

RecOR complex including RecR N-N dimer and RecO monomer displays a high affinity for ssDNA

Qun Tang¹, Pu Gao¹, Yan-Ping Liu¹, Ang Gao^{1,2}, Xiao-Min An¹, Shun Liu^{1,2},
Xiao-Xue Yan^{1,*} and Dong-Cai Liang^{1,*}

¹National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101 and ²Graduate University of Chinese Academy of Sciences, Beijing 100039, China

Received February 17, 2012; Revised August 30, 2012; Accepted September 3, 2012

ABSTRACT

RecR is an important recombination mediator protein in the RecFOR pathway. RecR together with RecO and RecF facilitates RecA nucleoprotein filament formation and homologous pairing. Structural and biochemical studies of *Thermoanaerobacter tengcongensis* RecR (TTERecR) and its series mutants revealed that TTERecR uses the N-N dimer as a basic functional unit to interact with TTERecO monomer. Two TTERecR N-N dimers form a ring-shaped tetramer via an interaction between their C-terminal regions. The tetramer is a result of crystallization only. Hydrophobic interactions between the entire helix-hairpin-helix domains within the N-terminal regions of two TTERecR monomers are necessary for formation of a RecR functional N-N dimer. The TTERecR N-N dimer conformation also affects formation of a hydrophobic patch, which creates a binding site for TTERecO in the TTERecR Toprim domain. In addition, we demonstrate that TTERecR does not bind single-stranded DNA (ssDNA) and binds double-stranded DNA very weakly, whereas TTERecOR complex can stably bind DNA, with a higher affinity for ssDNA than double-stranded DNA. Based on these results, we propose an interaction model for the RecOR:ssDNA complex.

INTRODUCTION

The repair of damaged DNA is required to maintain the genomic integrity of living organisms, and the non-homologous end-joining and recombination DNA repair mechanisms occur in all organisms (1,2). Two recombination DNA repair pathways exist in bacteria: the RecBCD pathway and the RecFOR pathway (3–5). The RecBCD pathway is responsible for repairing

double-stranded DNA (dsDNA) breaks, whereas the RecFOR pathway is mainly used for single-stranded DNA (ssDNA) gap repair (6–8). Recently, it was shown that the RecF pathway, which has many parallels with recombinational repair in eukaryotes, is important for recombinational repair of DNA breaks and gaps using a reconstituted system (9), and some work suggests that the functions of the proteins involved in the RecFOR pathway are conserved from bacteria to humans (10–14). RecR is the most conserved protein in the RecFOR pathway; RecR binds to RecO, and RecO interacts with SSBs (15–17), then the complex facilitates RecA filament formation on ssDNA-binding protein (SSB)-coated ssDNA (5,16,18–21). When associated with RecF, the RecOR complex is implicated in the recognition of dsDNA–ssDNA junctions (9,22,23).

RecR is a zinc metalloprotein (24), which consists of a helix-hairpin-helix (HhH) motif, zinc finger motif, Toprim domain and C-terminal hydrophobic region. At present, the HhH motif is thought to be required for binding DNA (21,25,26), the zinc finger motif plays a structural or DNA-binding role (21,24) and the RecR Toprim domain is the binding site for both RecO and RecF (23). RecR proteins from different species have different polymerization states in solution (18,21,23,27) and varying abilities to bind DNA (16,21,27,28). The crystal structure of *Deinococcus radiodurans* RecR (drRecR) indicated that a ring-shaped tetramer is the functional unit for DNA binding (21). From the NMR model of *Thermus thermophilus* RecR (ttRecR), Honda *et al.* (23) suggested that ttRecR forms a dimer in solution by N-terminal interactions. Based on the ring-shaped tetramer structure of drRecR and the drRecR: drRecO molecular ratio of 4:2 in solution, Timmins *et al.* (29) obtained a structural model of the drRecOR complex at a resolution of 3.8 Å and speculated a model of RecOR complex with dsDNA. Small-angle X-ray-scattering data have indicated that ttRecFR forms a globular structure consisting of four ttRecR and two ttRecF monomers, and the modularized interaction mechanisms have been speculated

*To whom correspondence should be addressed. Tel: +86 010 64888509; Fax: +86 010 64889867; Email: snow@moon.ibp.ac.cn
Correspondence may also be addressed to Dong-Cai Liang, Tel: +86 010 64888509; Fax: +86 010 64889867; Email: dcliang@sun5.ibp.ac.cn

for the RecOR complex and RecFR complex with DNA (23,30).

The RecR, RecF and RecO proteins have been intensively researched. The individual crystal structures of RecR, RecO and RecF have been solved, and the regions that are possibly involved in protein–protein and protein–DNA interactions have been identified (17,21,31–34). However, in contrast to the RecBCD pathway, the mechanism of recombination repair mediated by the RecFOR pathway is poorly understood. Establishing the assembly pattern of RecR, RecO and RecF with DNA will promote further study of the RecFOR pathway. In this study, we investigated the structures and biological function of *Thermoanaerobacter tengcongensis* RecR (TTERecR) and its series mutants. Based on these experiments, we proposed a novel interaction model for RecOR:ssDNA complex, which would provide further understanding of the mechanism of RecFOR repair pathway.

MATERIALS AND METHODS

Protein purification

The *recR* (TTE0041; GenBank: AAM23354.1), *recF* (TTE0004; GenBank: AAM23321.1) and *recO* (TTE0976; GenBank: AAM24231.1) genes were amplified from *T. tengcongensis* MB4 genomic DNA by PCR and individually cloned into the pETDuet plasmid (Novagen) for expression with an N-terminal hexahistidine tag. TTERecR site-specific mutants and deletion mutants were generated from the TTE-recR-pETDuet plasmid. All sequences were confirmed by sequencing. The proteins were over-expressed in *Escherichia coli* BL21 (DE3). The cells were cultured in LB media containing 100 mg/l ampicillin at 37°C for 8 h and induced with 0.4 mM isopropyl β -D-thiogalactoside for 10 h at 28°C. The recombinant proteins were purified by sonication and two-step column chromatography using a Ni-affinity column and Superdex200 gel-filtration column (GE Healthcare). RecR and the RecR mutants were concentrated to 20 mg/ml by ultrafiltration in 10 mM Tris, pH 7.5, 200 mM NaCl, and RecO was concentrated to 20 mg/ml by ultrafiltration in 10 mM Tris, pH 7.5, 1 M NaCl. All proteins were stored at –80°C.

Crystallization and structure determination

The crystals of TTERecR and its mutants were obtained at 20°C for a few days by the hanging drop vapor diffusion technique. TTERecR was crystallized in buffer containing 6% (w/v) PEG3350, 200 mM lithium citric tribasic tetrahydrate, 100 mM MES, pH 6.8. The TTERecR₁₆₋₁₉₆ deletion mutant was crystallized in 14% (w/v) PEG3350, 150 mM ammonium sulfate, 100 mM MES, pH 5.15. The site-specific mutant TTERecR_{K21G} was crystallized in 8% (w/v) PEG4000, 200 mM ammonium sulfate, 10% (v/v) 2-propanol, 100 mM HEPES sodium salt, pH 7.5. The TTERecR₁₋₁₆₆ deletion mutant was crystallized in 2.1 M sodium formate, 100 mM Bis-Tris propane, pH 7.0. The TTERecR₁₋₁₈₀ deletion mutant was crystallized in 10% (w/v) PEG6000, 2.0 M NaCl, pH 5.6. The crystals were

flash frozen by immersion in a reservoir of 15–25% glycerol followed by transferring to liquid nitrogen. The crystals were maintained at 100 K during X-ray diffraction data collection using the beamline BL17A (TTERecR, $\lambda = 0.9875$ Å) or beamline NW12 (TTERecR₁₆₋₁₉₆, $\lambda = 0.9875$ Å) at the Photon Factory (Tsukuba, Japan), beamline BL17U (TTERecR_{K21G}, $\lambda = 1.005$ Å) at Shanghai Synchrotron Radiation Facility (Shanghai, China) or beamline 1W2B (TTERecR₁₋₁₈₀ and TTERecR₁₋₁₆₆, $\lambda = 1.005$ Å) at Beijing Synchrotron Radiation Facility (Beijing, China). The diffraction images were indexed and integrated using HKL2000 (35). The data collection statistics are presented in Table 1.

Structure determination and refinement

The structure of full-length TTERecR was solved by the molecular replacement method using PHASER (36) with one monomer of drRecR [Protein Data Bank (PDB):1VDD] as the search model at 20–3 Å resolution. The model was completed and refined using the REFMAC5 refinement (37,38) and COOT (39) at 20–2.45 Å resolution. The structures of the mutants TTERecR₁₆₋₁₉₆ and TTERecR_{K21G} were solved using the model of TTERecR, and refined using the REFMAC5 or PHENIX refinement programs (40). All structural images were drawn using PyMOL (<http://www.pymol.org/>). Detailed crystallographic statistics are shown in Table 1. Coordinates have been deposited into PDB under the accession codes: 3VDP, 3VE5 and 3VDU.

Ultracentrifugation analysis

Sedimentation equilibrium experiments were performed using a Beckman Optima XL-A analytical ultracentrifuge (Beckman Instruments, Palo Alto, CA) equipped with absorbance optics and an An60 Ti rotor at 20°C for 6 h at 60 000 rpm for TTERecR and TTERecO and 55 000 rpm for the TTERecOR complex. The proteins were concentrated to 1 mg/ml in 5 mM Tris–HCl, pH 7.5, 50 mM NaCl. Data analysis was conducted using the $c(s)$ or $ls-g^*(s)$ method using Peter Shuck's software program SEDFIT (41).

Size-exclusion chromatography

Size-exclusion chromatography was performed using a fast protein liquid chromatography system (GE Healthcare) on a Superdex-200 HR 10/300 column at a flow rate of 0.5 ml/min. Full-length TTERecR, TTERecR mutants, TTERecO and the TTERecOR complexes (at concentration of 200 μ g/ml) were loaded onto the column equilibrated with 50 mM Tris–HCl, pH 7.5, 300 mM NaCl and eluted using the same buffer. Protein elution was monitored by measuring the absorbance at 280 nm. Data analysis was conducted using UNICORN version 5.11 software program.

Surface plasmon resonance assays of protein with protein

Surface plasmon resonance (SPR) experiments were performed using a Biacore3000 machine (GE Healthcare) at 25°C. One flow cell of the CM5 sensor chip was activated

Table 1. Data collection and refinement statistics for TTERecR and its series mutant

	TTERecR _{full-length}	TTERecR _{K21G}	TTERecR ₁₆₋₁₉₆
Data collection			
Space group	P21	F222	P21212
Cell dimensions			
a, b, c (Å)	90.3, 68.3, 94.1	66.8, 123.7, 135.2	84.4, 84.7, 72.6
α, β, γ (°)	$\beta = 92.95$		
Resolution (Å)	50–2.45 (2.54–2.45)	50–2.9 (3.2–2.8)	50–2.8 (3.0–2.8)
R_{merge}	6.2 (38.3)	9.7 (46.6)	13.5 (54.3)
$I/\sigma I$	36.7 (4.8)	15.1 (4.2)	39.4 (8.45)
Completeness (%)	97.56 (86.00)	98.21 (100.00)	98.00 (94.00)
Redundancy	7.4 (6.3)	5.8 (6.2)	13.1 (13.3)
Refinement			
Resolution (Å)	20–2.45	20–2.8	20–2.8
No. reflections	41 210	6949	13 020
$R_{\text{work}}/R_{\text{free}}$	19.6, 22.6	22.7, 26.0	20.1, 25.4
No. atoms			
Protein	6125	1488	2655
Zn	4	1	2
B-factors			
Protein	54.65	95.84	56.49
Zn	38.2	82.82	48.84
R.m.s. deviations			
Bond lengths (Å)	0.013	0.012	0.015
Bond angles (°)	1.56	1.49	1.43

with a 1:1 mixture of 0.2 M EDC and 0.05 M NHS in water, as described by the manufacturer, then 10 $\mu\text{g/ml}$ TTERecO was injected over the flow cell in 10 mM sodium acetate (pH 5.5) at a flow rate of 10 $\mu\text{l/min}$. The remaining binding sites were blocked using 1 M ethanolamine, pH 8.5, and then 4900 response units (RU) of TTERecO were immobilized. TTERecR and the mutants (ranging in concentration from 0.0125 μM to 4 μM) were injected for 2 min at a flow rate of 30 $\mu\text{l/min}$. The running buffer was the same as the protein buffer (20 mM HEPES, pH 7.4, 200 mM NaCl and 0.005% [v/v] Tween 20); 20 mM NaOH was injected for 1 min to regenerate the chip surface. Non-specific binding to a blank flow cell was subtracted to obtain corrected sensorgrams; all data were analysed using BIAevaluation software version 4.1 and fitted to a 1:1 (Langmuir) binding model (steady state affinity model) to obtain equilibrium constants.

SPR assays of protein with DNA

Binding of protein to DNA was investigated using the BIAcore 3000 SPR machine (GE Healthcare) at 25°C. The running buffer (20 mM HEPES, pH 7.4, 200 mM NaCl and 0.005% [v/v] Tween 20) was prepared, vacuum filtered and degassed immediately before use. Biotin-labeled 60 mer ssDNA (5'-[Bio] ATAAATATCgATgTTAAAgAggATAAgATTTATAAAATAgCTTCTTATTTTCCCTgTAgTA-3') and biotin-labeled 60 mer dsDNA (biotin-labeled 60 mer ssDNA mixed with the corresponding 5'-TACTACAggAAAATAAgAAgCTATTTTATAAATCTTATCCTCTTTAACATCgATATTTA T-3' oligonucleotide) were anchored to the surface of the SA chip (55 RU ssDNA and 89 RU dsDNA). Different concentrations (0.03–9 μM) of the proteins and mutants were injected onto the DNA surface or a blank flow cell

for 3 min at a flow rate of 40 $\mu\text{l/min}$. After a 3–4 min dissociation phase, the remaining proteins were removed using 30 μl 0.03% SDS (v/v). Non-specific binding to the blank flow cell was subtracted to obtain corrected sensorgrams. Equilibrium and kinetic constants were calculated by a global fit to 1:1 (Langmuir) binding model using BIAevaluation software version 4.1.

Electrophoretic gel mobility shift assays

A total of 2 μM substrate DNA (60 mer ssDNA, 60 mer dsDNA or plasmid DNA) was incubated with full-length TTERecR, TTERecR mutants, TTERecO or TTERecOR complex (1, 2, 4 and 10 μM) in 10 μl reaction buffer (20 mM HEPES, pH 7.4, 200 mM NaCl) at 25°C for 30 min. The ssDNA and dsDNA sequences were the same as those used for SPR. Proteins incubated with the 60 mer DNA complex were loaded onto 6% non-denaturing polyacrylamide gels with TBE buffer, and proteins incubated with the plasmid DNA complex were loaded onto 1% agarose gels with TBE buffer. The DNA and DNA–protein complexes were visualized using a SYBR Green EMSA (Molecular Probes EMSA Kit, Invitrogen).

RESULTS

TTERecR is a dimer in solution

The crystal structure of TTERecR was solved at a resolution of 2.45 Å by molecular replacement using the model of drRecR (PDB code: 1VDD). The structures of TTERecR and drRecR exhibited similar folds. Four TTERecR monomers (A/B/C/D) in the asymmetric unit formed a ring-shaped tetramer with a central hole of 30–36 Å diameter (Figure 1A). Two types of intersubunit

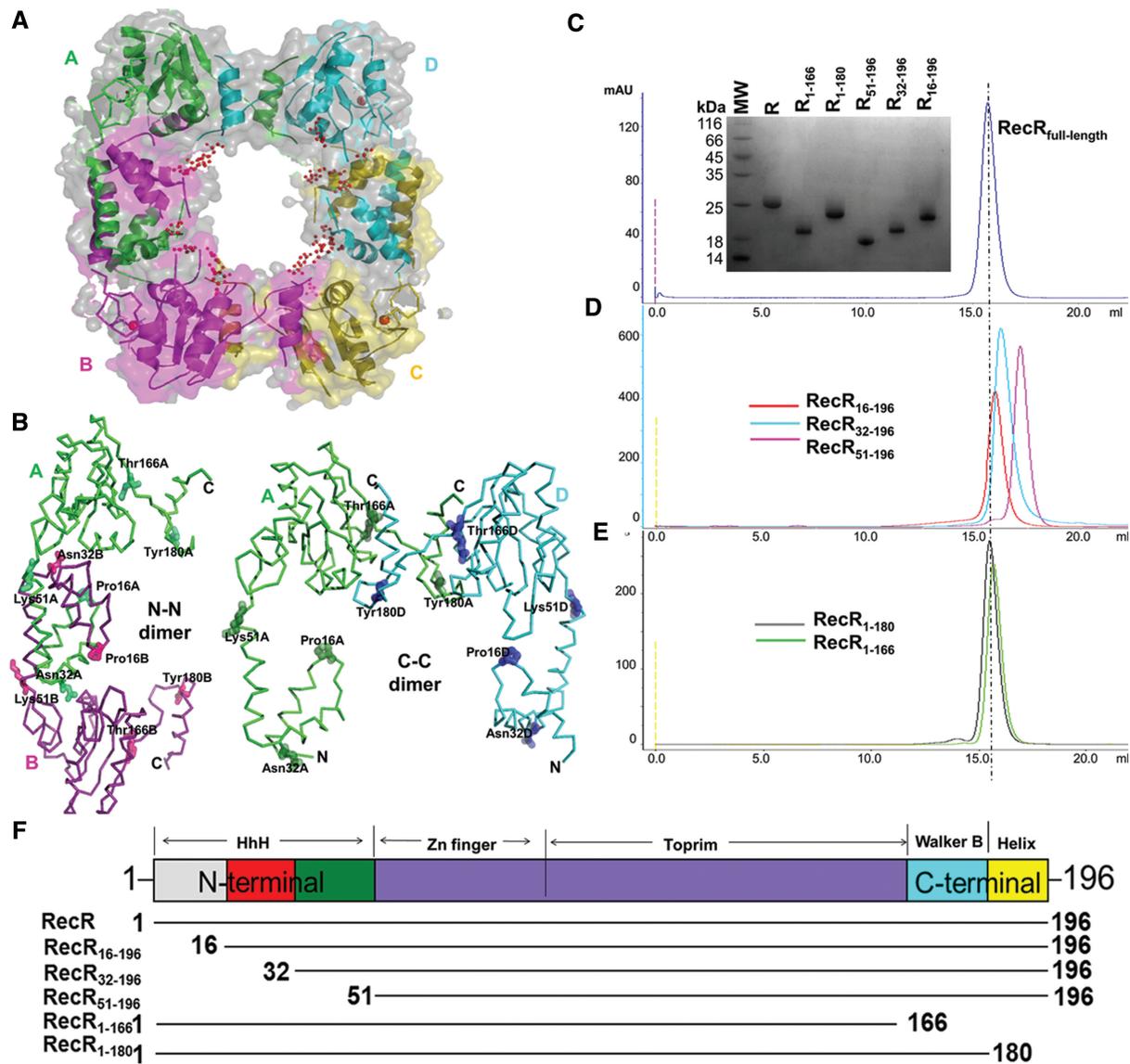


Figure 1. Dimerization interface analysis of TTERecR. (A) Sphere model structure of the TTERecR tetramer and the putative DNA-binding region of the central hole. (B) Ribbon model structure of the TTERecR A-B monomer dimer (N-terminal interface) and A-D monomer dimer (C-terminal interface). (C–E) Size-exclusion chromatography analysis of full-length TTERecR (C), the N-terminal deletion mutants TTERecR₁₆₋₁₉₆, TTERecR₃₂₋₁₉₆ and TTERecR₅₁₋₁₉₆ (D) and the C-terminal deletion mutants TTERecR₁₋₁₈₀ and TTERecR₁₋₁₆₆ (E). Size exclusion chromatography was performed using a HiLord 16/60 Superdex 200 column (GE Health Life Sciences) at 0.5 ml/min in 50 mM Tris-HCl, pH 7.5, 300 mM NaCl. Protein elution was monitored by measuring absorbance at 280 nm. The peak fractions of each sample were analysed on Coomassie Blue R-250-stained reducing 15% SDS-polyacrylamide (w/v) gels (inset) and compared with Low MW Protein Ladder (MW; Biomed). (F) Schematic domain organization of full-length TTERecR and the TTERecR mutants. The amino acid sequence boundaries of the TTERecR constructs are indicated below.

interactions were observed between the neighboring TTERecR: the N-N dimer (A-B; C-D) and the C-C dimer (A-D; B-C) (Figure 1B).

However, size-exclusion chromatography and ultracentrifugation analysis proved that TTERecR [molecular weight (MW): 23.5 kDa] forms a stable dimer (47 kDa) in solution (Figure 1C; Supplementary Figure S2A). Investigation of the RecR dimerization interface pattern is crucially important to study the biological function of this protein (21,27,29,30,42). We constructed a series of N-terminal and C-terminal deletion mutants to investigate

the dimerization interface of TTERecR (Figure 1B and F). The TTERecR₁₋₁₆₆ (MW: 18 kDa) and TTERecR₁₋₁₈₀ (MW: 20 kDa) C-terminal deletion mutants also formed stable dimers in solution, with molecular masses of 36 kDa and 40 kDa, respectively (Figure 1E). We obtained crystals of the TTERecR₁₋₁₆₆ and TTERecR₁₋₁₈₀ mutants (Supplementary Figure S1A and B). Unfortunately, it was difficult to solve these crystal structures because of dozens of TTERecR mutants in an asymmetric unit (Supplementary Figure S1C and D; Supplementary Table S1). Size-exclusion chromatography analysis of the

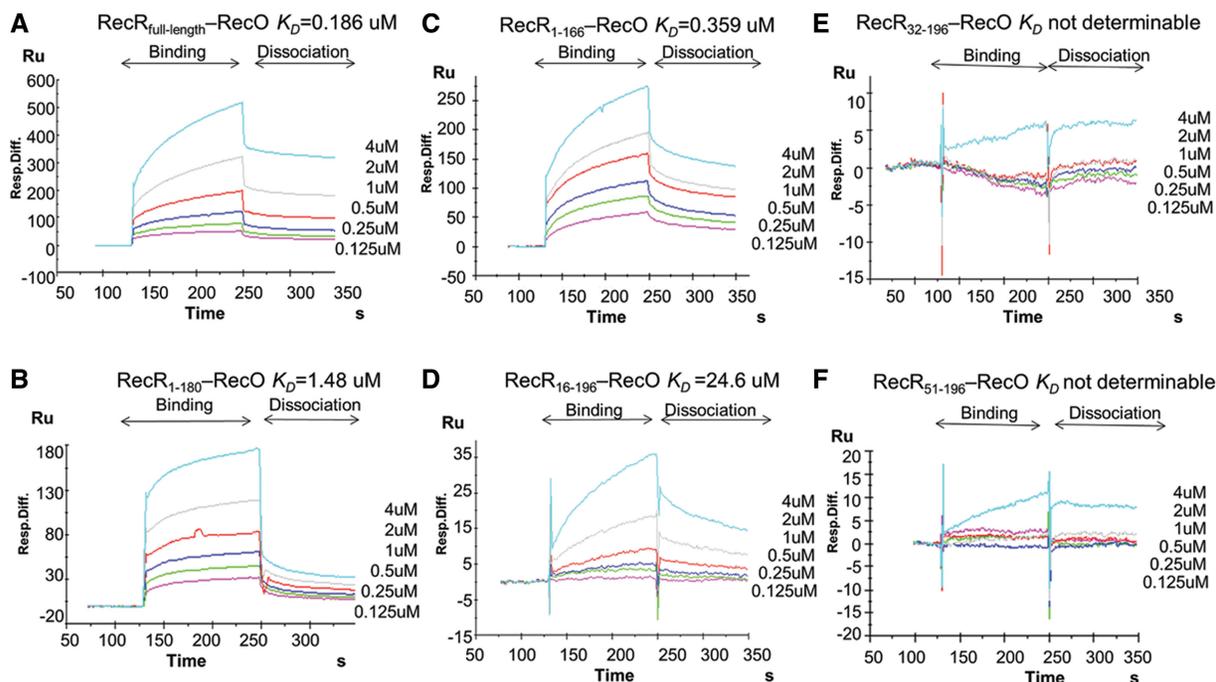


Figure 2. Contribution of the TTERecR N-N dimer to TTERecO binding. BIAcore biosensor analysis of full-length TTERecR (A); TTERecR₁₋₁₈₀ (B); TTERecR₁₋₁₆₆ (C); TTERecR₁₆₋₁₉₆ (D); TTERecR₃₂₋₁₉₆ (E); TTERecR₅₁₋₁₉₆ (F), binding to TTERecO at 25°C. Sensorgrams are shown for 0.125, 0.25, 0.5, 1, 2 and 4 μM TTERecR or TTERecR mutants injected over the TTERecO-coupled surface. The apparent K_D values were calculated from the kinetic K_D (M) = K_d/K_a .

N-terminal deletion mutants showed that TTERecR₁₆₋₁₉₆ (MW: 20 kDa) formed a 40 kDa dimer in solution, whereas both TTERecR₃₂₋₁₉₆ (MW: 19 kDa) and TTERecR₅₁₋₁₉₆ (MW: 16 kDa) were monomeric (Figure 1D). These results indicate that TTERecR cannot form a C-C dimer, when the N-terminus has been truncated in TTERecR₃₂₋₁₉₆ and TTERecR₅₁₋₁₉₆, even though the C-terminus is still intact. Furthermore, the crystal structure of TTERecR₁₆₋₁₉₆ revealed that TTERecR₁₆₋₁₉₆ formed an N-N dimer, not a C-C terminal interaction (Supplementary Figure S3A and B). Therefore, the biochemical and structural biology analysis demonstrated that the N-terminus is responsible for dimerization of TTERecR.

Molecular ratio of TTERecR interacting with TTERecO is 2R:1O

Recent reports indicated that the ring-shaped tetramer of RecR was the basic unit for interaction with two RecO monomers or two RecF monomers (23,29,30). We used size-exclusion chromatography and ultracentrifugation analysis to investigate formation of the TTERecR and TTERecO complex. TTERecO formed a monomer in solution, TTERecR formed a dimer in solution, and the calculated molecular mass of the TTERecOR complex was 78 kDa, corresponding to 2RecR:1RecO (Figure 3G; Supplementary Figure S2B and C).

To further analyse the interaction between the TTERecR dimer and TTERecO monomer, we used the SPR assay to investigate the interactions between TTERecO and the N-terminal and C-terminal deletion

mutants of TTERecR (Figure 1F). The binding affinities (K_D) of TTERecO for TTERecR₁₋₁₆₆ and TTERecR₁₋₁₈₀ were 0.359 μM and 1.48 μM, respectively; these values were not significantly different to that of binding affinity of full-length TTERecR for TTERecO ($K_D = 0.186$ μM; Figure 2A–C). These data indicate that the C-terminus hydrophobic helix region, which forms the C-terminal interaction, does not contribute to the binding of TTERecR with TTERecO. Size-exclusion chromatography assay indicated that TTERecR₃₂₋₁₉₆ and TTERecR₅₁₋₁₉₆ exist as a monomer in solution with no binding affinity for TTERecO, owing to the loss of the N-N terminal interaction (Figure 1D; Figure 2E and F). These results demonstrate that the N-N dimer of TTERecR is the basic functional unit for interacting with a single TTERecO.

The RecR N-N dimer is crucial for the interaction with RecO

Interestingly, TTERecR₁₆₋₁₉₆ formed a dimer in solution as full-length TTERecR, but did not form a detectable complex with TTERecO in size-exclusion chromatography or SPR ($K_D = 24.6$ μM; Figure 2D; Figure 3G). We resolved the crystal structure of TTERecR₁₆₋₁₉₆ at 2.8 Å resolution using molecular replacement. The space group of TTERecR₁₆₋₁₉₆ was P21212, and the packing of the molecules in the unit cell was significantly different to full-length TTERecR. Two TTERecR₁₆₋₁₉₆ monomers exist in a back-to-back conformation in the asymmetric unit, and two monomers from different asymmetric units form the TTERecR₁₆₋₁₉₆ dimer by their N-terminal

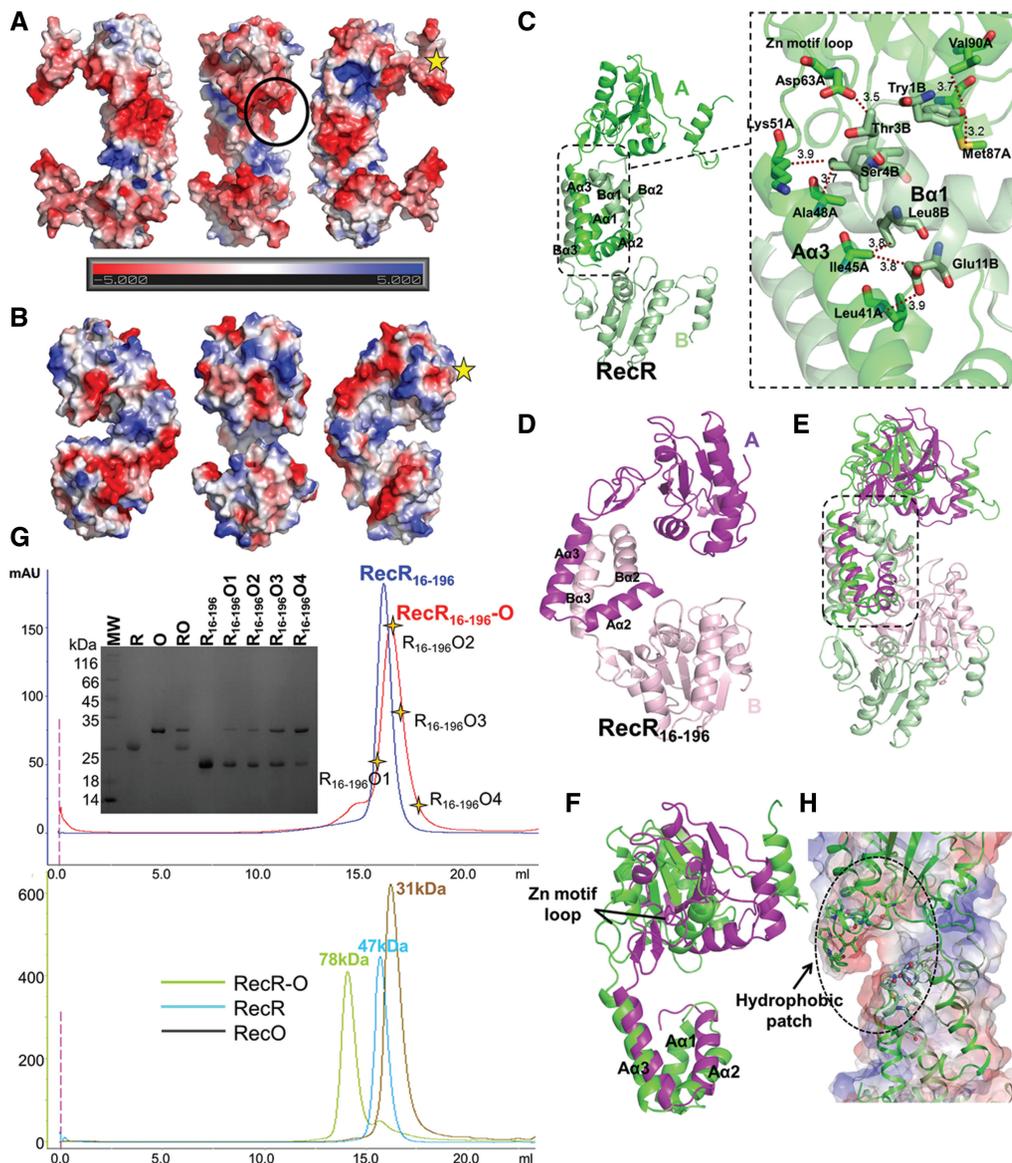


Figure 3. Structure of TTERecR₁₆₋₁₉₆. Molecular surface of full-length TTERecR (A) and TTERecR₁₆₋₁₉₆ (B) viewed in the same orientations. Electrostatic potential is indicated by different colours: blue is basic; red is acidic. (C) The structure of TTERecR N-N dimer (monomer A is green, monomer B is palegreen) is shown as a cartoon diagram. Expanded view of the circle, helix $\alpha 1$ of monomer B (palegreen) interacts with helix $\alpha 3$ and the zinc finger motif of monomer A (green). Interacting residues are shown as sticks. (D) The structure of the TTERecR₁₆₋₁₉₆ N-N dimer (monomer A is purple, monomer B is light pink). (E) Superimposed structures of TTERecR₁₆₋₁₉₆ N-N dimer (purple) and TTERecR N-N dimer (green). (F) Superimposed structures of the TTERecR₁₆₋₁₉₆ monomer (purple) and TTERecR monomer (green). (G) Size-exclusion chromatography analysis of the interaction between TTERecR₁₆₋₁₉₆ and TTERecO; TTERecR and TTERecO. (H) Hydrophobic patch (marked with a circle in Figure 3A) is formed by the hydrophobic loop06-121 and HhH domain of full-length TTERecR. The hydrophobic residues are shown as sticks.

interaction, with an interface area of about 5581 Å² (Supplementary Figure S3B).

The electrostatic surface potential of the TTERecR₁₆₋₁₉₆ dimer and the TTERecR dimer was totally different. The most notable difference was that the interface area of the TTERecR N-N dimer had a negative potential, whereas the interface area of the TTERecR₁₆₋₁₉₆ dimer had a positive potential (Figure 3A and B). Additionally, the C-terminus of TTERecR₁₆₋₁₉₆ interacted with the Toprim domain to form a globular C-terminal region, which is apparent different from the C-terminus of

TTERecR (Figure 3A and B). The TTERecR structure revealed that some residues of helix $\alpha 1$ (Ser4, Leu8, Leu12) in the HhH motif of one subunit had a hydrophobic interaction with some residues of helix $\alpha 3$ (Leu41, Ile45, Ala48, Lys51) in the HhH motif of the other TTERecR subunit (Figure 3C). Moreover, Thr3 interacted with Asp63 in the loop of the Zinc finger motif to stabilize the loop (Figure 3C). Superimposing the structures of TTERecR and TTERecR₁₆₋₁₉₆, we found that helix $\alpha 2$ of TTERecR₁₆₋₁₉₆ was 16° deviated to the helix $\alpha 2$ of TTERecR relatively, and helix $\alpha 3$ of TTERecR₁₆₋₁₉₆

correspondingly moved 3 Å, and the zinc finger motif moved 5 Å towards the HhH motif (Figure 3F).

These structures indicated that the hydrophobic interactions between helix α 1 of one monomer and helix α 3 of the other monomer play a key role in maintaining the structure of the TTERecR N-N dimer (Figure 3C–E). In the full-length TTERecR N-N dimer, the hydrophobic residues Val106, Phe109 and Ile110 of the loop₁₀₆₋₁₂₁ with the HhH domain created a hydrophobic patch for the interaction with TTERecO; however, this hydrophobic patch was destroyed in the TTERecR₁₆₋₁₉₆ mutant, preventing the binding with TTERecO (Figure 3A and H). The helix α 1 is essential for formation of the TTERecR N-N dimer and the TTERecOR complex, although helix α 1 is located far away from the other three functional domains.

A conserved basic Lys21 residue plays an important role in the RecR–RecO interaction

We investigated the ability of the TTERecR conserved residue mutants to interact with DNA or TTERecO. In contrast to reports, which have indicated that RecR_{K21G} has reduced or no ability to bind DNA (21,23,29), the TTERecR_{K21G} mutant had the same affinity to interact with plasmid DNA as full-length TTERecR (Supplementary Figure S4), and SPR demonstrated that the conserved residue mutant TTERecR_{K21G} and full-length TTERecR had an equally weak dsDNA-binding affinity (Supplementary Figure S5A and B). But TTERecR_{K21G} did not interact with TTERecO and could not form a TTERecO–TTERecR_{K21G} complex (Supplementary Figure S6). This result has not previously been reported.

The Lys21 residue of RecR is conserved in different species. It is hypothesized that the main function of Lys21 is to clamp DNA, and Lys21 has no influence on the binding to RecO (21). To investigate the significant difference observed in TTERecO binding after mutation of the TTERecR Lys21 residue to Gly, we solved the 2.8 Å crystal structure of the TTERecR_{K21G} mutant. The space group of TTERecR_{K21G} was F222, with only one monomer in the asymmetric unit (Figure 4A). Four crystallographic symmetric TTERecR_{K21G} mutants formed a ring-shaped tetramer similar to the full-length TTERecR structure (Figure 4B). The main difference between TTERecR and TTERecR_{K21G} was the positional change of loop₁₀₆₋₁₂₁. The offset angle of TTERecR_{K21G} reached 90° (Figure 4A and B), suggesting that loop₁₀₆₋₁₂₁ plays an important role in the binding of TTERecR to TTERecO. Moreover, we found that mutation of Lys21 to Gly in TTERecR_{K21G} only lead to slight deviations in the polypeptide chains. However, this slight offset resulted in the helix- α (residues 143–163) moving 3–4 Å away from loop₁₀₆₋₁₂₁, breaking the interactions between the helix- α (residues 143–163) and loop₁₀₆₋₁₂₁. The distance between Tyr153 and Pro117 increased from 3.25 Å to 20.88 Å, and the distance between Glu146 and Ile112 increased from 3.45 Å to 13.01 Å; these shifts induced the loop₁₀₆₋₁₂₁ to offset by 90° (Figure 4C). We superimposed the structural models of TTERecR, TTERecR_{K21G} and the drRecOR

complex (PDB code: 2V1C), which revealed that loop₁₀₆₋₁₂₁ was very similar in TTERecR and drRecR; but the loop₁₀₆₋₁₂₁ of TTERecR_{K21G} completely deviated from the drRecR loop (Figure 4D). In the crystal structure of drRecOR complex, loop₁₀₆₋₁₂₁ was like a wedge tightly bound to the hydrophobic pocket of drRecO, allowing RecR and RecO to form a stable complex (Figure 4D and E). In the crystal structure of TTERecR_{K21G}, the loop (residues 117–122) flipped 90° away from RecO and lost their interaction with the link region (residues 78–88), the helix- α (residues 87–99) and the helix- α (residues 228–234), which are all in the hydrophobic pocket of RecO (Figure 4E and F). Based on this analysis, we speculated that the Lys21 is required for the interaction between TTERecR and TTERecO, and the loop₁₀₆₋₁₂₁ in the Toprim domain is the most crucial site for the interaction with TTERecO.

TTERecOR complex displays a clear preference for ssDNA

We used the EMSA assay to detect the interactions between TTERecR (10 μ M), TTERecO (10 μ M) and the TTERecOR (2 μ M) complex with 60 mer ssDNA (2 μ M), and found that only the TTERecOR complex had the capability to interact with ssDNA (Figure 5A and B). The DNA-binding affinities were determined by SPR. TTERecR did not interact with 60 mer ssDNA (Figure 5C), and only interacted weakly with 60 mer dsDNA ($K_D = 24.8 \mu$ M; Figure 5D). TTERecO displayed high non-specific binding to the SA chip; therefore, the binding affinity of RecO for DNA could not be determined using this method. TTERecOR interacted strongly with 60 mer dsDNA ($K_D = 26.4$ nM) and ssDNA ($K_D = 0.931$ nM; Figure 5E and F). Thus, EMSA and SPR demonstrated that the TTERecOR complex stably bound to linear DNA, whereas the TTERecR and TTERecO did not stably bind linear DNA. Notably, TTERecOR had a much stronger binding affinity for ssDNA ($K_D = 0.931$ nM; Figure 5E) than dsDNA ($K_D = 26.4$ nM; Figure 5F), suggesting that the TTERecOR complex tends to bind ssDNA. We also obtained a stable sample of the TTERecOR complex with 27 mer ssDNA for further structural and functional research (Figure 5G). In addition, TTERecR_{P16G} and TTERecR_{R25G} had no binding affinity for 60 mer dsDNA. Surprisingly, TTERecR_{P16G} or TTERecR_{R25G} bound to TTERecO had an equally strong affinity for dsDNA or ssDNA as the full-length TTERecOR complex (Supplementary Figure S7). As a consequence, we speculated that the residues Pro16 and Arg25 of TTERecR are not part of the DNA-binding site, but maybe the recognition sites of DNA damage instead.

DISCUSSION

It is notable that the HhH domain plays an important role in the interaction of RecR with RecO; however, it has been reported that only the HhH domain is responsible for DNA binding (21,25,26). Using structural and biochemical analysis of the TTERecR₁₆₋₁₉₆ deletion mutant

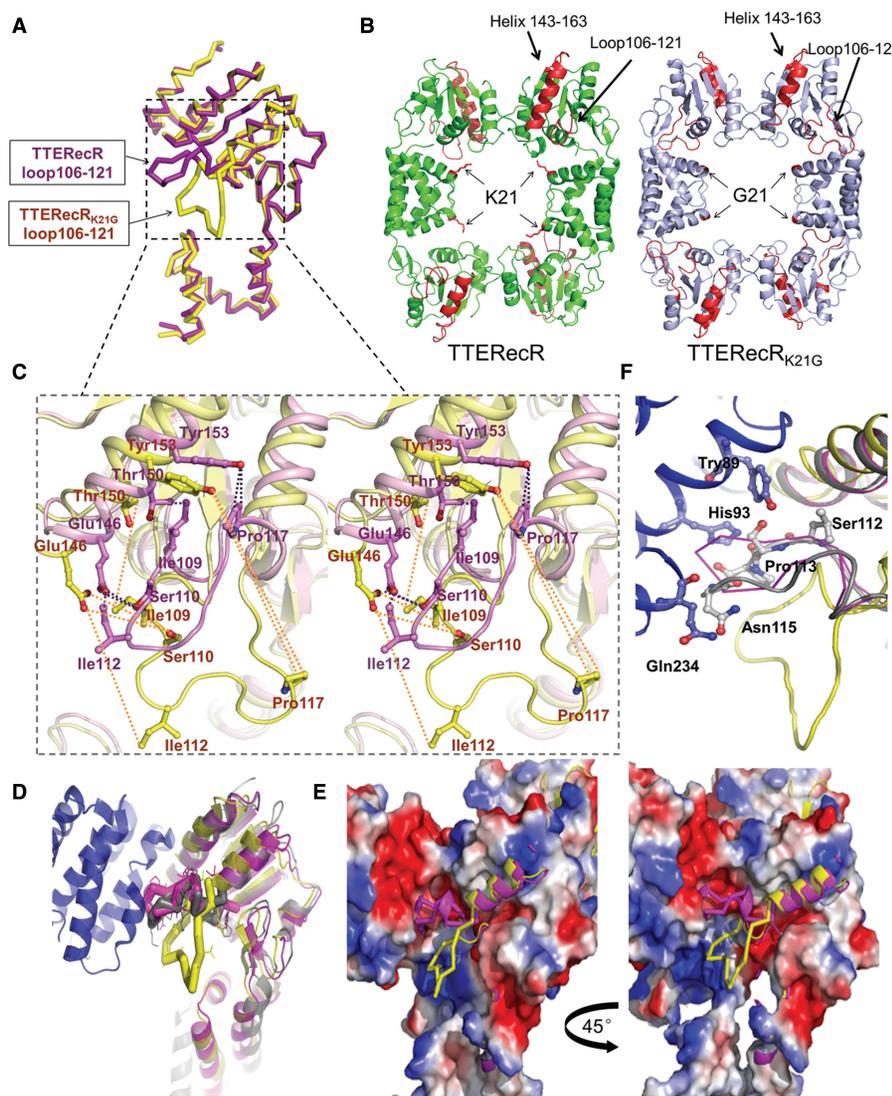


Figure 4. Structure of TTERecR_{K21G}. (A) Superimposed TTERecR_{K21G} monomer (yellow) and TTERecR monomer (purple) structures as a ribbon diagram. (B) Comparison of the TTERecR_{K21G} tetramer structure (light blue) with TTERecR tetramer structure (green). Lys 21 of TTERecR and Gly 21 of TTERecR_{K21G} are shown as sticks in red. Loop 106-121 and helix- α 143-163 is in red. (C) Superimposed TTERecR_{K21G} (yellow) and TTERecR (purple) structures, wall-eyed view showing the interactions between the helix- α (residues 143-163) and loop₁₀₆₋₁₂₁ are broken in the TTERecR_{K21G} structure. Interacting residues are shown as sticks. (D) Superimposed structures of TTERecR_{K21G}, TTERecR and drRecOR; drRecOR is coloured grey, other colours are as in A. (E) Electrostatic properties of the binding sites formed between TTERecR_{K21G} or TTERecR and drRecO. The complexes are shown as solvent-accessible surfaces coloured by electrostatic potential contoured at ± 5 Kt/e (red, acidic; blue, basic). (F) Loop 106-121 of TTERecR (purple) and drRecR (grey) interact with the hydrophobic pocket of drRecO (blue). Interacting residues are shown as sticks.

and TTERecR_{K21G} mutant, we demonstrated that the interaction between the entire HhH domains within the N-terminal regions of two TTERecR monomers is necessary for formation of a functional N-N RecR dimer. Slight movement of any helix in the HhH domain will firstly alter the electrostatic surface potential of TTERecR; secondly, lead to an enormous conformational change in the key loop₁₀₆₋₁₂₁ (which is considerably far away from the TTERecR_{K21G} mutation site and forms the site of interaction between TTERecR and TTERecO); and finally, destroy the hydrophobic patch in the Toprim domain of TTERecR, which interacts with RecO. Our research also

shows that the TTERecR dimer forms via an N-N terminal interaction, providing the basic functional unit for the interaction with TTERecO. Two N-N dimers of RecR form a ring-shaped tetramer via an interaction between their C-terminal regions. The ring-shaped TTERecR tetramer is preferred during crystal packing and is predicted to stabilize the RecR structure. The TTERecR_{K21G} crystal formed a similar ring-shaped tetramer to full-length TTERecR, but lost the ability to bind TTERecO. Therefore, we speculate that besides the interaction between 4 RecR:2 RecO in RecOR complex and 4 RecR:2 RecF in RecFR complex of the RecFOR

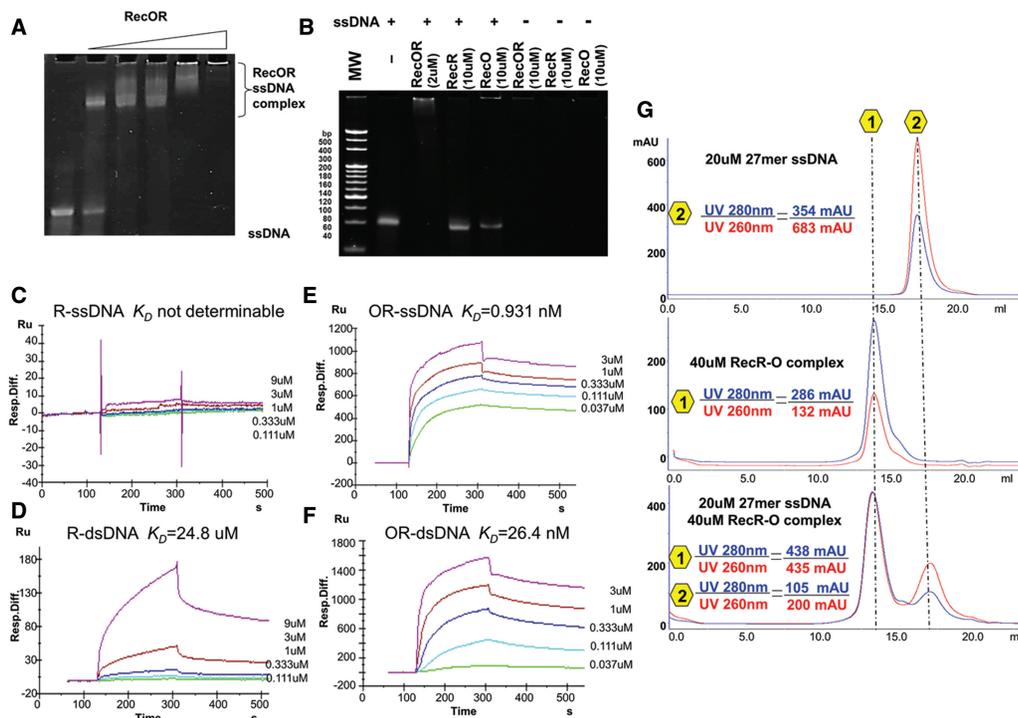


Figure 5. The TTERecOR complex displays a clear preference for ssDNA. (A) Analysis of the interaction between 60 mer ssDNA (2 μ M) and different concentrations of TTERecOR (Lines 1–6: 0, 1, 2, 4, 10 and 50 μ M, respectively) by native PAGE. (B) Analysis of the interaction between 60 mer ssDNA (2 μ M) and TTERecR (10 μ M), TTERecO (10 μ M) or TTERecOR (2 μ M) by native PAGE. (C–F) BIAcore biosensor analyses of TTERecR or TTERecOR binding to immobilized 60 mer ssDNA or 60 mer dsDNA at 25°C. The apparent K_D values were calculated from the kinetic K_D (M) = K_d/K_a . (G) Size-exclusion chromatography analysis of the interaction between 27 mer ssDNA and TTERecOR.

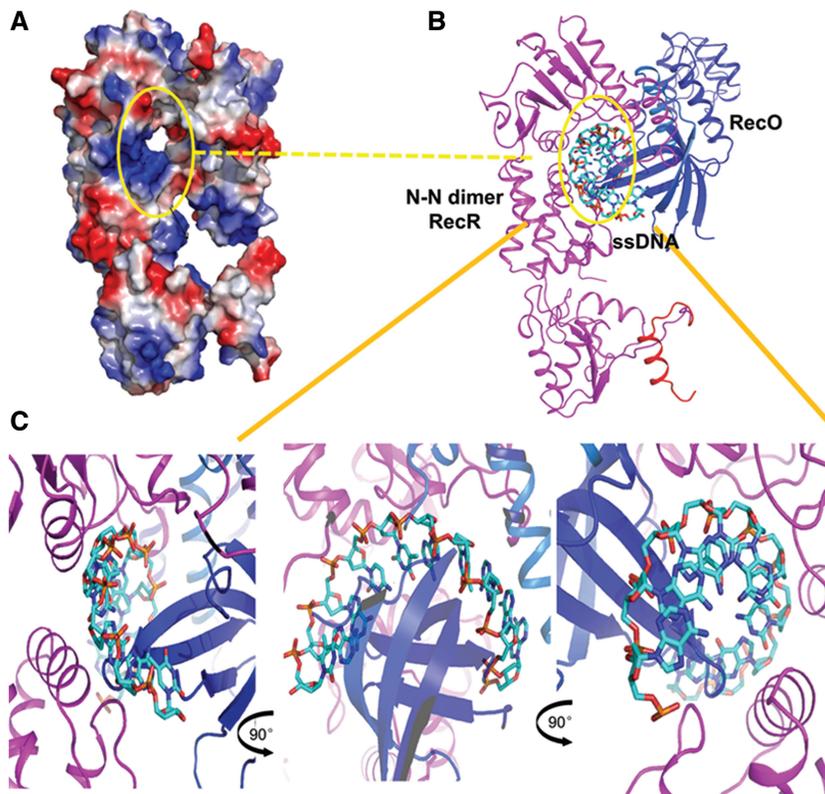


Figure 6. Model of ssDNA binding to TTERecOR. (A) Electrostatic properties of TTERecOR. The complex is shown as a solvent-accessible surface coloured by electrostatic potential contoured at ± 5 Kt/e (red, acidic; blue, basic). (B) Hydrophobic channel (marked with a circle) formed between 1-23 aa of RecO and the N-terminal interaction sites of the TTERecR dimer. (C) ssDNA bound to the basic and hydrophobic channel of TTERecOR.

repair pathway (29,30), it exists another mode of interaction, which is that the RecR N-N dimer interacts with RecO monomer.

We also observed that TTERecR did not bind linear ssDNA and very weakly bound linear dsDNA; but when TTERecR interacted with TTERecO, the TTERecOR complex could bind linear DNA. The TTERecOR complex had a much stronger affinity for ssDNA. Reported models of the interactions of RecR or the RecOR complex with dsDNA show that dsDNA walks through the centre of the RecR ring (21,23,29,30). In our analysis, when the ring-shaped tetramer of RecR was bound to two RecO monomers, it was difficult to find a channel, which interacts with dsDNA in the centre of RecR ring and also interacts with the hydrophobic residues in the HhH domain (Supplementary Figure S8). We have constructed a structure model of the RecR N-N dimer and RecO monomer complex (Figure 6A) and RecOR and ssDNA complex (Figure 6B) by superimposing structures of the drRecOR complex and the TTERecR N-N dimer based on our structural and functional analysis of TTERecR. This model shows that the N-terminus of RecO (α 18-22 aa) interacts with the N-terminus of the RecR N-N dimer (1-30 aa), and this complex forms a channel within which a large number of positive charges and hydrophobic residues are distributed (Figure 6A). The diameter of this channel is about 15 Å, which is the same as the diameter of ssDNA. Therefore, we speculate that damaged ssDNA could pass through this channel (Figure 6A and B). Based on our structural and biochemical analysis, we constructed structure models of the RecOR:ssDNA complex (Figure 6B and C).

ACCESSION NUMBERS

The structure factors and coordinates of TTERecR, TTERecR₁₆₋₁₉₆ and TTERecR_{K21G} have been deposited into PDB under the accession codes: 3VDP, 3VE5 and 3VDU, respectively.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1 and Supplementary Figures 1–8.

ACKNOWLEDGEMENTS

We thank the staff of the Photon Factory, KEK (beamline NW12, Project No. 2009G188 and No. 2011G035) in Japan, Beijing Synchrotron Radiation Facility (beamline 1W2B) and Shanghai Synchrotron Radiation Facility (beamline BL17U, Project No. 10sr0133) for their kind help in the data collection. We also thank Yuan-Yuan Chen of the IBP Core Facilities Center for advice with SPR.

FUNDING

Major State Basic Research Development Program of China (973 Program) [No. 2011CB910302, 2011CB966303 and No. 2011CB911101]; National Natural Science Foundation of China [No. 31070684]. Funding for open access charge: Institute of Biophysics, Chinese Academy of Sciences.

Conflict of interest statement. None declared.

REFERENCES

- Kaina,B., Christmann,M., Tomacic,M.T. and Roos,W.P. (2003) Mechanisms of human DNA repair: an update. *Toxicology*, **193**, 3–34.
- West,S.C. (2003) Molecular views of recombination proteins and their control. *Nat. Rev. Mol. Cell. Bio.*, **4**, 435–445.
- Kowalczykowski,S.C. (2000) Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem. Sci.*, **25**, 156–165.
- Cox,M.M. (2007) Regulation of bacterial RecA protein function. *Crit. Rev. Biochem. Mol. Biol.*, **42**, 41–63.
- Rocha,E.P., Cornet,E. and Michel,B. (2005) Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genet.*, **1**, e15.
- Kowalczykowski,S.C., Dixon,D.A., Eggleston,A.K., Lauder,S.D. and Rehrauer,W.M. (1994) Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.*, **58**, 401–465.
- Cox,M.M. (2001) Recombinational DNA repair of damaged replication forks in *Escherichia coli*: questions. *Annu. Rev. Genet.*, **35**, 53–82.
- Misra,H.S., Khairnar,N.P., Kota,S., Shrivastava,S., Joshi,V.P. and Apte,S.K. (2006) An exonuclease I-sensitive DNA repair pathway in *Deinococcus radiodurans*: a major determinant of radiation resistance. *Mol. Microbiol.*, **59**, 1308–1316.
- Handa,N., Morimatsu,K., Lovett,S.T. and Kowalczykowski,S.C. (2009) Reconstitution of initial steps of dsDNA break repair by the RecF pathway of *E. coli*. *Genes Dev.*, **23**, 1234–1245.
- Luisdeluca,C. (1995) Homologous pairing of single-stranded-DNA and superhelical double-stranded DNA catalyzed by reco protein from *Escherichia coli*. *J. Bacteriol.*, **177**, 566–572.
- Yokoyama,S., Kagawa,W., Kurumizaka,H., Ikawa,S. and Shibata,T. (2001) Homologous pairing promoted by the human Rad52 protein. *J. Biol. Chem.*, **276**, 35201–35208.
- Kowalczykowski,C., Kantake,N., Madiraju,M.V.V.M. and Sugiyama,T. (2002) *Escherichia coli* RecO protein anneals ssDNA complexed with its cognate ssDNA-binding protein: a common step in genetic recombination. *Proc. Natl Acad. Sci. USA*, **99**, 15327–15332.
- Kowalczykowski,S.C. (2005) Cancer: catalyst of a catalyst. *Nature*, **433**, 591–592.
- Yang,H., Li,Q., Fan,J., Holloman,W.K. and Pavletich,N.P. (2005) The BRCA2 homologue Brh2 nucleates RAD51 filament formation at a dsDNA-ssDNA junction. *Nature*, **433**, 653–657.
- Umez,K., Chi,N.W. and Kolodner,R.D. (1993) Biochemical interaction of the *Escherichia coli* RecF, RecO, and RecR proteins with RecA protein and single-stranded DNA binding protein. *Proc. Natl Acad. Sci. USA*, **90**, 3875–3879.
- Umez,K. and Kolodner,R.D. (1994) Protein interactions in genetic recombination in *Escherichia coli*. Interactions involving RecO and RecR overcome the inhibition of RecA by single-stranded DNA-binding protein. *J. Biol. Chem.*, **269**, 30005–30013.
- Ryzhikov,M., Koroleva,O., Postnov,D., Tran,A. and Korolev,S. (2011) Mechanism of RecO recruitment to DNA by single-stranded DNA binding protein. *Nucleic Acids Res.*, **39**, 6305–6314.
- Hegde,S.P., Qin,M.H., Li,X.H., Atkinson,M.A., Clark,A.J., Rajagopalan,M. and Madiraju,M.V. (1996) Interactions of RecF protein with RecO, RecR, and single-stranded DNA binding

- proteins reveal roles for the RecF-RecO-RecR complex in DNA repair and recombination. *Proc. Natl Acad. Sci. USA*, **93**, 14468–14473.
19. Shan, Q., Bork, J.M., Webb, B.L., Inman, R.B. and Cox, M.M. (1997) RecA protein filaments: end-dependent dissociation from ssDNA and stabilization by RecO and RecR proteins. *J. Mol. Biol.*, **265**, 519–540.
 20. Webb, B.L., Cox, M.M. and Inman, R.B. (1997) Recombinational DNA repair: the RecF and RecR proteins limit the extension of RecA filaments beyond single-strand DNA gaps. *Cell*, **91**, 347–356.
 21. Lee, B.I., Kim, K.H., Park, S.J., Eom, S.H., Song, H.K. and Suh, S.W. (2004) Ring-shaped architecture of RecR: implications for its role in homologous recombinational DNA repair. *EMBO J.*, **23**, 2029–2038.
 22. Morimatsu, K. and Kowalczykowski, S.C. (2003) RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair. *Mol. Cell*, **11**, 1337–1347.
 23. Honda, M., Inoue, J., Yoshimasu, M., Ito, Y., Shibata, T. and Mikawa, T. (2006) Identification of the RecR Toprim domain as the binding site for both RecF and RecO. A role of RecR in RecFOR assembly at double-stranded DNA-single-stranded DNA junctions. *J. Biol. Chem.*, **281**, 18549–18559.
 24. Ayora, S., Stiege, A.C. and Alonso, J.C. (1997) RecR is a zinc metalloprotein from *Bacillus subtilis* 168. *Mol. Microbiol.*, **23**, 639–647.
 25. Doherty, A.J., Serpell, L.C. and Ponting, C.P. (1996) The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA. *Nucleic Acids Res.*, **24**, 2488–2497.
 26. Ayora, S., Rojo, F., Ogasawara, N., Nakai, S. and Alonso, J.C. (1996) The Mfd protein of *Bacillus subtilis* 168 is involved in both transcription-coupled DNA repair and DNA recombination. *J. Mol. Biol.*, **256**, 301–318.
 27. Webb, B.L., Cox, M.M. and Inman, R.B. (1995) An interaction between the *Escherichia coli* RecF and RecR proteins dependent on ATP and double-stranded DNA. *J. Biol. Chem.*, **270**, 31397–31404.
 28. Alonso, J.C., Stiege, A.C., Dobrinski, B. and Lurz, R. (1993) Purification and properties of the RecR protein from *Bacillus subtilis* 168. *J. Biol. Chem.*, **268**, 1424–1429.
 29. Timmins, J., Leiros, I. and McSweeney, S. (2007) Crystal structure and mutational study of RecOR provide insight into its mode of DNA binding. *EMBO J.*, **26**, 3260–3271.
 30. Honda, M., Fujisawa, T., Shibata, T. and Mikawa, T. (2008) RecR forms a ring-like tetramer that encircles dsDNA by forming a complex with RecF. *Nucleic Acids Res.*, **36**, 5013–5020.
 31. Aono, S., Hartsch, T. and Schulze-Gahmen, U. (2003) Crystallization of a member of the recFOR DNA repair pathway, RecO, with and without bound oligonucleotide. *Acta. Crystallogr. D. Biol. Crystallogr.*, **59**, 576–579.
 32. Makharashvili, N., Koroleva, O., Bera, S., Grandgenett, D.P. and Korolev, S. (2004) A novel structure of DNA repair protein RecO from *Deinococcus radiodurans*. *Structure*, **12**, 1881–1889.
 33. Leiros, I., Timmins, J., Hall, D.R. and McSweeney, S. (2005) Crystal structure and DNA-binding analysis of RecO from *Deinococcus radiodurans*. *EMBO J.*, **24**, 906–918.
 34. Koroleva, O., Makharashvili, N., Courcelle, C.T., Courcelle, J. and Korolev, S. (2007) Structural conservation of RecF and Rad50: implications for DNA recognition and RecF function. *EMBO J.*, **26**, 867–877.
 35. Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.
 36. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
 37. Winn, M.D., Murshudov, G.N. and Papiz, M.Z. (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol.*, **374**, 300–321.
 38. 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.
 39. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta. Crystallogr. D. Biol. Crystallogr.*, **60**, 2126–2132.
 40. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta. Crystallogr. D. Biol. Crystallogr.*, **66**, 213–221.
 41. Dam, J. and Schuck, P. (2004) Calculating sedimentation coefficient distributions by direct modeling of sedimentation velocity concentration profiles. *Methods Enzymol.*, **384**, 185–212.
 42. Honda, M., Rajesh, S., Nietlispach, D., Mikawa, T., Shibata, T. and Ito, Y. (2004) Backbone 1H, 13C, and 15N assignments of a 42 kDa RecR homodimer. *J. Biomol. NMR*, **28**, 199–200.