metabolically labelled with BrdU (10 μM, 4 h), trypsinized, and fixed with 70% ethanol. Nuclei were isolated and stained with propidium iodide and FITC-conjugated anti-BrdU antibodies (Becton Dickinson, USA), as described[7]. Flow cytometry was performed on a FACS Sorter (Becton Dickinson). All analysed events were gated to remove debris and aggregates.

Cell death assays

TUNEL assay for DNA fragmentation was done using an In Situ Cell Death Detection kit (BMB), according to manufacturer's protocol. Alternatively, cells were stained with Annexin-V-FLUOR (BMB) and propidium iodide, and analysed by fluorescence microscopy.

Telomere length assay

Genomic DNA (10 μg), isolated from cultured cells, was digested with restriction enzymes *Rsa*I and *Hin*fI and then separated in a 0.5% agarose gel. DNA was transferred to Hybond-N^+ membrane (Amersham, UK). Blots were probed with 5′-end-labelled oligonucleotide $(TTAGGG)_6$ and exposed to a PhosphoImager plate to detect the telomeric ends. An average telomere length was determined using ImageQuant software (Molecular Dynamics, USA).

Quantitative *in situ* hybridization with telomeric probe

In situ hybridization of telomere-specific peptide nucleic acid probe (Telomere PNA FISH Kit/Cy3, DAKO) to metaphase chromosome spreads was performed according to the manufacturer's protocol. Images were captured by a CCD camera attached to a Nikon TE300 microscope and analysed using IPLab Spectrum (Scanalytics Inc.). Background was subtracted and fluorescence signal was integrated in segments corresponding to individual telomeres[15].

Senescence-associated β-galactosidase assay

Senescence-associated β-galactosidase was detected in fixed cells using a described protocol[10]. When staining was fully developed, the cells were washed with PBS and incubated with propidium iodide (1 $\mu g\,ml^{-1}$ in PBS) and with DNase-free RNase A (5 $\mu g\,ml^{-1}$). Both phase-contrast and fluorescent microscopy were performed to identify senescent cells and their nuclei.

Western analysis

Whole-cell extracts were fractionated in gradient (4–20%) polyacrylamide gels (FMC) and transferred to Hybond-P (Amersham) membrane. The following antibodies were used: mouse monoclonal anti-human $p16^{ink4a}$(NeoMarkers, AB-1), mouse anti-p53 (Santa Cruz, DO-1), mouse anti-p21(WAF1) (Calbiochem, Ab-1), rabbit anti-$p14^{ARF}$ (NeoMarkers, Ab-1) and mouse anti-α-actin (Sigma); HRP-conjugated goat-anti-mouse antibody (GibcoBRL) and goat anti-rabbit antibody (Calbiochem). Staining was developed using ECL-detection protocol.

Received 21 September; accepted 16 November 2000.

1. Shay, J. W., Wright, W. E. & Werbin, H. Toward a molecular understanding of human breast cancer: a hypothesis. *Breast Cancer Res. Treatment.* **25** 83–94 (1993).
2. Hayflick, L. The limited *in vitro* lifetime of human diploid cell strains. *Exp. Cell Res.* **37,** 614–636 (1965).
3. Hammond, S. L., Ham, R. G. & Stampfer, M. R. Serum-free growth of human mammary epithelial cells: rapid clonal growth in defined medium and extended passage with pituitary extract. *Proc. Natl Acad. Sci. USA* **81,** 5435–5439 (1984).
4. Foster, S. A. & Galloway, D. A. Human papillomavirus type 16 E7 alleviates a proliferation block in early passage human mammary epithelial cells. *Oncogene* **12,** 1773–1779 (1996).
5. Huschtscha, L. I. *et al.* Loss of $p16^{INK4}$ expression by methylation is associated with lifespan extension of human mammary epithelial cells. *Cancer Res.* **58,** 3508–3512 (1998).
6. Kiyono, T. *et al.* Both Rb/$p16^{INK4a}$ inactivation and telomerase activity are required to immortalize human epithelial cells. *Nature* **396,** 84–88 (1998).
7. Meyer, K. M., Hess, S. M., Tlsty, T. D. & Leadon, S. A. Human mammary epithelial cells exhibit a differential p53-mediated response following exposure to ionizing or UV light. *Oncogene* **18,** 5792–57805 (1999).
8. Walen, K. H. & Stampfer, M. R. Chromosome analyses of human mammary epithelial cells at stages of chemical-induced transformation progression to immortality. *Cancer Genet. Cytogenet.* **37,** 249–261 (1989).
9. Brenner, A. J., Stampfer, M. R. & Aldaz, C. M. Increased p16 expression with first senescence arrest in human mammary epithelial cells and extended growth capacity with p16 inactivation. *Oncogene* **17,** 199–205 (1998).
10. Dimri, G. P. *et al.* A biomarker that identifies senescent human cells in culture and in aging skin *in vivo*. *Proc. Natl Acad. Sci. USA* **92,** 9363–9367 (1995).
11. Pignolo, R. J., Rotenberg, M. O. & Cristofalo, V. J. Alterations in contact and density-dependent arrest state in senescent WI-38 cells. *In Vitro Cell. Dev. Biol. Anim.* **30A,** 471–476 (1994).
12. Shay, J. W. & Wright, W. E. Quantitation of the frequency of immortalization of normal human diploid fibroblasts by SV40 large T-antigen. *Exp. Cell Res.* **184,** 109–118 (1989).
13. Taylor-Papadimitriou, J. *et al.* Keratin expression in human mammary epithelial cells cultured from normal and malignant tissue: relation to *in vivo* phenotypes and influence of medium. *J. Cell Sci.* **94,** 403–413 (1989).
14. Foster, S. A., Wong, D. J., Barrett, M. T. & Galloway, D. A. Inactivation of p16 in human mammary epithelial cells by CpG island methylation. *Mol. Cell. Biol.* **18,** 1793–1801 (1998).
15. Lansdorp, P. M. *et al.* Heterogeneity in telomere length of human chromosomes. *Hum. Mol. Genet.* **5,** 685–691 (1996).
16. van Steensel, B., Smogorzewska, A. & de Lange, T. TRF2 protects human telomeres from end-to-end fusions. *Cell* **92,** 401–413 (1998).
17. Stampfer, M. R. *et al.* Gradual phenotypic conversion associated with immortalization of cultured human mammary epithelial cells. *Mol. Biol. Cell* **8,** 2391–2405 (1997).
18. Karlseder, J., Broccoli, D., Dai, Y., Hardy, S. & de Lange, T. p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. *Science* **283,** 1321–1325 (1999).
19. Artandi, S. E. *et al.* Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. *Nature* **406,** 641–645 (2000).
20. Chin, L. *et al.* p53 Deficiency rescues the adverse effects of telomere loss and cooperates with telomere dysfunction to accelerate carcinogenesis. *Cell* **97,** 527–538 (1999).
21. Alcorta, D. A. *et al.* Involvement of the cyclin-dependent kinase inhibitor p16 (INK4A) in replicative senescence of normal human fibroblasts. *Proc. Natl Acad. Sci. USA* **92,** 13742–13747 (1996).
22. Hara, E. *et al.* Regulation of p16CDKN2 expression and its implications for cell immortalization and senescence. *Mol. Cell. Biol.* **16,** 859–867 (1996).
23. Burbano, R. R. *et al.* Cytogenetics of epithelial hyperplasias of the human breast. *Cancer Genet. Cytogenet.* **119,** 62–66 (2000).
24. Pandis, N. *et al.* Chromosome abnormalities in bilateral breast carcinomas. Cytogenetic evaluation of the clonal origin of multiple primary tumors. *Cancer* **76,** 250–258 (1995).
25. Berg, J. W. & Hutter, R. V. Breast cancer. *Cancer* **75,** 257–269 (1995).
26. Stoeber, K. *et al.* Cdc6 protein causes premature entry into S phase in a mammalian cell-free system. *EMBO J.* **17,** 7219–7229 (1998).
27. Wei, W. & Sedivy, J. M. Differentiation between senescence (M1) and crisis (M2) in human fibroblasts cultures. *Exp. Cell Res.* **253,** 519–522 (1999).
28. Tlsty, T. D. *et al.* Potentiation of genomic instability in normal human mammary epithelial cells by an epigenetic event. *J. Mammary Gland Biol. Neoplasia* (in the press).

Supplementary information is available on *Nature*'s World-Wide Web site (http://www.nature.com) or as paper copy from the London editorial office of *Nature*.

Acknowledgements

We thank E. H. Blackburn, I. Herskowitz and J. Li for comments, criticism and reading the manuscript. Stimulating discussion and thoughtful critique were provided by Y. Crawford, G. Whitworth, M. Heiman, M. Springer, D. Crawford, P. Hein and J. Anderson. We thank S. Gilbert for assistance with the figures; P. Ortiz for library support; and G. Williams for MCM2 antibodies. This work was supported by NIH and NIH/NASA grants to T.D.T. and a DOE and NIH grant to M.R.S. C.R.H. is supported by a Howard Hughes Pre-doctoral Fellowship.

Correspondence and requests for materials should be addressed to T.D.T. (e=mail: ttlsty@itsa.ucsf.edu).

The bacterial conjugation protein TrwB resembles ring helicases and F_1-ATPase

F. Xavier Gomis-Rüth*, Gabriel Moncalián†, Rosa Pérez-Luque*, Ana González‡§, Elena Cabezón†, Fernando de la Cruz† & Miquel Coll*

** Institut de Biologia Molecular de Barcelona, CSIC, Jordi Girona, 18-26, 08034 Barcelona, Spain*
† Departamento de Biología Molecular (Laboratorio asociado al CIB, CSIC), Universidad de Cantabria, Herrera Oria, 39011 Santander, Spain
‡ EMBL Hamburg Outstation, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany

The transfer of DNA across membranes and between cells is a central biological process; however, its molecular mechanism remains unknown. In prokaryotes, *trans*-membrane passage by bacterial conjugation, is the main route for horizontal gene transfer. It is the means for rapid acquisition of new genetic information, including antibiotic resistance by pathogens. *Trans*-kingdom gene transfer from bacteria to plants[1] or fungi[2] and even bacterial sporulation[3] are special cases of conjugation. An integral membrane DNA-binding protein, called TrwB in the *Escherichia coli* R388 conjugative system, is essential for the conjugation process. This large multimeric protein is responsible for recruiting the relaxosome DNA–protein complex, and participates in the

§ Present address: Stanford Linear Accelerator Center, 2575 Sand Hill Road, Mail Stop 99, Menlo Park, California 94025, USA.

transfer of a single DNA strand during cell mating. Here we report the three-dimensional structure of a soluble variant of TrwB. The molecule consists of two domains: a nucleotide-binding domain of α/β topology, reminiscent of RecA and DNA ring helicases, and an all-α domain. Six equivalent protein monomers associate to form an almost spherical quaternary structure that is strikingly similar to F_1-ATPase. A central channel, 20 Å in width, traverses the hexamer.

Bacterial conjugation is mediated by proteins encoded by two gene regions of conjugative plasmids in Gram-negative bacteria. The first region encodes the DNA transfer and replication proteins (*Dtr*) and the second one encodes the mating pair formation proteins (*Mpf*). R388 is a 33-kilobase enterobacterial conjugative plasmid that contains genes that code for sulphonamide and trimethoprim resistance. It displays the simplest known *Dtr* region[4] encoding only three proteins. TrwC acts as a relaxase as well as a DNA helicase. TrwA is a transcriptional repressor. These two proteins comprise, together with the origin of transfer (*oriT* DNA) and the host-encoded integration host factor, the relaxosome. The third protein encoded by the R388 *Dtr* region, TrwB, is a basic (isoelectric point (pI) 10) integral membrane protein of 507 residues[5] with a nucleotide-binding site (NBS). This feature is essential for conjugation, because a single point mutation affecting the Walker A motif (K136T) is completely deficient in gene transfer[5].

TrwB binds single-stranded and double-stranded DNA nonspecifically and independently of NTP binding. The protein belongs to the family of TraG-like coupling proteins responsible for recruiting the relaxosome and coupling it with the structures involved in the cell-to-cell contact[6,7]. This family displays similarity with the SpoIIIE/FtsK protein family[5]. SpoIIIE is an essential sporulation membrane-bound protein in *Bacillus subtilis* that is responsible for DNA translocation from the mother cell to the prespore[3], whereas FtsK is involved in the resolution of chromosomal dimers during

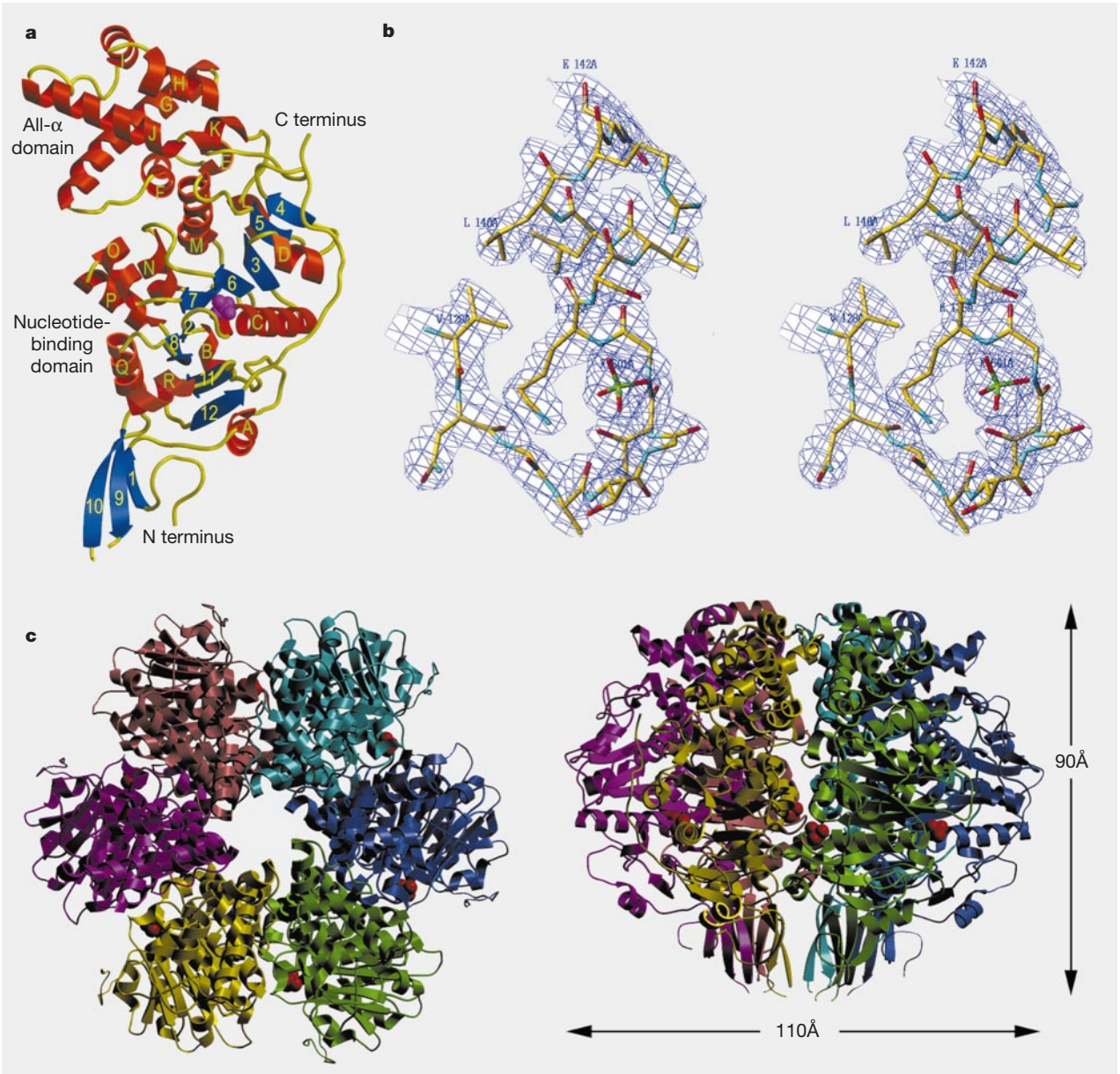

Figure 1 Structural features of TrwB. **a**, Ribbon diagram of a TrwBΔN70 protomer. Strands are numbered (1–12), helices are labelled with capital letters (A–R). **b**, Stereo representation of the final ($2F_{obs} - F_{calc}$) map of the $P3_121$ crystal at the nucleotide-binding site with the final model coordinates superimposed. A sulphate anion is visible at the position of a β-phosphate group of the substrate. **c**, Hexamer viewed from the cytoplasmic side. Left, view along the local six-fold axis; right, side view.

bacterial cell division[8]. SpoIIIE and FtsK are believed to function as motors that drive DNA across the membrane annulus during sporulation and cell division, respectively. In addition, TraG-like proteins are also related to VirD4, a protein from the plant tumour-inducing bacterial pathogen *Agrobacterium tumefaciens*, which is needed to deliver transforming DNA into the plant cell.

We have solved the structure of a soluble fragment of TrwB lacking the transmembrane part (TrwBΔN70) in two crystal forms. The TrwBΔN70 monomer displays an 'orange-segment' shape of approximate dimensions 90×45×40 Å (Figs 1a, 2a). It can be subdivided into two domains: an all-α helical domain (AAD); and an α/β twisted open-sheet domain containing the nucleotide-binding domain (NBD), which is proximal to the membrane. The NBD runs from the amino terminus, which is the anchor point of the excised 70-residue transmembrane domain, to residue Lys 183 and continues from position Asp 298 to the carboxy terminus located at the surface. The NBD is composed of a central, highly twisted (about 90°), nine-stranded β-pleated sheet of mixed parallel/antiparallel topology (strands 2–8, 11 and 12; see Fig. 1a), which is flanked on its concave side by four helices (A–D) and on its convex side by seven helices (L–R). The convex side faces the interior channel of the particle. On the membrane-proximal edge of this central β-sheet, a small three-stranded antiparallel sheet (strands 1, 9 and 10) is inserted, almost perpendicular to the former. Strands 9 and 10 are only partially defined by electron density (see Methods), probably owing to the absence of interactions with the (excised) transmembrane domain preceding strand 1. On top of the NBD, the smaller AAD is inserted between strand β4 and helix αL, comprising residues Gly 184 to Gly 297 (Fig. 1a). This domain contains seven helices: a 3_{10}-helix (helix E), followed by a two-helix hairpin (helices F and G), followed at the C-terminal end by a second, curved, helical hairpin segment (helices H–K), which mainly protects and interacts with the central helix G.

In both crystal forms, six TrwBΔN70 protomers oligomerize using a local six-fold axis (Fig. 1c). This agrees both with biochemical data showing that the native protein migrates as a hexamer in gel-filtration column chromatography, and with electron microscopy images (data not shown). The hexamer has overall dimensions of 110 Å in diameter and 90 Å in height (Fig. 1c), and a roughly spherical shape, which is slightly flattened at both poles. A central channel runs from the cytosol pole (made up by the AADs) to the membrane pole (formed by the NBDs; Fig. 2a, b). The channel entrance is plugged by a ring of asparagine residues (Asn 271 of loop αI–αJ from each subunit) and restricted to ~8 Å in diameter. This is the narrowest part of the channel—which is ~20 Å wide along its whole length and ends at its membrane side with an opening of ~22 Å—although a narrower section may occur at the transmembrane part to avoid membrane permeability.

The interaction between TrwBΔN70 monomers is mainly hydrophilic (Fig. 2a), and buries a common surface of 4,588 $Å^2$ (about 25% of the total surface of a monomer) with the central twisted β-pleated sheets arranged almost in parallel (C-edge to N-edge; Fig. 1c). The total intermolecular contact area of a monomer (for example, A with both B and G), is 9,100 $Å^2$, which is almost 50% of the total surface. A notable interaction is observed between strands 9 and 10 (partially interrupted at their ends) and residues of the N-terminal segments of each monomer to form a 12-stranded β-barrel. This barrel might coat the interior of the transmembrane channel formed by the helical segments of the transmembrane domain (Fig. 2c).

The NBS is located at the C-terminal edge of the sheet and is shaped by loops β2–αC and β6–αN. These two segments contain

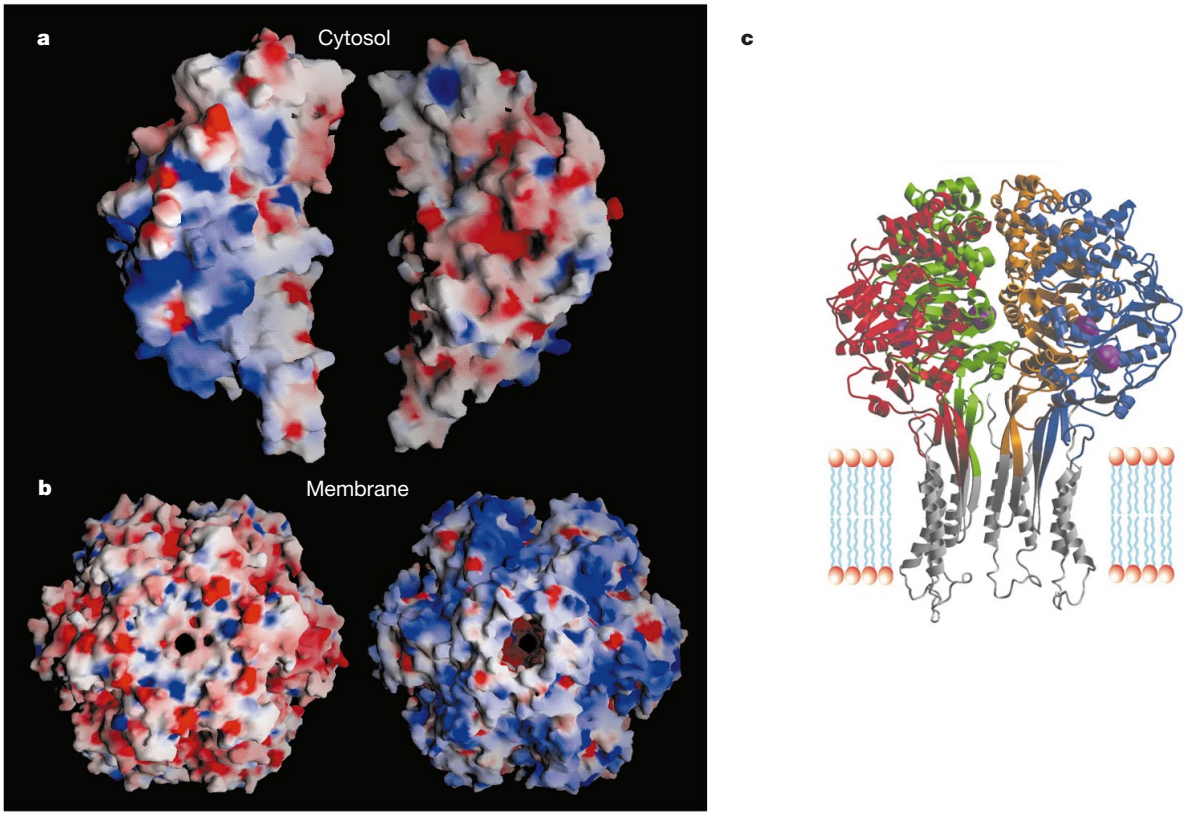

Figure 2 Electrostatic surface representations. **a**, Lateral view. Only two (opposed) protomers of the hexamer are shown to highlight the surface complementarity and the central channel. **b**, View along the six-fold axis from the cytosolic pole (left) and from the membrane pole (right). **c**, Ribbon representation of the complete TrwB, including a model for the transmembrane part (in grey), based on two consecutive transmembrane helices of the photosynthetic reaction centre structure (PDB access code 1mps). Only four of the six protomers are shown for clarity.

the Walker A motif or P-loop (G130-A131-T132-G133-T134-G135-K136-S137) and the Walker B motif (D356-E357-L358-A359) found in NTP-binding proteins[9]. The six nucleotide-binding sites are located on superficial cavities, about 32 Å apart from each other at interfaces between vicinal protomers. We superimposed the structure of F_1-ATPase[10], with which TrwB has strong structural similarity, to deduce the residues involved in nucleotide binding. Possible residues are Thr 132, whose Oγ atom would bind the γ-phosphate oxygens; Lys 136, which binds the β-phosphate oxygens; and Ser 137, which would interact through a Mg^{2+} cation with oxygens of both the β- and γ-phosphates. The proposed catalytic residues in F_1-ATPase do not match any residues in the present structure, suggesting that it has a different enzymatic mechanism. The C terminus of TraG-like proteins has been implicated in the efficiency and specificity of their interaction with the relaxosome[31]. It might be significant that the TrwB C-terminal part is directly linked to the NBS, and thus conceivably may modulate NTP hydrolysis through interaction with components of the relaxosome.

Co-crystallization and soaking trials of trigonal or monoclinic crystals in solutions containing standard NTP analogues always resulted in partially occupied NBSs, making the precise definition of inhibitor binding impossible, in particular of the base moiety. It might be that the active sites function in a sequential manner, as has been suggested for bacteriophage T7 gene 4 helicase[11], and that a particle with all six sites occupied is not viable. This could lead to disordered low-occupancy sites in the complexes. In the trigonal crystals, however, a strongly bound sulphate anion (100% occupancy) was found in the nucleotide-binding pocket, at the position where the nucleotide β-phosphate should be located (Fig. 1b). When we superimposed a protomer from this trigonal crystal with one of a monoclinic crystal grown from tartrate (root mean square deviation (r.m.s.d.) 0.47 Å) that has empty NBSs, we found that some regions deviate strongly, despite the close overall similarity of the hexamers. In particular, segments Gly 130 to Arg 141 (r.m.s.d. 1.1 Å) encompassing the P-loop, Gly 384 to Gln 390 (r.m.s.d. 1.0 Å), and Gly 414 to Gly 415 (r.m.s.d. 1.6 Å) diverge greatly, including their main chain atoms. Notably, the main chain of Walker motif B is not affected. These differing regions are not involved in any surface crystal contact, and should therefore be considered to be a consequence of the occupation of the NBSs. An internal salt bridge made up by Glu 357 (in Walker B motif) with the crucial Lys 136 (in Walker A motif) in the unliganded structure is disrupted in the SO_4^{2-} complex. The latter side chain is rotated to approach the substrate β-phosphate position. This is accompanied by a series of rearrangements transmitting the movement to the interior channel at position Ser 289, which is located at the beginning of helix O.

Topological searches of TrwB against proteins of known tertiary structure reveal, despite negligible sequence similarity, a high structural similarity of its NBD with the equivalent domain of RecA[12] (PDB access code 2reb) and other RecA-like core encompassing enzymes. Among these are bacteriophage T7 gene 4 helicase[11,13] (PDB access codes 1cr0, 1eok and 1eoj), the δ′-subunit of the clamp loader complex of *E. coli* DNA polymerase III (ref. 14) (1a5t), and the *N*-ethylmaleimide-sensitive fusion protein (NFS), a cytosolic ATPase required for intracellular vesicle fusion reactions[15] (1d2n). Also, there is strong structural similarity with both α- and β-subunits of F_1-ATPase, part of the membrane-associated F_oF_1-ATPase complex responsible for energy conversion through axial rotation movement in mitochondria, chloroplasts and bacteria[10] (1bmf). This structural analogy is much higher with the β-subunits in the occupied conformation (β_{DP} and β_{TP}) than with the empty one (β_E).

When comparing TrwB with all the hexameric ring quaternary structures, helicases and NFS appear to be slightly wider and more flat-topped, although this effect is overemphasized because of missing domains in the solved structures. TrwB is almost spherical and remarkably similar, in shape and dimensions, to the F_1-ATPase $\alpha_3\beta_3$ heterohexamer, with which it shares the property of being a membrane-associated protein. In contrast, TrwB is a homohexamer. Several representatives of the helicase/ATPase-like proteins display an additional AAD, besides the core NBD, as shown for RecA, NSF, *Bacillus stearothermophilus* DExx box DNA helicase[16], δ′-subunit of the clamp loader complex of *E. coli* DNA polymerase III and α- and β-subunits of F_1-ATPase, among others. These are arranged in different ways with respect to the NBD; however, detailed topological inspections reveal that TrwB AAD bears significant structural similarity only with N-terminal domain 1 of the site-specific recombinase, XerD, of the λ-integrase family (PDB access code 1a0p; ref. 17). In XerD, this all-α domain 1 has been proposed to contribute partially to DNA binding, which is achieved mainly by domain 2. Proteins in the TrwB superfamily share the NBD but have non-homologous AADs. This argues in favour of these domains being added later on in the evolution of this protein family to modulate ATP hydrolysis as a result of AAD-specific interactions.

DNA helicases use the energy from NTP hydrolysis to unwind double-stranded DNA. Those that act at the replication fork are ring-shaped hexameric enzymes that move along one DNA strand that passes through the central hole, and that catalyse the displacement of the complementary strand[13,18–20]. The crystal structures of the two ring helicases T7 gene 4 (ref. 11) and RepA (see ref. 21) led the former authors to propose a sequential NTP hydrolytic mechanism. The strong structural resemblance of TrwB with DNA ring helicases suggests, by analogy, that the single DNA T-strand might pass through the central channel of the particle, thus being introduced into the mating/translocation apparatus that connects donor and recipient cells. ATP hydrolysis would provide the energy to pump the single-stranded DNA, much in the same way as it does in helicases for their processive movement along the DNA[22,23]. The structure of TrwB with a sulphate ion bound at the NBS shows that this binding results in structural changes that are transmitted to the central channel. This gives an idea on how the protein might function: ATP binding/hydrolysis could trigger a molecular switch mechanism affecting the channel, and therefore the DNA binding and displacement through it. One obstacle to the hypothesis that the single-stranded DNA traverses the central channel is the asparagine plug at the cytoplasmic side, which narrows the channel to ~8 Å (Fig. 2a); elsewhere the channel is an appropriate diameter (~20 Å) to accommodate a single DNA strand. If the channel hypothesis is correct, our structure would represent a closed conformation and the cytoplasmic gate would open on activation—a movement that might be triggered by attachment of relaxosome components. □

Table 1 Final refinement and model statistics

Structure	SO_4^{2-} complex	Apo
Space group	$P3_121$	$P2_1$
Resolution range (Å)	50–2.4	50–2.5
No. of reflections used	133,805	176,026
Crystallographic R_{factor} (free R_{factor})*	21.0 (24.8)	20.9 (26.7)
No. of protein atoms (a.u.)	19,693	41,537
No. of solvent molecules	965	1,559
No. of sulphate anions	15	–
R.m.s. deviation from target values		
Bonds (Å)	0.008	0.008
Angles (°)	1.393	1.387
Bonded *B* factors (Å^2)	1.826	1.959
Average *B* factors for protein atoms (Å^2)	43.0	40.8

*$R_{factor} = \Sigma_{hkl}||F_{obs}| - k|F_{calc}||/\Sigma_{hkl}|F_{obs}|$; free R_{factor}, same for a test set of 7% reflections (>500) not used during refinement (until last-but-one cycle).

Methods

Crystallization and heavy-atom derivation

The soluble 437-residue variant of TrwB (sequence accession number SWALL:Q04230), TrwBΔN70, was prepared as described[5]. The protein was crystallized from drops initially containing 1.5 M ammonium sulphate, 0.1 M NaCl, 0.1 M HEPES pH 7.5. Trigonal crystals belonging to space group $P3_121$ ($a = b = 151.3$ Å, $c = 258.2$ Å) were obtained. These crystals diffract to about 2.0 Å resolution and contain six protomers per asymmetric unit.

When 0.9 M potassium/sodium tartrate, 0.1 M HEPES pH 7.5 was used as precipitant, similar trigonal crystals appeared in the drops together with a second, monoclinic ($P2_1$) crystal form (a = 107.4 Å, b = 153.4 Å, c = 162.5 Å, β = 94.2°). These crystals diffract as well as the trigonal although they contain 12 protomers in the asymmetric unit. A tantalum bromide ($Ta_6Br_{12}^{2+}$) derivative of the trigonal crystal form was obtained by overnight soaking of native crystals resulting in a unit cell with an enlarged c axis parameter ($a = b$ = 151.2 Å, c = 262.8 Å).

Data collection and phasing

Complete diffraction datasets were collected from a single crystal each at EMBL beam lines BW7A and BW7B (Outstation Hamburg, DESY) and ID14-2 and ID14-4 (Outstation Grenoble, ESRF). Three datasets were recorded from the $Ta_6Br_{12}^{2+}$ derivative (peak, inflection point and hard remote) around the theoretical Ta L-III edge. All data were processed with MOSFLM 6.01 (ref. 24) and scaled, merged and reduced with programs of the CCP4 suite[25]. A self-rotation calculation performed for the trigonal data indicated the presence of a local six-fold axis at Eulerian angles α = 12.8°, β = 56.6° and γ = 12.8°, in accordance with the presence of six monomers displaying local C6 symmetry. The structure was solved by multiple-wavelength anomalous diffraction (MAD) using the tantalum bromide derivative of the trigonal crystal form using SHELXS[26] and MLPHARE[25]. Six sites were found and refined, and phases were computed to 4.5 Å. A subsequent density modification step using DM within CCP4 (ref. 25) rendered an F_{obs} electron density map that allowed us to build a monomer mask and localize the local six-fold axis which was consistent with the self-rotation function. The original MAD-phases to 4.5 Å were used with the native structure factor amplitudes and a density modification run with non-crystallographic-symmetry (NCS) averaging and phase extension to 2.7 Å was computed. This led to a straightforward traceable native map. The model was built with Turbo-Frodo.

Structure refinement

Successive cycles of maximum-likelihood positional and temperature factor refinement with CNS version 1.0 (ref. 27) using progressively all data up to the full resolution of 2.4 Å and keeping NCS restraints followed. Computation of phased-combined maps and manual model building permitted the gradual completion of the model. At the final stages, solvent molecules and sulphate ions were introduced. Table 1 provides a summary for final model refinement. The structure of the monoclinic crystal form was solved by molecular replacement with AMoRe[28]. A whole TrwBΔN70 hexamer was used as a searching model and two clear solutions were obtained with correlation coefficient and R_{factor} equal to 67.4/36.3% (second highest solution 41.2/47.1%). The model was inspected and the new electron-density-based differences with the trigonal structure were corrected. Solvent molecule position assignment and model refinement with CNS proceeded similarly and applying NCS restraints. No ions were localized in these crystals. All residues excepting His 125 of each chain, clearly defined by electron density and involved in β-sheet interactions, are in allowed regions of the Ramachandran plot. Superimpositions were calculated with Turbo-Frodo, Lsqkab of the CCP4 suite[25] and Lsqman of the RAVE package (Uppsala Software Factory; http://alpha2.bmc.uu.se/~gerard/manuals). Figures were computed with Turbo-Frodo, SETOR[29] and GRASP[30]. Lee & Richards buried surface accessibility calculations (probe radius 1.4 Å) and close contacts (< 4 Å) were ascertained with CNS version 1.0. Structural similarity searches were performed with the DALI server (http://www.ebi.ac.uk/dali).

Received 21 September; accepted 8 November 2000.

1. Stachel, S. E. & Zambryski, P. *Agrobacterium tumefaciens* and the susceptible plant cell: a novel adaptation of extracellular recognition and DNA conjugation. *Cell* **47,** 155–157 (1986).
2. Heinemann, J. A. & Sprague, G. F. J. Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature* **340,** 205–209 (1989).
3. Wu, L. J., Lewis, P. J., Allmansberger, R., Hauser, P. M. & Errington, J. A conjugation-like mechanism for prespore chromosome partitioning during sporulation in *Bacillus subtilis*. *Gene Dev.* **9,** 1316–1326 (1995).
4. Llosa, M., Bolland, S. & de la Cruz, F. Genetic organization of the conjugal DNA processing region of the IncW plasmid R388. *J. Mol. Biol.* **235,** 448–464 (1994).
5. Moncalián, G. *et al.* Characterization of ATP and DNA binding activities of TrwB, the coupling protein essential in plasmid R388 conjugation. *J. Biol. Chem.* **274,** 36117–36124 (1999).
6. Zechner, E. L. in *The Horizontal Gene Pool: Bacterial Plasmids and Gene Spread* (ed. Thomas, C. M.) 87–173 (Harwood Academic, London, 2000).
7. Cabezón, E., Sastre, J. I. & de la Cruz, F. Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation. *Mol. Gen. Genet.* **254,** 400–406 (1997).
8. Begg, K. J., Dewar, S. J. & Donachie, W. D. A new *Escherichia coli* cell division gene, *ftsK*. *J. Bacteriol.* **177,** 6211–6222 (1995).
9. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1,** 945–951 (1982).
10. Abrahams, J. P., Leslie, A. G. W., Lutter, R. & Walker, J. E. Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria. *Nature* **370,** 621–628 (1994).
11. Singleton, M. R., Sawaya, M. R., Ellenberger, T. & Wigley, D. B. Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* **101,** 589–600 (2000).
12. Story, R. M., Weber, I. T. & Steitz, T. A. The structure of the *E.coli* recA protein monomer and polymer. *Nature* **355,** 318–325 (1992).
13. Sawaya, M. R., Guo, S., Tabor, S., Richardson, C. C. & Ellenberger, T. Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7. *Cell* **99,** 167–177 (1999).
14. Guenther, B., Onrust, R., Sali, A., O'Donnell, M. & Kuriyan, J. Crystal structure of the δ′ subunit of the clamp-loader complex of *E.coli* DNA polymerase III. *Cell* **91,** 335–345 (1997).
15. Lenzen, C. U., Steinmann, D., Whiteheart, S. W. & Weis, W. I. Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. *Cell* **94,** 525–535 (1998).
16. Subramanya, H. S., Bird, L. E., Brannigan, J. A. & Wigley, D. B. Crystal structure of a DExx box DNA helicase. *Nature* **384,** 379–383 (1996).
17. Subramanya, H. S. *et al.* Crystal structure of the site-specific recombinase, XerD. *EMBO J.* **16,** 5178–5187 (1997).
18. Egelman, E. H., Yu, X., Wild, R., Hingorani, M. M. & Patel, S. M. Bacteriophage T7 helicase/primase proteins form rings around single-stranded DNA that suggest a general structure for hexameric helicases. *Proc. Natl Acad. Sci. USA* **92,** 3869–3873 (1995).
19. Hacker, K. J. & Johnson, K. A. A hexameric helicase encircles one DNA strand and excludes the other during DNA unwinding. *Biochemistry* **36,** 14080-14087 (1997).
20. Yu, X., Hingorani, M. M., Patel, S. S. & Egelman, E. H. DNA is bound within the central hole to one or two of the six subunits of the T7 DNA helicase. *Nature Struct. Biol.* **3,** 740–743 (1996).
21. Soultanas, P. & Wigley, D. B. DNA helicases: 'inching forward'. *Curr. Opin. Struct. Biol.* **10,** 124–128 (2000).
22. Raney, K. D. & Benkovic, S. J. Bacteriophage T4 DDA helicase translocates in an unidirectional fashion on single-stranded DNA. *J. Biol. Chem.* **270,** 22236–22242 (1995).
23. Kaplan, D. L. The 3′-tail of a forked-duplex sterically determines whether one or two DNA strands pass through the central channel of a replication-fork helicase. *J. Mol. Biol.* **301,** 285–299 (2000).
24. Leslie, A. G. W. in *Crystallographic computing V* (eds Moras, D., Podjarny, A. D. & Thierry, J. C.) 27–38 (Oxford Univ. Press, Oxford, 1991).
25. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
26. Sheldrick, G. M. Patterson superposition and *ab initio* phasing. *Methods Enzymol.* **276,** 628–641 (1997).
27. Brünger, A. T. *et al.* Crystallography & NMR System: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
28. Navaza, J. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50,** 157–163 (1994).
29. Evans, S. V. SETOR: hardware lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graphics* **11,** 134–138 (1993).
30. Nicholls, A., Bharadwaj, R. & Honig, B. GRASP: graphical representation and analysis of surface properties. *Biophys. J.* **64,** A166–A166 (1993).
31. Sastre, J. I., Cabezon, E. & de la Cruz, F. The carboxyl terminus of protein TraD adds specificity and efficiency to F-plasmid conjugative transfer. *J. Bacteriol.* **180,** 6039–6042 (1998).

Acknowledgements

We are most grateful to R. Huber for making tantalum bromide available to us, and to I. Usón for help with SHELX. This work was supported by grants from the Ministerio de Educación y Cultura of Spain, the Generalitat de Catalunya and the European Union. Synchrotron data collection was supported by EU grants and the ESRF.

Correspondence and requests for materials should be addressed to M.C. (e-mail: mcccri@ibmb.csic.es). The co-ordinates of TrwBΔN70 have been deposited with the Protein Data Bank (access codes 1e9r and 1e9s).

Three key residues form a critical contact network in a protein folding transition state

Michele Vendruscolo*, Emanuele Paci*†, Christopher M. Dobson* & Martin Karplus*†‡

** Oxford Centre for Molecular Sciences, New Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QT, UK*
† Laboratoire de Chimie Biophysique, ISIS, Université Louis Pasteur, 4 rue Blaise Pascal, 67000 Strasbourg, France
‡ Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, USA

Determining how a protein folds is a central problem in structural biology. The rate of folding of many proteins is determined by the transition state, so that a knowledge of its structure is essential for understanding the protein folding reaction. Here we use mutation measurements—which determine the role of individual residues in stabilizing the transition state[1,2]—as restraints in a Monte Carlo sampling procedure to determine the ensemble of structures that make up the transition state. We apply this approach to the experimental data for the 98-residue protein acylphosphatase[3], and obtain a transition-state ensemble with the