Bacterial conjugation is an example of macromolecular trafficking between cells and responsible for the spreading of antibiotic resistance among bacteria. It involves translocation of single-stranded DNA across membranes through a type IV secretion system. A coupling protein links the DNA-processing nucleoprotein complex, the relaxosome, with the transport apparatus during cell mating. In *Escherichia coli* plasmid R388, such a protein is TrwB, a basic integral inner-membrane nucleoside-triphosphate-binding protein. TrwB is the structural prototype for the type IV secretion system coupling proteins, a family of proteins essential for macromolecular transport between cells and export. The structure of a soluble TrwB variant unveils an elongated molecule with six equivalent protein units featuring a spherical quaternary structure, leaving a central channel. The structures of the non-ligated protein and four different complexes with substrate analogues and products allow the precise description of the active site architecture. The active sites are located at the interface between protomers, each of them shaped mainly by residues of one monomer, but including two crucial arginine residues belonging to the adjacent molecule. Upon substrate binding and putative hydrolysis, conformational changes are transferred from the external surface to the interior central channel.

Transport of DNA across cell membranes occurs in different biological processes such as viral infection, cell division, sporulation, or conjugation. The latter is mediated by conjugative plasmids aimed to transfer DNA unidirectionally between bacteria, from a donor to a recipient cell, which then becomes a potential donor. It is the main process by which bacterial pathogens acquire antibiotic resistance and constitutes a widespread tool to shuttle genes between different species of bacteria. Even trans-kingdom gene transfer can be achieved through a conjugation-like system, from bacteria to fungal or plant cells (1–5).

The transfer utilizes a type IV secretion system (6), a multiprotein transmembrane organelle, and requires physical contact between donor and recipient cells. Genes contained in the plasmid transference (*tra*) region, further subdivisible into *dtr* and *mpf*, control conjugation (5, 7). Mpf proteins are required for mating pair formation and assembly of the conjugative or sex-pilus. The region of DNA transfer and replication genes, *dtr*, includes *oriT*, the origin of transfer, and genes encoding proteins involved in plasmid DNA processing. Establishment of stable intercellular contacts may trigger a signal that is transmitted to a specialized nucleoprotein complex, the relaxosome (1), made up by the DNA region at *oriT*, and Dtr proteins that cleave and unwind DNA. The relaxosome moves to the transport site, and the type IV DNA transport apparatus (formed by Mpf proteins) transfers single-stranded (ss) DNA or T-strand to the recipient cell.

*Escherichia coli* plasmid R388 (8) has a size of 33 kb and contains genes that code for sulfonamide and trimethoprim resistance (9, 10). Its *mpf* region encodes a functional type IV secretion system, featuring 11 *trw* genes (*trwD* to *trwN*), which are similar to other conjugative *mpf* genes and to *Agrobacterium tumefaciens* virB genes (11, 12). R388 displays the shortest *dtr* region known (12, 13), just made up by *oriT*, *trwA*, *trwB*, and *trwC*. *TrwC* acts as a relaxase and helicase. It is responsible for both nick cleavage at *oriT* and T-strand unwinding (14, 15). *TrwA* is a small-sized tetrameric protein, which binds two sites at *oriT* and enhances TrwC relaxase activity, while repressing transcription of the *trwABC* operon (16). These two proteins constitute, together with *oriT* DNA and the host-encoded integration host factor, the relaxosome of plasmid R388 (12). Integration host factor has a modulator role in

*This work was supported in part by Grant PB98-1106 from Dirección General de Investigación-Ciencia y Tecnología (Spain). Data collection was performed at protein crystallography beamlines BW7A, BW7B, X31, and X11 (all EMBL) and BW6 (Max-Planck-Gesellschaft) at DESY (Hamburg) and EMBL ID14 at ESRF (Grenoble) within a Block Allocation Group (BAG Barcelona) as well as at ESRF BM14. It was financially supported by the ESRF and by Grant HPRI-CT-1999-00022 of the European Union Improving Human Potential Program awarded to EMBL Outstation Grenoble. Further help was granted by the European Union (Grants HPRI-CT-1999-00017 and EBFMGMCECT980134) to EMBL Outstation Hamburg. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1e9r, 1e9s, 1gki, 1gks, 1glt) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ Supported by Grants PB98-1631, PB98-1106, BIO2000-1659, and 2FD97-0518 from the Ministerio de Educación y Cultura and the Ministerio de Ciencia y Tecnología (Spain) and by Grant 1999SGR188 and help provided by the Centre de Referència en Biotecnologia (Generalitat de Catalunya). To whom correspondence may be addressed: Tel: 34-93-400-61-44/49; Fax: 34-93-204-59-04; E-mail: xgorri@ibmb.csic.es (for F. X. G.-R.) and nccrii@ibmb.csic.es (for M. Coll).

1 The abbreviations used are: ss, single-stranded; AAD, all-α domain; ds, double-stranded; NBS, NTP-binding site; T4CP, type IV secretion system coupling protein; NBD, NTP-binding domain; ADPNP, adenosine-5′-[(β,γ-imido)triphosphate]; GDPNP, guanosine-5′-[(β,γ-imido)triphosphate]; DEsy, Deutches Elektronensynchrotron; ESRF, European Synchrotron Radiation Facility; TMD, transmembrane domain; NsSP, N-ethylmaleimide-sensitive fusion protein; r.m.s.d., root mean square deviation.

2 F. X. Gomis-Ruth, G. Moncalián, F. de la Cruz, and M. Coll, unpublished results.

Received for publication, October 31, 2001, and in revised form, December 12, 2001. Published, JBC Papers in Press, December 17, 2001, DOI 10.1074/jbc.M110462200

F. Xavier Gomis-Ruth‡§, Gabriel Moncalián‡, Fernando de la Cruz‡, and Miquel Coll‡§

From the ‡Institut de Biologia Molecular de Barcelona, Consejo Superior de Investigaciones Científicas, c/Jordi Girona, 18–26, 08034 Barcelona and the Departamento de Biología Molecular, Universidad de Cantabria, c/Herrera Oria, s/n, 39001 Santander, Spain
Structure of Type IV Conjugative Coupling Protein TrwB

TABLE I
Data collection and processing

| Data set                  | TrwBAN70 (native) | TrwBAN70 + GDPNP complex | TrwBAN70 + ADPNP complex | TrwBAN70 + ADP/Mg\(^{2+}\) complex | TrwBAN70 + SO\(_4\)\(^{2-}\) complex |
|---------------------------|-------------------|--------------------------|--------------------------|------------------------------------|----------------------------------|
| Space group               | P2\(_1\)          | P3,21                    | P3,21                    | P3,21                              | P3,21                            |
| Cell constants (\(\AA\)^3) | 107.4, 153.4, 182.5, 94.2 | 151.6, 252.0            | 152.7, 252.1            | 151.1, 251.6                      | 151.3, 258.2                     |
| Wavelength (\(\AA\) in \(\AA\)) | 1.053             | 0.842                    | 1.053                    | 0.842                              |                                  |
| No. of measurements       | 728,622           | 433,599                  | 236,811                  | 360,835                            | 1,205,165                        |
| No. of unique reflections | 176,179           | 81,160                   | 67,095                   | 67,169                             | 133,904                          |
| Whole resolution range (\(\AA\)) | 50.0–2.50         | 50.0–2.80                | 52.0–2.91                | 51.3–3.00                          | 72.6–2.40                        |
| Completeness (%)          | 97.3              | 97.7                     | 89.5                     | 99.9                               | 99.9                             |
| \(R_{merge}\) (%)         | 5.4               | 9.3                      | 10.7                     | 11.8                               | 5.7                              |
| Average intensity (I/\(I_{0}\)) | 9.9               | 6.1                      | 5.3                      | 5.9                                | 9.0                              |
| Last resolution shell (\(\AA\)) | 2.64–2.50         | 2.95–2.80                | 3.07–2.91                | 3.16–3.00                          | 2.53–2.40                        |
| Completeness (%)          | 90.0              | 92.1                     | 66.0                     | 99.9                               | 99.6                             |
| \(R_{merge}\) (%)         | 16.9              | 45.3                     | 45.3                     | 41.8                               | 28.6                             |
| Average intensity (I/\(I_{0}\)) | 4.3               | 1.7                      | 1.6                      | 1.5                                | 2.6                              |
| \(B\)-factor (Wilson) (\(\AA^2\)) | 51.3              | 76.7                     | 61.0                     | 66.6                               | 46.0                             |
| Average multiplicity      | 4.1               | 5.3                      | 3.5                      | 5.4                                | 9.0                              |

\(^{a}\) Values for \(a = b = c\) (trigonal) or \(a, b, c, \) and \(\beta\) (monoclinic).

\(^{b}\) The high resolution data collection was limited by the diameter of the detector and the large cell constants.

\(R_{merge} = \frac{\sum_{i} \sum_{hkli} \sum_{j} I_{hkli} - \langle I_{hkli} \rangle}{\sum_{i} \sum_{hkli} \sum_{j} I_{hkli}} \times 100\), where \(I_{i}\) is the \(i\)-th measurement of reflection \(hkl\), including symmetry equivalent ones, and \((hkli)\) its mean intensity.

Conjugation (17), because it inhibits TrwC nic cleavage, probably by affecting the topology of the TrwC DNA target site (18).

The third protein encoded by R388 dtr region, TrwB, is a basic (pI = 10) integral inner-membrane protein of 507 residues (12, 19). It features a nucleosyl-triphosphate (NTP)-binding motif and a similar arrangement of possible transmembrane segments, and its proximal residues (two transmembrane helices and a small integral inner-membrane protein of 507 residues (12, 19). This fragment has been shown to bind ATP (19). Furthermore, TrwB interacts with both ss and double-stranded (ds) DNA with approximately the same affinity as supercoiled dsDNA (19). This DNA binding is nonspecific, in-
FIG. 1. TrwB structure. a, domain structure of representatives of the T4CP superfamily and related proteins. The first family, exemplified by TrwB, presently contains six members of conjugative plasmids that also contain conjugative helicases. A second family group comprises a broad set of conjugative plasmids (exemplified by TraG of RP4), as well as VirD4 of Ti plasmids involved in transfer-DNA transfer to plant cells. Finally, a further group of proteins is involved in bacterial cell division (exemplified by FtsK of E. coli) and in sporulation of Gram-positive bacteria (related to SpoIIIE of B. subtilis). Proteins within a family are 30% identical in amino acid sequence and share the domain organization. TrwB and VirD4 families (27) are more closely related between them than they are to the FtsK/SpoIIIE-proteins (30, 73). Transmembrane segments (TM) are represented in red and NBDs in blue with A and B walker motifs depicted. AADs present in the TrwB and VirD4 families are colored in green.

b, topology scheme displaying the regular secondary structure elements of TrwB N70. Green arrows represent α-strands (labeled α1 through α12), and rods stand for α-helices (labeled aA through aD). The N and C termini, the relative position of the central channel, the inner-membrane and the AAD, the NTP-binding site and the NBD are further displayed, as is the amino acid positions that make up each secondary structure element.

c, stereo ribbon plot of a TrwB N70 monomer. Helices are magenta, strands are cyan arrows, and coils and turns are yellow ropes. A bound GDPNP (substrate complex) molecule is shown as a green stick model to position the NBS.

d, surface representation of the complete TrwB hexamer, including the modeled transmembrane parts (lower stem).

Table II

| Data set             | TrwBΔN70 (native) | TrwBΔN70 + GDPNP (Subs. GDPNP complex) | TrwBΔN70 + ADPNP (Subs. ADPNP complex) | TrwBΔN70 + ADP/Mg2+/SO42− (Prod. ADP complex) | TrwBΔN70 + SO42− (Prod. P complex) |
|----------------------|-------------------|---------------------------------------|----------------------------------------|-----------------------------------------------|-----------------------------------|
| Resolution range (Å) | 50–2.5            | 50–2.8                                | 50–2.9                                 | 50–3.0                                        | 50–2.4                           |
| No. of reflections   | 176,026           | 81,037                                | 65,903                                 | 67,164                                        | 133,805                           |
| Crystallographic R-factor (free R-factor) | 20.9 (26.7) | 21.4 (25.8) | 21.9 (26.5) | 21.0 (23.2) | 21.0 (24.8) |
| No. of protein atoms (a.u.) | 41,537 | 19,624 | 19,525 | 19,548 | 19,693 |
| No. of solvent molecules | 1,559 | 696 | 215 | 452 | 965 |
| No. of anions/cations | 1 (Cl−) | 1 (Cl−) | 5 (Mg2+) | 15 (SO42−) |
| Other molecules      | 1 (HEPES), 7 (GDPNP) | 6 (ADPNP) | 2 (HEPES), 7 (ADP) | &compound; |
| r.m.s.d. from target values (excluding inhibitors/products) | 0.008 | 0.009 | 0.009 | 0.011 | 0.008 |
| Bonds (Å)            | 1.39              | 1.42                                  | 1.42                                   | 1.65                                          | 1.39                             |
| Angles (°)           | 2.09              | 2.63                                  | 2.33                                   | 1.83                                          |                                |
| Bonded B-factors (Å²) | 40.8             | 56.0                                  | 64.3                                   | 49.6                                          | 43.0                             |

* R-factor = (Σhk|Fhk|− k|Fhk|Σhk|Fhk|) × 100, with Fobs and Fcalc as the observed and calculated structure factor amplitudes; free R-factor same for a test set of 7% reflections (>500) not used during refinement (until the penultimate cycle). Data sets of same space group share the test set in the common resolution range.
anion, and the ADP complex revealed an ion of $\text{Mg}^{2+}$ in each NBS (except in the NBS of chain F), besides two HEPES molecules at the molecular surface. Model building and computational refinement were assessed as mentioned. The final refinement statistics are presented in Table II.

The polypeptide chains of the different monomers within each hexamer are defined from their N termini at residues Asn$^{72}$ to Glu$^{78}$ until the C-terminal amino acids Glu$^{504}$ to Ile$^{507}$. Residues Arg$^{437}$-Thr$^{444}$ to Thr$^{450}$-Glu$^{455}$ from a surface-exposed region (strands $\beta 9$ and $\beta 10$) probably interacting with the inner part of the channel-off transmembrane region, are flexible and have not been traced. All residues, excepting His$^{225}$ ($\Phi/\Psi = -70/1$) of each polypeptide chain in all structures, are placed in allowed regions of the Ramachandran plot. This residue is, however, clearly defined by electron density and located at the beginning of strand $\beta 2$, because its (and the preceding residue's) carbonyl oxygen is involved in $\beta$ sheet interactions with amide nitrogen atoms of residues Arg$^{228}$ and Ser$^{229}$.

**RESULTS AND DISCUSSION**

**The Monomer Structure—**The TrwB$\text{N70}$ monomer consists of two domains: a nucleotide-binding domain (NBD) attached to the inner-membrane and a membrane-distal all-α domain (AAD; Fig. 1, b and c). This NBD displays an $\alpha/\beta$ P-loop containing NTP-hydrolase core structure. This domain runs from the membrane-proximal N terminus, anchor point of the excised transmembrane domain (TMD), to residue Lys$^{183}$ and continues from position Asp$^{298}$ to the surface-located C terminus. It is composed of a central, highly twisted (~90°) nine-stranded pleated $\beta$-sheet of mixed parallel/antiparallel topology (strands $\beta 2$–$\beta 8$, $\beta 11$, and $\beta 12$; Fig. 1b), flanked on its concave side by four helices (termed $\alpha$ through $\alpha D$) and on the other side by seven helices (ol through $\alpha R$). The convex side faces an interior channel. On the membrane-proximal edge of the central $\beta$-sheet, a small three-stranded antiparallel sheet (strands $\beta 1$, $\beta 9$, and $\beta 10$; see Fig. 1b) is inserted, almost perpendicular. Strands $\beta 9$ and $\beta 10$ are only partially defined by electron density, probably due to the absence of interactions with the (excised) TMD preceding strand $\beta 1$.

On top of NBD, the smaller AAD is inserted between strand $\beta 4$ and helix $\alpha L$, comprising residues Gly$^{184}$-Gly$^{297}$ (Fig. 1, b and c). This domain contains seven helices: a $\alpha_{10}$ helix (helix $\alpha E$), followed by a two-helix hairpin (helices $\alpha F$ and $\alpha G$) and by a second, curved, helical hairpin segment (helices $\alpha H$ through $\alpha K$). The latter mainly protects and interacts with the central helix $\alpha G$. The interaction between NBD and AAD is mainly of hydrophilic nature and encompasses a common surface of 2314 Å$^2$ (about 36% of the total surface of AAD). It comprises mainly segments Leu$^{175}$-Tyr$^{178}$, Thr$^{182}$-Lys$^{183}$, Asp$^{298}$-Glu$^{306}$, Ala$^{324}$-Ser$^{342}$, Glu$^{362}$-Ala$^{365}$, and Asp$^{393}$-Val$^{394}$ of NBD and Gly$^{184}$-Tyr$^{185}$, Arg$^{199}$-Lys$^{209}$, and Asp$^{298}$-Gly$^{297}$ of AAD. We analyzed the 56 close contacts between domains and observed 18 hydrogen bonds, 4 salt bridges, and 6 van der Waals interactions.

**TruB Forms Hexamers—**As derived from both crystal forms obtained (see “Experimental Procedures”), six TrwB$\text{N70}$ protomers intimately associate employing a local 6-fold axis (see Figs. 1d and 2b). One hexamer is observed in the asymmetric unit of the trigonal crystals (monomers A to G), and two are present in the monoclinic system (monomers A to G and H to M, respectively). This agrees with the native protein migrating as a hexamer in gel filtration column chromatography (data not shown). The protomers are practically indistinguishable from each other within a hexamer. The hexamer, with overall dimensions of 110-Å diameter and 90-Å height (see Fig. 2a), displays an orange-like shape, somewhat flattened at both poles along the axis. A central channel runs from the cytosolic pole (made up by the AADs) to the membrane pole (formed by...
the NBDs), ending at the (tentatively modeled) transmembrane pore shaped by the 12 transmembrane helices, rendering an overall mushroom-like structure (Fig. 1d). This channel is delimited at its entrance mainly by gatekeeping residues Asn271 of loop α10βJ from each subunit and is restricted to a diameter of ~7–8 Å. This is the closest point of the channel that, at its exit to the TMD, has an opening of ~22 Å. The interaction between TrwBΔN70 monomers is mainly of hydrophilic nature with the central twisted β-sheets arranged almost in parallel (C-edge to N-edge). It is characterized by a common surface of 4588 Å² as calculated for monomers A and B (about 25% of the total surface of a monomer). This contact surface is lined up mainly by segments Gly80-Phe83, Ala131-Thr134, Tyr178-Leu209-Trp216, Phe243-Phe251, Asn271-Asp286, Arg318-Ala324, Ser360-Gln366-Val397, Gly415-Lys421, Arg437-Arg439, Glu455-Asp472 (monomer A) and Leu48-Thr91, Arg124-Leu127, Tyr195, Arg225, Thr261-Ala276, Thr338-Asp342, Asp369-Lys373, Lys398, Leu410, Asp455-Arg458, Ala453-Arg458, Gly478-Ile485 (monomer B). The main interaction is made between vicinal NBDs (9100 Å²/H9251 and both protease HsIU-HsIV (PDB 1e94 (58)). AAA ATPase p97 is involved in homo-meric membrane fusion and ATP-dependent translocation (PDB 1d2n (56)). AAA ATPases, an inner-membrane-associated part of the bacterial F1-ATPase (see Fig. 3, a), a membrane protein (membrane-associated) like our TrwB reveals a stringent structural similarity of its NBD with the equivalent part of RNA and DNA helicases, molecular motors involved in DNA metabolism that use the energy from NTP hydrolysis for nucleic acid unwinding or strand separation/translocation (45). RecA (Protein Data Bank access code (PDB) 2reb; see Fig. 3a), essential for genetic recombination and repair, is the structural prototype for these enzymes that can adopt varying (functional) oligomerization states (46–48). The family further includes toroidal hexameric DNA ring helicases, like T7 phage gene 4 replicative helicase/primase (PDB 1e0j (49) and 1c0r (50)) and RepA (PDB 1g8y (51)), and monomeric enzymes like 3D oligomerization states (46–48). The family further includes toroidal hexameric DNA ring helicases, like T7 phage gene 4 replicative helicase/primase (PDB 1e0j (49) and 1c0r (50)) and RepA (PDB 1g8y (51)), and monomeric enzymes like 3D-ADPNP-bound aβ3, F1-ATPase heterohexamer (NBD in cyan, C-terminal AAD containing six α-helices in blue, and N-terminal β-barrel domain in dark blue). The AADs are arranged differently in RecA and F1-ATPase than in TrwB. In F1-ATPase, the crown-forming β-barrel domain occupies the place of AAD in TrwB. P-loop Lys192 of TrwB corresponds to Lys345 in F1-ATPase β-subunites and to Lys72 in RecA.

Striking structural similarity is further found between TrwB and both α and β subunits of F1-ATPase, part of the membrane-associated F,F'-ATPase complex responsible for energy conversion through axial rotation movement in mitochondria, chloroplasts, and bacteria (see Fig. 3b; PDB 1hmnf (59)). Finally, the recently reported six-clawed, grapple-shaped structure of homohexameric Helicobacter pylori Cag525/HP0525 traffic ATPase, an inner-membrane-associated part of the bacterial type IV secretion system involved in pathogenic protein CagA export, also reveals structural similarity in the NBD core (60).

Nevertheless, on examining the hexameric toroidal quaternary structures of the mentioned protein families (Fig. 2), helicases, AAA ATPases, and Cag525 appear more flat-topped than TrwB, with much less interaction surface between the constituting proteins. This weaker interaction is required in helicases to permit interchanging aggregation stages and helicoidal protein filament formation (50, 61, 62). The oligomeric structure of TrwB displays an almost spherical shape, which, together with the overall hexamer dimensions, is much more reminiscent of the F1-ATPase aβ3 heterohexamer (Fig. 2, a and b), a membrane protein (membrane-associated) like our bacterial protein. Furthermore, both particles share the presence of a crown made up by β-strands, although placed on opposite faces of the central NBD (see Fig. 3b); i.e. cytosolic in ATPase, membrane-proximal in TrwBΔN70.

Is the AAD a DNA-binding Domain?—Several representatives of the helicase/ATPase-like proteins display, besides a NBD, also an AAD, as shown for RecA, AAA ATPases, helicases, and the α and β subunits of F1-ATPase (see Fig. 3a, 3, and b), among others, although arranged in distinct ways with respect to their NBDs. However, detailed topological inspections reveal that TrwB AAD bears only significant structural similarity with N-terminal domain 1 of the site-specific recombinase, XerD, of the λ integrate family (see Fig. 4e; PDB access code 1a0p (63)), with an r.m.s.d. of 3.4 Å over a common stretch of 75 residues (out of 114 in TrwB). In XerD, this domain 1

\[ \text{Structure of Type IV Conjugative Coupling Protein TrwB}\]
displays four helices arranged such that there are two parallel helix hairpins arranged at 90° to each other. It has been proposed to contribute to the shape of the DNA-binding site with two of its helices, equivalent to helices $\alpha G$ and $\alpha J$ in TrwB, being considered for interaction with the major groove of the inner part of the recombinase specific site. This domain is positioned over the DNA binding region blocking the access to it, so that a large conformational rearrangement has to occur for target DNA binding. Such a rearrangement, which could also happen in TrwB, could be implicated in a putative working mechanism.

Xer site-specific recombination has been associated with FtsK activity (64), a protein that displays functional and structural similarities with TrwB and other T4CPs (see above).

TrwB AAD displays further topological similarity in a 40-residue segment encompassing three $\alpha$-helices with the recently reported solution structure of the 56-residue DNA-binding domain of TraM, the first of a (putative) endogenous component of the relaxosome encoded by the $dtr$ region ((65); see Fig. 4, b and c). This protein enhances relaxase activity in plasmid R1 and is capable of binding DNA (66). TraM specifically interacts with the TrwB orthologue TraD of plasmid R1, suggesting that TraM links the relaxosome with the DNA transfer apparatus (67). Plasmid R388 lacks such a TraM orthologue. Therefore, an appealing hypothesis is that TrwB uses a DNA-binding domain similar to that of TraM, but imbedded in its structure, to directly recruit the relaxosome. Unfortunately, experimental proof of sequence-specific interactions between TrwB and oriT DNA is still unavailable.

**NTP-binding Site Location**—We managed to obtain the protein structure in four states, non-ligated (monoclinic crystal form), in complexes with the NTP analogues adenosine- and guanosine-5′-(β,γ-imido)triphosphate (ADPNP and GDPNP; trigonal crystal form), with ADP in the presence of Mg$^{2+}$ ions and with a sulfate anion (both trigonal crystal form; see Figs. 5–7). By analogy to PcrA (52, 68), we suggest that the protein is trapped in states that emulate its apo form, two equivalent substrate complexes and two distinct, and possibly alternative or mutually excluding, product complexes (ADP and phosphate (P) complexes).

As predictable from the topology of the $\beta$-sheet and confirmed by the ability of TrwB to bind a fluorescent ATP analogue (19), the active-site crevice (in this case the NBS) is located at the C-terminal edge of the sheet and is shaped by loops $\beta 2eC$ and $\beta 6oN$ (see Fig. 1b). These two segments feature helicase superfamily motif I (Walker box A; Gly$^{130}$-Ala$^{131}$-Thr$^{132}$-Gly$^{133}$-Thr$^{134}$-Gly$^{135}$-Lys$^{136}$-Ser$^{137}$) and motif II (Walker box B or DEXX box; Trp$^{352}$-Leu$^{353}$-Phe$^{354}$-Ile$^{355}$-Asp$^{356}$-)

---

**Fig. 4. Structural similarity with XerD and TraM.**

*a*, superimposition of TrwB AAD (magenta) with the N-terminal domain 1 of XerD (blue); *b*, superimposition of the DNA-binding domain of TraM (yellow) with region Tyr$^{255}$-Gly$^{256}$- of TrwB AAD (blue); *c*, partial cytosolic view of a TrwB particle (corresponding to the region shown in the upper left corner) with the region (green) showing homology to the TraM DNA-binding domain (magenta) pointing to the exterior surface. The Connolly surface showing the electrostatic potential of the protein without the compared segments is displayed, ranging from $-15$ kT/e (red) to $-15$ kT/e (blue).
Glu357-Leu358-Ala359) described for NTP-binding proteins (20, 21), so as a putative Asp box (Met154-Val155-Ile156-Val157-Asp158-Pro159-Asn160-Gly161). If compared with other NTP-binding pockets of RecA-like family members, including F1-ATPase or PcrA (47, 49, 52, 59), the NBS is not very deep and readily accessible to bulk solvent. Considering the whole particle, the six NBSs form a belt around the hexamer at half height of the particle, and they are located 32 Å apart on superficial cavities at interfaces between vicinal protomers. Therefore, residues coming from two vicinal subunits contribute to feature each NBS. There are no significant differences between the NBSs, contrary to what has been described for helicases or F1-ATPases (49, 59); therefore, they can be considered as equivalent.

Active Site Architecture: A Model for the Substrate Complex—The presently described structures of TrwBΔN70, non-liganded and with ADPNP and GDPNP, make possible the definition of the residues involved in the binding and catalysis of the substrate. No unambiguous Mg2+ ion could be assigned in the electron density maps of either complex despite it having been added to the crystallization mixture. Both complexes (see Figs. 5–7) are very similar (r.m.s.d. for all common Ca atoms is 0.34 Å), thus only the GDPNP complex structure will be discussed in detail. GDPNP binding provokes only minor rearrangements when compared with the apo form. The overall r.m.s.d. between A monomers is 0.25 Å; no significant deviation is observed in the overall hexamer structure either (see Fig. 7b).

The main interaction of the bound nucleotide analogue is made with the P-loop, located between strand β2 and helix αC. Protein–substrate interactions involve the main-chain amide groups from Gly135 to Val138 of the beginning of helix αC accommodating the substrate β-glycerate oxygen atoms. This phosphate group interacts further with the main-chain nitrogen of Gly133 and P with Thr132 O. The substrate base moiety resides on a small hydrophobic pillow made up by the side chains of Ile492, Val138, and Leu473. The less well defined base (located on the molecular surface, see Figs. 5b, 5c, and 6a) may further interact with Gln494 (Figs. 5–7). Contacts between the substrate and a vicinal protomer include the side chains of Arg375 (with P oxygen atoms) and Arg124 (with the nucleotide sugar moiety O3* atom and P oxygen atoms). Therefore, the side chains of both Arg124 and Arg375 (the latter putatively imbedded in a helicase superfamily II motif VI (20)) both rotate on a complex formation to...
accommodate and bind the substrate Pγ group (see below). These two residues do not have their net charges compensated by surrounding acidic residue side chains in the non-ligated structure (although they are both partially accessible to bulk solvent).

Interestingly, PcrA displays also a pair of basic residues in charge of liganding Pγ oxygen atoms, Arg^{610} (from a helicase superfamily motif VI (20, 52)) and Arg^{207} (of motif IV). In the structures of both TrwB and PcrA, the NBSs are located at the interface between two domains, duplicated in PcrA (1A and 2A) and corresponding to two vicinal protomers in TrwB. Similarly, the arginine residues are provided by the vicinal domain (2A in PcrA, highly homologous to the P-loop containing domain 1A and, putatively, a product of a gene duplication) or protomer (TrwB). Arg^{610} of PcrA and Arg^{207} of TrwB are localized at the end of a topologically equivalent α helix (αN in TrwB), with their Ca atoms just 2.3 Å apart. Arg^{207} (PcrA) and Arg^{124} (TrwB) are somewhat farther away (7 Å for their Ca atoms), but their side chains come closer together and are localized on the same side of the NBS. Cag525 also displays a cluster of arginine residues contiguous to the ADP-binding site, also made up by two vicinal subunits, at whose interface the NBS is located (60). Furthermore, basic residues in the proximity of the phosphate oxygen atoms have been associated with transition state stabilization in G proteins and F_{1}-ATPase (59, 69, 70) and, putatively a product of gene duplication) or protomer

Significant changes were observed when comparing the substrate complex with the product P complex, where a strongly bound sulfate anion (100% occupation; see Fig. 5f) may be interpreted as a leftover phosphate. This complex would constitute an alternative to the ADP complex, because during the hydrolytic mechanism one of the two products, ADP or phosphate, must leave the NBS first. When molecules A of each complex are superimposed, some strongly deviating regions appear despite the close overall similarity (r.m.s.d. of 0.47 Å; note negligible differences in the overall hexamer structure). In particular and besides the N- and C-terminal regions, segments Gly^{130}-Arg^{141} (encompassing the P-loop; r.m.s.d. of 1.1 Å), Gly^{384}-Gln^{390} (1.0 Å), and Gly^{414}-Gly^{415} (1.6 Å) strongly diverge, even for their main chains (Fig. 6c). The sulfate anion is located 2 Å away from the Pγ position of the substrate (Fig. 7, a and c). Vicinal monomer residues Arg^{124} and Arg^{375} move away, when compared with the GDPNP complex structure. The P-loop suffers a major rearrangement from Gly^{130} to Arg^{141} as does the surface-located C-terminal strand between Leu^{190} and Phe^{195}. The side chain of Ser^{137} rotates, so that its Oγ atom points in the direction of the sulfate. Most importantly, the side chain of DEXX box Gln^{386} rotates around its Ca-Cβ bond. This is accompanied by a main-chain and side-chain rearrangement of loop β7αO. The region of TrwB mainly involved in these conformational changes and facing the interior channel displays a striking similarity in topology, but not in sequence, including the same regular secondary structure elements, with a structural feature described for AAA proteins, the N-terminal part of the second region of homology (72), or AAA minimum consensus region (57). This –20-residue region mainly distinguishes AAA ATPases from other Walker-type ATPases (53, 72, 74) and is surface-accessible from the central channel, as ob-
served in NSF (56). This segment is proposed to play a role in ATP hydrolysis but not in binding (72). In TrwB this region runs from the end of strand $\beta 7$ to the beginning of helix $\alpha P$ (see Fig. 1b). Although TrwB does not show any of the conserved residues, it is, however, noteworthy that this region appears to be involved in transmitting a movement from the exterior surface to the interior channel, activated by nucleotide-binding/hydrolysis. These changes could play a role in moving the ssDNA through the central channel during conjugation.

Considerations About the Hydrolytic Mechanism—ATP hydrolysis in DEXX box and other NTP-hydrolases implies a covalent bond formation between $P\gamma$ (or $P\beta$) and an attacking solvent molecule, giving rise to a trigonal bipyramidal, negatively charged transition state species, and subsequent breakage of the $O3\gamma$--$P\gamma$ (or $O3\beta$--$P\beta$) scissile bond. This intermediate may be stabilized by a $Mg^{2+}$ cation, coordinated to the $P\gamma$ and/or $P\beta$ oxygen atoms of the substrate, and possibly the main-chain amide groups of P-loop motif I (20, 21, 75). Further basic residues in the pocket may act additionally as $P\beta$ sensors (see above). Two acidic residues may act as a general base, polarizing the attacking solvent molecule, and may coordinate the catalytic cation, respectively. Two acidic residues may act as a general base, polarizing the attacking solvent molecule, and may coordinate the catalytic cation, respectively (56, 76–78). These two residues are possibly Glu$^{96}$ and Asp$^{144}$, the latter of which if from the DEXX box
motif I (20), in the archetypal RecA enzyme (46, 47). This motif is also essential in monomeric eIF4A RNA helicase ATPase domain, where Asp<sup>169</sup> of DEAD box (Asp<sup>169</sup>Asp<sup>172</sup>) is shown to bind the Mg<sup>2+</sup> cation and proposed to act also as the general base (76), and in the structurally related monomeric DNA helicases Rep and PcrA (here the DEXX box encompasses Asp<sup>233</sup> and Glu<sup>224</sup>). In these latter cases, the aspartate residues are binding the cation and the glutamate residues are suggested to function as the general base (52, 68, 79), although spatially at a different position from that of Glu<sup>96</sup> in RecA.

Our protein displays a DEXX box motif II (including Asp<sup>356</sup> and Glu<sup>357</sup>), which is topologically equivalent to this motif in PcrA. Both residues, however, do not point their side chains toward the bound GDPNP P<sub>γ</sub> group moiety. A rotation around the Co-Cβ and Cβ-Cγ bonds of Asp<sup>356</sup> would position its Oδ2 atom 6 Å away from a Pβ oxygen. The same distance (~6 Å) holds for the immediately vicinal Glu<sup>357</sup>, involved, prior to NTP-binding, in a salt bridge with the tip of Lys<sup>136</sup> of the bromide, Ana González, Carazo, F., and Coll, M. (2001) Nature 409, 637–641.

Gomis-Ruth, F. X., and Coll, M. (2001) Acta Crystalogr. Sect. D Biol. Crystalogr. 57, 800–806.

Leslie, A. G. W. (1991) in Crystallographic Computing V (Moras, D., Podjarny, A. D., and Thierry, J. C., eds) pp. 27–38, Oxford University Press, Oxford.

Evans, P. (1991) in Data Collection and Processing: Proceedings of the CCP4 Study Weekend, 29–30 January 1993 (Sawyer, L., Isaacs, N., and Bailey, S., eds) pp. 114–122, SERC Daresbury Laboratory, Warrington.

Navaza, J. (1994) Acta Crystalogr. Sect. A 50, 157–163.

Roussel, A., and Cambillau, C. (1989) in Silicon Graphics Geometry Partners Directory, pp. 77–78, Silicon Graphics, Mountain View, CA.

Brünger, 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., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystalogr. Sect. D Biol. Crystalogr. 54, 905–921.

Collaborative Computational Project Number 4 (1994) Acta Crystalogr. Sect. D Biol. Crystalogr. 50, 760–763.

Nicholls, A., Bharadwaj, R., and Henig, B. (1993) Biophys. J. 64, A16–A166.

Ecclestone, R. M. (1997) J. Mol. Graph. 15, 133–138.

Merritt, E. A., and Murphy, M. E. P. (1994) Acta Crystalogr. Sect. D Biol. Crystalogr. 50, 869–872.

Laskowsky, R. A., McArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystalogr. 26, 434–435.

Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) Prot. Eng. 8, 127–134.

Waksman, G., Lank, E., and Caruso, J.-M. (2000) Nat. Struct. Biol. 7, 20–22.

Story, R. M., Weber, I. T., and Steitz, T. A. (1992) Nature 355, 318–325.

Story, R. M., and Steitz, T. A. (1992) Nature 355, 374–376.

Yu, X., and Egelman, E. H. (1997) Nat. Struct. Biol. 4, 101–104.

Singleton, M. R., Sawaya, M. R., Ellenberger, T., and Wigley, D. B. (2000) Cell 101, 689–690.

Sawaya, M. R., Guo, S., Tabor, S., Richardson, C. C., and Ellenberger, T. (1999) Cell 99, 167–171.

Niedenzu, T., Roleke, D., Bains, G., Scherzinger, E., and Saenger, W. (1997) Mol. Biol. 306, 479–487.

Velankar, S. S., Soutlanas, P., Dillingham, M. S., Subramanya, H. S., and Wrigley, D. B. (1999) Cell 97, 75–84.

Dalal, S., and Hanson, P. F. (1999) Cell 104, 5–8.

Vale, R. D. (2000) Cell 100, F13–F19.

Guenter, B., Orust, R., Sali, A., O’Donnell, M., and Kuriyan, J. (1997) Cell 91, 335–345.

Lenzen, C. U., Steinemann, D., and Huber, S. W., and Weiss, W. I. (1998) Cell 94, 525–535.

Zhang, X., Shaw, A., Bates, P. A., Newman, R. H., Gowen, B., Orlova, E., Gorman, M. A., Kondo, H., Dokunova, F., Lilly, J., Leonard, G., Meyer, H., van Heel, M., and Freemont, P. S. (2000) Mol. Cell 6, 1473–1484.

Bochler, M., Hartmann, C., Song, H. K., Bourenkov, G. P., Bartonik, H. D., and Huber, R. (2000) Nature 403, 800–805.

Acharya, J. P., Lesie, A. G. W., Lutter, R., and Walker, J. E. (1994) Nature 370, 621–628.

Yee, H.-J., Savvides, S. N., Herr, A. B., Lanka, E., and Waksman, G. (2000) Mol. Cell 6, 1461–1472.

Soutlanas, P., and Wigley, D. B. (2000) Curr. Opin. Struct. Biol. 10, 124–128.

Yu, X., Shibata, T., and Egelman, E. H. (1998) J. Mol. Biol. 283, 985–992.

Subramanya, H. S., Arciszewska, L. K., Baker, E. A., Bird, L. E., Sherratt, D. J., and Wigley, D. B. (1997) EMBO J. 16, 5178–5187.

Recchioni, G. D., Arroyo, M. D., Wolf, D., Blakely, G., and Sherratt, D. J. (1999) EMBO J. 18, 5724–5734.

Stockner, T., Plucagiu, C., Keramini, G., Higenauer, G., Berml, W., Prittulla, H., and Sper, H. (2000) Biochemistry 39, 3707–3737.

Kapilwieser, G., Schwaab, M., Higenauer, G., Keramini, G., and Zechner, E. (1998) J. Mol. Biol. 275, 81–94.

Disque-Kochum, C., and DischeiKelam, B. (1997) J. Biol. Chem. 179, 355–361.

Subramanya, H. S., Bird, L. E., Brannigan, J. A., and Wigley, D. B. (1996) Nature 384, 379–383.

Coleman, D. E., Berghuis, A. M., Lee, E., Linder, M. E., Gilman, A. G., and Sprang, S. R. (1994) Science 265, 1405–1408.
Structure of Type IV Conjugative Coupling Protein TrwB

70. Sondek, J., Lambright, D. G., Noel, J. P., Hamm, H. E., and Sigler, P. B. (1994) *Nature* **372**, 276–279
71. Karata, K., Inagawa, T., Wilkinson, A. J., Tatsuta, T., and Ogura, T. (1999) *J. Biol. Chem.* **274**, 26225–26232
72. Karata, K., Verma, C. S., Wilkinson, A. J., and Ogura, T. (2001) *Mol. Microbiol.* **39**, 890–893
73. Begg, K. J., Dewar, S. J., and Donachie, W. D. (1995) *J. Bacteriol.* **177**, 6211–6222
74. Neuwald, A. F., Aravind, L., Spouge, J. L., and Koonin, E. V. (1999) *Genome Res.* **9**, 27–43
75. Maegley, K. A., Admiraal, S. J., and Herschlag, D. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 8160–8166
76. Benz, J., Trochael, H., and Baumann, U. (1999) *Structure* **7**, 671–679
77. Smith, C., and Raymont, I. (1996) *Biophys. J.* **70**, 1590–1602
78. Bird, L. E., Subramanya, H. S., and Wigley, D. B. (1998) *Curr. Opin. Struct. Biol.* **8**, 14–18
79. Korolev, S., Hsieh, J., Gauss, G. H., Lohman, T. M., and Waksman, G. (1997) *Cell* **90**, 635–647
80. Schweins, T., Geyer, M., Scheffzek, K., Warshel, A., Kalbitzer, H. R., and Wittinghofer, A. (1995) *Nat. Struct. Biol.* **2**, 36–44
Conjugative Plasmid Protein TrwB, an Integral Membrane Type IV Secretion System Coupling Protein: DETAILED STRUCTURAL FEATURES AND MAPPING OF THE ACTIVE SITE CLEFT
F. Xavier Gomis-Rüth, Gabriel Moncalián, Fernando de la Cruz and Miquel Coll

J. Biol. Chem. 2002, 277:7556-7566.
doi: 10.1074/jbc.M110462200 originally published online December 17, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M110462200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 77 references, 18 of which can be accessed free at http://www.jbc.org/content/277/9/7556.full.html#ref-list-1