F_o-F_c Fourier maps. Another cycle of restrained refinement and map inspection revealed the position of DNA residue A26, which was subsequently added. Residue 27, however, could not be unambiguously located and was not modelled. Final refinement statistics are shown in Table 1.

Crystallization and structure determination of TrwC–DNA25 metal ion complexes

In order to prepare TrwC–DNA binary complexes, a solution of 0.179 mM TrwC in 450 mM NaCl, 50 mM Hepes (pH 7.6), was mixed with a solution of 0.179 mM of annealed DNA oligonucleotide, encompassing the *OriT* sequence 25 nucleotides upstream from the *nic* site (DNA25=5′-GCGCACCGAAAGGTGCGTATTGTCT-3′) dissolved with similar buffer and salt. The resulting TrwC–DNA25 complex was purified by size-exclusion chromatography and concentrated to 6.8 mg/ml of protein. Crystals were grown as described.[12] These crystals were used to prepare different metal ion complexes by soaking them in metal ion solutions and X-ray diffraction data were collected at the ESRF (Grenoble), as follows: (1) crystals soaked in 5 mM $ZnSO_4$ and 10 mM $MgCl_2$ for 24 h were measured at the K absorption edge of Zn (data set 25ZNpkMG) and 24 eV below this edge (data set 25ZNrmMG), on beamline ID14-4; (2) crystals soaked in 5 mM $ZnSO_4$, 10 mM $MnCl_2$ for 24 h were measured at both the K absorption edges of Zn (data set 25ZNpkMN) and Mn (data set 25ZNMNpk), on beamline BM16; (3) crystals soaked in 5 mM $MnCl_2$ for 24 h (data set 25MN) were measured at the K absorption edge of Mn, on beamline BM16; (4) crystals soaked in 5 mM $CuCl_2$ for 24 h (data set 25CU) were measured below the K absorption edge of Cu (0.97855 Å), at beamline ID23EH1; and (5) crystals soaked in 5 mM $NiCl_2$ for 24 h (25NI) were measured below the K absorption edge of Ni, at beamline BM14. Data collection statistics are summarized in Table 1.

Refinement against datasets 25ZNpkMG and 25ZNpkMN was performed with REFMAC5.2 program in the CCP4 program suite,[25] starting from a rigid body refinement using the previously deposited Zn^{2+}-bound structure of TrwC complexed with the same DNA25 fragment[12] (PDB code 1QX0), followed by a restrained refinement. For the 25MN structure a similar protocol was used, starting from the deposited metal-free structure[12] (PDB code 1OMH). Anomalous maps were calculated from a combination of the corresponding phases and the anomalous differences associated with each dataset. The metal-free protein structure (PDB code 1OMH) was also used as initial model for refinement against data of the Cu^{2+} and Ni^{2+}-soaked crystals. Initial inspection of the density maps clearly indicated the presence of metal ions at the active site. Subsequent simulated annealing and

Table 1. Data collection and refinement statistics

Data set[a]	27NAT	25ZNpkMG	25ZNrmMG	25ZNpkMN	25ZNMNpk	25MN	25CU	25NI
A. *Data collection*								
Unit cell (Å)	$a=b=148.42$, $c=75.52$	$a=b=90.70$, $c=204.22$	$a=b=90.75$, $c=204.36$	$a=b=90.69$, $c=204.22$	$a=b=90.53$, $c=203.77$	$a=b=90.24$, $c=202.08$	$a=b=91.088$, $c=205.55$	$a=b=92.7312$, $c=208.8551$
Space group	$P6_5$	$P6_122$	$P6_122$	$P6_122$	$P6_122$	$P6_122$	$P6_122$	$P6_122$
Wavelength (Å)	0.97930	1.2820	1.2852	1.2823	1.89223	1.8924	0.97855	1.00100
Resolution range (Å)[b]	25.0–2.7 (2.9–2.7)	19.8–2.7 (2.9–2.7)	19.8–2.7 (2.9–2.7)	51.4–3.16 (3.2–3.0)	45.3–2.90 (3.1–2.9)	43.9–2.6 (2.7–2.6)	24.7–2.5 (2.6–2.5)	43.9–2.3 (2.4–2.3)
No. of observed reflections[b]	165,382 (22,961)	169,473 (24,851)	146,870 (21,542)	120,010 (17,326)	130,924 (18,815)	171,348 (25,127)	386,179 (56,283)	114,701 (13,061)
No. of unique reflections[b]	27,451 (3996)	143,70 (2037)	14,393 (2039)	10,633 (1496)	11,442 (1598)	15,710 (2226)	18,252 (2592)	28,857 (3506)
Completeness[b]	99.7 (100)	99.9 (100)	99.8 (100)	100 (100)	98.9 (98.0)	99.9 (99.8)	99.9 (100)	98.8 (98.8)
R_{merge} [b,c]	10.9 (41.3)	9.0 (22.4)	5.9 (15.6)	14.1 (42.7)	9.9 (36.1)	12.7 (45.5)	10.4 (51.6)	4.9 (14.6)
$I/\sigma(I)$[b]	5.4 (1.5)	5.7 (3.2)	9.1 (4.6)	4.8 (1.7)	7.1 (2.0)	4.6 (1.5)	5.6 (1.4)	10.2 (4.8)
Multiplicity[b]	6.0 (5.7)	11.8 (12.2)	10.2 (10.6)	11.3 (11.6)	11.4 (11.8)	10.9 (11.3)	21.2 (21.7)	4.7 (4.1)
Anomalous completeness[b]	99.0 (99.9)	99.9 (100)	99.9 (100)	100 (100)	99.4 (98.6)	100 (100)	100(100)	98.3 (88.9)
Anomalous multiplicity[b]	3.1 (2.9)	6.5 (6.6)	5.6 (5.6)	6.4 (6.3)	6.4 (6.4)	5.9 (6.0)	11.7(11)	2.5 (2.2)
B. *Refinement*								
Protein atoms[d]	2213+2213	2288	Not refined[e]	2288	Not refined[e]	2226	2314	2275
DNA atoms[d]	492+475	512		512		512	512	512
Metal atom[d]	–	Zn^{2+}		Zn^{2+}		–	Cu^{2+}	Ni^{2+}
Water O atoms[d]	239	178		178		256	207	162
No. of reflections used in refinement[f]	26,050	13,597	Not refined[e]	10,055	Not refined[e]	10,614	18,200	24,598
R factor (%)[f]	20.5	21.8		19.9		20.4	23.3	20.6
R_{free} (%)[f]	24.7	26.0		24.0		25.0	27.6	24.5
rmsd from target values								
Bonds (Å)	0.009	0.009		0.011		0.010	0.005	0.013
Angles (deg.)	1.37	1.31		1.44		1.35	1.52	1.35
Chiral (Å)	0.10	0.075		0.079		0.074	0.080	0.086
Mean B factor (Å^2)	49.2	34.0		32.4		31.6	37.1	38.2

[a] See Materials and Methods for an explanation.
[b] Outermost resolution shell values are in parenthesis.
[c] $R_{merge} = [\Sigma_{hkl}\Sigma_i|I_i(hkl) - <I(hkl)>|/\Sigma_{hkl}\Sigma_i I_i(hkl)] \times 100$.
[d] Per asymmetric unit.
[e] Structures were not refined against data sets 25ZNrmMG and 25 NMNpk.
[f] $R_{factor} = [\Sigma_{hkl}||F_{obs}| - k|F_{calc}||/\Sigma_{hkl}|F_{obs}|] \times 100$; R_{free}, same as for a test set of reflections not used during refinement.

positional and overall *B*-factor refinement were performed using CNS.[26] The correction of the model, through visual inspection of the electron density maps using TURBO–FRODO,[27] was alternated with solvent molecule addition and refinement, including atomic *B*-factor refinement, using CNS. Final refinement statistics are shown in Table 1. Anomalous difference Fourier maps (Bijvoet maps) were calculated from various datasets as indicated in the legend to Figure 4. All pictures were prepared using PyMol†, Bobscript extention of Molscript[28] and the Raster3D package.[29]

Protein Data Bank accession codes

Atomic coordinates have been deposited with the RCSB PDB under accession codes 2CDM (TrwC^{Y18F}–DNA27), 1S6M (TrwC-Ni^{2+}–DNA25) and 1ZM5 (TrwC-Cu^{2+}–DNA25).

Acknowledgements

This study was supported by the Ministerio de Educación y Ciencia of Spain (grants BIO2002-03964, GEN2003-20642 and BFU2005-06758/BMC to M.C.; grant BMC2002-00379 to F.C.) and by the Generalitat de Catalunya (grant 2001SGR-346 to M.C.). Synchrotron data collection was supported by the ESRF and the EU. S.R. is the recipient of a predoctoral fellowship from the Generalitat de Catalunya. M.L. acknowledges a fellowship from the Fundación Marqués de Valdecilla IFIMAV.

References

1. Ochman, H., Lawrence, J. G. & Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature*, **405**, 299–304.
2. de la Cruz, F. & Davies, J. (2000). Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends Microbiol.* **8**, 128–133.
3. Llosa, M., Gomis-Ruth, F. X., Coll, M. & de la Cruz Fd, F. (2002). Bacterial conjugation: a two-step mechanism for DNA transport. *Mol. Microbiol.* **45**, 1–8.
4. Zechner, E. L., de la Cruz, F., Eisenbrandt, T., Grahn, A. M., Koraimann, G., Lanka, E. *et al.* (2000). Conjugative DNA transfer processes. In *The Horizontal Gene Pool: Bacterial Plasmids and Gene Spread* (Thomas, C. M., ed.), pp. 87–174, Harwood Academic Publishers, London.
5. Francia, M. V., Varsaki, A., Garcillan-Barcia, M. P., Latorre, A., Drainas, C. & de la Cruz, F. (2004). A classification scheme for mobilization regions of bacterial plasmids. *FEMS Microbiol. Rev.* **28**, 79–100.
6. Lanka, E. & Wilkins, B. M. (1995). DNA processing reactions in bacterial conjugation. *Annu. Rev. Biochem.* **64**, 141–169.
7. Pansegrau, W. & Lanka, E. (1996). Mechanisms of initiation and termination reactions in conjugative DNA processing. Independence of tight substrate binding and catalytic activity of relaxase (TraI) of IncPalpha plasmid RP4. *J. Biol. Chem.* **271**, 13068–13076.
8. Llosa, M., Grandoso, G., Hernando, M. A. & de la Cruz, F. (1996). Functional domains in protein TrwC of plasmid R388: dissected DNA strand transferase and DNA helicase activities reconstitute protein function. *J. Mol. Biol.* **264**, 56–67.
9. Llosa, M., Grandoso, G. & de la Cruz, F. (1995). Nicking activity of TrwC directed against the origin of transfer of the IncW plasmid R388. *J. Mol. Biol.* **246**, 54–62.
10. Hanai, R. & Wang, J. C. (1993). The mechanism of sequence-specific DNA cleavage and strand transfer by phi X174 gene A* protein. *J. Biol. Chem.* **268**, 23830–23836.
11. Grandoso, G., Avila, P., Cayon, A., Hernando, M. A., Llosa, M. & de la Cruz, F. (2000). Two active-site tyrosyl residues of protein TrwC act sequentially at the origin of transfer during plasmid R388 conjugation. *J. Mol. Biol.* **295**, 1163–1172.
12. Guasch, A., Lucas, M., Moncalian, G., Cabezas, M., Perez-Luque, R., Gomis-Ruth, F. X. *et al.* (2003). Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nature Struct. Biol.* **10**, 1002–1010.
13. Datta, S., Larkin, C. & Schildbach, J. F. (2003). Structural insights into single-stranded DNA binding and cleavage by F factor TraI. *Structure (Camb)*, **11**, 1369–1379.
14. Larkin, C., Datta, S., Harley, M. J., Anderson, B. J., Ebie, A., Hargreaves, V. & Schildbach, J. F. (2005). Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the f factor relaxase. *Structure (Camb)*, **13**, 1533–1544.
15. Dudev, T. & Lim, C. (2003). Principles governing Mg, Ca, and Zn binding and selectivity in proteins. *Chem. Rev.* **103**, 773–788.
16. Bock, C. W., Katz, A. K., Markham, G. D. & Glusker, J. P. (1999). Manganese as a replacement for magnesium and zinc: functional comparison of the divalent ions. *J. Am. Chem. Soc.* **121**, 7360–7372.
17. Garcillan-Barcia, M. P., Bernales, I., Mendiola, M. V. & de la Cruz Fd, F. (2005). IS91 rolling-circle transposition. In *Mobile DNA II* (Craig, N. L., Craigie, R., Gellert, M. & Lambowitz, A., eds), pp. 891–904, ASM Press, Washington, DC.
18. Ronning, D. R., Guynet, C., Ton-Hoang, B., Perez, Z. N., Ghirlando, R., Chandler, M. & Dyda, F. (2005). Active site sharing and sub-terminal hairpin recognition in a new class of DNA transportases. *Mol. Cell*, **20**, 143–154.
19. Draper, O., César, C., Machón, C., de la Cruz, F. & Llosa, M. (2005). Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. *Proc. Natl Acad. Sci. USA*, **102**, 16385–16390.
20. Byrd, D. R. & Matson, S. W. (1997). Nicking by transesterification: the reaction catalysed by a relaxase. *Mol. Microbiol.* **25**, 1011–1022.
21. Grandoso, G., Llosa, M., Zabala, J. C. & de la Cruz, F. (1994). Purification and biochemical characterization of TrwC, the helicase involved in plasmid R388 conjugal DNA transfer. *Eur. J. Biochem.* **226**, 403–412.
22. Miroux, B. & Walker, J. E. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260**, 289–298.

† http:/www.pymol.org

23. Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry*, **6**, 1948–1954.
24. Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J. & Huber, R. (1995). High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. *Eur J Biochem.* **230**, 788–796.
25. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
26. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al*. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
27. Roussel, A. & Cambilleau, C. (1989). Turbo-Frodo. In *Silicon Graphics Geometry Partners Directory*, pp. 77–79. Silicon Graphics, Mountain View, CA.
28. Esnouf, R. M. (1999). Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. *Acta Crystallog. sect. D*, **55**, 938–940.
29. Merritt, E. & Bacon, D. (1997). Raster3D: photorealistic molecular graphics. In *Methods in Enzymology* (Carter, C., Jr & Sweet, R., eds), vol. 277, pp. 505–524, Elsevier, San Diego.

Edited by R. Huber

(Received 27 October 2005; received in revised form 3 February 2006; accepted 8 February 2006)
Available online 28 February 2006

PLASMIDS

CURRENT RESEARCH AND FUTURE TRENDS

Edited by:

Georg Lipps
University of Bayreuth
Department of Biochemistry II
Universitätsstr. 30
D-95440 Bayreuth
Germany

Caister Academic Press

www.caister.com

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN: 978-1-904455-35-6

Printed and bound in Great Britain

PLASMIDS

CURRENT RESEARCH AND FUTURE TRENDS

Edited by:

Georg Lipps
University of Bayreuth
Department of Biochemistry II
Universitätsstr. 30
D-95440 Bayreuth
Germany

Caister Academic Press

Chapter 6

Molecular Machinery for DNA Translocation in Bacterial Conjugation

Silvia Russi, Roeland Boer and Miquel Coll

ABSTRACT

Whatever the route used, horizontal transfer of DNA requires elaborated multi-protein machinery to enable the long and charged nucleic acid polymer to cross the cell envelope barriers. The best-studied system for cell-to-cell DNA translocation is bacterial conjugation. This system can be divided in two discrete specialized modules: the relaxosome, which triggers and takes part in plasmid DNA processing and replication, and a type IV secretion system (T4SS), which impels protein and single-stranded DNA through the membranes. In addition, a coupling protein (CP), linking both modules, and a number of ancillary proteins are needed. Over the last decades research efforts in the field have resulted in the clarification of many aspects of this system and its machinery assembly. In particular, structural biology has provided details of the molecular architecture of several of the pieces involved in this intricate scenario.

INTRODUCTION

Single-stranded DNA (ssDNA) transfer in conjugation, DNA uptake in transformation, and double-stranded DNA (dsDNA) transfer -via bacteriophages- in transduction, are all different routes by which a DNA molecule can be translocated across the bacterial cell envelope (Boucher et al., 2003; Lederberg, 2000; Ochman et al., 2000). During the translocation process the DNA molecule has to cross several barriers of different physicochemical properties. For example, in Gram-negative bacteria, the hydrophobicity of the inner membrane, the presence of nucleases in the periplasm and the negative charge of the outer membrane are some of the hurdles associated to this transport (Dreiseikelmann, 1994). To overcome these barriers, sophisticated DNA translocation molecular machinery is employed.

Two of the DNA transfer routes, conjugation (including T-DNA transfer from bacteria to plants) and transformation, have some functional and structural elements in common (Chen et al., 2005). The third route, transduction, is not related to the others and requires the dsDNA packaging and injection machinery of a phage (Johnson and Chiu, 2007). In conjugation and transformation the transferred DNA

molecule is single-stranded. Still, there are marked differences in the two systems (Chen and Dubnau, 2004).

Among the different DNA translocation systems, conjugative DNA transfer is the best studied (Lanka and Wilkins, 1995; Silverman, 1997). The ability of bacteria to conjugate, had already been discovered in *E. coli* cultures in 1946 (Lederberg and Tatum, 1946). Gram-positive bacteria have also been found to harbour conjugative plasmids (Grohmann et al., 2003; Scott and Churchward, 1995; Ton-That and Schneewind, 2003). This has attracted much interest because of the clinical problems associated with it (Alekshun and Levy, 2007). The spread of antibiotic resistance genes in just a few decades since antibiotics started to be widely used, is largely due to this fast and powerful horizontal gene transfer mechanism.

The transfer of conjugative DNA is mediated by a multi-component molecular machinery, whose protein components are usually encoded on the (plasmid) DNA involved in the conjugation process (Llosa et al., 2002). Bacterial conjugation transmembrane transport channels are a subfamily of the type IV secretion system (T4SS) involved in the transfer of macromolecules, toxins and DNA-protein complexes across the cell envelope (Backert and Meyer, 2006; Burns, 2003; Cascales and Christie, 2003; Chen et al., 2005; Christie, 2004; Llosa and O'Callaghan, 2004; Sexton and Vogel, 2002). T4SS are therefore not restricted to the translocation of DNA. Example of non-DNA transporting T4SS are those from *Bordetella pertussis* (Ptl), responsible for whooping cough (Weiss et al., 1993); *Helicobacter pylori* (Cag), related to peptic ulcers, MALT-lymphoma and adenocarcinoma of the stomach (Graham, 2000; Wotherspoon, 1998); and *Bartonella henselae* (VirB/VirD4) causing the cat scratch disease (Cascales and Christie, 2003). Bacterial conjugation of DNA may be, in fact, an adaptation of ancestral translocation machinery that transfers proteins (Christie et al., 2005). These transport systems are functionally diverse, not only in terms of the transported substrate, but also with respect to the recipient cell, which can be bacteria of the same or different species, fungi, plants or mammalian cells (Christie, 2001; Christie and Cascales, 2005; Christie and Vogel, 2000; Ding et al., 2003; Seubert et al., 2003; Waters, 2001). In general, the substrate delivery to the recipient cell occurs via direct cell-to-cell contact, although there are examples of contact-independent protein export and DNA release (Dillard and Seifert, 2001; Hofreuter et al., 2001).

In recent years, X-ray crystallography and NMR spectroscopy studies have provided valuable structural information on various proteins involved in the conjugation process (Gomis-Ruth and Coll, 2006; Remaut and Waksman, 2004; Yeo and Waksman, 2004). This information, revealing the molecular architecture of some of the central pieces of the mechanism, has answered some of the fundamental questions about T4SS-mediated DNA transport machinery. Here we will review

relaxosome and T4SS structures determined by X-ray crystallography and NMR spectroscopy.

COMPONENTS OF THE TYPE IV SECRETION SYSTEM

The T4SS components can be grouped into two modules according to their function: i) the relaxosome (Lanka and Wilkins, 1995; Wilkins and Lanka, 1993), ii) the trans-membranal conduit or channel. In addition, a coupling protein (CP) links both modules (Gomis-Ruth et al., 2004; Llosa and de la Cruz, 2005). The relaxosome is a nucleoprotein complex that prepares the T-DNA for transfer by specifically nicking the DNA at the origin of transfer (*oriT*) (Ziegelin et al., 1989) after the donor and recipient cells have come into contact (Durrenberger et al., 1991). It also is responsible for the delivery of the DNA to the T4SS secretion system (Fekete and Frost, 2000) through the action of the coupling protein (T4CP). Its protein subunits are encoded on a gene region called the DNA transfer replication region (*Dtr*). It consists of a relaxase, which also works as the pilot protein, and two or more accessory proteins.

The T4SS trans-membranal conduit is a membrane-spanning multi-protein complex that comprises copies of around 12 different proteins, which are encoded on the mating pair formation (Mtf) region (Backert and Meyer, 2006; Christie et al., 2005; Silverman, 1997). Its extracellular part, called the pilus, appears as an elongated tubular appendage of the donor cell in micrographs. The T4S system mediates the association between the host and recipient cells through the pilus, which binds to the surface of the recipient cell and brings both cell envelopes into close contact. What happens to the pilus after establishing the mating junction is not clear. In F plasmids, which encode for a group of proteins with little homology to most other conjugation machines, the pilus retracts after establishing cell-to-cell contact (Lawley et al., 2003). However, pilus retraction has not been demonstrated in other systems (Christie et al. 2005). Some systems lack the subunits dedicated to pilus retraction, which include plasmids RP4, R388 and pKM101 and the Gram-negative bacterium *Agrobacterium tumefaciens* (Harris and Silverman, 2004). Whether DNA passes through the extracellular part of the pilus during conjugation is not yet clear (Kalkum et al., 2004).

T4S systems can be classified into types IV-A, IV-B and into an as of yet unclassified group, according to their homology among the functional subunits. The IV-A systems are homologous to that of the *Agrobacterium tumefaciens* VirB/D system, which is the archetype of type IV-A secretion systems and has been studied extensively at biochemical and structural levels. This Gram-negative bacterium introduces an oncogenic DNA fragment, the T-DNA fragment, into plant cells (Christie, 2004). The T4SS components identified and mapped on the genome are found in the VirB and VirD genes, which correspond to the transmembrane conduit and relaxosome

respectively. The *virB* gene products comprise eleven proteins (VirB1 to VirB11) encoded by the *virB* operon. So far, two of the *virD* gene products, VirD2 and VirD4, have been shown to be involved in T4SS DNA translocation. VirD2 is responsible for correct DNA processing and is transported through the T4SS as a covalent complex with the T-DNA (Howard et al., 1992; Vergunst et al., 2000). Its function is comparable to that of TrwC of the relaxosome of plasmid R388 (Draper et al., 2005). VirD4 is the coupling protein in this system, analogous to TrwB of plasmid R388 (Christie, 1997).

Further along this paper the structural organization and the reaction mechanisms of the relaxosome components and the T4SS will be discussed in more detail, based on the available structural information. The section on the relaxosome will focus on the systems studied by us and others in the Gram-negative *E. coli* plasmids R388 and F and will include a discussion on the structure of the coupling protein, determined in our laboratory. In a subsequent section, the description of the structural information available for the secretion apparatus will be based on proteins of the DNA translocation membrane conduit of *Agrobacterium tumefaciens*, encoded on the bacterial non-translational chromosome. Many of the structures of the T4SS constituents of *A. tumefaciens* or closely related organisms have been determined in the laboratory of G. Waksman.

THE RELAXOSOME: OVERVIEW

The relaxosome is a nucleoprotein structure involved in the initiation of DNA-processing reactions that occur during bacterial conjugation. The proteins of the relaxosome bind specifically to the origin of transfer (*oriT*), a short DNA sequence where the process starts and ends. It comprises a relaxase/helicase protein, responsible for nicking the DNA and, in many cases, for unwinding the supercoiled plasmid DNA for further processing. This protein is also responsible for the reconstitution of a single-stranded circular DNA by transesterification of the 3' and 5' ends of the translocated DNA strand, presumably, in the recipient cell. In addition to the helicase, host factors such as the integration host factor (IHF) may form part of the complex and a transcriptional auto-regulatory protein is often included. Furthermore, an ancillary protein (TraM) is required in several plasmids, like F and R1 (Fig. 1a).

The architecture of the relaxosome of the *E. coli* plasmid R388 (Fig. 1b), one of the representative plasmids of the plasmid incompatibility group IncW (Moncalian et al., 1997; Moncalian et al., 1999b), has been studied in detail. In this plasmid, two of the encoded proteins (TrwA and TrwC) form the relaxosome, and a third protein (TrwB) is responsible for coupling the relaxosome to the base of the secretion channel. TrwC is the relaxase/helicase protein of this system. TrwA and IHF are the transcription regulator and the integration host factor, respectively.

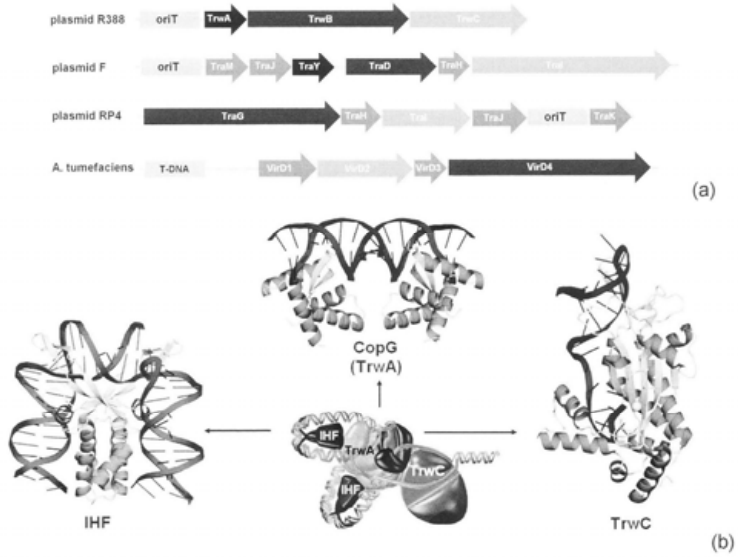

Fig. 1. a) Gene organization and nomenclature of the relaxosome proteins of conjugative plasmids R388, F, RP4 and Ti. b) Structural models of the relaxosome components in the *E. coli* plasmid R388 system: the TrwA-related transcriptional regulator CopG dimer in complex with its cognate 19-bp dsDNA (PDB code 1b01, upper central panel), *E. coli* host-encoded IHF in complex with a 35-bp dsDNA (PDB code 1ihf, left), relaxase/helicase TrwC relaxase domain in complex with a 27-base DNA hairpin (PDB code 2cdm, right). The central panel shows a modeled scheme of the assembly of the relaxosome complex.

The R388 system lacks a TraM ortholog and the relaxosome is assembled at the origin of transfer (*oriT*) DNA. The helicase activity of TrwC on the plasmid DNA separates the nicked parental strand from the complementary strand and allows the regeneration of duplex plasmid DNA through replicative synthesis of the latter by host polymerases (Kingsman and Willetts, 1978). The mechanism of this process is a variant of the rolling-circle replication mechanism (RCR) (Furuya and Komano, 2000; Kramer et al., 1997).

DNA BINDING FACTORS

TrwA is a 53 kDa tetrameric DNA-binding protein that specifically recognizes the oriT sequence of plasmid R388. The two binding sites (*sbaA* and *sbaB*) are not close to the *nic* site. Site *sbaA* and *sbaB* are located 125 to 150 bp and 67 to 96 bp downstream from *nic*, respectively. The relative location of these sites is similar to

the recognition sites *sbyA* and *sbyB* of the homologous protein TraY of plasmid F (Moncalian et al., 1997). TrwA has two biochemically independent roles in R388 conjugation. First, the binding of TrwA to a specific recognition sequence in *oriT*, different from that of TrwC, results in a transcriptional repression of the *trwABC* operon that encodes TrwA, TrwB, and TrwC. Secondly, TrwA has a direct role in *oriT* processing, enhancing the relaxation reaction catalyzed by TrwC on supercoiled DNA R388 *oriT* (Moncalian and de la Cruz, 2004). Analogous behaviour has been observed in the TrwC homolog TraY with the enhancement of the relaxase/helicase activity of TraI and the regulation of the expression of the *tra* genes (Lum et al., 2002). Since TrwA expression is autoregulated, it may be expected to play an important role on the regulation of the frequency of plasmid conjugation.

The N-terminal domain of TrwA and TraY is a DNA-binding domain that recognises the two palindromic sites around *oriT*. The C-terminal is a dimerization domain. The secondary structure prediction of the DNA-binding domain suggests a ribbon-helix-helix motif, as found in the Arc/MetJ/CopG superfamily of transcriptional repressors (Suzuki, 1995). The crystal structure of a representative member of this family of transcriptional regulators, CopG of plasmid pMV158, in complex with its cognate double-stranded DNA has been solved (Gomis-Ruth et al., 1998). It was found to recognize the DNA as a dimer of dimers (Fig. 1b). An antiparallel β-ribbon, formed by the ribbon-helix-helix motif of each of the protomers of the dimer, enters the DNA major groove. The cognate DNA is curved over 120° per protein tetramer. The homology between the N-terminal domain of CopG and that of TrwA and TraY, suggests that TrwA and TraY have a similar mode of action and help the DNA acquire the topology necessary for the optimal cleavage by way of this relaxase (del Solar et al., 2002).

The integration host factor (IHF) is a heterodimeric protein that binds specifically to the R388 oriT region at two different sites (*ihfA* and *ihfB*). Site *ihfA* lies between *nic* and *sbaB*, whereas site *ihfB* is located between *sbaA* and *sbaB* (Moncalian et al., 1999b). This protein is known to assist in many prokaryotic processes, such as replication, transcriptional regulation, and several site-specific recombination events, in which special DNA axis distortions are required. In the plasmid F system, IHF facilitates the cleavage reaction of the relaxase by affecting the topology of the *nic* DNA site (Howard et al., 1995), whereas it was reported to inhibit the TrwC *nic* cleavage reaction (Moncalian et al., 1999b). The crystal structure of an IHF-DNA complex (Rice et al., 1996) shows that the two 10 kDa component subunits of IHF, which share about 30% sequence identity, are structurally equivalent (Fig. 1b). An N-terminal segment, formed by two consecutive α-helices and three β-strands featuring an antiparallel sheet, is responsible for the dimerization of the protein. The β–ribbon shape of the DNA-binding domain allows it to penetrate the DNA minor groove and curl around the DNA producing a sharp bend into the double

helix which results in the formation of a ~160° U-turn (Rice et al., 1996). Two large kinks in the phosphate backbone, mediated by prolines at the tip of each β-ribbon, are responsible for most of the bending.

INITIATION AND PREPARATION

TrwC is a large protein (966 residues long) of 108 kDa harbouring two functionally different domains (Llosa et al., 1996). The recognition of the cleavage point at *oriT* and endonuclease, relaxase and transesterification activities take place at the N-terminal domain of TrwC, whereas the C-terminal domain displays a 5' to 3' DNA helicase activity. TrwC contains two active-site tyrosyl residues (instead of the single one in P-family relaxases) that catalyze the initiation and termination steps in conjugative DNA replication and a highly conserved histidine triad, a site for the binding of a metal ion, also required for the relaxase activity of the protein (Grandoso et al., 2000). Biochemical studies demonstrated that *in vitro*, the N-terminal relaxase domain, like the full-length TrwC, cleaves single-stranded oligonucleotides containing the *nic* site and performs complete strand transfer reactions in a way that resembles type I topoisomerases (Champoux, 2001; Wang, 1996). Furthermore, these studies stressed the importance of metal ions for driving catalysis.

TrwC recognizes the cleavage point at the *nic* site because the distinctive feature of the *nic* sequence: a six-base pair inverted repeat that forms a cruciform structure (Guasch et al., 2003). The supercoiled DNA plasmid is locally deformed and stabilized by TrwA, the *trw* gene promoter repressor, and the IHF protein. After the initial cleavage of the T-strand, its 3' end to *nic* remains covalently bound to a catalytic TrwC tyrosine by a 5'-phosphotyrosyl linkage during transfer to the recipient cell, whereas the 3' end is released (Llosa et al., 2002; Wilkins and Lanka, 1993). TrwC, acting as a helicase, now migrates processively in the 5' to 3' direction (away from the *trw* genes) on the displaced strand, unwinding the DNA. The accompanying host DNA polymerase III binds the uncleaved strand for complementary strand synthesis. TrwC mediated helicase activity has been shown to depend on an external ssDNA-dependent ATPase activity (Grandoso et al., 1994). There is evidence that TrwC stays bound to the T-DNA strand until it reaches the reconstituted *oriT* hairpin, upon which a second cleavage occurs, followed by a termination reaction resulting in the formation of the single stranded circular T-DNA. In this model, the termination reaction would occur in the host cell and the circular ssDNA would subsequently be transported through the T4SS. However, it was demonstrated that the relaxase protein is also transported to the recipient cell by the T4SS, where it is anticipated to track the incoming DNA until the passage of the *oriT* hairpin and then to carry out the termination reaction (Draper et al., 2005; Garcillan-Barcia et al., 2007). A model was proposed in which the initial nicking is performed by a TrwC dimer, which dissociates afterwards, as shown in Fig. 2. The

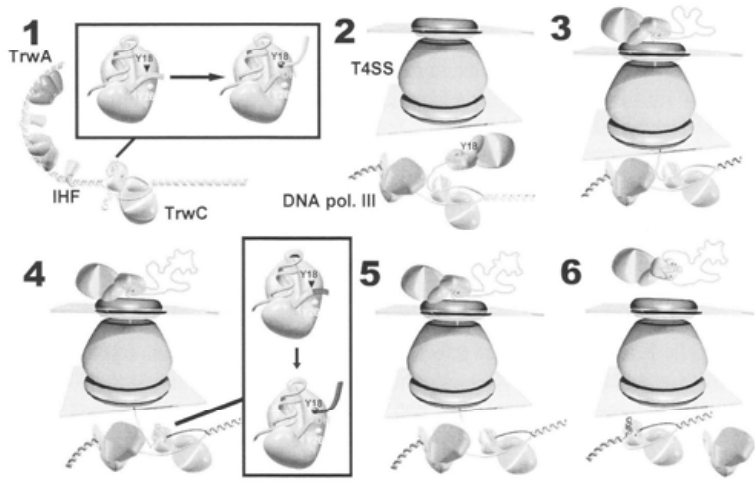

Fig. 2. A schematic representation of the proposed sequence of events during bacterial conjugation of plasmid R388, encoding the *trw* genes. A discontinuous process is assumed which allows conjugative translocation of a single copy of the parental strand. Initiation requires TrwA, IHF and dimeric TrwC to bind to the *oriT* region of the plasmid. Subsequent nicking of the parental strand releases the 3' end (1) and one of the TrwC monomers, covalently linked to the 5' end, passes through the T4SS (2). Simultaneously, the free 3' end is elongated by a host polymerase (new DNA shown in dark grey). The parental strand of the plasmid DNA passes to the recipient cell where it is scanned by the translocated TrwC monomer (3). After processing the full plasmid, the donor-cell TrwC again reaches the *oriT* hairpin and a second cut occurs (4), which frees the original parental strand for passage to the host cell (5). The host polymerase duplicates the template strand up to the *nic* site and dissociates from the complex (6). At both the donor and recipient side, the TrwC-DNA complexes recognise the reconstituted hairpins and finish their tasks by closure of the two parental DNA strands.

monomer, that is covalently bound to the T-DNA, is passed on to the recipient cell, whereas the other monomer stays in the host to separate the plasmid strands.

MECHANISM OF TrwC AND HOMOLOGS

The crystal structures of the relaxase N-terminal domain of TrwC (Guasch et al., 2003) and its homolog TraI from plasmid F (Datta et al., 2003) show a two-layer α/β plate or open sandwich core (Fig. 3a). The core structure bears resemblance to DNA polymerases and looks like a right hand, hence the nickname 'palm' (Fig. 3b). A recent structure of a 186 residue N-terminal domain of MobA of plasmid R1162 by X-ray crystallography shows a similar fold (Monzingo et al., 2007). The main secondary structure elements of the 'palm' are a central β-sheet and two parallel α-helices (α5 and α7), diagonally crossing the sheet on one of its sides. Two other

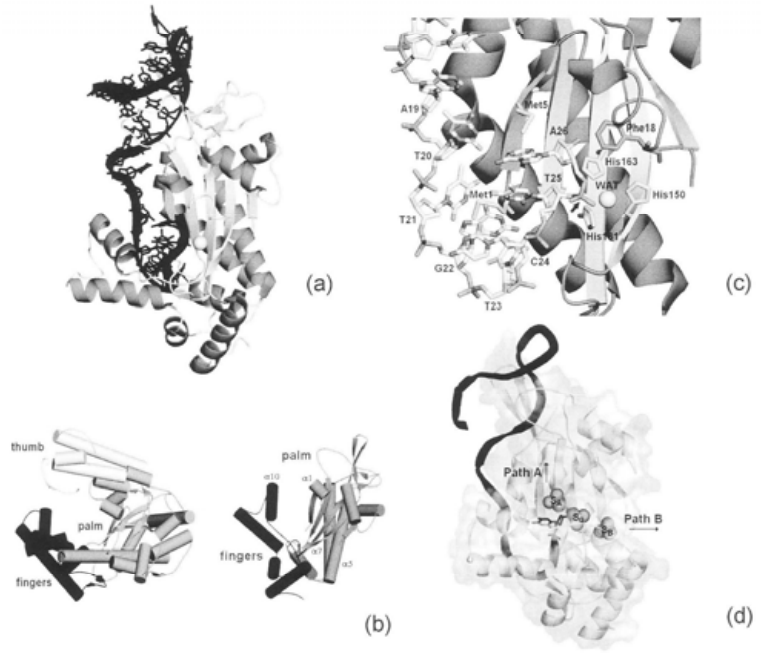

Fig. 3. a) *E. coli* R388 TrwC relaxase domain Y18F mutant in complex with a 27-base ssDNA that mimics the recognised *oriT* region (PDB code 2cdm). The hairpin is recognized by the top part of the protein. The active site is located at the lower end of the β-sheet, after the point where the DNA makes a sharp turn upwards. The direction of the DNA follows one of the possible exit paths of the DNA (see panel d and text). b) Overall topological similarity between DNA polymerase I Klenow fragment (left) and TrwC (right), showing the domain characteristics of polymerases, namely the thumb, the palm and the fingers, but only palm and fingers in the relaxase c) Ribbon representation of the TrwCY18F-DNA27 complex showing the U-turn of the DNA near the active site. The interactions of the G22:T25 mismatch pairing and bases G22 and T23 are shown. The interactions of Met1 at the hydrophobic cage formed by T21, G22 and T23 are shown, as well as the T25/A26/Met5 stack. The black arrow indicates the scissile bond. The position of the metal ion is occupied, in this structure, by a water molecule labelled WAT. d) Surface representation of the TrwCY18F-DNA27 complex. The 27-mer DNA oligonucleotide is depicted as a ribbon. Three sulphate ions, labelled SA, SO and SB, that are present in the active site crevice of the TrwC-DNA25 complex are also shown. Two putative exit paths for the DNA (A and B) are indicated with black arrows.

helices (α1 and α10) are at the opposite side of the β-sheet. They are arranged in a parallel fashion that gives rise to a narrow and long crevice that sits on the central β-sheet. The DNA runs in a direction parallel to the major longitudinal axis of the

'palm' until the 'fingers' engulf it. In TrwC, the active site is located at the crevice and is formed by the catalytic tyrosine (Tyr18) in charge of the T-strand DNA cleavage and the triad of histidines (His 150, 161 and 163) coordinating a metal ion. The interactions between the protein and the DNA are essentially electrostatic in nature, where most of the TrwC surface in contact with the DNA is electropositive. Specific interactions –including many hydrogen bonds- also occur between the DNA and the protein. A β-ribbon, which is an extension of a loop connecting two strands of the central β-sheet, enters the major groove of the dsDNA stem of the hairpin, while another loop segment, flanked by two turns, enters the opposite minor groove. After the hairpin, the DNA strand follows until it forms a peculiar U-turn just before entering the active site pocket. The U-turn serves as an anchoring point to TrwC via Met1, while the DNA continues antiparallel to an extended (G17–T21) stretch previous to the U-turn (Fig. 3c). Bases G22 and T25, conserved in different *nic* sequences of the F-family of conjugative plasmids, were found to form a mismatch pair. It is likely that this interaction serves to orient the DNA to the active site entrance and to place the scissile phosphate within proximity of the metal binding site, important for the cleavage reaction. More recent complex studies of TrwC and TraI confirmed that the scissile phosphate is directly coordinated to the metal ion (Boer et al., 2006; Larkin et al., 2005), which suggests that it may polarize the scissile phosphate facilitating the nucleophilic attack of the hydroxyl group of the catalytic tyrosine, and stabilize the subsequent penta-coordinate intermediate.

In the first structure of TrwC in complex with a 25 nucleotide *oriT* mimicking a ssDNA fragment, three sulphate ions were found near the cleavage site. Since they can mirror phosphate groups, their positions could indicate the directions of two possible functional paths, bifurcating at the metal binding site (Guasch et al., 2003). One path is nearly along the major longitudinal axis of TrwC, whereas the other runs perpendicular to it (Fig. 3d). A second structure of a T18F mutated TrwC in complex with an elongated DNA fragment is in agreement with the initial proposal for the DNA exit pathways (Boer et al., 2006), where the DNA could follow either path according to the timeline of the transfer process. For the initiation reaction, the DNA follows the longitudinal path. This configuration might be the preferred one in the constrained binding to supercoiled DNA, where only Tyr18 can cleave the DNA. After the first cleavage, Tyr18 forms a covalent bond with the DNA 5' to the cleavage site (see Fig. 2). The DNA hairpin containing the cleaved 3'-OH leaves the active site after which the RCR mechanism follows its course as described above and the TrwC monomer with the 5' end of the parental bound DNA is transported to the recipient cell. Passage of the DNA through the perpendicular pathway of TrwC as it enters the recipient cell allows the protein to scan the incoming DNA for the *oriT* sequence in order to complete the termination reaction. Upon passage of the *oriT* hairpin, the DNA is locked into place within TrwC (Garcillan-Barcia et al., 2007), forming a structure similar to that described previously with the additional hairpin

bound. The 3' end of the hairpin is subsequently covalently linked to the 5' end of the bound DNA, which reconstitutes the parental strand in the recipient cell.

The overall fold of the relaxase domain of TrwC/TraI is similar to that observed in viral Rep proteins (Campos-Olivas et al., 2002; Hickman et al., 2002), the DNA-binding domain of the replication initiation protein E1 from papillomavirus (Enemark et al., 2002) and the origin-of-replication DNA-binding domain of SV40 T-antigen (Meinke et al., 2006). All these proteins are grouped into the same topological family (Murzin et al., 1995), and exhibit the typical 'origin-of-replication-binding domain' (OBD) fold, which comprises a central antiparallel five-folded β-sheet, which acts as support to two diagonally crossing helices. TrwC also bears some resemblance to DNA polymerases. Although the folds are similar, there are considerable differences between the proteins of this family, both in the domain constitution and size, as well as in function (Fig. 4). Rep proteins and TrwC/TraI are transesterification enzymes with a catalytic tyrosine and a HXH metal-binding motif, while the DNA-binding domains of E1 and the SV40 T-antigen have only DNA-binding function and lack these catalytic residues and an active-site pocket. A comparison of relaxases and viral Rep proteins shows that a circular permutation has occurred between them, resulting in the positioning of the catalytic tyrosine close to the N-terminus in the relaxases, but close to the C-terminus in the Reps (Fig. 4). More distantly, TrwC/TraI folding resembles the DNA polymerases folding, although in the former only the 'palm' and 'finger' subdomains are recognisable (Fig. 3b). In spite of the fact that the function of these proteins is different, the catalytic residues of the relaxase at the 'palm' spatially coincide with the corresponding residues of the Klenow fragment of the DNA polymerase I (Ollis et al., 1985).

ACTIVE SITE AND REACTION MECHANISM

The nicking and transesterification reactions are catalyzed through the action of a metal ion that is coordinated by the histidine triad, composed of His150,161 and 163. It has been shown that TrwC-bound Mn^{2+}, Ni^{2+}, Zn^{2+} and Cu^{2+} are able to cleave DNA at low concentration, and that Mg^{2+} and Ca^{2+} infer cleavage activity at high concentrations. X-ray structures of complexes of TrwC with Zn^{2+}, Ni^{2+} and Cu^{2+} have been resolved and show that these bind to the active site pocket. However, the nature of the endogenous metal ligand is as of yet unknown. The interaction of the metal ion at the active site is stabilized through a network of hydrogen bonds with origins at the nitrogen atoms of the histidines distal to the metal ion. Nitrogen Nδ of His161 is coordinated to the amide carboxyl oxygen of Gln159 and Nε of His150 to the carboxyl oxygen atom of Tyr18. Nδ of His163 is hydrogen bonded to a oxygen atom of the carboxyl group of Asp85, which in turn interacts with the hydroxyl group of Tyr18. It is the oxygen atom of the Tyr18 hydroxyl group that effectuates the nucleophilic attack on the DNA scissile phosphate. The fact that mutation of Asp85 impairs cleavage and that it occupies a pivotal location in the

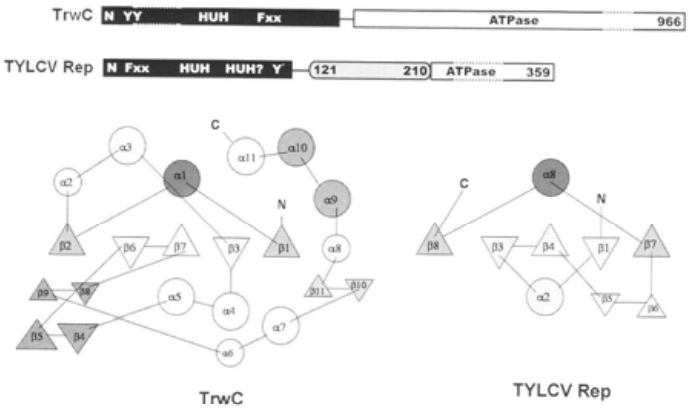

Fig. 4. Topological organization of the functional domains and secondary structural elements of two representative proteins of the origin-binding domain (OBD) family, *i.e.* relaxase/helicase TrwC of plasmid R388 and the replication initiator of tomato yellow leaf virus (TYLCV Rep). The top panel shows the domains present in the sequence of each protein and the relative location of the metal-binding (HUH) motifs and active site tyrosines (Y' or YY) in the OBD (solid black bars). The lower panels shows topology diagrams of the OBD of both proteins. The similarity in fold is evidenced by the homology between the central five β–sheets and the two contacting α-helices on both sides. The catalytic tyrosines are located in α1 and α8 in the relaxase and Rep protein, respectively. The order of the TrwC structural elements HUH and Y can be reconstituted in TYLCV Rep by applying a circular permutation, that is by breaking the connection between β6 and β7 and connecting the C to N terminal ends.

reaction site, suggests that it plays a role in the activation of Tyr18 for cleavage. A plausible function is that it acts as a general base through the abstraction of the Tyr18 hydroxyl proton in order to facilitate the nucleophilic attack of its hydroxyl oxygen on the phosphate. The phosphor atom of the scissile phosphate, in turn, is polarized by the coordination to the metal ion, which makes it more susceptible for cleavage. The presence of the DNA in the active site is stabilized by positively charged residues in the vicinity of the active site, *i.e.* Arg14, Arg154 and Arg266.

A continuing issue that has recently been resolved is the role and significance of a second tyrosine in the relaxases involved in the conjugational process. Originally, it had been proposed that two tyrosines are required for the nicking and transesterification reactions through a flip-flop mechanism as mutation of either tyrosine in TrwC led to impairment of conjugation (Grandoso et al., 2000). In this model, the DNA stays covalently attached to a first tyrosine after nicking, followed by transesterification through a second tyrosine. In the structures available, however, these tyrosines are spatially separated. Thus, the second tyrosine can reach

the active site in the monomer only by extensive movement of the loop in which it is integrated. Evidence exists that TrwC forms a dimer in solution, which would allow the Tyr26 of a second molecule to come into close proximity of the active site. This model, however, seems to contradict the recent finding that TrwC passes to the recipient cell (see above). A recent paper confirmed that Tyr26 is responsible for the transesterifcation reaction in a N-terminal truncated form of R388 TrwC (Gonzalez-Perez et al., 2007). In addition, the authors show that this reaction occurs intramolecularly and that Y26 supplied 'in trans' to a Y26F mutant is not capable of transesterification. Interestingly, the initial cleavage reaction, but not the Tyr26 transesterification activity, could be detected for the dimeric full-length protein. This is consistent with the notion that transesterification takes place in the recipient cell by a monomer of TrwC (see above).

RELAXASE-BASED ANTIBIOTIC DESIGN

An exciting development in the field of structural biology of plasmid relaxases is exemplified by a recent report that describes the identification of biphosphonate inhibitors of F plasmid TraI (Lujan et al., 2007). The authors present a structure of a phosphate-bound Mg^{2+} ion intermediate that was obtained by soaking imidobisphosphate into crystals of an Y16F mutant of TraI. The resulting structure has been refined with the phosphate group attached to the Mg^{2+} ion. The covalent bond between the Mg^{2+} ion and the phosphate group sheds new light on the role of the metal ion in the reaction, hitherto assumed to be auxiliary through the polarization of the phoshodiester of the DNA. Efficacy assays of the compounds on the inhibition of conjugation demonstrated that some biphosphonates are highly effective *in vitro*. The presented structure of the impaired complex marks the starting point of the design of a new class of antibiotic (helper) agents that attack at the very heart of the bacterium's defence system against antibiotics.

INTERACTION WITH THE T4SS

To function as a DNA translocation apparatus, the type IV secretion system requires the assistance of a component called the type IV coupling protein or T4CP (Cabezon et al., 1997; Firth et al., 1996; Gomis-Ruth et al., 2004; Moncalian et al., 1999a; Santini and Stanisich, 1998). T4S systems dedicated exclusively to protein export generally lack T4CPs, as in the *Bordetella pertussis* Ptl system (Weiss et al., 1993). However, examples exist where coupling proteins mediate protein substrate translocation, as in *Agrobacterium tumefaciens* and *Helicobacter pylori* (Covacci et al., 1999; Christie, 2001; Vergunst et al., 2000). T4CPs are integral inner membrane proteins, with an N-terminal transmembrane moiety and a C-terminal cytoplasmic domain. They possess the typical Walker A and B motifs present in NTP-binding proteins (Walker et al., 1982) and their NTPase activity probably contributes to energize the secretion machinery (Cabezon and de la Cruz, 2006). The T4CPs family comprises members like TrwB from plasmid R388, TraD and TraG from various gram-negative plasmids, VirD4 from *Agrobacterium tumefaciens*, and other

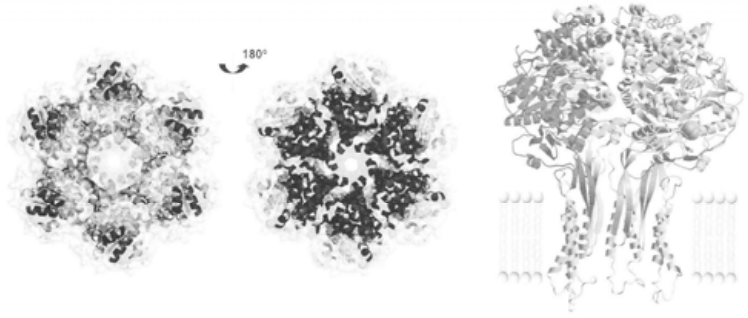

Fig. 5. a) The crystal structure of the conjugative T4CP TrwBΔN70 (PDB code 1e9r) viewed along the central channel on both sides of the hexamer. b) Ribbon representation of the full-length TrwB, including a model for the transmembrane part, based on two consecutive transmembrane helices of the photosynthetic reaction centre structure (PDB code 1mps). Only four of the six protomers are shown for clarity.

related proteins. The structure of TrwB has been determined and will be discussed in detail below (Gomis-Ruth et al., 2001).

A REPRESENTATIVE CRYSTAL STRUCTURE OF THE T4CP

TrwB is a 507-residue protein that is anchored to the inner membrane by a 70-residue N-terminal transmembrane domain. It binds ssDNA and dsDNA non-specifically and independently of NTP-binding. The crystal structure of the soluble part of the TrwB protein (TrwBΔN70) (Gomis-Ruth et al., 2002; Gomis-Ruth et al., 2001) consists of six equivalent monomers associated in an almost spherical structure, flattened at both poles, with a central channel (Fig. 5). Each monomer has two domains: a cytosol-oriented all-α helical C-terminal domain, AAD, comprised of seven α-helices; and an N-terminal membrane-proximal nucleotide-binding domain, NBD, a central twisted β-sheet flanked by helices on both sides. The central channel that connects the cytoplasm with the periplasm for the full-length protein, is 20Å wide along its entire length, except at the cytosol pole. A ring of asparagine residues plug the cytoplasmic channel entrance and restrict its diameter to ~8Å (Fig. 5a). The nucleotide-binding site is located at the C-terminal edge of the β-sheet. The architecture of this site, where the Walker A and B motif are located, is determined by two segments comprising the β2-αC and β6-αN loops. The six nucleotide-binding sites are 32Å apart from each other, located in superficial cavities at the interfaces among vicinal protomers.

TrwB has structural homology with the dsDNA translocase FtsK protein, a hexameric helicase involved in the resolution of chromosomal dimers during bacterial cell division, which supposedly functions as a molecular motor in the transport of DNA across the membrane. The X-ray structure of the soluble FtsK (Massey et al., 2006) shows a hexameric arrangement, similar to TrwB, as shown in Fig. 6b. Each protein monomer has three domains (α, β and γ). Due to structural disorder around the γ domain region, this domain could not be observed in the structure. There is no counterpart for the FtsK α-domain in the sequence of TrwB, but the β domain, which

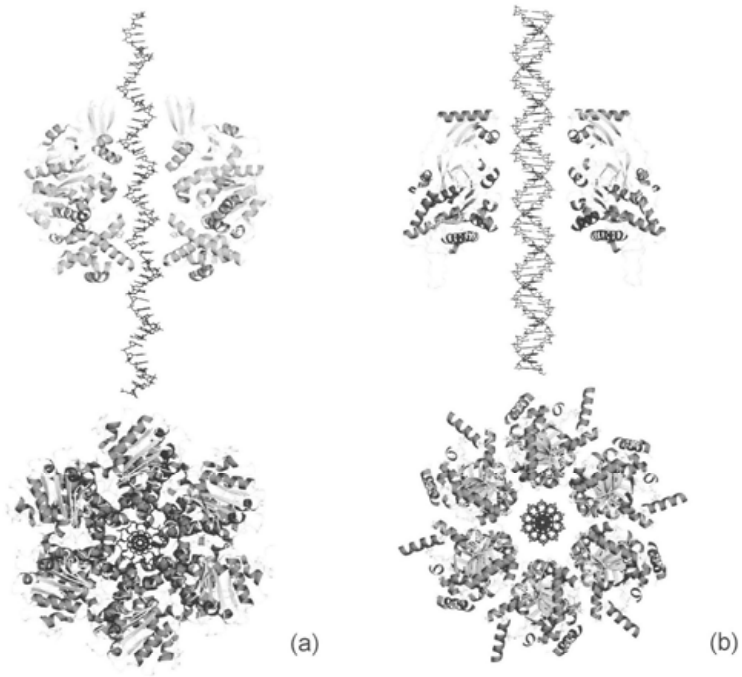

Fig. 6. a) The crystal structure of the conjugative T4CP TrwBΔN70 (PDB code 1e9r) with a ssDNA modelled along the central channel based on the 13-base ssDNA described in the papillomavirus E1 hexameric helicase ssDNA complex (PDB code 2gxa). b) FtsK DNA translocase (PDB code 2iuu) with a standard dsDNA modelled along the central channel. Only two opposite monomers flanking the DNA are shown in the upper panels. ATPase domains are depicted in equivalent orientations after their superimposition. Both DNA translocator proteins share a common central ATPase domain (DALI Z-score: TrwB/FtsK=14.2, DALI URL: http://www.ebi.ac.uk/dali), but the adjacent domains are different. The corresponding hexamers, viewed along the channel, are depicted in the lower panels.

is the NTP-binding domain, is structurally similar to the TrwB NBD. The central channel of the FtsK hexamer is ~30Å wide and, as shown by electron micrographs, is large enough to accommodate a dsDNA helix (Massey et al., 2006). In addition, there is a strong structural similarity between TrwB and the α- and β-subunits of the heterohexameric F_1-ATPase protein, a RecA-like ATPase protein related to the mitochondrial proton-driven molecular machine for ATP synthesis (Abrahams et al., 1994). The structural analogy is much stronger with the β-subunits in the occupied conformation (β_{DP} and β_{TP}) than with the empty one (β_E). The superposition of the NTP-binding domains shows the same relative orientation with respect to the membrane, although TrwB has six equally potential catalytic sites (like hexameric helicases), instead of the three catalytic subunits of F_1-ATPase.

DNA TRANSLOCATION BY T4CP

Structural and biochemical data suggest that a single-strand DNA might pass through the TrwB central channel (Fig. 6a), as in the distantly related E1 ATPase from bovine papillomavirus I (Enemark and Joshua-Tor, 2006). However, the asparagine plug at the cytosol pole narrows the entrance of the channel and could impede the passage of DNA. This apparent contradiction can be explained by assuming that the TrwBDN70 structure represents a closed conformation of the protein and the cytoplasmic gate opens on activation when the protein couples with the relaxosome. The DNA-dependent ATPase activity of TrwB (Tato et al., 2005) and the structural changes of the central channel observed in TrwB structures due to the binding of phosphate or nucleotides at the active site (Gomis-Ruth et al., 2002), reinforce the idea that the coupling protein is a molecular motor. Based on the structure, a model was proposed in which active DNA transport is achieved by cyclic conformational changes of the ring subunits that would push the ssDNA across the inner membrane during conjugal transfer. The conformational changes associated with pumping occur at the nucleotide binding site and are conferred to the inner lining of the central core of the hexamer. The region involved in the transfer has homology to a conserved sequence of AAA-ATPases and is suggested to aid in DNA displacement.

The ATPase driven pumping model of ssDNA through the TrwB hexamer can only occur after the relaxase, covalently bound to the T-strand, has entered the pilus. Therefore, the TrwB hexamer somehow needs to be assembled on the T-DNA, downstream of the TrwC-DNA complex. In a process proposed by Cabezon and de la Cruz (2006), the nucleoprotein complex is guided towards the inner membrane by a monomeric form of TrwB. After establishing the connection between TrwC and the secretion apparatus, the relaxase is transported to the recipient cell, piloting the trailing DNA. The energy needed for this first step could be provided by VirB11 and/or VirB4 ATPases from the T4SS. Once the proteic part of the relaxase-DNA complex has crossed the inner membrane, TrwB wraps around the DNA and oligomerizes as a hexamer. The hexameric assembly of TrwB in the

inner membrane allows the pumping of DNA across the inner membrane through its ATPase activity and cyclic conformational changes. TrwB can be considered an adapter to the T4S system that changes the substrate specificity from proteins to DNA, as it establishes the connection between the *oriT*-specific relaxase TrwC and the T4S system, responsible for TrwC translocation.

TraM, AN ANCILLARY PROTEIN

TraM is a small 14 kDa cytoplasmic DNA-binding protein coded in the *tra* region of a number of plasmids, like F and R1, but not R388 (see Fig. 1a). It has been associated with various ancillary functions required for T-strand processing and complementing relaxase/helicase, IHF, and the transcriptional regulator. TraM binds to three sites within *oriT*, repressing its own gene promoter. It could be a constituent of the relaxosome and/or could signal the cell that mating pair formation is completed; however, it is not essential for relaxosome-dependent nicking. Other investigations showed that TraM enhances relaxase activity and that it specifically interacts with the coupling protein of plasmid R1, TraD (Beranek et al., 2004). Thus, TraM may be a helper protein in the coupling of the relaxosome with the DNA-transfer apparatus (Disque-Kochem and Dreiseikelmann, 1997; Kupelwieser et al., 1998; Stockner et al., 2001). Consistent with this idea, the C-terminal domain of coupling protein TraD that interacts with TraM, is missing in the TrwB, although the cytoplasmic domains of TraD and TrwB bear strong sequence similarity (Beranek et al., 2004; Llosa et al., 1994).

A N-terminal stretch of 54 residues of the 127-residue long TraM from R1 was studied by NMR- spectroscopy (Stockner et al., 2001). This portion of TraM fully retains the specific recognizing and binding capacity to DNA. The NMR study revealed that the domain is all α-helical and consists of three helical regions. The first two helical regions of TraM are connected by a loop while the last two are linked through a glycine residue. Overall, the N-terminal domain of TraM displays topological similarity with a region within the C-terminal all-α domain of TrwB (Gomis-Ruth et al., 2002). It can be speculated that, similar to the N-terminal domain of TraM, the C-terminal DNA-binding domain TrwB comes into contact with the relaxosome.

A structural analysis by X-ray crystallography of the C-terminal fragment of TraM from plasmid F containing the TraD-binding segment revealed that four protomers interact in a compact eight-helical bundle (Lu et al., 2006). Here, the N-terminal helices from each protomer interact to form a central, parallel four-stranded coiled-coil, whereas each C-terminal helix packs in an antiparallel arrangement around the periphery of the structure. This oligomerization creates a central tunnel lined by the inner N-terminal helices, where four protonated glutamate residues provided by each of the protomers create a central solvent-rich ring. Deprotonation of this acidic residue relaxes the TraM structure and this affects interactions with TraD.

T4SS COMPONENTS: STRUCTURE AND ASSEMBLY

As mentioned before, the *Agrobacterium tumefaciens* VirB/VirD4 secretion system has been extensively characterized, both biochemically and structurally and it is considered the archetype of T4SS (Fig. 7). Its components can be classified into three groups, according to function (mostly putative) and cellular location: i) the cytoplasm- or inner membrane associated ATPases: VirB4, VirB11 and VirD4; ii) the core complex: VirB6, VirB7, VirB8, VirB9 and VirB10; and iii) the pilus components: VirB2 and VirB5 (Baron et al., 2002; Christie, 2001). The VirB2 channel has its origin in the periplasmic space, then spans the outer membrane, ending in the extracellular milieu. The translocation pathway for DNA through this secretory apparatus (Cascales and Christie, 2004b) has been shown to roughly

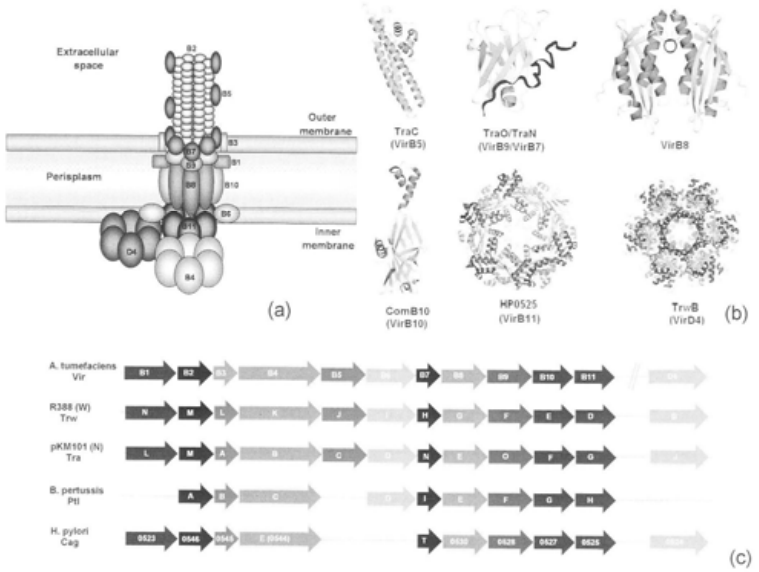

Fig. 7. a) Tentative assembly model of the type IV secretion system of *A. tumefaciens* showing the cytoplasm- or inner membrane associated ATPases: VirB4, VirB11 and VirD4; the core complex: VirB6, VirB7, VirB8, VirB9 and VirB10; and the pilus components: VirB2 and VirB5. b) Resolved structures of components of the T4SS: X-ray structures of the VirB5 homologue, TraC (pdb code 1r8i), VirB8 (PDB code 2cc3), the VirB10 homologue, ComB10 (PDB code 2bhv), the VirB11 homologue, HP0525 (PDB code 1nlz), and the VirD4 homologue, TrwB (PDB code 1e9r) and the NMR-spectroscopy structure of the interacting domains of the complex of VirB7/VirB9 homologous TraN/TraO (PDB code 2ofq). c) Gene organization and nomenclature of *A. tumefaciens* VirB/VirD4, *E. coli* R388 Trw, *E. coli* pKM101 Tra, *B. pertussis* Ptl and *H. pylori* Cag, secretion systems.

follow the order of the subcomponents mentioned here. All the components of the VirB operon and the VirD2 and VirD4 gene products are essential for the secretion of DNA and/or protein substrates, with the exception of VirB1, which is involved in the localized lysis of the peptidoglycan layer (Berger and Christie, 1994).

The structures of several of these proteins, or their homologous, are nowadays known (Fig. 7b). These include X-ray structures of a VirB5 homologue, TraC (Yeo et al., 2003), VirB8 from *Brucella suis* (Terradot et al., 2005) and *Agrobacterium tumefaciens* (Bailey et al., 2006), a VirB10 homologue, ComB10 (Terradot et al., 2005), a VirB11 homologue, HP0525 (Yeo et al., 2000), and a VirD4 homologue, TrwB (Gomis-Ruth et al., 2001). In addition, NMR-spectroscopy structures of the interacting domains of the complex of VirB7/VirB9 homologous TraN/TraO are available (Bayliss et al., 2007).

INNER MEMBRANE ATPases

VirB4, VirB11 and the coupling protein VirD4 are all associated inner membrane ATPases that have been shown to mutually interact (Atmakuri et al., 2004). They are believed to provide the energy to the T4SS, based on the observation that they all contain the conserved Walker A and B sequence motifs. The structure of a VirD4 analogue, TrwB from the R388 plasmid, was discussed in the section on relaxosomes. Crystal structures of the VirB11 homologue protein HP0525 from *Helicobacter pylori* have been determined in an ADP-bound form (Yeo et al., 2000) as well as in an ATPγS-bound and a free form (Savvides et al., 2003). In addition, a structure of a homologue from *Brucella suis* has been determined by X-ray crystallography, showing an interesting topological domain swap (Hare et al., 2006). HP0525 consists of a hexameric ring formed by monomers consisting of a C-terminal RecA-like α/β domain and a grapple-shaped N-terminal domain. Conformational differences between the ADP-bound form and the free form occur at the N-terminal domain, which rotates away from the central edge. These motions are different for each of the monomers and result in an asymmetric ring. Based on these nucleotide-induced conformational changes a mechanism of sequential binding and hydrolysis of ATP by HP0525 has been proposed (Savvides et al., 2003), in which the protein alternates between symmetrical and asymmetrical hexamer forms upon substrate binding, hydrolysis and product release.

CORE COMPLEX

The components that span the periplasmic space consist of components VirB7 to VirB10, of which VirB8 and VirB10 appear to be the major constituents of the periplasmic channel. VirB7 and VirB9 are generally considered to be outer-membrane associated proteins. Atomic structures are available for (analogues of) both VirB8 and VirB10.

PERIPLASMIC SPACE-SPANNING COMPONENTS

VirB8 is an inner membrane protein with a short N-terminal domain, exposed to the cytoplasm, and a large C-terminal domain, exposed to the periplasm, linked by a transmembrane helix (Buhrdorf et al., 2003; Das and Xie, 1998; Thorstenson and Zambryski, 1994). It has been proposed to function as a periplasmic nucleation centre recruiting VirB9 and VirB10 to the sites of the complex assembly (Kumar et al., 2000) In addition to VirB9 and VirB10, VirB8 is able to interact with itself (Das and Xie, 2000), which was confirmed by structural studies using X-ray crystallography of the *Agrobacterium tumefaciens* VirB8 (Bailey et al., 2006) and its *Brucella suis* homologue (Terradot et al., 2005). The structures of the periplasmic domain of VirB8 represent dimers in which each monomer has five α-helices interacting with one side of a four-strand antiparallel β-sheet which extends over the lengths of the molecule and curves around the first helix (Fig. 7b). A highly conserved surface region extends over the surface of the central region of the β-sheet at a groove opposite to the dimer interface. It was suggested that this region corresponds to a protein-protein interaction site for other T4SS proteins.

VirB10 seems to alternate between two states depending on the cellular ATP level of the system. These states are controlled through interactions with VirB11 and VirD4, but not VirB4 (Cascales and Christie, 2004a). Only the 'energised' conformation of VirB10 is able to form complexes with the outer membrane component VirB9 (Terradot et al., 2005). In the reported crystal structure of the C-terminal, periplasmic domain of the VirB10-like protein ComB of *Helicobacter pylori*, the protein crystallizes as a dimer in which each monomer folds as a C-terminal distorted β-barrel with a large connected N-terminal α-helical antenna (Fig. 7b). The dimer forms through mutual interactions between the C-terminal and the N-terminal domains of the crystallized fragment. The authors propose that the dimer in the crystal structure represents the unenergised conformation of the protein, which dissociates in presence of ATP, releasing the protruding helices for interaction with VirB9 (Cascales and Christie, 2004a). The N-terminal domains of the VirB10 dimer found in the crystal structure are positioned at opposite ends, which is consistent with the finding that VirB10 localizes to both membranes (Fernandez et al., 1996a; Kumar et al., 2000). The thickness of the periplasmic space (20-30 nm) of *H. pylori* (Wen et al., 1997) and the distance between the two N-terminal membrane spanning regions (~11 nm) of only the C-terminal domain of VirB10 that was crystallized, do not contradict this possibility.

THE OUTER MEMBRANE COMPONENTS

The outer membrane components consist of VirB7 and VirB9. VirB7 is a small lipoprotein localized in the periplasmic side of the outer membrane and at the extracellular milieu, which is involved in stabilising the integrity of the pilus (Beaupre et al., 1997; Fernandez et al., 1996a; Sagulenko et al., 2001). VirB7 is able to associate with itself and with VirB9 through the formation of intra- and

intermolecular disulfide bridges (Anderson et al., 1996; Baron et al., 1997; Das et al., 1997; Farizo et al., 1996; Fernandez et al., 1996b; Harris et al., 2001; Spudich et al., 1996). The NMR-spectroscopy structure of the C-terminal domain of the plasmid pKM101 encoded TraO homologue of VirB9 in complex with the corresponding TraN homologue of VirB7 has recently been reported (Bayliss et al., 2007). TraO adopts an immunoglobulin-like β-sandwich folding, comprising nine β-strands and a short stretch of 3_{10}-helix located between the two first strands (Fig. 7b). TraN wraps around TraO perpendicularly to the direction of the β-sandwich and introduces one of its residues between two β-sheets. The part of the protein, that has been demonstrated to span the outer membrane, protrudes from the structure supporting the general model that VirB9-like proteins are a component of an outer membrane channel complex.

PILUS COMPONENTS

Pili are formed by the association of small subunits called pilins (VirB2-like proteins), which assemble into a filamentous pilus structure at the cell surface (Eisenbrandt et al., 1999). Two additional proteins are also localized to the pilus, VirB5 and VirB7 (Sagulenko et al., 2001). The minor component VirB5-like proteins (Schmidt-Eisenlohr et al., 1999a; Schmidt-Eisenlohr et al., 1999b), are ~220 residue proteins containing an N-terminal signal sequence that targets the protein to the periplasmic space. The crystal structure of the VirB5-like protein TraC of IncN plasmid pKM101 revealed an α-helical elongated structure, as shown in Fig. 7b (Yeo et al., 2003). Although the protein dimerizes in the crystal, structured-based mutagenesis studies of activity and characterization of the oligomeric state of TraC by sedimentation equilibrium methods, demonstrated that this dimeric state was an artefact. The extended protein surface region involved in these crystalline contacts is probably the region of protein-protein interactions with other component of the T4SS system. Since the connection between the pilus of the host cell and the outer membrane of the recipient cell is achieved through interactions between VirB5 of T4SS and adhesins on the recipient membrane (Soto and Hultgren, 1999), it has been proposed that VirB5 proteins have an adhesin-like function (Yeo et al., 2003).

T4SS ASSEMBLY AND MECHANISM

A four-stage model of the assembly pathway for the VirB/D4 secretion system has been proposed by Christie and collaborators based on the available literature on the subject (Christie et al., 2005). It comprises (i) the formation of the core complex, (ii) the recruitment of pilus associated proteins, (iii) the recruitment of the VirB11 ATPase and (iv) the formation of either the T pilus or a secretion channel. In the first stage VirB4, VirB7, VirB8, VirB9 and VirB10 assemble across the cell envelope forming the core complex. Once the core complex is assembled, the recruitment of the pilus proteins VirB2, VirB3 and VirB5 starts through VirB4, which functions as a scaffold. In the next stage, VirB11 ATPase is recruited to assist in the completion of the inner-membrane basal structure, the formation of a transenvelope structure

comprising a VirB2 polymer and the outer-membrane pore structure. Finally, the last stage of the assembly pathway results in the formation of either a secretion channel or a pilus.

In order to describe the DNA pathway across the inner membrane, different models have been proposed, which differ in the role attributed to the ATPases VirB4, VirB11 and the T4CP. In theory, both the T DNA-relaxase and the T-DNA substrates could be translocated by any of the ATPases. However, VirB4 is generally not considered important as supplier of translocation energy, a role which is reserved for VirB11. In the models that have been proposed, VirB11 unfolds the relaxase and then it either 1) passes the unfolded peptide chain and DNA substrate through the inner membrane (channel model), or 2) it passes the unfolded peptide chain to the T4CP for passage through the inner membrane (ping-pong model, Atmakuri et al., 2004), or 3) it passes the peptide chain through the inner membrane and the DNA is then pumped through by T4CP ('shoot-and-pump', Llosa et al., 2002). The shoot-and-pump model is consistent with the origin of T4S systems as protein translocation machines and the function of the coupling protein in DNA substrate recognition and delivery. In this model, the coupling protein acts as an DNA adaptor to the protein secretion machine. The assembly of VirD4 monomers around the DNA-relaxase has been proposed to account for the structural reorganization required for this mode of action (Cabezon and de la Cruz, 2006). The ping-pong model also agrees well with the available biochemical and functional data and relies on recent findings that, in some systems, both protein and DNA translocation requires the T4CP. Future experiments will determine how the T4SS functions on a molecular level.

CONCLUSION

A large body of experimental data and literature has accumulated concerning DNA translocation in bacterial conjugation. As a result, the location of the genes important for these processes and the assembly topology of the functional machinery have been accurately mapped. Furthermore, a discrete number of specialised sub-systems, such as the relaxosome and the type IV secretion system (T4SS), have been identified. The recent increase in the number of sequenced genomes has helped identify these sub-systems in different bacteria and has allowed the determination of similarities and differences between them. In addition, the particular function of many of the protein components has been ascertained, although the details of many remain unknown.

Several protein components have been analyzed at the structural level, which has laid the foundation for the analysis of the function and molecular mechanisms of the molecular machines of the bacterial conjugation mechanism at a (near-)atomic level. *E.g.*, many relaxosome components have been structurally characterised. These include ternary complexes of the relaxase domains of TrwC from the R388

plasmid and TraI from F plasmid with DNA and the catalytic metal ions (Datta et al., 2003; Guasch et al., 2003). This revealed important details of the initial steps of bacterial conjugation, especially about the recognition of the *nic* site and the possible routes by which DNA is processed by the relaxase at different stages of conjugation. Other components of the relaxosome that have been studied at the structural level are the CopG homologue of the *trw* gene repressor TrwA, the integration host factor (IHF) of *Escherichia coli* (Rice et al., 1996) and DNA binding domain of the ancillary transfer protein (TraM) of plasmid R1 (Stockner et al., 2001). Furthermore, the atomic structure of a coupling protein, *i.e.* R388 plasmid TrwB, has been determined, along with several related helicases that are somehow involved in DNA translocation processes. These structures allowed the postulation of a step-wise mechanism of DNA translocation mediated through this family of ATPases.

Structures of many representative components for the type IV secretion system (T4SS) have also been determined. Full-length inner membrane associated ATPase VirB11 has been crystallized both in the apo- and ATP-bound forms (Savvides et al., 2003; Yeo et al., 2000). These crystal structures provided further insight into the mode of action of the unit, which, via nucleotide hydrolysis, powers the transport of substrates from the cytoplasmic/inner-membrane end of the secretion channel. Of the proteins that form the core, the NMR spectroscopy solution structure of full-length TraN (a homologue of VirB7 from plasmid pKM101) in complex with the C-terminal domain of TraO (a VirB9 homologue from that plasmid) shed some light on how two pieces of the core form a stable association (Bayliss et al., 2007). The elucidation of this structure, together with the structures of VirB8, the periplasmic nucleation centre for complex assembly, and VirB10, the sensor of the energetic state of VirB11, allowed the spatial mapping of several protein-protein interactions on the protein surface. The structure of a homologue of VirB5, TraC from plasmid pKM101, served to speculate about its role in attaching the pilus to the membrane (Yeo et al., 2003).

FUTURE TRENDS

Overall, many of the full-length or partial individual components of the relaxosome and T4SS have now been structurally characterized. From these puzzle pieces, a crude picture of the fully assembled pilus apparatus can be put together. However, the full understanding of the T4SS-mediated bacterial conjugation process is still far away. First of all, several individual components have not, or only partially, been structurally characterized to date. For instance, the structure of VirB4, one of the engines of the power plant of the complex, remains unresolved. The same is true for VirB3, the outer-membrane component of the complex, and for VirB1, the disruptor of the peptidoglycan layer for facilitating complex assembly.

Although the structural analysis of the proteins involved in the conjugal transfer of DNA does provide valuable information, they nonetheless represent individual components of a integrated multicomponent assembly. With the exception of the complex of the VirB7 and VirB9 analogues determined by NMR spectroscopy, little is known about the way the individual components interact with one another from a structural point of view. A structural analysis of multiprotein macromolecular assemblies needs to be performed. As some of the protein-protein interactions can be expected to be transient and often weak, difficulties arising from technical problems, such as the stabilization of interactions among a variety of protein components sufficient for structural characterization and the crystallization of components with trans-membrane segments, will have to be overcome. Electron microscopy and SAXS are likely to contribute importantly to the characterization of these large assemblies. If electron microscopy 3D reconstructions of the complexes were available, the individual proteins or domains determined by X-ray or NMR at high resolution could be fitted in the electron microscopy envelopes yielding quasi-atomic models of the whole assemblies, as has been demonstrated in other systems.

Another important aspect of the study of large and dynamic machinery of proteins conglomerates lies in the inherent flexibility and rearrangements of components that are associated with function. The individual proteins and multi-component complexes can be expected to change conformation, oligomerisation state, interaction partners and reaction mechanisms depending on the processing stage on the conjugational timeline. As a consequence, interpretation of the available structural information is complicated by the fact that they represent still images of very dynamic and constantly varying structures. For a full understanding of the conjugational process, structural biologists will need to focus on the resolution of multicomponent complexes with well-defined functions at various points of the conjugational process.

Some of the fundamental questions appear elusive to be answered. What is the *in vivo* factor that triggers *oriT* cleavage and DNA transfer? How does the pilus sense and attach to the recipient cell? How is the recipient cell membrane barrier penetrated? What is the structure of the helicase domain of the relaxase like and how does it move processively along the DNA? What are the precise mutual interactions of the components of the relaxosome with those of the coupling protein? What is the mechanism by which the T4SS translocates the substrate? In the case of the VirD/VirB system of *Agrobacterium tumefaciens*, what are the specific functions of each of the three ATPases involved in the process (VirD4, VirB4 and VirB11)?

The answer to these questions will require a time-resolved structural analysis of complicated arrangements of proteins and protein-DNA complexes. The

achievement of this formidable task will likely require the integrated use of all techniques available to the structural biologist, including SAXS, NMR, EM and X-ray crystallography. Elucidation of the interactions that the components form could yield a satisfactory explanation of the observed biochemical data, while at the same time may trigger further questions at the academic as well as the applied scientific areas.

ACKNOWLEDGEMENTS

Funding provided by research grants BFU2005-06758/BMC and GEN2003-20642 from the Ministerio de Educación y Ciencia of Spain, as well as 2005SGR-00280 from the Generalitat de Catalunya are acknowledged. We are most grateful to F. de la Cruz in whose laboratory the R388 Trw proteins were prepared and for many appealing ideas on the functioning of the conjugation machinery described in this chapter. We thank A.G. Blanco for making part of Fig. 1 and 2.

REFERENCES

Abrahams, J.P., Leslie, A.G., Lutter, R., and Walker, J.E. (1994). Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria. Nature *370*, 621-628.

Alekshun, M.N., and Levy, S.B. (2007). Molecular mechanisms of antibacterial multidrug resistance. Cell *128*, 1037-1050.

Anderson, L.B., Hertzel, A.V., and Das, A. (1996). Agrobacterium tumefaciens VirB7 and VirB9 form a disulfide-linked protein complex. Proc. Natl. Acad. Sci. U. S. A. *93*, 8889-8894.

Atmakuri, K., Cascales, E., and Christie, P.J. (2004). Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion. Mol. Microbiol. *54*, 1199-1211.

Backert, S., and Meyer, T.F. (2006). Type IV secretion systems and their effectors in bacterial pathogenesis. Curr. Opin. Microbiol. *9*, 207-217.

Bailey, S., Ward, D., Middleton, R., Grossmann, J.G., and Zambryski, P.C. (2006). Agrobacterium tumefaciens VirB8 structure reveals potential protein-protein interaction sites. Proc. Natl. Acad. Sci. U. S. A. *103*, 2582-2587.

Baron, C., D, O.C., and Lanka, E. (2002). Bacterial secrets of secretion: EuroConference on the biology of type IV secretion processes. Mol. Microbiol. *43*, 1359-1365.

Baron, C., Thorstenson, Y.R., and Zambryski, P.C. (1997). The lipoprotein VirB7 interacts with VirB9 in the membranes of Agrobacterium tumefaciens. J. Bacteriol. *179*, 1211-1218.

Bayliss, R., Harris, R., Coutte, L., Monier, A., Fronzes, R., Christie, P.J., Driscoll, P.C., and Waksman, G. (2007). NMR structure of a complex between the VirB9/VirB7 interaction domains of the pKM101 type IV secretion system. Proc. Natl. Acad. Sci. U. S. A. *104*, 1673-1678.

Beaupre, C.E., Bohne, J., Dale, E.M., and Binns, A.N. (1997). Interactions between VirB9 and VirB10 membrane proteins involved in movement of DNA from Agrobacterium tumefaciens into plant cells. J. Bacteriol. *179*, 78-89.

Beranek, A., Zettl, M., Lorenzoni, K., Schauer, A., Manhart, M., and Koraimann, G. (2004). Thirty-eight C-terminal amino acids of the coupling protein TraD of the F-like conjugative resistance plasmid R1 are required and sufficient to confer binding to the substrate selector protein TraM. J. Bacteriol. *186*, 6999-7006.

Berger, B.R., and Christie, P.J. (1994). Genetic complementation analysis of the Agrobacterium tumefaciens virB operon: virB2 through virB11 are essential virulence genes. J. Bacteriol. *176*, 3646-3660.

Boer, R., Russi, S., Guasch, A., Lucas, M., Blanco, A.G., Perez-Luque, R., Coll, M., and de la Cruz, F. (2006). Unveiling the molecular mechanism of a conjugative relaxase: The structure of TrwC complexed with a 27-mer DNA comprising the recognition hairpin and the cleavage site. J Mol Biol *358*, 857-869.

Boucher, Y., Douady, C.J., Papke, R.T., Walsh, D.A., Boudreau, M.E., Nesbo, C.L., Case, R.J., and Doolittle, W.F. (2003). Lateral gene transfer and the origins of prokaryotic groups. Annu. Rev. Genet. *37*, 283-328.

Buhrdorf, R., Forster, C., Haas, R., and Fischer, W. (2003). Topological analysis of a putative virB8 homologue essential for the cag type IV secretion system in Helicobacter pylori. Int. J. Med. Microbiol. *293*, 213-217.

Burns, D.L. (2003). Type IV transporters of pathogenic bacteria. Curr. Opin. Microbiol. *6*, 29-34.

Cabezon, E., and de la Cruz, F. (2006). TrwB: an F(1)-ATPase-like molecular motor involved in DNA transport during bacterial conjugation. Res. Microbiol. *157*, 299-305.

Cabezon, E., Sastre, J.I., and de la Cruz, F. (1997). Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation. Mol. Gen. Genet. *254*, 400-406.

Campos-Olivas, R., Louis, J.M., Clerot, D., Gronenborn, B., and Gronenborn, A.M. (2002). The structure of a replication initiator unites diverse aspects of nucleic acid metabolism. Proc. Natl. Acad. Sci. U. S. A. *99*, 10310-10315.

Cascales, E., and Christie, P.J. (2003). The versatile bacterial type IV secretion systems. Nat. Rev. Microbiol. *1*, 137-149.

Cascales, E., and Christie, P.J. (2004a). Agrobacterium VirB10, an ATP energy sensor required for type IV secretion. Proc. Natl. Acad. Sci. U. S. A. *101*, 17228-17233.

Cascales, E., and Christie, P.J. (2004b). Definition of a bacterial type IV secretion pathway for a DNA substrate. Science *304*, 1170-1173.

Covacci, A., Telford, J.L., Del Giudice, G., Parsonnet, J., and Rappuoli, R. (1999). Helicobacter pylori virulence and genetic geography. Science *284*, 1328-1333.

Champoux, J.J. (2001). DNA topoisomerases: structure, function, and mechanism. Annu. Rev. Biochem. *70*, 369-413.

Chen, I., Christie, P.J., and Dubnau, D. (2005). The ins and outs of DNA transfer in bacteria. Science *310*, 1456-1460.

Chen, I., and Dubnau, D. (2004). DNA uptake during bacterial transformation. Nat. Rev. Microbiol. *2*, 241-249.

Christie, P.J. (1997). The Agrobacterium T-complex transport apparatus: a paradigm for a new family of multifunctional transporters in eubacteria. J. Bacteriol. *179*, 3085-3304.

Christie, P.J. (2001). Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. Mol. Microbiol. *40*, 294-305.

Christie, P.J. (2004). Type IV secretion: the Agrobacterium VirB/D4 and related conjugation systems. Biochim. Biophys. Acta *1694*, 219-234.

Christie, P.J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S., and Cascales, E. (2005). Biogenesis, architecture, and function of bacterial type IV secretion systems. Annu. Rev. Microbiol. *59*, 451-485.

Christie, P.J., and Cascales, E. (2005). Structural and dynamic properties of bacterial type IV secretion systems (review). Mol. Membr. Biol. *22*, 51-61.

Christie, P.J., and Vogel, J.P. (2000). Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. Trends Microbiol. *8*, 354-360.

Das, A., Anderson, L.B., and Xie, Y.H. (1997). Delineation of the interaction domains of Agrobacterium tumefaciens VirB7 and VirB9 by use of the yeast two-hybrid assay. J. Bacteriol. *179*, 3404-3409.

Das, A., and Xie, Y.H. (1998). Construction of transposon Tn3phoA: its application in defining the membrane topology of the Agrobacterium tumefaciens DNA transfer proteins. Mol. Microbiol. *27*, 405-414.

Das, A., and Xie, Y.H. (2000). The Agrobacterium T-DNA transport pore proteins VirB8, VirB9, and VirB10 interact with one another. J. Bacteriol. *182*, 758-763.

Datta, S., Larkin, C., and Schildbach, J.F. (2003). Structural insights into single-stranded DNA binding and cleavage by F factor TraI. Structure *11*, 1369-1379.

del Solar, G., Hernandez-Arriaga, A.M., Gomis-Ruth, F.X., Coll, M., and Espinosa, M. (2002). A genetically economical family of plasmid-encoded transcriptional repressors involved in control of plasmid copy number. J. Bacteriol. *184*, 4943-4951.

Dillard, J.P., and Seifert, H.S. (2001). A variable genetic island specific for Neisseria gonorrhoeae is involved in providing DNA for natural transforation and is found more often in disseminated infection isolates. Mol. Microbiol. *41*, 263-277.

Ding, Z., Atmakuri, K., and Christie, P.J. (2003). The outs and ins of bacterial type IV secretion substrates. Trends Microbiol. *11*, 527-535.

Disque-Kochem, C., and Dreiseikelmann, B. (1997). The cytoplasmic DNA-binding protein TraM binds to the inner membrane protein TraD *in vitro*. J. Bacteriol. *179*, 6133-6137.

Draper, O., Cesar, C.E., Machon, C., de la Cruz, F., and Llosa, M. (2005). Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. Proc. Natl. Acad. Sci. U. S. A. *102*, 16385-16390.

Dreiseikelmann, B. (1994). Translocation of DNA across bacterial membranes. Microbiol. Rev. *58*, 293-316.

Durrenberger, M.B., Villiger, W., and Bachi, T. (1991). Conjugational junctions: morphology of specific contacts in conjugating Escherichia coli bacteria. J. Struct. Biol. *107*, 146-156.

Eisenbrandt, R., Kalkum, M., Lai, E.M., Lurz, R., Kado, C.I., and Lanka, E. (1999). Conjugative pili of IncP plasmids, and the Ti plasmid T pilus are composed of cyclic subunits. J. Biol. Chem. *274*, 22548-22555.

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

Enemark, E.J., Stenlund, A., and Joshua-Tor, L. (2002). Crystal structures of two intermediates in the assembly of the papillomavirus replication initiation complex. EMBO J. *21*, 1487-1496.

Farizo, K.M., Cafarella, T.G., and Burns, D.L. (1996). Evidence for a ninth gene, ptlI, in the locus encoding the pertussis toxin secretion system of Bordetella pertussis and formation of a PtlI-PtlF complex. J. Biol. Chem. *271*, 31643-31649.

Fekete, R.A., and Frost, L.S. (2000). Mobilization of chimeric oriT plasmids by F and R100-1: role of relaxosome formation in defining plasmid specificity. J. Bacteriol. *182*, 4022-4027.

Fernandez, D., Dang, T.A., Spudich, G.M., Zhou, X.R., Berger, B.R., and Christie, P.J. (1996a). The Agrobacterium tumefaciens virB7 gene product, a proposed component of the T-complex transport apparatus, is a membrane-associated lipoprotein exposed at the periplasmic surface. J. Bacteriol. *178*, 3156-3167.

Fernandez, D., Spudich, G.M., Zhou, X.R., and Christie, P.J. (1996b). The Agrobacterium tumefaciens VirB7 lipoprotein is required for stabilization of VirB proteins during assembly of the T-complex transport apparatus. J. Bacteriol. *178*, 3168-3176.

Firth, N., Ippen-Ihler, K., and Skurray, R.A. (1996). Structure and function of the F factor and mechanism of conjugation, In Escherichia Coli and Salmonella: Cellular and Molecular Biology, N. FC, ed. (American Sociaty of Microbiology), pp. 2377-2401.

Furuya, N., and Komano, T. (2000). Initiation and termination of DNA transfer during conjugation of IncI1 plasmid R64: roles of two sets of inverted repeat sequences within oriT in termination of R64 transfer. J. Bacteriol. *182*, 3191-3196.

Garcillan-Barcia, M.P., Jurado, P., Gonzalez-Perez, B., Moncalian, G., Fernandez, L.A., and de la Cruz, F. (2007). Conjugative transfer can be inhibited by blocking relaxase activity within recipient cells with intrabodies. Mol. Microbiol. *63*, 404-416.

Gomis-Ruth, F.X., and Coll, M. (2006). Cut and move: protein machinery for DNA processing in bacterial conjugation. Curr. Opin. Struct. Biol. *16*, 744-752.

Gomis-Ruth, F.X., Moncalian, G., de la Cruz, F., and Coll, M. (2002). Conjugative plasmid protein TrwB, an integral membrane type IV secretion system coupling protein. Detailed structural features and mapping of the active site cleft. J. Biol. Chem. *277*, 7556-7566.

Gomis-Ruth, F.X., Moncalian, G., Perez-Luque, R., Gonzalez, A., Cabezon, E., de la Cruz, F., and Coll, M. (2001). The bacterial conjugation protein TrwB resembles ring helicases and F1-ATPase. Nature *409*, 637-641.

Gomis-Ruth, F.X., Sola, M., Acebo, P., Parraga, A., Guasch, A., Eritja, R., Gonzalez, A., Espinosa, M., del Solar, G., and Coll, M. (1998). The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. EMBO J. *17*, 7404-7415.

Gomis-Ruth, F.X., Sola, M., de la Cruz, F., and Coll, M. (2004). Coupling factors in macromolecular type-IV secretion machineries. Curr. Pharm. Des. *10*, 1551-1565.

Gonzalez-Perez, B., Lucas, M., Cooke, L.A., Vyle, J.S., de la Cruz, F., and Moncalian, G. (2007). Analysis of DNA processing reactions in bacterial conjugation by using suicide oligonucleotides. Embo J. *26*, 3847-3857.

Graham, D.Y. (2000). Helicobacter pylori infection is the primary cause of gastric cancer. J. Gastroenterol. *35 Suppl 12*, 90-97.

Grandoso, G., Avila, P., Cayon, A., Hernando, M.A., Llosa, M., and de la Cruz, F. (2000). Two active-site tyrosyl residues of protein TrwC act sequentially at the origin of transfer during plasmid R388 conjugation. J. Mol. Biol. *295*, 1163-1172.

Grandoso, G., Llosa, M., Zabala, J.C., and de la Cruz, F. (1994). Purification and biochemical characterization of TrwC, the helicase involved in plasmid R388 conjugal DNA transfer. Eur. J. Biochem. *226*, 403-412.

Grohmann, E., Muth, G., and Espinosa, M. (2003). Conjugative plasmid transfer in gram-positive bacteria. Microbiol. Mol. Biol. Rev. *67*, 277-301, table of contents.

Guasch, A., Lucas, M., Moncalian, G., Cabezas, M., Perez-Luque, R., Gomis-Ruth, F.X., de la Cruz, F., and Coll, M. (2003). Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. Nat. Struct. Biol. *10*, 1002-1010.

Hare, S., Bayliss, R., Baron, C., and Waksman, G. (2006). A large domain swap in the VirB11 ATPase of Brucella suis leaves the hexameric assembly intact. J. Mol. Biol. *360*, 56-66.

Harris, R.L., Hombs, V., and Silverman, P.M. (2001). Evidence that F-plasmid proteins TraV, TraK and TraB assemble into an envelope-spanning structure in Escherichia coli. Mol. Microbiol. *42*, 757-766.

Harris, R.L., and Silverman, P.M. (2004). Tra proteins characteristic of F-like type IV secretion systems constitute an interaction group by yeast two-hybrid analysis. J. Bacteriol. *186*, 5480-5485.

Hickman, A.B., Ronning, D.R., Kotin, R.M., and Dyda, F. (2002). Structural unity among viral origin binding proteins: crystal structure of the nuclease domain of adeno-associated virus Rep. Mol. Cell *10*, 327-337.

Hofreuter, D., Odenbreit, S., and Haas, R. (2001). Natural transformation competence in Helicobacter pylori is mediated by the basic components of a type IV secretion system. Mol. Microbiol. *41*, 379-391.

Howard, E.A., Zupan, J.R., Citovsky, V., and Zambryski, P.C. (1992). The VirD2 protein of A. tumefaciens contains a C-terminal bipartite nuclear localization signal: implications for nuclear uptake of DNA in plant cells. Cell *68*, 109-118.

Howard, M.T., Nelson, W.C., and Matson, S.W. (1995). Stepwise assembly of a relaxosome at the F plasmid origin of transfer. J. Biol. Chem. *270*, 28381-28386.

Johnson, J.E., and Chiu, W. (2007). DNA packaging and delivery machines in tailed bacteriophages. Curr. Opin. Struct. Biol. *17*, 237-243.

Kalkum, M., Eisenbrandt, R., and Lanka, E. (2004). Protein circlets as sex pilus subunits. Curr. Protein Pept. Sci. *5*, 417-424.

Kingsman, A., and Willetts, N. (1978). The requirements for conjugal DNA synthesis in the donor strain during flac transfer. J. Mol. Biol. *122*, 287-300.

Kramer, M.G., Khan, S.A., and Espinosa, M. (1997). Plasmid rolling circle replication: identification of the RNA polymerase-directed primer RNA and requirement for DNA polymerase I for lagging strand synthesis. EMBO J. *16*, 5784-5795.

Kumar, R.B., Xie, Y.H., and Das, A. (2000). Subcellular localization of the Agrobacterium tumefaciens T-DNA transport pore proteins: VirB8 is essential for the assembly of the transport pore. Mol. Microbiol. *36*, 608-617.

Kupelwieser, G., Schwab, M., Hogenauer, G., Koraimann, G., and Zechner, E.L. (1998). Transfer protein TraM stimulates TraI-catalyzed cleavage of the transfer origin of plasmid R1 *in vivo*. J. Mol. Biol. *275*, 81-94.

Lanka, E., and Wilkins, B.M. (1995). DNA processing reactions in bacterial conjugation. Annu. Rev. Biochem. *64*, 141-169.

Larkin, C., Datta, S., Harley, M.J., Anderson, B.J., Ebie, A., Hargreaves, V., and Schildbach, J.F. (2005). Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the F factor relaxase. Structure *13*, 1533-1544.

Lawley, T.D., Klimke, W.A., Gubbins, M.J., and Frost, L.S. (2003). F factor conjugation is a true type IV secretion system. FEMS Microbiol. Lett. *224*, 1-15.

Lederberg, J. (2000). Infectious history. Science *288*, 287-293.

Lederberg, J., and Tatum, E. (1946). Gene recombination in E. coli. Nature *158*, 558.

Llosa, M., Bolland, S., and de la Cruz, F. (1994). Genetic organization of the conjugal DNA processing region of the IncW plasmid R388. J. Mol. Biol. *235*, 448-464.

Llosa, M., and de la Cruz, F. (2005). Bacterial conjugation: a potential tool for genomic engineering. Res. Microbiol. *156*, 1-6.

Llosa, M., Gomis-Ruth, F.X., Coll, M., and de la Cruz Fd, F. (2002). Bacterial conjugation: a two-step mechanism for DNA transport. Mol. Microbiol. *45*, 1-8.

Llosa, M., Grandoso, G., Hernando, M.A., and de la Cruz, F. (1996). Functional domains in protein TrwC of plasmid R388: dissected DNA strand transferase and DNA helicase activities reconstitute protein function. J. Mol. Biol. *264*, 56-67.

Llosa, M., and O'Callaghan, D. (2004). Euroconference on the Biology of Type IV Secretion Processes: bacterial gates into the outer world. Mol. Microbiol. *53*, 1-8.

Lu, J., Edwards, R.A., Wong, J.J., Manchak, J., Scott, P.G., Frost, L.S., and Glover, J.N. (2006). Protonation-mediated structural flexibility in the F conjugation regulatory protein, TraM. EMBO J. *25*, 2930-2939.

Lujan, S.A., Guogas, L.M., Ragonese, H., Matson, S.W., and Redinbo, M.R. (2007). Disrupting antibiotic resistance propagation by inhibiting the conjugative DNA relaxase. Proc. Natl. Acad. Sci. U. S. A. *104*, 12282-12287.

Lum, P.L., Rodgers, M.E., and Schildbach, J.F. (2002). TraY DNA recognition of its two F factor binding sites. J. Mol. Biol. *321*, 563-578.

Massey, T.H., Mercogliano, C.P., Yates, J., Sherratt, D.J., and Lowe, J. (2006). Double-stranded DNA translocation: structure and mechanism of hexameric FtsK. Mol. Cell *23*, 457-469.

Meinke, G., Bullock, P.A., and Bohm, A. (2006). Crystal structure of the simian virus 40 large T-antigen origin-binding domain. J. Virol. *80*, 4304-4312.

Moncalian, G., Cabezon, E., Alkorta, I., Valle, M., Moro, F., Valpuesta, J.M., Goni, F.M., and de La Cruz, F. (1999a). Characterization of ATP and DNA binding activities of TrwB, the coupling protein essential in plasmid R388 conjugation. J. Biol. Chem. *274*, 36117-36124.

Moncalian, G., and de la Cruz, F. (2004). DNA binding properties of protein TrwA, a possible structural variant of the Arc repressor superfamily. Biochim. Biophys. Acta *1701*, 15-23.

Moncalian, G., Grandoso, G., Llosa, M., and de la Cruz, F. (1997). oriT-processing and regulatory roles of TrwA protein in plasmid R388 conjugation. J. Mol. Biol. *270*, 188-200.

Moncalian, G., Valle, M., Valpuesta, J.M., and de la Cruz, F. (1999b). IHF protein inhibits cleavage but not assembly of plasmid R388 relaxosomes. Mol. Microbiol. *31*, 1643-1652.

Monzingo, A.F., Ozburn, A., Xia, S., Meyer, R.J., and Robertus, J.D. (2007). The structure of the minimal relaxase domain of MobA at 2.1 A resolution. J. Mol. Biol. *366*, 165-178.

Murzin, A.G., Brenner, S.E., Hubbard, T., and Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. J. Mol. Biol. *247*, 536-540.

Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000). Lateral gene transfer and the nature of bacterial innovation. Nature *405*, 299-304.

Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G., and Steitz, T.A. (1985). Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. Nature *313*, 762-766.

Remaut, H., and Waksman, G. (2004). Structural biology of bacterial pathogenesis. Curr. Opin. Struct. Biol. *14*, 161-170.

Rice, P.A., Yang, S., Mizuuchi, K., and Nash, H.A. (1996). Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. Cell *87*, 1295-1306.

Sagulenko, V., Sagulenko, E., Jakubowski, S., Spudich, E., and Christie, P.J. (2001). VirB7 lipoprotein is exocellular and associates with the Agrobacterium tumefaciens T pilus. J. Bacteriol. *183*, 3642-3651.

Santini, J.M., and Stanisich, V.A. (1998). Both the fipA gene of pKM101 and the pifC gene of F inhibit conjugal transfer of RP1 by an effect on traG. J. Bacteriol. *180*, 4093-4101.

Savvides, S.N., Yeo, H.J., Beck, M.R., Blaesing, F., Lurz, R., Lanka, E., Buhrdorf, R., Fischer, W., Haas, R., and Waksman, G. (2003). VirB11 ATPases are dynamic hexameric assemblies: new insights into bacterial type IV secretion. EMBO J. *22*, 1969-1980.

Scott, J.R., and Churchward, G.G. (1995). Conjugative transposition. Annu. Rev. Microbiol. *49*, 367-397.

Schmidt-Eisenlohr, H., Domke, N., Angerer, C., Wanner, G., Zambryski, P.C., and Baron, C. (1999a). Vir proteins stabilize VirB5 and mediate its association with the T pilus of Agrobacterium tumefaciens. J. Bacteriol. *181*, 7485-7492.

Schmidt-Eisenlohr, H., Domke, N., and Baron, C. (1999b). TraC of IncN plasmid pKM101 associates with membranes and extracellular high-molecular-weight structures in Escherichia coli. J. Bacteriol. *181*, 5563-5571.

Seubert, A., Hiestand, R., de la Cruz, F., and Dehio, C. (2003). A bacterial conjugation machinery recruited for pathogenesis. Mol. Microbiol. *49*, 1253-1266.

Sexton, J.A., and Vogel, J.P. (2002). Type IVB secretion by intracellular pathogens. Traffic *3*, 178-185.

Silverman, P.M. (1997). Towards a structural biology of bacterial conjugation. Mol. Microbiol. *23*, 423-429.

Soto, G.E., and Hultgren, S.J. (1999). Bacterial adhesins: common themes and variations in architecture and assembly. J. Bacteriol. *181*, 1059-1071.

Spudich, G.M., Fernandez, D., Zhou, X.R., and Christie, P.J. (1996). Intermolecular disulfide bonds stabilize VirB7 homodimers and VirB7/VirB9 heterodimers during biogenesis of the Agrobacterium tumefaciens T-complex transport apparatus. Proc. Natl. Acad. Sci. U. S. A. *93*, 7512-7517.

Stockner, T., Plugariu, C., Koraimann, G., Hogenauer, G., Bermel, W., Prytulla, S., and Sterk, H. (2001). Solution structure of the DNA-binding domain of TraM. Biochemistry *40*, 3370-3377.

Suzuki, M. (1995). DNA recognition by a beta-sheet. Protein Eng. *8*, 1-4.

Tato, I., Zunzunegui, S., de la Cruz, F., and Cabezon, E. (2005). TrwB, the coupling protein involved in DNA transport during bacterial conjugation, is a DNA-dependent ATPase. Proc. Natl. Acad. Sci. U. S. A. *102*, 8156-8161.

Terradot, L., Bayliss, R., Oomen, C., Leonard, G.A., Baron, C., and Waksman, G. (2005). Structures of two core subunits of the bacterial type IV secretion system, VirB8 from Brucella suis and ComB10 from Helicobacter pylori. Proc. Natl. Acad. Sci. U. S. A. *102*, 4596-4601.

Thorstenson, Y.R., and Zambryski, P.C. (1994). The essential virulence protein VirB8 localizes to the inner membrane of Agrobacterium tumefaciens. J. Bacteriol. *176*, 1711-1717.

Ton-That, H., and Schneewind, O. (2003). Assembly of pili on the surface of Corynebacterium diphtheriae. Mol. Microbiol. *50*, 1429-1438.

Vergunst, A.C., Schrammeijer, B., den Dulk-Ras, A., de Vlaam, C.M., Regensburg-Tuink, T.J., and Hooykaas, P.J. (2000). VirB/D4-dependent protein translocation from Agrobacterium into plant cells. Science *290*, 979-982.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. *1*, 945-951.

Wang, J.C. (1996). DNA topoisomerases. Annu. Rev. Biochem. *65*, 635-692.

Waters, V.L. (2001). Conjugation between bacterial and mammalian cells. Nat. Genet. *29*, 375-376.

Weiss, A.A., Johnson, F.D., and Burns, D.L. (1993). Molecular characterization of an operon required for pertussis toxin secretion. Proc. Natl. Acad. Sci. U. S. A. *90*, 2970-2974.

Wen, M., Yamada, N., Zhang, Y., and Matsuhisa, T. (1997). Morphological changes of Helicobacter pylori after antibacterial therapy: an electron microscope study. Med. Electron. Microsc. *30*, 131-137.

Wilkins, B., and Lanka, E. (1993). DNA processing and replication during plasmid transfer between Gram-negative bacteria (New York: Plenum Press).

Wotherspoon, A.C. (1998). Gastric lymphoma of mucosa-associated lymphoid tissue and Helicobacter pylori. Annu. Rev. Med. *49*, 289-299.

Yeo, H.J., Savvides, S.N., Herr, A.B., Lanka, E., and Waksman, G. (2000). Crystal structure of the hexameric traffic ATPase of the Helicobacter pylori type IV secretion system. Mol. Cell *6*, 1461-1472.

Yeo, H.J., and Waksman, G. (2004). Unveiling molecular scaffolds of the type IV secretion system. J. Bacteriol. *186*, 1919-1926.

Yeo, H.J., Yuan, Q., Beck, M.R., Baron, C., and Waksman, G. (2003). Structural and functional characterization of the VirB5 protein from the type IV secretion system encoded by the conjugative plasmid pKM101. Proc. Natl. Acad. Sci. U. S. A. *100*, 15947-15952.

Ziegelin, G., Furste, J.P., and Lanka, E. (1989). TraJ protein of plasmid RP4 binds to a 19-base pair invert sequence repetition within the transfer origin. J. Biol. Chem. *264*, 11989-11994.