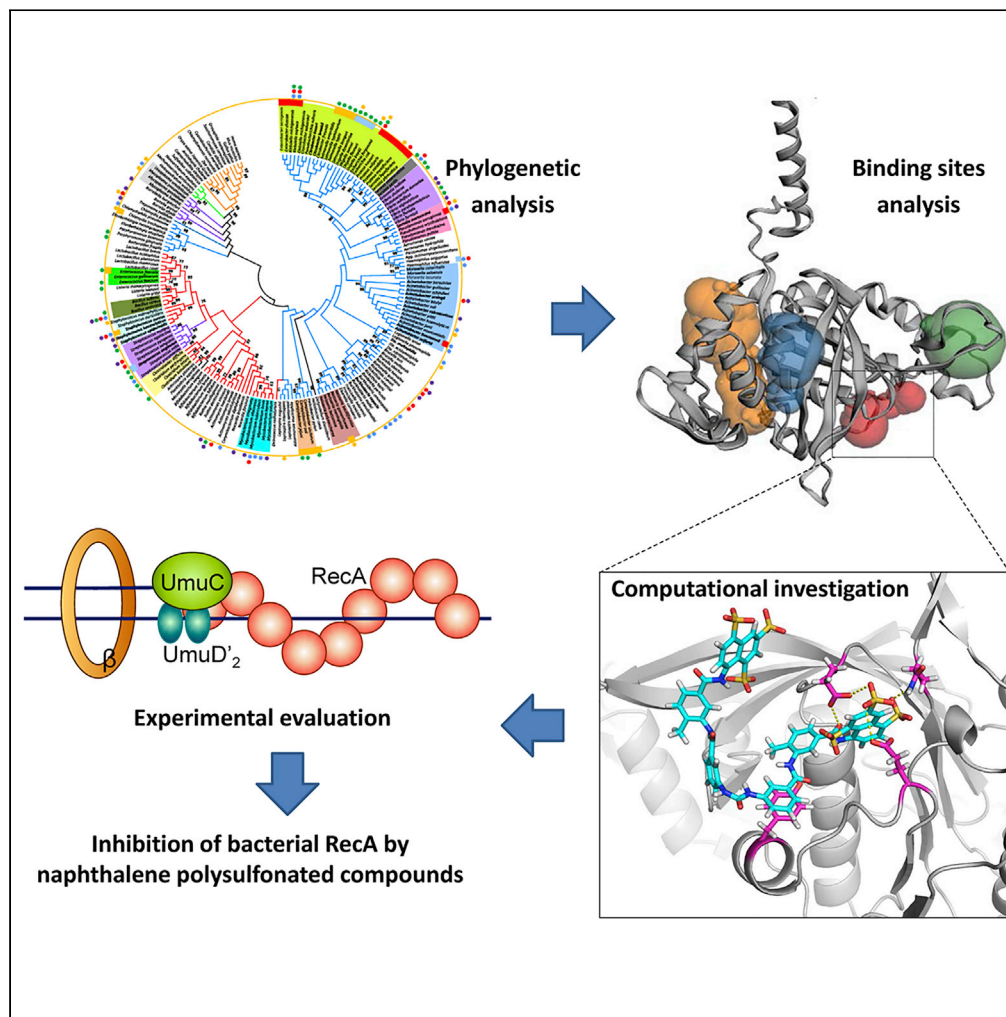# iScience

Article

Structural insights into the inhibition of bacterial RecA by naphthalene polysulfonated compounds

Ziyuan Zhou, Qing Pan, Xinchen Lv, ..., Rui Liu, Zigang Li, Nai-Kei Wong

liur6@mail.sustech.edu.cn (R.L.)
lizg@pkusz.edu.cn (Z.L.)
wongnksz@163.com (N.-K.W.)

Highlights

RecA is conserved among major bacterial pathogens but deviates from mammalian Rad51

Naphthalene polysulfonated compounds (NPS) block RecA ATPase and SOS response

Selected NPS bind recombinant RecA protein with K_d values in micromolar range

Tyr103 is an amino acid residue critical to RecA-NPS interactions

Article

Structural insights into the inhibition of bacterial RecA by naphthalene polysulfonated compounds

Ziyuan Zhou,[1,2,11] Qing Pan,[5,11] Xinchen Lv,[7,8,11] Jing Yuan,[2,11] Yang Zhang,[6] Ming-Xia Zhang,[2] Ming Ke,[10] Xiao-Mei Mo,[2] Yong-Li Xie,[2] Yingxia Liu,[2] Ting Chen,[6] Mingchan Liang,[3] Feng Yin,[3,4] Lei Liu,[2] Yiqing Zhou,[9] Kun Qiao,[2] Rui Liu,[7,8,*] Zigang Li,[3,4,*] and Nai-Kei Wong[1,2,12,*]

Summary

As a promising target for alternative antimicrobials, bacterial recombinase A (RecA) protein has attracted much attention for its roles in antibiotic-driven SOS response and mutagenesis. Naphthalene polysulfonated compounds (NPS) such as suramin have previously been explored as antibiotic adjuvants targeting RecA, although the underlying structural bases for RecA-ligand interactions remain obscure. Based on our *in silico* predictions and documented activity of NPS *in vitro*, we conclude that the analyzed NPS likely interact with Tyr103 (Y103) and other key residues in the ATPase activity center (pocket A). For validation, we generated recombinant RecA proteins (wild-type versus Y103 mutant) to determine the binding affinities for RecA protein interactions with suramin and underexamined NPS in isothermal titration calorimetry. The corresponding dissociation constants (K_d) ranged from 11.5 to 18.8 μM, and Y103 was experimentally shown to be critical to RecA-NPS interactions.

Introduction

Over the past few decades, a draining antimicrobial pipeline and rapid emergence of multidrug-resistant (MDR) organisms have precipitated a global crisis in public health. To galvanize response, the World Health Organization (WHO) published a "global priority pathogens" list for R&D of new antibiotics (2017) (Tacconelli et al., 2018a). Within this list, carbapenem-resistant Gram-negative "superbugs" including MDR *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Escherichia coli,* and *Klebsiella pneumoniae* have already gained notoriety in health care-associated infections (HAIs) (Tacconelli et al., 2018b) by virtue of their extraordinary ability to develop resistance against broad-spectrum antibiotics (Tacconelli et al., 2018b). These nosocomial pathogens cause significant morbidity and mortality among patients with severe injury or compromised immunity. Their infections also greatly increase medical costs and burdens by complicating or foiling fundamental health care practice such as mechanical ventilation, organ transplantation, HIV/AIDS therapy, anti-cancer chemotherapy, and child delivery, among others. There is thus a pressing need for alternative anti-infective solutions less prone to induce selective pressure, in addition to new antibiotics (Brown and Wright, 2016; Seifert et al., 2018).

One alternative anti-infective strategy involves the use of adjuvants and anti-virulence agents, which impose minimized selective pressure in bacterial pathogens compared with conventional antibiotics. The adjuvant approach seeks to exploit novel agents or repurpose existing drugs as antibiotic adjuvants, which themselves do not kill bacteria but enhance the effectiveness of a co-administered antibiotic by blocking antimicrobial resistance (AMR) mechanisms (Wright, 2016; Brown, 2015; Zheng et al., 2016). Examples that have entered into clinical use include a closely allied approach that employs anti-virulence agents to limit a pathogen's infectivity and thus the symptoms caused by its virulence factors (Zheng et al., 2016). Recently, our group has successfully developed small-molecule inhibitors targeting the eukaryotic-like Ser/Thr phosphatase (Stp1) in *Staphylococcus aureus* (Zheng et al., 2016). Notably, the Stp1 inhibitor aurintricarboxylic acid evidently blunted the expression of staphylococcal virulence in mice, without affecting *in vitro* growth of *S. aureus* (Zheng et al., 2015).

The bacterial recombinase A (RecA) protein represents an attractive target for adjuvant development in MDR pathogens, as challenges by exogenous stressors such as UV irradiation, sublethal-dose antibiotics,

[1]Department of Pharmacology, Shantou University Medical College, Shantou 515041, China

[2]National Clinical Research Center for Infectious Diseases, Shenzhen Third People's Hospital, The Second Hospital Affiliated to Southern University of Science and Technology, Shenzhen 518112, China

[3]Pingshan Translational Medicine Center, Shenzhen Bay Laboratory, Shenzhen 518055, China

[4]State Key Laboratory of Chemical Oncogenomics, School of Chemical Biology and Biotechnology, Shenzhen Graduate School of Peking University, Shenzhen 518055, China

[5]Shenzhen Key Laboratory of Microbial Genetic Engineering, College of Life Sciences and Oceanology, Shenzhen University, Shenzhen 518055, China

[6]CAS Key Laboratory of Tropical Marine Bio-resources and Ecology, Guangdong Provincial Key Laboratory of Applied Marine Biology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, 164 West Xingang Road, Guangzhou 510301, China

[7]Key Laboratory of Molecular Design for Plant Cell Factory of Guangdong Higher Education Institutes, Institute of Plant and Food Science, School of Life Sciences, Southern University of Science and Technology, Shenzhen 518055, China

[8]National Key Laboratory of Plant Molecular Genetics & Shanghai Center for Plant Stress Biology, CAS Center for Excellence in Molecular

Continued

and host respiratory bursts often converge on the generation of DNA-damaging reactive oxygen/nitrogen species (Hu et al., 2015, 2016; Van Acker and Coenye, 2017). RecA critically regulates an elaborate ATPase-dependent DNA repair process known as SOS response. In the case of unresolved antibiotic stress, RecA hyperactivity could give rise to aberrant expression of inducible error-prone repair systems such as polymerase V (Schlacher et al., 2005, 2006) that promote mutations (Bellio et al., 2017), genomic instability, and ultimately mutagenesis-fueled resistance acquisition (Kohanski et al., 2007). Among naphthalene polysulfonated compounds (NPS), suramin has been found to potently inhibit the ATPase activity of bacterial RecA protein, with IC_{50} values in the low micromolar range (Nautiyal et al., 2014). Experimentally, suramin displayed efficacy in inhibiting RecA activity and SOS response in *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* (Nautiyal et al., 2014). Despite efforts to develop several RecA inhibitors, the binding sites for these inhibitors remain obscure due to the functional versatility and multiple binding positions within RecA. In addition, research on the adjuvant mode of action by NPS has been hampered by a paucity of understanding about the structural details of RecA-suramin interactions.

To provide a theoretical foundation for drug discovery of RecA-targeting adjuvants with novel scaffolds, it is logically desirable to attempt to bridge the abovementioned gap via a combinatorial approach that encompasses molecular docking, dynamics simulation, and experimental validation. The utility of computational approaches to predicting binding sites can be further augmented by a judicious choice of energy-based detection algorithms, 3D structure- or geometry-based algorithms, etc. (Leis et al., 2010). Recently, our group successfully applied these approaches to map out RecA protein-binding determinants and predict binders among curcuminoid compounds (Zhou et al., 2019). In this present study, the phylogenetically relevant features within the RecA protein sequence, probabilistically favorable pocketness in the RecA protein structure, and predicted mechanisms underlying interactions between RecA and NPS of interest were scrutinized. To illustrate, the commercially available suramin was predicted to be a generic RecA inhibitor among untested bacterial pathogens and verified by quantitative real-time PCR (RT-qPCR) in terms of RecA-controlled transcription in *E. coli*. Additionally, we also performed isothermal titration calorimetry (ITC) to assess binding affinities for interactions between recombinant RecA proteins (wild-type versus mutant) and selected NPS including suramin. The amino acid residue Tyr103 (Y103) in *E. coli* RecA was critically implicated in RecA-NPS interactions. This work should prove valuable to future studies on the construction and application of structure-based virtual screening protocols that aim to accelerate the discovery of potent small-molecule candidate inhibitors against bacterial RecA, as a potential key regulator of AMR development.

Results and discussion
Phylogenetic analysis

Thus far, no specific study has been devoted to analyzing RecA protein sequences within bacterial species, especially clinically significant pathogens in AMR contexts. As a first step toward evaluating the feasibility and generality of RecA as a drug target, we performed phylogenetic analysis on the evolutionary relationship and sequence features of RecA-like proteins (that is, RecA/Rad51/RadA) among eukaryotic and prokaryotic species, with particular attention to MDR bacterial pathogens high in importance in the WHO "global priority pathogens" list (Tacconelli et al., 2018a). Extensive searches for RecA/Rad51/RadA proteins were performed with sequences extracted from the NCBI database, which were then subjected to multiple phylogenetic analyses. In this study, Rad51 protein sequences from 10 eukaryotic species and RadA protein sequences from four archaebacterial species were included for comparison.

Phylogenetic relationships depicted in Figure 1 were inferred by the neighbor-joining (NJ) method (Saitou and Nei, 1987). A bootstrap consensus tree inferred from 1,000 replicates is taken to represent the evolutionary history of the taxa analyzed (Felsenstein 1985). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. Evolutionary distances were computed by using the Poisson correction method and are denoted in units as number of amino acid substitutions per site. The analysis involved amino acid sequences of 185 RecA and RecA-like proteins (see sequence details in Tables S4, S5, and S6). All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted with MEGA7 (Kumar et al., 2016). Maximum-likelihood, minimum-evolution, and NJ consensus trees for RecA protein sequences are topologically congruent on most clades (as depicted in the Supplemental information, Figures S3 and S4). Percent bootstrap values above 70 are given. Importantly, we annotated the pathogens' infection sites, based on extensive literature searches (See details in the Tables S2, S3, and S4), with an emphasis on species known to be causative agents of nosocomial infections. Attributes

Plant Sciences, Chinese Academy of Sciences, Shanghai 201602, China

[9]School of Biotechnology and Food Engineering, Changshu Institute of Technology, Changshu, Jiangsu 215500, China

[10]BGI-Shenzhen, Shenzhen 518083, China

[11]These authors contributed equally

[12]Lead contact

*Correspondence:
liur6@mail.sustech.edu.cn (R.L.),
lizg@pkusz.edu.cn (Z.L.),
wongnksz@163.com (N.-K.W.)
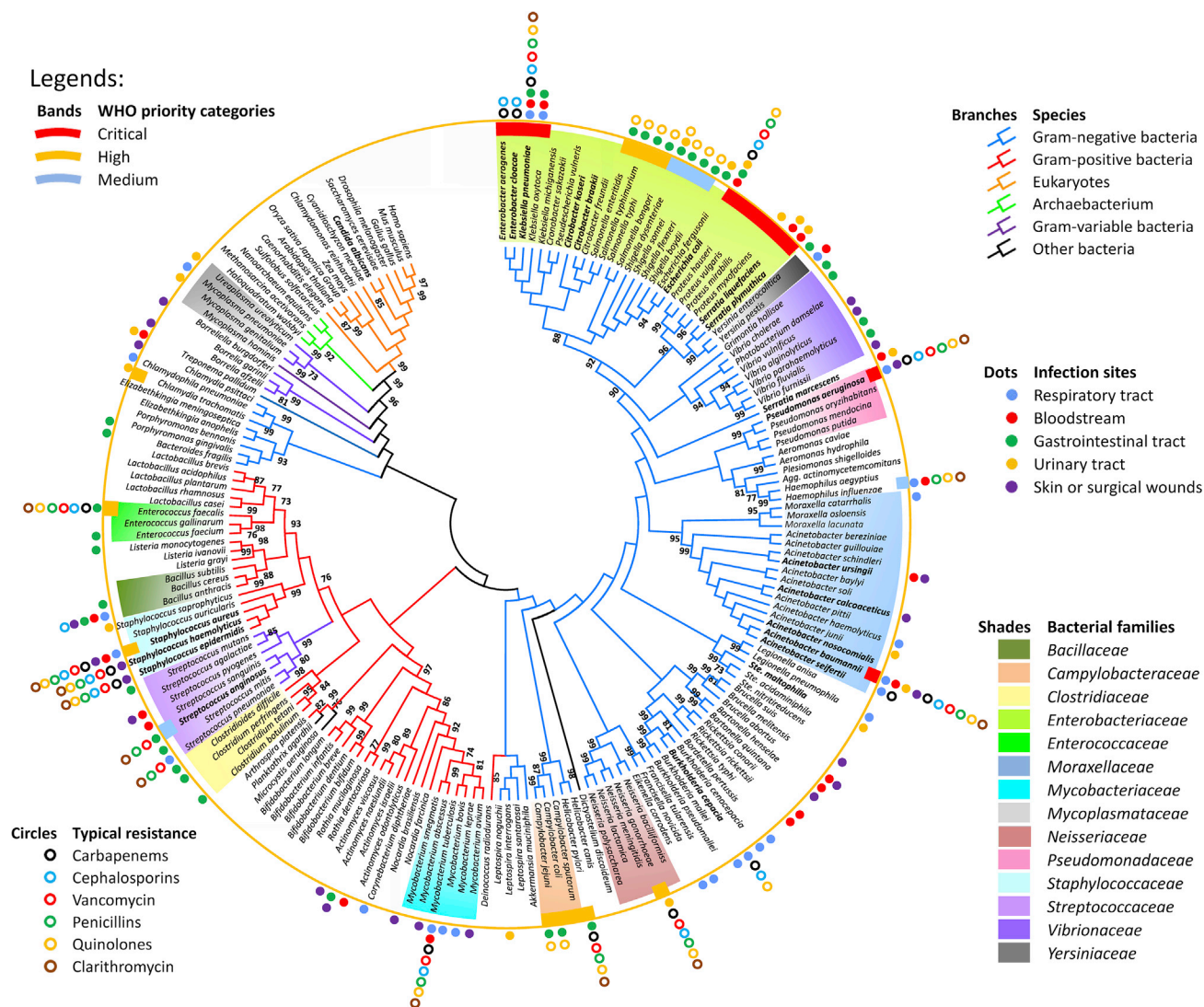
https://doi.org/10.1016/j.isci.2020.101952

Figure 1. Phylogenetic tree of RecA and RecA-like proteins among bacteria and eukaryotes, with annotations on antibiotic resistance and infection modes for bacterial pathogens

The broad categories of species, relative importance on the WHO "global priority pathogens" list, typical antibiotic resistance, and documented infection sites are indicated with different color branches, bands, closed circles, and dots, respectively. Major bacterial families were indicated as different color shades. Bacteria labeled in bold are known nosocomial pathogens. Detailed information used for annotations is summarized in the Supplemental information (Tables S4, S5, S6, and S9).

such as Gram status and families are phylogenetically grouped together, whereas other attributes such as infection sites and antibiotic susceptibilities are dispersed. In addition, we analyzed representative antibiotic susceptibility profiles of clinical isolates of major HAI pathogens tested in this study, whose results are presented in two heatmaps (Figures S6 and S7).

Phylogenetically, the evolution of RecA/Rad51 proteins includes both vertical and horizontal gene transfers, whereas the eukaryotic Rad51 proteins are evolutionarily close to the archaebacterial RecA-like protein and a branch of Gram-variable bacterial RecA. RecA proteins from several families of pathogenic bacteria (including Enterobacteriaceae, Moraxellaceae, and Pseudomonadaceae) with critical and high WHO priorities were principally clustered into a major branch of Gram-negative bacteria, which is very distant from the eukaryotic Rad51s. RecA-like protein-mediated recombination was reported to be a major reason for the maintenance of large plant chloroplast and mitochondrial genomes, as efficient DNA repair by homologous recombination presumably reduces deleterious effects of mutations.(Xu et al., 2016). On the other

hand, the mitochondrial RecA was believed to be lost in the ancestor of animals and fungi (Lin et al., 2006). Phylogenetic relationships of RecA protein between different species indicate that the development of bacterial RecA-selective inhibitor with little or no off-target effects on human Rad51 is feasible. A lack of mitochondrial RecA in human and animals also means that potential adjuvants targeting RecA should be safe for human use.

Structural alignment of RecA-like proteins

To illustrate the structural details of the ATPase core and evaluate the possibility of selective inhibition therein, we performed multiple alignment of amino acid sequences (Figure 2A) and structural alignment of RecA and Rad51 X-ray crystal structures from selected species (Figures 2B, 2C, and S1 in the Supplemental information). In multiple alignment analysis of amino acid sequence, as shown in Figure 2A, Walker A (amino acid residues 66–73), Walker B (amino acid residues 139–144), L1 (amino acid residues 157–164) (Malkov and Camerini-Otero, 1995), and L2 (amino acid residues 195–209) (Malkov and Camerini-Otero, 1995) were marked with red, blue, green, and black boxes, respectively. RecA proteins of 20 representative pathogenetic eubacterial species share a highly conserved core domain (yellow in Figure 2A), in particular, with respect to Walker A and Walker B, which are two highly conserved consensus motifs within the RecA/ RAD51 ATPase domain (Walker et al., 1982). Walker A is nestled within the ATPase activity center of RecA (Nahrstedt et al., 2005). Key residues Glu68 and Lys72 in Walker A are pivotal in conferring the RecA its ATP-binding ability and interacting with γ-phosphate of the nucleotide, thus critically regulating ATP hydrolysis (Story and Steitz, 1992). Likewise, Walker B also containing resides within RecA's ATPase activity core and its key residue Asp144 reportedly determines ATPase activity for it involved in ATP-Mg^{2+} binding (Story and Steitz, 1992). Walker A is relatively conserved from bacteria to eukaryotic species, whereas a considerable variability exists in the Walker B motif of RecA/Rad51 across bacterial species, with a stretch of bulky, hydrophobic amino acids followed by a negatively charged residue as an invariant feature (Koonin, 1993).

Ideally, RecA-specific inhibitors should exhibit adjuvant activity by blocking the ATPase domain of bacterial RecA without inhibiting the eukaryotic RecA homolog Rad51. For eukaryotic RadA and Rad51 and prokaryotic RecA filament, crystal structures are available for the extended form (active state) containing an ATP site at the subunit interface bound with ATP or ATP analogs. Also available are crystal structures corresponding to the compressed form (inactive state) containing an ATP site bound with ADP (Xing and Bell, 2004a, 2004b). The ATPase activity centers within the X-ray structures of *E. coli* RecA (gray, PDB 3CMT) (Chen et al., 2008) and *M. tuberculosis* RecA (cyan, PDB 1MO3) (Datta et al., 2003) are almost identical, with all the key residues conserved (Figure 2B). In contrast, due to evident disparity between RecA (gray, PDB 3CMT) and Rad51 (yellow, PDB 5H1B) (Xing and Bell, 2004a, 2004b) in 3D structures, as depicted in Figure 2C, bacterial RecA behaves differently from human Rad51 (Tombline and Fishel, 2002). For instance, unlike bacterial RecA, human Rad51 lacks catalytic efficiency and magnitude of ATP-induced cooperativity (Tombline et al., 2002a). Another difference is that the RecA-ssDNA complex could be stabilized by ATP-γS and destabilized by ADP. Conversely, Rad51-ssDNA complex can be stabilized by ADP or ATP and destabilized by ATP-γS (Tombline et al., 2002b). RecA in the active state binds ssDNA or dsDNA in the presence of ATP (Egelman and Stasiak, 1986). Segment of RecA-DNA-ATP-γS filaments identified by Edward et al. has been suggested to represent two distinct states of RecA (Yu et al., 2001), but later these two states was elaborated to be separable for the slow hydrolysis of ATP-γS, randomly occurring within a filament (VanLoock et al., 2003). Mapping of functional sites on RecA based on degree of conservation of sequence homology as analyzed by ConSurf and the data are depicted with the NGL viewer (Glaser et al., 2003). Yellow parts stand for lack of information (Figure 2D). The locations of ligand-binding sites on RecA are predicted by the prediction server CASTp 3.0 (Dundas et al., 2006; Chen et al., 2018) (Figure 2E). The ATPase activity center (pocket A) is the only binding site predicted by 3DLigandSite (Wass et al., 2010). Figure 2F depicts the view from different directions of the same structure.

Binding mechanisms of naphthalene polysulfonated compounds

Thus far, there have been no attempts to decipher mechanisms underlying the interactions between RecA and NPS via structural or *in silico* approaches. To unravel details of how NPS interact with RecA, the binding sites on RecA were first computationally predicted by using various algorithms. Compounds that bind RecA nucleoprotein filament in its active conformation can serve as competitive inhibitors of RecA ATPase hydrolysis, whereas compounds that bind the inactive RecA conformation can facilitate the dissociation of DNA from RecA (Wigle et al., 2006). The identification of novel small molecules or peptides targeting RecA could
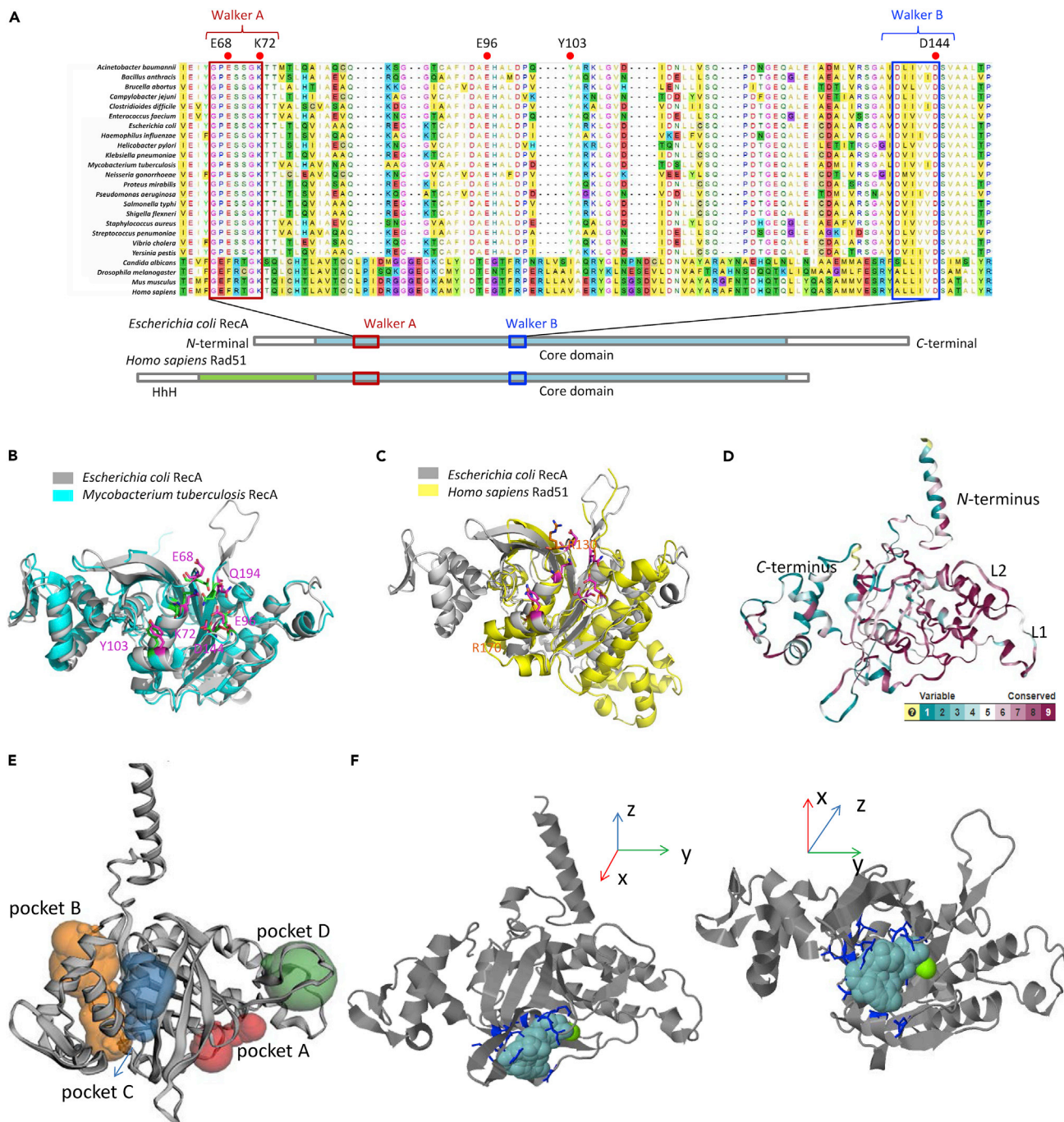
Figure 2. Structural comparisons of RecA and Rad51 crystal structures, and multiple alignment analysis on pivotal regions of amino acid sequences in the RecA/Rad51 superfamily

(A) Multiple alignment of RecA amino acid sequences from 20 bacterial species and Rad51 from 4 eukaryotic species. Domain structures of representative RecA proteins are schematically depicted.

(B) 3D protein structure alignment of *E. coli* RecA (gray) and *M. tuberculosis* RecA (cyan).

(C) 3D protein structure alignment of *E. coli* RecA (gray) and human Rad51 (yellow).

(D) Degree of conservation of sequence homology.

(E) The predicted binding regions are shown as colored spheres: pocket A (red), pocket B (orange), pocket C (blue), pocket D (green).

(F) The ATPase activity center predicted by 3DLigandSite.

Table 1. Summary of reported RecA inhibitors

Compounds	Classification	M.W. (Da)	Target organism	Methods	RecA inhibition (IC$_{50}$)	Ref.
Congo red (cpd03)	NPS	1290	E. coli	Fluorescent ATPase assay	2 μM	(Wigle and Singleton, 2007)
bis-ANS (cpd02)	NPS	594	E. coli	Fluorescent ATPase assay	ND	(Wigle and Singleton, 2007)
Suramin (cpd01)	NPS	652	M. tb	In vitro RecA ATPase; in bacteria	~1–2 μM	(Nautiyal et al., 2014; Wigle and Singleton, 2007)
Fe/Cu-PcTs (cpd04)	PTA	~900	E. coli	In vitro RecA ATPase; in miceND		(Alam et al., 2016)
Curcumin	Natural phenols	368	E. coli	PCR assay	ND	(Bellio et al., 2014)
Sphaerophorins	Lichen metabolites	~300	E. coli	Fluorescent ATPase assay	≥ 14.2 μM	(Bellio et al., 2017)
4 × 10^1	Peptides	~1-2k	E. coli	In vitro RecA ATPase	ND	(Yakimov et al., 2017)
INPEP	Peptides	~1-2k	E. coli	In vitro RecA ATPase	3 μM	(Cline et al., 2007)
Ara-adenosines	Nucleotide analogs	~400	E. coli	In vitro RecA ATPase	ND	(Wigle et al., 2006)
N6-(1-naphthyl)-ADP	Nucleotide analogs	~400	E. coli	PCR assay	ND	(Lee et al., 2005)
Zn^{2+}, Cu^{2+}, Hg^{2+}	Metal ions	64–200	E. coli	In vitro RecA ATPase	ND	(Lee and Singleton, 2004)

NPS, naphthalene polysulfonated compounds; M.W., molecular weight; PTA, phthalocyanine tetrasulfonic acid analogs; ND, not determined; M. tb, Mycobacterium tuberculosis.

lead to discovery of potential antibiotic adjuvants and a better understanding of the structural determinants for binding between RecA and its conformationally selective ligands (Wigle et al., 2006).

The loop part of RecA in the inactive state is highly flexible in the RecA filament, which has proved difficult to resolve (Yu et al., 2004). Previously, the loop region of much of the DNA-free RecA X-ray crystal structures in the ssDNA binding part has been omitted (Figure S1) (Story and Steitz, 1992; Xing and Bell, 2004a, 2004b; Singleton et al., 2002). In this study, we chose the RecA crystal structure that represents the active state of RecA resolved by Chen et al. (2008) for in silico studies. As noted, Walker A and Walker B are two highly conserved motifs in pocket A. Pocket A is the ATPase center of RecA (Figure 2E), and the underlined residues in Table S8 are the key residues in ATPase function (Story and Steitz, 1992; Chen et al., 2008). Specifically, the highly conserved consensus motifs Walker A (containing Glu68, Ser69, Ser70, Gly71, Lys72, Thr73) and Walker B (containing Asp144) all lie within pocket A. In addition, the catalytic Glu96 (responsible for activating water molecule for nucleophilic attack on γ-phosphate), Gln194 (a key allosteric residue), and Tyr103 (required for stabilization of ATP molecule) are all located within pocket A (Story and Steitz, 1992; VanLoock et al., 2003; Xing and Bell, 2004a, 2004b; Singleton et al., 2002). Pocket C putatively participates in dsDNA binding (Chen et al., 2008; Prabu et al., 2008), whereas pocket D has been found to be an interaction site for ssDNA binding (Shinohara et al., 2015).

Since the discovery of RecA protein, a number of RecA inhibitors (Table 1), including several NPS such as suramin (Figure S2), Congo red, and bis-ANS (4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid), have been reported. Most of them are small molecule inhibitors, whereas several peptide inhibitors have also been proved to be effective RecA inhibitors. However, the structural determinants in RecA that govern its interactions with putative RecA inhibitors have not been examined in detail.

To appraise the robustness of currently available docking algorithms for estimating binding free energies that characterize binding between RecA and its ligands, we selected AutoDock and LeDock, and scrutinized their performance with respect to various pockets on RecA. Our results support a good correlation between the experimental inhibitory activities and binding free energies over a training set tested (Figure 3). Several reported RecA inhibitors (cpd01–cpd04 and cpd12) were predicted by the algorithms as the most potent inhibitors (displayed in bold in Table I of Figure 3). In contrast, those compounds determined to lack inhibitory activities against RecA (cpd05–cpd08) have the lowest binding scores. In addition, compounds cpd09–cpd11 were reported to be active in a screening study for RecA targeting inhibitors
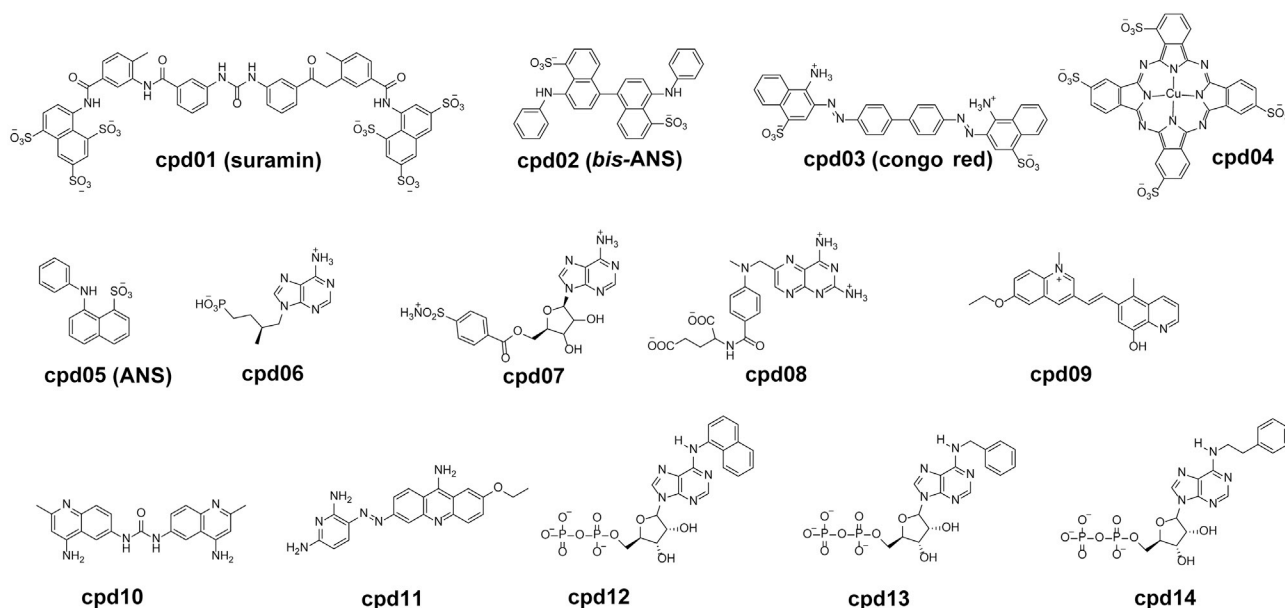
cpd01 (suramin)   **cpd02 (*bis*-ANS)**   **cpd03 (congo red)**   **cpd04**

cpd05 (ANS)   **cpd06**   **cpd07**   **cpd08**   **cpd09**

cpd10   **cpd11**   **cpd12**   **cpd13**   **cpd14**

Table I. Estimated energies for the binding of reported RecA inhibitors or inactive analogues with *E. coli*

Entry	Estimated binding energies with *E. coli* RecA (kcal/mol)[a]			
	pocket A	pocket B	pocket C	pocket D
cpd01	**-10.5 (-8.51)**	-8.6 (-6.93)	-9.0 (-6.16)	-8.5 (-5.51)
cpd02	**-9.4 (-6.00)**	-7.4 (-4.28)	-8.1 (-5.54)	-7.2 (-5.12)
cpd03	**-10.2 (-8.11)**	-8.5 (-7.85)	-8.7 (-7.86)	-8.5 (-9.24)
cpd04	**-9.5 (-6.80)**	-8.8 (-4.70)	-8.8 (-3.83)	-8.8 (-3.91)
cpd05	-7.4 (-3.59)	-5.6 (-3.47)	-5.9 (-3.65)	-5.5 (-3.46)
cpd06	-7.0 (-5.11)	-5.1 (-4.70)	-5.8 (-4.84)	-5.1 (-4.46)
cpd07	-8.7 (-5.87)	-6.9 (-5.24)	-7.3 (-4.77)	-6.9 (-4.66)
cpd08	-8.1 (-4.89)	-7.4 (-4.56)	-7.4 (-5.07)	-6.8 (-4.70)
cpd09	-8.5 (-4.71)	-7.9 (-4.55)	-7.0 (-4.37)	-6.4 (-4.43)
cpd10	**-9.3 (-7.25)**	-8.8 (-7.22)	-7.8 (-6.44)	-7.7 (-6.80)
cpd11	-8.2 (-6.98)	-8.1 (-6.13)	-7.2 (-5.70)	-6.5 (-6.16)
cpd12	**-10.0 (-7.60)**	-9.3 (-4.58)	-8.1 (-4.42)	-7.7 (-4.50)
cpd13	-8.8 (-5.51)	-8.8 (-4.88)	-7.3 (-4.16)	-6.8 (-4.58)
cpd14	-8.3 (-5.10)	-8.5 (-4.82)	-6.9 (-4.39)	-6.9 (-4.95)

[a] Binding energies estimated by Autodock vina (and Ledock)

Table II. Energy components for the binding between RecA protein and representative compounds

Energy terms [b]	RecA-inhibitor complexes			
	RecA + cpd01	RecA + cpd02	RecA + cpd05	RecA + cpd06
ΔG_{vdW}	-97.09	-68.21	-30.81	-18.06
ΔG_{ele}	-26.44	2.57	1.185	1.185
$\Delta G_{solv,p}$	78.81	38.51	16.04	28.85
$\Delta G_{solv,np}$	-10.74	-8.47	-3.75	-3.86
ΔG_{MM}	-123.53	-65.63	-29.62	-46.91
ΔG_{solv}	68.08	30.04	12.29	32.72
ΔG_{bind}	-55.45	-35.59	-17.33	-14.11

[b] Definition of the energy terms:
ΔG_{vdW} = van der Waals contribution from MM.
ΔG_{ele} = EEL, electrostatic energy as calculated by the MM force field.
$\Delta G_{solv,p}$ = EPB/EGB, which is the electrostatic contribution to the solvation free energy calculated by MMPBSA or MMGBSA respectively.
$\Delta G_{solv,np}$ = ESURF (or ECAVITY / ENPOLAR) for MMGBSA (MMPBSA) calculation. Nonpolar contribution to the solvation free energy.
ΔG_{MM} = ΔG_{ele} + ΔG_{vdW}
ΔG_{solv} = $\Delta G_{solv,p}$ + $\Delta G_{solv,np}$
ΔG_{bind} = ΔG_{MM} + ΔG_{solv}, final estimated binding free energy calculated from the terms above. (kcal/mol), which is the binding free energies in the absence of entropic contribution. This is described as equation (6) in the SI.

Figure 3. Docking analysis with reported RecA inhibitors or inactive compounds as a training set
Table I depicts the binding free energies estimated by AutoDock Vina (LeDock). Table II depicts the binding free energies between the RecA protein and representative compounds.

(Wigle et al., 2009), whereas cpd13 and cpd14 were reported to have less potency than cpd12 in blocking RecA-mediated SOS response (Lee et al., 2005).

To further characterize the binding potency of key molecules, we optimized the complexes of RecA and representative compounds with Amber16 and then carried on a short MD simulation with these complexes;

finally, the binding energies were predicted with molecular mechanics/generalized Born surface area (MM-GBSA) calculations. Detailed MM-GBSA results are displayed in Table II of Figure 3. A significant difference was observed between the known RecA inhibitors (cpd01 and cpd02) and known compounds (cpd05 and cpd06) that show no inhibitory activities against RecA. Results for docking and the MM-GBSA calculations are in good agreement, which suggests that these computational approaches could be useful as prediction tools for explaining the ATPase activities of these reported compounds. The MM-GBSA calculation procedure used here is similar to that of our previous work on identification of small-molecule inhibitors against a known protein target (Zhou et al., 2017). The protocol is a multiple step procedure, including format conversion of small molecules by Open Babel, 2,000 steps of energy minimization, 100 ps of molecular dynamics (MD) simulation, and 4,000 steps of energy minimization, and then MM-GBSA method was employed for calculating solvation free energy, using an Amber 16 software package. This procedure is similar to the established binding estimation after refinement (BEAR) protocol. More details about the calculation are listed in the Supplemental information.

In general, binding scores alone are not sufficiently conclusive due to the low accuracy of scoring functions in some docking software. Important aspects of binding free energy such as entropic contributions, long-range electrostatics, and desolvation process upon binding may not be adequately defined in conventional scoring functions built within AutoDock Vina and other algorithms of molecular docking (Boyer and Bryan, 2012; Rastelli et al., 2010). To further delineate details of binding between RecA and NPS, MD simulations were performed to evaluate the stability of RecA-NPS complexes, using suramin, bis-ANS, and Congo red as examples.

Generally, short MD simulations reveal that the predicated binding modes between RecA and suramin and selected NPS are stable within the first 10 ns of MD simulations (Figure 4), although hydrogen bonds could be transiently disrupted and reformed during the simulations. A case in point is that the hydrogen bonds between Tyr103 and Congo red could break and reform, but the overall binding mode remains stable during the process (Figures 4E and 4F).

We next validated the predicted biological effects of RecA inhibition on SOS response in an untested model organism, *E. coli*, whose RecA protein crystal structure was well resolved and computationally exploited in this study. Ciprofloxacin, a typical quinolone antibiotic, was used in combined treatment with suramin as a commercially available representative NPS. To evaluate the effects of suramin on RecA protein in terms of SOS response, *E. coli* (strain ATCC 25922) was treated with suramin, ciprofloxacin, or ciprofloxacin plus suramin, according to methods as described previously (Alam et al., 2016). For quantitative assessment by real-time PCR (RT-qPCR) (Figure 5), mRNA levels of four selected SOS response genes including *recA* (DNA recombination/repair protein), *umuD* (DNA polymerase V protein), *umuC* (DNA polymerase V catalytic protein), *dinI* (DNA damage-inducible protein), and one non-SOS response gene *soxR* (superoxide response protein) were assayed. In all treatments, *soxR* expression remained almost unchanged. However, transcriptional expression of SOS response-associated genes was differentially modulated upon treatment of the RecA inhibitor. In the presence of 50 μM suramin alone, transcript levels of *recA*, *umuD*, *umuC*, and *dinI* moderately decreased. Under antibiotic stress by ciprofloxacin, transcript levels of *recA* and *dinI* were elevated, affirming activation of the SOS pathway. Upon addition of 50 μM suramin to ciprofloxacin-challenged cells, up-regulated transcription of *recA*, *umuD*, *umuC*, and *dinI* was reduced to basal levels, indicating that the interaction of suramin and RecA protein effectively blocked bacterial SOS response, including the transcription of RecA itself.

Prediction of binding free energies in RecA-drug interactions

In conjunction with experimental and theoretical approaches discussed above, *in silico* prediction of binding free energies for binding between RecA and its ligands is essential for optimizing or discovering RecA inhibitors as potential antibiotic adjuvants. In addition to compounds with experimentally demonstrated RecA inhibitory activities such as cpd01 (suramin), cpd02 (bis-ANS), and cpd03 (Congo red) (Figure 3), we also explored other untested compounds and predicted their RecA binding affinities. For example, sulfanilamide (s01) is one of the first oral antibiotics in chemotherapeutic treatment for bacterial infectious diseases (Al-Omari et al., 2014). Besides sulfanilamide, other antibacterials classified as sulfonamides include sulfapyridine, sulfathiazole, and sulfadiazine (Al-Omari et al., 2014). Chicago sky blue (s05), Evans blue (s06), and direct yellow 50 (s07) are anionic dyes structurally related to suramin. Suramin and these three compounds have been found to be biologically active against infectivity and replication of human T cell
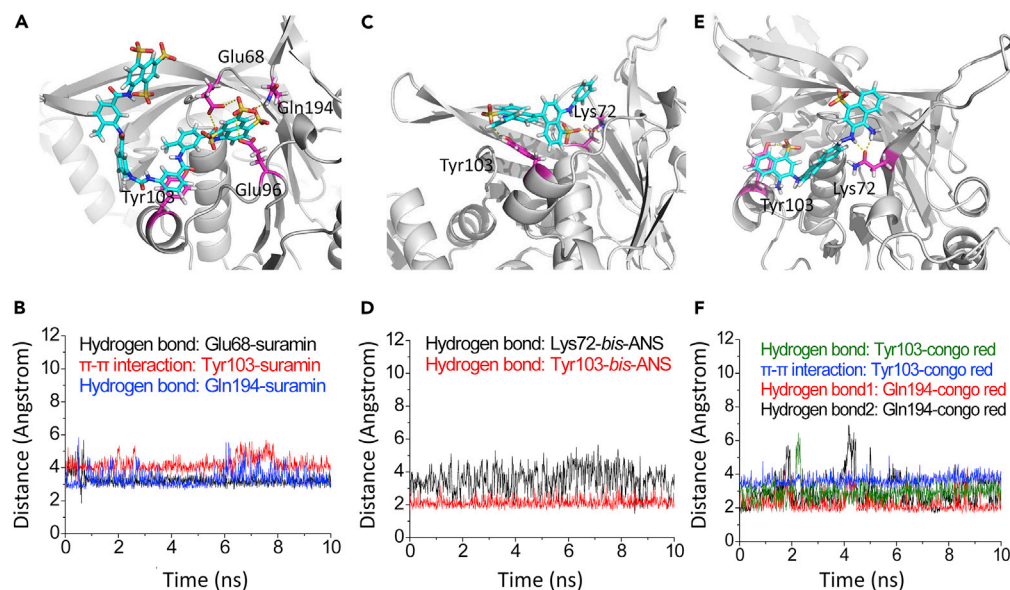
Figure 4. MD simulations of the RecA-inhibitor complexes

(A) Key residues (colored in pink: Glu68, Glu96, Tyr103, Gln194) around cpd01 (suramin) (colored in cyan) in the RecA-suramin complex are shown in stick style. Suramin is colored in cyan.

(B) Tracked distance changes in key hydrogen bonds. Tracked distance changes in π-π interactions between Tyr103 and benzene ring of suramin.

(C) Key residues (Tyr103, Lys72) around cpd02 (bis-ANS) in the RecA-bis-ANS complex are shown in stick style. The small-molecule RecA inhibitor bis-ANS is colored in cyan.

(D) Tracked distance changes in hydrogen bonds.

(E) Key residues (Tyr103, Gln194) around cpd03 (Congo red) in the RecA-Congo red complex are shown in stick style. Congo red is colored in cyan.

(F) Tracked distance changes in hydrogen bonds and tracked distance changes in π-π interactions between Tyr103 and naphthalene ring of Congo red.

lymphtropic virus (Balzarini et al., 1986). Elsewhere, the suramin analogs NF299 (s02), NF301 (s03), and NF307 (s04) reportedly activated the ryanodine receptor in skeletal muscles via a calmodulin binding site (Klinger et al., 1999). Additional NPS have been explored for their antiproliferative and angiostatic activities in mammalian cells. These analogs include NF037 (s08), NF059 (s09), NF061 (s10), NF062 (s11), NF063 (s12), NF064 (s13), NF013 (s14), NF289 (s15), NF031 (s16), NF130 (s17), NF136 (s18), NF291 (s19), NF326 (s20), NF357 (s21), NF036 (s22), NF520 (s23), NF015 (s24), NF033 (s25), and NF066 (s26), as listed in Scheme 1 (Firsching et al., 1995). Another source of reported NPS is the inhibitors and activators of protein tyrosine phosphatases, which include NF504 (s27), NF506 (s28), NF110 (s29), NF250 (s30), NF201 (s31), NF290 (s32), NF336 (s33), NF339 (s34), NF067 (s35), NF069 (s36) in Scheme 1 (Nishimura et al., 2015). Compounds in black boxes in Scheme 1 are NPs that have symmetric urea structures, whereas those in blue boxes are NPs that are asymmetric. Of the 36 structures 11 were denoted in bold in Figure 6, which were predicted to be more potent RecA inhibitors than suramin.

Protein-ligand interactions can be strongly influenced by intrinsic characteristics of small-molecule compounds. Along this line of reasoning, we asked what factors contribute to NPS binding to RecA protein. In our analysis, strong correlations seem to exist between binding affinity and intrinsic molecular properties of NPS, including molecular charge, molecular weight, and molecular span (Table III of Figures 6A–6D). The molecular span was measured with GaussView with the geometry optimized by B3LYP method at the 6-31G(d) level. Robust inter-algorithmic correlations were also observed between binding activity values generated by AutoDock Vina and LeDock (Figure 6A). Grid boxes used for AutoDock Vina and LeDock are identical, and their detailed parameters are given in the Supplemental information. To elaborate, the size of a ligand can be described by the radius of gyration (R_g), which is defined as the root mean squared distance of all elemental scattering volumes from either a given axis or their center of mass weighted by scattering densities (Jacques and Trewhella, 2010). R_g values provide information with respect to the mass distribution within a molecule. Dynamically, proteins can undergo conformational changes
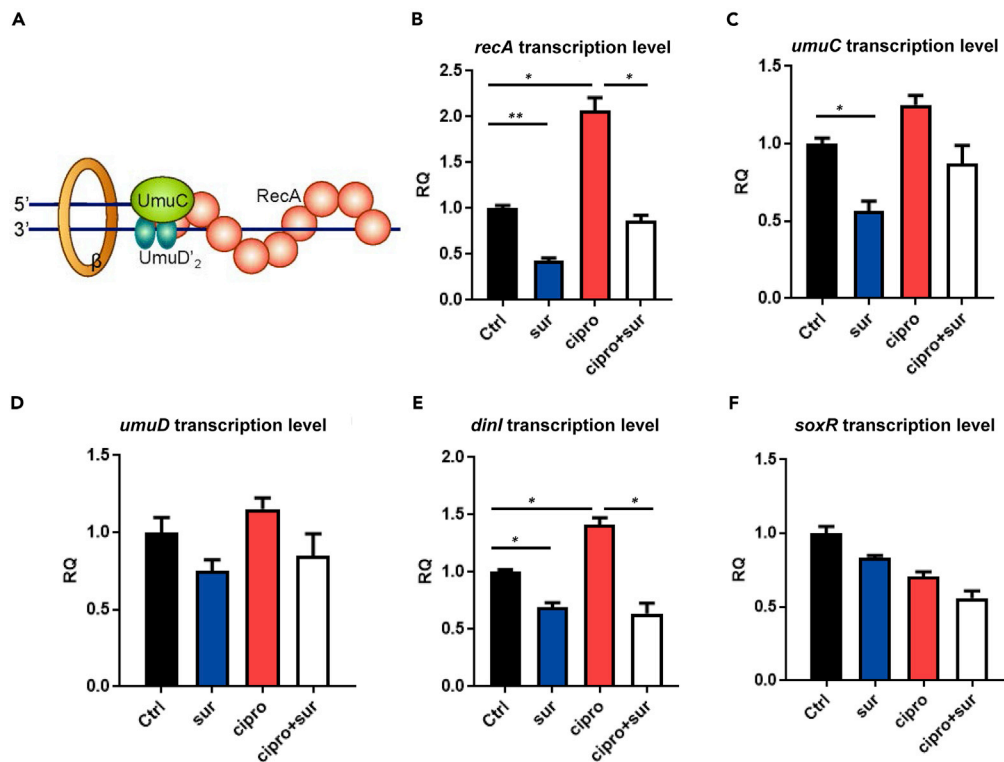
Figure 5. Effects of blockage of RecA protein on transcriptional expression of SOS genes in *E. coli* under antibiotic stress, as assessed by RT-qPCR

(A–F) (A) Schematic representation for polymerase switching that occurs in the transition from a DNA damage checkpoint to translational biosynthesis and replication (Schlacher et al., 2005), (B) *recA* transcription level, (C) *umuC* transcription level, (D) *umuD* transcription level, (E) *dinI* transcription level, and (F) *soxR* transcription level. Results are representative of at least three independent experiments. *β*, beta-sliding clamp. Statistical difference was determined at *$p < 0.05$, and **$p < 0.01$.

upon ligand binding, whereas ligands could also adapt different conformations in the same process. For a target protein, AutoDock Vina was reported to give the best predictions, when the box size is around $2.857R_g$ (Feinstein and Brylinski, 2015). However, calculations for the box size by this approach are not feasible for RecA, as this protein possesses four possible binding sites with distinct physiological functions.

In general, the structure-activity relationship analysis and docking calculations on these NPS shed light on several trends in RecA-ligand interactions. First, NPS need to be large enough to be stably anchored within the binding pockets in RecA. Second, the negatively charged sulfonate groups formed hydrogen bonds with the conserved key residues (including Glu68, Lys72, Glu96, Asp144, Gln196) in the ATP-binding site of pocket A. Third, a benzene ring, urea group, or other linkers at the center of NPS seems to contribute less to its overall binding affinity and may be a favorable position for further modifications.

Experimental validation of computationally predicted binding affinities for RecA/RecA[Y103A] and NPS analogs

To experimentally verify the strength of binding between RecA and NPS predicted *in silico*, we ventured to express and purify recombinant wild-type RecA protein, as well as its Tyr103 (Y103) mutant RecA[Y103A], to test their interactions with NPS *in vitro*. Subsequently, ITC measurements were performed to determine the dissociation constant (K_d) for interactions between RecA/RecA[Y103A] and suramin, along with three selected NPS for comparison. The binding affinity of RecA and suramin was first assessed in titration, wherein K_d was calculated as 17.6 μM (Figure 7A). We next utilized RecA[Y103A] to appraise the importance of the key residue Tyr103 in RecA binding by potentials inhibitors, including suramin. Consistently, the K_d for binding between RecA[Y103A] and suramin was 40.2 μM, much higher than the case of wild-type RecA

|	Table III. Estimated RecA binding energies and molecular properties for the compounds studied							
Entry	Binding energies estimated (kcal/mol)[a]				Relative binding energies (suramin=1)	Molecular properties		
	Pocket A	Pocket B	Pocket C	Pocket D		Net charge	Molecular weight	Molecular span
s01	-5.5 (-3.03)	-5.1 (-3.60)	-4.4 (-3.08)	-4.8 (-3.68)	0.52 (0.36)	+2	174	6
s02	-10.2 (-6.17)	-8.4 (-5.57)	-8.6 (-5.37)	-7.5 (-5.17)	0.97 (0.73)	-6	1402	36
s03	-10.0 (-5.56)	-7.0 (-3.34)	-7.1 (-3.87)	-7.2 (-3.52)	0.95 (0.65)	-6	1374	41
s04	**-11.2 (-8.46)**	-11.1 (-6.10)	-9.5 (-6.04)	-7.5 (-9.49)	1.07 (0.99)	-6	1374	41
s05	-9.4 (-6.29)	-6.6 (-4.27	-7.8 (-5.92)	-7.1 (-4.79)	0.90 (0.74)	-2	902	26
s06	-9.9 (-8.04)	-6.3 (-7.55)	-7.3 (-7.14)	-8.4 (-7.20)	0.94 (0.94)	-2	870	27
s07	-10.4 (-5.94)	-9.1 (-5.17)	-8.9 (-5.00)	-8.3 (-4.27)	0.99 (0.70)	-2	864	28
s08	**-11.0 (-6.64)**	-8.7 (-6.36)	-6.7 (-6.08)	-9.0 (-5.58)	1.05 (0.78)	-6	1262	31
s09	**-10.7 (-5.69)**	-9.8 (-4.73)	-8.0 (-4.84)	-7.3 (-5.82)	1.02 (0.67)	-6	1394	35
s10	**-10.8 (-8.01)**	-10.2 (-7.48)	-8.8 (-6.26)	-10.0 (-6.05)	1.03 (0.94)	-6	1346	34
s11	**-10.9 (-8.07)**	-9.2 (-7.42)	-9.3 (-6.93)	-9.0 (-6.80)	1.04 (0.95)	-2	874	25
s12	-9.7 (-6.25)	-9.3 (-6.13)	-9.3 (-6.36)	-8.1 (-6.37)	0.92 (0.73)	-4	1032	24
s13	**-10.6 (-7.82)**	-9.2 (-6.09)	-9.6 (-5.95)	-8.9 (-6.34)	1.01 (0.92)	-4	1032	24
s14	**-11.1 (-8.01)**	-10.9 (-3.99)	-9.1 (-4.34)	-7.6 (-4.66)	1.06 (0.94)	-6	1023	23
s15	-10.4 (-7.05)	-9.9 (-5.08)	-8.4 (-5.22)	-8.4 (-4.87)	0.99 (0.83)	-4	866	24
s16	-10.5 (-6.99)	-10.1 (-4.73)	-9.7 (-5.11)	-7.9 (-5.49)	1.00 (0.82)	-6	1262	37
s17	-10.5 (-7.62)	-7.4 (-4.56)	-8.4 (-4.73)	-7.5 (-4.97)	1.00 (0.90)	-6	1176	35
s18	**-11.1 (-7.43)**	-8.6 (-4.39)	-8.1 (-3.91)	-7.8 (-4.00)	1.06 (0.87)	-6	1176	30
s19	**-11.1 (-7.49)**	-10.0 (-5.46)	-9.0 (-6.73)	-8.3 (-5.09)	1.06 (0.88)	-4	866	29
s20	-10.5 (-6.76)	-10.2 (-6.21)	-10.1 (-6.99)	-8.9 (-5.65)	1.00 (0.79)	-6	898	30
s21	-10.4 (-6.43)	-8.1 (-5.06)	-8.2 (-6.01)	-7.4 (-5.62)	0.99 (0.76)	-6	1270	31
s22	-8.7 (-4.31)	-8.1 (-3.52)	-6.9 (-3.01)	-7.4 (-3.89)	0.83 (0.51)	-3	541	14
s23	-10.2 (-7.98)	-8.7 (-5.25)	-8.5 (-5.49)	-7.9 (-5.19)	0.97 (0.94)	-2	633	17
s24	-10.1 (-8.31)	-8.4 (-9.07)	-8.6 (-9.20)	-6.9 (-9.01)	0.96 (0.98)	-2	619	16
s25	-9.4 (-5.56)	-8.3 (-3.27)	-9.3 (-4.29)	-7.5 (-3.60)	0.90 (0.65)	-3	648	16
s26	-10.4 (-6.62)	-9.2 (-5.48)	-8.3 (-6.10)	-6.9 (-5.08)	0.99 (0.78)	-6	1344	34
s27	**-11.8 (-8.55)**	-10.4 (-6.54)	-8.5 (-7.15)	-9.0 (-7.05)	1.12 (1.00)	-4	998	35
s28	-10.7 (-7.93)	-10.4 (-6.98)	-9.6 (-5.64)	-9.1 (-5.98)	1.02 (0.93)	-6	1256	34
s29	-9.8 (-7.58)	-9.5 (-6.22)	-8.8 (-5.78)	-8.5 (-6.31)	0.93 (0.89)	-4	1004	28
s30	-10.4 (-7.98)	-9.9 (-5.95)	-9.1 (-5.22)	-9.1 (-4.74)	0.99 (0.94)	-2	636	25
s31	**-11.2 (-8.14)**	-9.7 (-4.26)	-8.8 (-4.63)	-8.6 (-3.99)	1.07 (0.96)	-6	1176	29
s32	**-10.8 (-5.43)**	-9.4 (-4.62)	-8.5 (-4.80)	-8.4 (-4.15)	1.03 (0.64)	-4	894	25
s33	-9.9 (-5.06)	-9.9 (-3.92)	-7.8 (-3.54)	-8.6 (-3.98)	0.94 (0.59)	-4	732	23
s34	**-11.2 (-6.93)**	-7.8 (-5.13)	-9.0 (-5.30)	-8.5 (-4.91)	1.07 (0.81)	-4	970	28
s35	-10.5 (-7.00)	-9.5 (-5.89)	-9.1 (-6.34)	-7.8 (-5.74)	1.00 (0.82)	-2	608	23
s36	-10.4 (-6.18)	-9.5 (-5.12)	-8.8 (-5.17)	-8.4 (-5.87)	0.99 (0.73)	-2	608	23
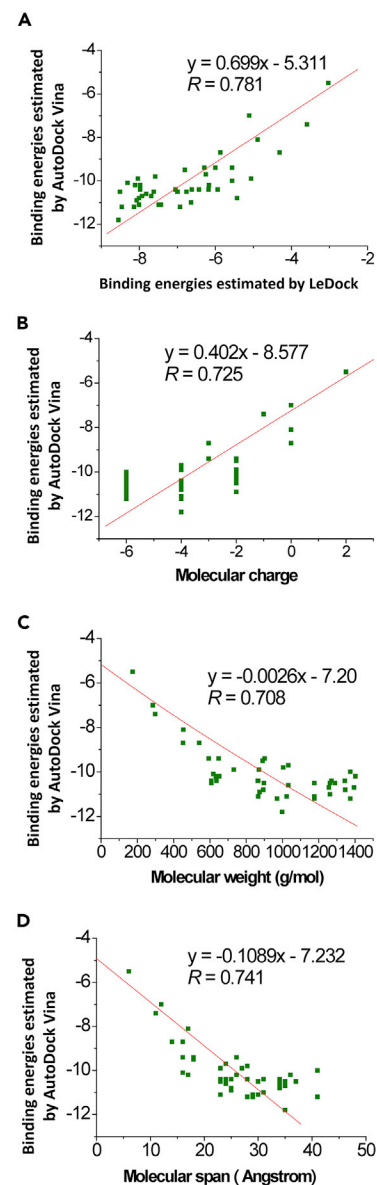
Figure 6. Binding energies for interactions between RecA and underexamined NPS, and their trends
(A) Correlation between the binding energies predicted by AutoDock Vina and binding energies by LeDock for NPS.
(B) Correlation between the binding energies predicted by AutoDock Vina and molecular charge of analyzed NPS.
(C) Correlation between the binding energies by AutoDock Vina and molecular weight of analyzed NPS.
(D) Correlation between the binding energies by AutoDock Vina and molecular span of analyzed NPS.

(Figure 7B). This confirms Tyr103 as a significant contributor in the interactions between RecA and suramin. In contrast, the binding affinity of RecA and ANS (8-(phenylamino)naphthalene-1-sulfonate), as a negative control, was undetectable (Figure 7C). Furthermore, the binding assay was also carried out with RecA and three other NPS compounds, s01, s05, and s06. As NPS, s05 and s06 were two compounds predicted as potential RecA inhibitors, with binding properties comparable to that of suramin (Table III). In ITC, the binding affinities of s05 and s06 to RecA were experimentally determined to be 18.8 μM and 11.5 μM, respectively (Figures 7D and 7E), which suggests that s05 is similar to suramin, whereas s06 binds more strongly than suramin in terms of binding affinity. On the other hand, our computational prediction suggests that the NPS s01 (sulfanilamide) might not bind RecA (Table III). Its ITC results corroborated this proposition, with no binding affinity detected for s01 and RecA (Figure 7F). Taken as a whole, our experimental results
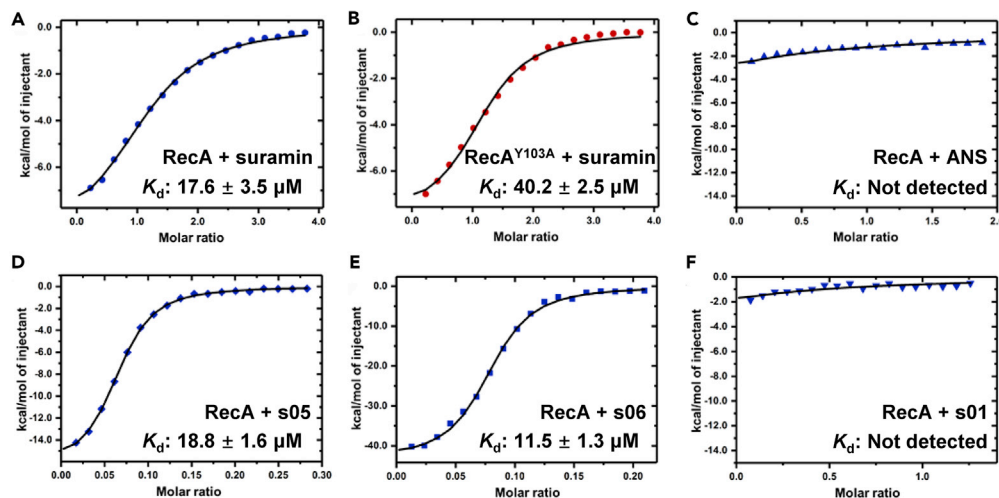
Figure 7. ITC measurements for experimental validation of computationally predicted binding affinities for interactions between RecA/RecA^{Y103A} and suramin/NPS

(A–F) (A) RecA and suramin, (B) RecA^{Y103A} and suramin, (C) RecA and ANS, (D) RecA and s05, (E) RecA and s06, (F) RecA and s01. RecA: recombinant wild-type RecA protein; RecA^{Y103A}: recombinant mutant RecA protein.

are in good agreement with our *in silico* predictions, indicating the reliability and utility of our docking approaches to assessing RecA-NPS interactions in general.

Conclusion

AMR among fast evolving MDR pathogens is an urgent, albeit underestimated, global health problem. To cope with an imminent crisis of antibiotic depletion, it is imperative to exploit alternative anti-infective strategies such as antibiotic adjuvants targeting bacterial RecA. Structurally, integration of phylogenetic information and protein crystal structures defining the active and inactive states could yield valuable insights into how bacterial RecA may be pharmacologically targeted. In this respect, we performed detailed phylogenetic analysis on RecA/Rad51/RadA proteins, which confirms that RecA is highly conserved among a large number of clinical bacterial pathogens including *M. tuberculosis* and *E. coli* but sequence-wise deviates substantially from mammalian homologs. In detailed crystal structure comparison, the ATP-binding region (pocket A) was demonstrated to be the major binding site. NPS are unlikely to bind RecA in pocket D, precluding interference at the interface between RecA and ssDNA. Molecular modeling and MD simulations consistently demonstrate the possible binding modes of NPS with RecA protein, in which π-π interactions with Tyr103 are anticipated to be critical for sustaining the stability of protein-ligand binding. Additionally, RT-qPCR results also confirmed that suramin as a representative NPS blocked transcription of SOS response genes in *E. coli*. Extending this knowledge, AutoDock Vina and LeDock were used to screen potential active compounds from untested NPS, based on predicted conformations and binding energies of RecA-NPS complex structures. Again, the ATPase core (pocket A) was predicted to be a primary binding site in RecA. Amid 36 uncharacterized NPS of unreported RecA inhibitory activities, 13 were predicted to be more potent than suramin in inhibiting RecA. Experimentally, suramin and selected NPS with computationally high binding affinities were demonstrated to bind recombinant wild-type RecA protein in ITC with K_d values in the range 11.5–18.8 μM. The amino acid residue Tyr103 was shown to be critical to RecA-NPS interactions, as the recombinant mutant RecA^{Y103A} protein displayed a much higher K_d value than that of wild-type RecA. Taken as a whole, our findings here offer structural and computational insights into how RecA protein inhibition can be achieved by known and under-examined NPS. Our work also furnishes an analytical framework for designing or re-engineering compounds with enhanced adjuvant activity against bacterial RecA in the urgent contexts of AMR.

Limitations of the study

Computational investigation on RecA-NPS interactions was largely based on monomeric RecA protein and its ATPase activity center defined by pocket A. In our previous study on curcuminoid compounds as potential RecA inhibitors (Zhou et al., 2019), we discussed the filamentous nature of RecA complexes. However,
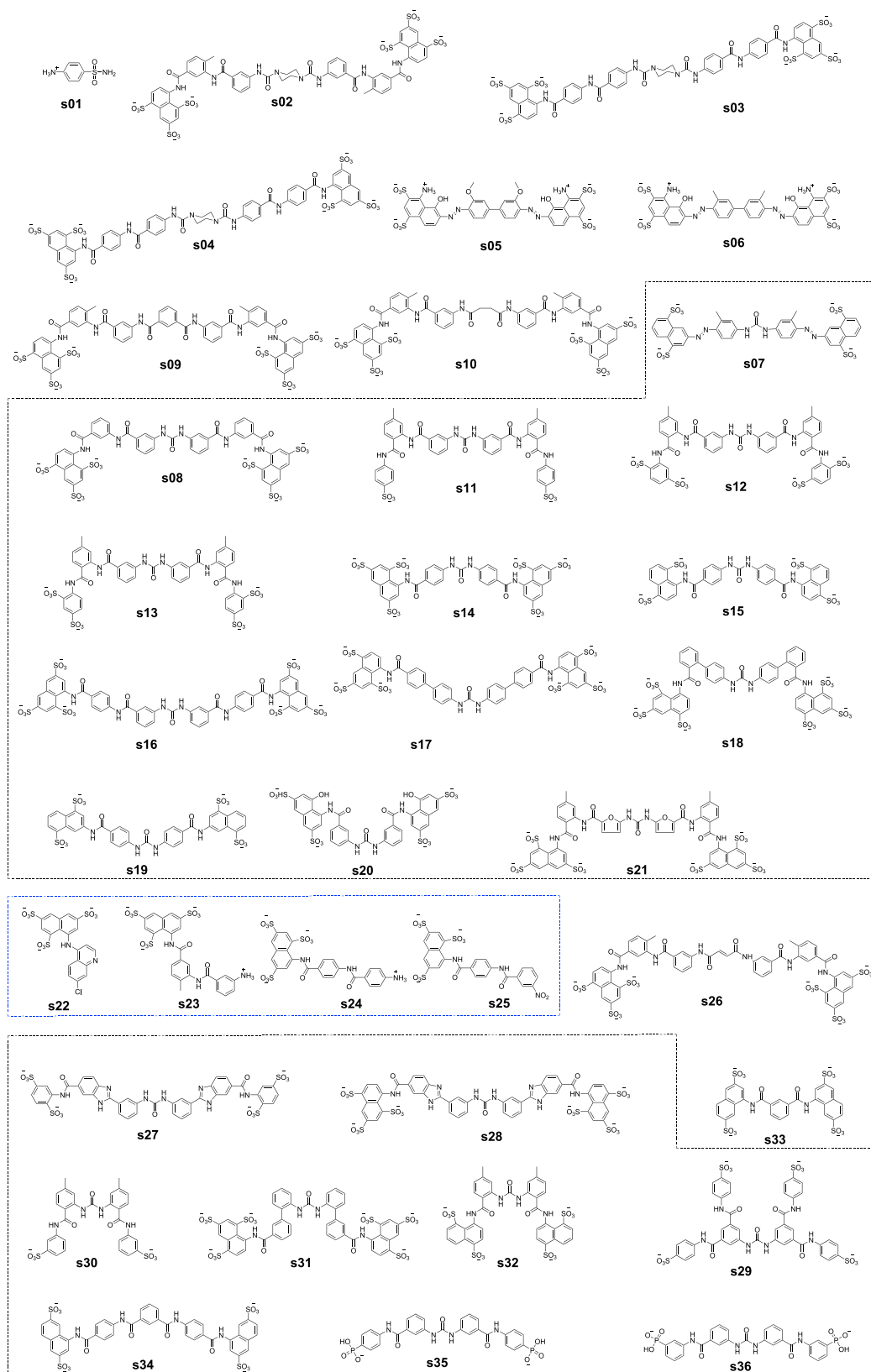
Scheme 1. Structures of reported naphthalene polysulfonated compounds (NPS) s1–s36 included in this study for prediction of their binding potency with RecA

protein-protein interactions within the RecA hexamer may potentially digress from the focus of our current study such as intra-molecular interactions and their determinants. For biological validation, we endeavored to show only Tyr103 as an amino acid residue crucial to RecA ATPase activity, without elaborating on other amino acid residues co-identified as key determinants. For example, as illustrated in Table S8 and Figure S5, both Tyr103 and Glu96 in RecA are likely to be crucial for ATP binding. NPS such as suramin may interfere with the formation of the RecA hexamer by firmly binding Glu96, the validation of which would require another set of elaborate molecular constructs and protein tools, which is beyond the scope of this study. Similarly, in addition to the ATPase activity center (pocket A), binding events at neighboring pockets such as pocket B and pocket D could modulate RecA protein conformation, with yet unclarified physiological functions. In our previous study (Zhou et al., 2019), structural attributes of the residues in pocket B were analyzed by multiple algorithms. However, none of such residues in pocket B have so far been reported to play significant roles in RecA-related protein-protein interactions.

Resource availability

Lead contact

Further information, requests, and inquiries should be directed to and will be fulfilled by the Lead Contact, Prof. Nai-Kei Wong (wongnksz@163.com).

Materials availability

Compounds were purchased from commercial suppliers including Sigma-Aldrich, Shenzhen Tenglong Logistics Co. or Energy Chemical Co. and were used without further purification unless otherwise stated. This study did not generate new specimens or materials. All images are included in the text and Supplemental information.

Data and code availability

The published article includes all data generated or analyzed during this study.

Methods

All methods can be found in the accompanying Transparent methods supplemental file.

Supplemental information

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101952.

Author contributions

N.-K.W., J.Y., and Z.L. conceived the project. Z.Z. conducted computational studies. Q.P., M.K., and T.C. performed the phylogenetic analysis. Z.Z., N.-K.W., and T.C. finalized presentation of the figures. X.-M.M., Y.-L.X., and M.-X.Z. handled clinical isolates and performed antimicrobial susceptibility tests (AST) and subsequent analysis. Q.P., R.L., and F.Y. designed the molecular experiments. Q.P., X.M., and Y.X. performed pharmacological studies, RT-qPCR and other molecular experiments. R.L., Q.P., and X.L. performed molecular cloning and protein expression and purification of recombinant RecA. X.L. and R.L.

carried out experimental validation of RecA-NPS binding in ITC. Q.P., Y. Zhang, R.L., K.Q., M.L., and N.-K.W. analyzed the resultant data. Z.Z., Q.P., Y.L., F.Y., L.L., Y. Zhang, Z.L., and N.-K.W. drafted the manuscript. N.-K.W., Z.L., Y. Zhou, and J.Y. guided the analysis of this work.

Declaration of interests

The authors declare no competing interests.

References

Van Acker, H., and Coenye, T. (2017). The role of reactive oxygen species in antibiotic-mediated killing of bacteria. Trends Microbiol. 25, 456–466.

Al-Omari, A., Cameron, D.W., Lee, C., and Corrales-Medina, V.F. (2014). Oral antibiotic therapy for the treatment of infective endocarditis: a systematic review. BMC. Infect. Dis. 14, 1–11.

Alam, Md K., Alhhazmi, A., DeCoteau, John F., Luo, Y., and Geyer, C.R. (2016). RecA inhibitors potentiate antibiotic activity and block evolution of antibiotic resistance. Cell Chem. Biol. 23, 381–391.

Balzarini, J., Mitsuya, H., De Clercq, E., and Broder, S. (1986). Comparative inhibitory effects of suramin and other selected compounds on the infectivity and replication of human T-cell lymphotropic virus (HTLV-III)/lymphadenopathy-associated virus (LAV). Int. J. Cancer 37, 451–457.

Bellio, P., Brisdelli, F., Perilli, M., Sabatini, A., Bottoni, C., Segatore, B., Setacci, D., Amicosante, G., and Celenza, G. (2014). Curcumin inhibits the SOS response induced by levofloxacin in Escherichia coli. Phytomedicine 21, 430–434.

Bellio, P., Di Pietro, L., Mancini, A., Piovano, M., Nicoletti, M., Brisdelli, F., Tondi, D., Cendron, L., Franceschini, N., Amicosante, G., et al. (2017). SOS response in bacteria: inhibitory activity of lichen secondary metabolites against Escherichia coli RecA protein. Phytomedicine 29, 11–18.

Boyer, R.D., and Bryan, R.L. (2012). Fast estimation of solvation free energies for diverse chemical species. J. Phys. Chem. B 116, 3772–3779.

Brown, D. (2015). Antibiotic resistance breakers: can repurposed drugs fill the antibiotic discovery void? Nat. Rev. Drug Discov. 14, 821–832.

Brown, E.D., and Wright, G.D. (2016). Antibacterial drug discovery in the resistance era. Nature 529, 336–343.

Chen, Z., Yang, H., and Pavletich, N.P. (2008). Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. Nature 453, 489–494.

Chen, C., Tian, W., Lei, X., Liang, J., and Zhao, J. (2018). CASTp 3.0: computed atlas of surface topography of proteins. Nucleic Acids Res. 46, 363–367.

Cline, D.J., Holt, S.L., and Singleton, S.F. (2007). Inhibition of Escherichia coli RecA by rationally redesigned N-terminal helix. Org. Biomol. Chem. 5, 1525–1528.

Datta, S., Ganesh, N., Chandra, N.R., Muniyappa, K., and Vijayan, M. (2003). Structural studies on MtRecA-nucleotide complexes: insights into DNA and nucleotide binding and the structural signature of NTP recognition. Proteins 50, 474–485.

Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. (2006). CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res. 34, 116–118.

Egelman, E.H., and Stasiak, A. (1986). Structure of helical RecA-DNA complexes. Complexes formed in the presence of ATP-gamma-S or ATP. J. Mol. Biol. 191, 677–697.

Feinstein, W.P., and Brylinski, M. (2015). Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets. J. Cheminf. 7, 1–10.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Firsching, A., Nickel, P., Mora, P., and Allolio, B. (1995). Antiproliferative and angiostatic activity of suramin analogues. Cancer Res. 55, 4957–4961.

Glaser, F., Pupko, T., Paz, I., Bell, R., Bechor-Shental, D., Martz, E., and Ben-Tal, N. (2003). ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics 19, 163–164.

Hu, J.J., Wong, N.-K., Ye, S., Chen, X., Lu, M.-Y., Zhao, A.Q., Guo, Y., Ma, A.C.-H., Leung, A.Y.-H., Shen, J., and Yang, D. (2015). Fluorescent probe HKSOX-1 for imaging and detection of endogenous superoxide in live cells and in vivo. J. Am. Chem. Soc. 137, 6837–6843.

Hu, J.J., Wong, N.-K., Lu, M.-Y., Chen, X., Ye, S., Zhao, A.Q., Gao, P., Kao, R.Y.-T., Shen, J., and Yang, D. (2016). HKOCl-3: a fluorescent hypochlorous acid probe for live-cell and in vivo imaging and quantitative application in flow cytometry and a 96-well microplate assay. Chem. Sci. 7, 2094–2099.

Jacques, D.A., and Trewhella, J. (2010). Small-angle scattering for structural biology-expanding the frontier while avoiding the pitfalls. Protein Sci. 19, 642–657.

Klinger, M., Freissmuth, M., Nickel, P., Stäbler-Schwarzbart, M., Kassack, M., Suko, J., and Hohenegger, M. (1999). Suramin and suramin analogs activate skeletal muscle ryanodine receptor via a calmodulin binding site. Mol. Pharmacol. 55, 462–472.

Kohanski, M.A., Dwyer, D.J., Hayete, B., Lawrence, C.A., and Collins, J.J. (2007). A common mechanism of cellular death induced by bactericidal antibiotics. Cell 130, 797–810.

Koonin, E.V. (1993). A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. Nucleic Acids Res. 21, 2541–2547.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

Lee, A.M., and Singleton, S.F. (2004). Inhibition of the Escherichia coli RecA protein: zinc(II), copper(II) and mercury(II) trap RecA as inactive aggregates. J. Inorg. Biochem. 98, 1981–1986.

Lee, A.M., Ross, C.T., Zeng, B.-B., and Singleton, S.F. (2005). A molecular target for suppression of the evolution of antibiotic resistance: inhibition of the Escherichia coli RecA protein by N6-(1-Naphthyl)-ADP. J. Med. Chem. 48, 5408–5411.

Leis, S., Schneider, S., and Zacharias, M. (2010). In silico prediction of binding sites on proteins. Curr. Med. Chem. 17, 1550–1562.

Lin, Z., Kong, H., Nei, M., and Ma, H. (2006). Origins and evolution of the recA/RAD51 gene family: evidence for ancient gene duplication and endosymbiotic gene transfer. Proc. Natl. Acad. Sci. U S A 103, 10328–10333.

Malkov, V.A., and Camerini-Otero, R.D. (1995). Photocrosslinks between single-stranded DNA and Escherichia coli RecA protein map to loops L1 (amino acid residues 157-164) and L2 (amino acid residues 195-209). J. Biol. Chem. 270, 30230–30233.

Nahrstedt, H., Schröder, C., and Meinhardt, F. (2005). Evidence for two RecA genes mediating

DNA repair in Bacillus megaterium. Microbiology *151*, 775–787.

Nautiyal, A., Patil, K.N., and Muniyappa, K. (2014). Suramin is a potent and selective inhibitor of Mycobacterium tuberculosis RecA protein and the SOS response: RecA as a potential target for antibacterial drug discovery. J. Antimicrob. Chemother. *69*, 1834–1843.

Nishimura, Y., McLaughlin, N.P., Pan, J., Goldstein, S., Hafenstein, S., Shimizu, H., Winkler, J.D., and Bergelson, J.M. (2015). The suramin derivative NF449 interacts with the 5-fold vertex of the enterovirus A71 capsid to prevent virus attachment to PSGL-1 and heparan sulfate. PLoS Pathog. *11*, e1005184.

Prabu, J.R., Manjunath, G.P., Chandra, N.R., Muniyappa, K., and Vijayan, M. (2008). Functionally important movements in RecA molecules and filaments: studies involving mutation and environmental changes. Acta Crystallogr. D Biol. Crystallogr. *64*, 1146–1157.

Rastelli, G., Del Rio, A., Degliesposti, G., and Sgobba, M. (2010). Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. J. Comput. Chem. *31*, 797–810.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. *4*, 406–425.

Schlacher, K., Leslie, K., Wyman, C., Woodgate, R., Cox, M.M., and Goodman, M.F. (2005). DNA polymerase V and RecA protein, a minimal mutasome. Mol. Cell *17*, 561–572.

Schlacher, K., Pham, P., Cox, M.M., and Goodman, M.F. (2006). Roles of DNA polymerase V and RecA protein in SOS damage-induced mutation. Chem. Rev. *106*, 406–419.

Seifert, H., Blondeau, J., and Dowzicky, M.J. (2018). *In vitro* activity of tigecycline and comparators (2014-2016) among key WHO 'priority pathogens' and longitudinal assessment (2004-2016) of antimicrobial resistance: a report from the T.E.S.T. study. Int. J. Antimicrob. Agents *52*, 474–484.

Shinohara, T., Ikawa, S., Iwasaki, W., Hiraki, T., Hikima, T., Mikawa, T., Arai, N., Kamiya, N., and Shibata, T. (2015). Loop L1 governs the DNA-binding specificity and order for RecA-catalyzed reactions in homologous recombination and DNA repair. Nucleic Acids Res. *43*, 973–986.

Singleton, S.F., Simonette, R.A., Sharma, N.C., and Roca, A.I. (2002). Intein-mediated affinity-fusion purification of the Escherichia coli RecA protein. Protein Expr. Purif. *26*, 476–488.

Story, R.M., and Steitz, T.A. (1992). Structure of the recA protein-ADP complex. Nature *355*, 374–376.

Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D.L., Pulcini, C., Kahlmeter, G., Kluytmans, J., Carmeli, Y., et al. (2018a). Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. Lancet Infect. Dis. *18*, 318–327.

Tacconelli, E., Sifakis, F., Harbarth, S., Schrijver, R., van Mourik, M., Voss, A., Sharland, M., Rajendran, N.B., Rodríguez-Baño, J., Bielicki, J., et al. (2018b). Surveillance for control of antimicrobial resistance. Lancet Infect. Dis. *18*, 99–106.

Tombline, G., and Fishel, R. (2002). Biochemical characterization of the human RAD51 protein. I. ATP hydrolysis. J. Biol. Chem. *277*, 14417–14425.

Tombline, G., Shim, K.-S., and Fishel, R. (2002a). Biochemical characterization of the human RAD51 protein. II. Adenosine nucleotide binding and competition. J. Biol. Chem. *277*, 14426–14433.

Tombline, G., Heinen, C.D., Shim, K.-S., and Fishel, R. (2002b). Biochemical characterization of the human RAD51 protein. III. Modulation of DNA binding by adenosine nucleotides. J. Biol. Chem. *277*, 14434–14442.

VanLoock, M.S., Yu, X., Yang, S., Lai, A.L., Low, C., Campbell, M.J., and Egelman, E.H. (2003). ATP-Mediated conformational changes in the RecA filament. Structure *11*, 187–196.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). 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.

Wass, M.N., Kelley, L.A., and Sternberg, M.J.E. (2010). 3DLigandSite: predicting ligand-binding sites using similar structures. Nucleic Acids Res. *38*, 469–473.

Wigle, T.J., and Singleton, S.F. (2007). Directed molecular screening for RecA ATPase inhibitors. Bioorg. Med. Chem. Lett. *17*, 3249–3253.

Wigle, T.J., Lee, A.M., and Singleton, S.F. (2006). Conformationally selective binding of nucleotide analogues to escherichia coli RecA: a ligand-based analysis of the RecA ATP binding site. Biochemistry *45*, 4502–4513.

Wigle, T.J., Sexton, J.Z., Gromova, A.V., Hadimani, M.B., Hughes, M.A., Smith, G.R., Yeh, L.-A., and Singleton, S.F. (2009). Inhibitors of RecA activity discovered by high-throughput screening: cell-permeable small molecules attenuate the SOS response in *Escherichia coli*. J. Biomol. Screen. *14*, 1092–1101.

Wright, G.D. (2016). Antibiotic adjuvants: rescuing antibiotics from resistance. Trends Microbiol. *24*, 862–871.

Xing, X., and Bell, C.E. (2004a). Crystal structures of *Escherichia coli* RecA in a compressed helical filament. J. Mol. Biol. *342*, 1471–1485.

Xing, X., and Bell, C.E. (2004b). Crystal structures of *Escherichia coli* RecA in complex with MgADP and MnAMP−PNP. Biochemistry *43*, 16142–16152.

Xu, J., Zhao, L., Xu, Y., Zhao, W., Sung, P., and Wang, H.-W. (2016). Cryo-EM structures of human RAD51 recombinase filaments during catalysis of DNA-strand exchange. Nat. Struct. Mol. Biol. *24*, 40.

Yakimov, A., Pobegalov, G., Bakhlanova, I., Khodorkovskii, M., Petukhov, M., and Baitin, D. (2017). Blocking the RecA activity and SOS-response in bacteria with a short α-helical peptide. Nucleic Acids Res. *45*, 9788–9796.

Yu, X., Jacobs, S.A., West, S.C., Ogawa, T., and Egelman, E.H. (2001). Domain structure and dynamics in the helical filaments formed by RecA and Rad51 on DNA. Proc. Natl. Acad. Sci. U S A *98*, 8419.

Yu, X., VanLoock, M.S., Yang, S., Reese, J.T., and Egelman, E.H. (2004). What is the structure of the RecA-DNA filament? Curr. Protein Pept. Sci. *5*, 73–79.

Zheng, W., Liang, Y., Zhao, H., Zhang, J., and Li, Z. (2015). 5,5'-Methylenedisalicylic acid (MDSA) modulates SarA/MgrA phosphorylation by targeting Ser/Thr phosphatase Stp1. ChemBioChem *16*, 1035–1040.

Zheng, W., Cai, X., Xie, M., Liang, Y., Wang, T., and Li, Z. (2016). Structure-based identification of a potent inhibitor targeting Stp1-mediated virulence regulation in *Staphylococcus aureus*. Cell Chem. Biol. *23*, 1002–1013.

Zhou, Z., Yuan, Y., Zhou, S., Ding, K., Zheng, F., and Zhan, C.-G. (2017). Selective inhibitors of human mPGES-1 from structure-based computational screening. Bioorg. Med. Chem. Lett. *27*, 3739–3743.

Zhou, Z., Yuan, J., Pan, Q., Mo, X.-M., Xie, Y.-L., Yin, F., Li, Z., and Wong, N.-K. (2019). Computational elucidation of the binding mechanisms of curcumin analogues as bacterial RecA inhibitors. RSC Adv. *9*, 19869–19881.

Supplemental Information

Structural insights into the inhibition

of bacterial RecA by naphthalene

polysulfonated compounds

Ziyuan Zhou, Qing Pan, Xinchen Lv, Jing Yuan, Yang Zhang, Ming-Xia Zhang, Ming Ke, Xiao-Mei Mo, Yong-Li Xie, Yingxia Liu, Ting Chen, Mingchan Liang, Feng Yin, Lei Liu, Yiqing Zhou, Kun Qiao, Rui Liu, Zigang Li, and Nai-Kei Wong

Grid box parameters

Table S1. Grid box parameters applied in Autodock Vina calculations in this study, Related to Figure 3 and Figure 6.

Pocket	Center coordinates[a] (x, y, z)	Box sizes[a] (x, y, z)
A	(42.0, 19.0, 5.0)	(20.0, 20.0, 24.0)
B	(55.0, 24.0, -17.0)	(32.0, 24.0, 24.0)
C	(53.0, 45.0, -5.0)	(30.0, 22.0, 22.0)
D	(56.0, 18.0, 23.0)	(24.0, 26.0, 24.0)

[a] Coordinate values or size values are all in Å.

Table S2. Grid box parameters applied in LeDock calculations in this study, Related to Figure 3 and Figure 6.

Pocket	box boundaries		
	x	Y	z
A	(32.0, 52.0)	(9.0, 29.0)	(-7.0, 19.0)
B	(39.0, 71.0)	(12.0, 36.0)	(-39.0, -5.0)
C	(38.0, 68.0)	(34.0, 56.0)	(-16.0, 6.0)
D	(44.0, 68.0)	(5.0, 31.0	(11.0, 35.0)

Table S3. Primers used in this study, Related to Figure 5 and Figure 7.

Primer	Sequence (5'→3')	Reference
qPCR-gapA-F	GCTCGTAAACACATCACCGC	This study
qPCR-gapA-R	CGATGTCCTGGCCAGCATAT	This study
qPCR-recA-F	GCGTGAAAGTGGTGAAGAAC	This study
qPCR-recA-R	CCTGACCGATCTTCTCACCT	This study
qPCR-umuD-F	TGTTGACGGTGAGTTTACGG	This study
qPCR-umuD-R	CACCAAAGACATCCAGCGTA	This study
qPCR-umuC-F	TGGGGGATTTCTTCAGTCAG	This study
qPCR-umuC-R	TTCCTCTGCCCTCTTTAGCA	This study
qPCR-dinI-F	CTTTCCCGCCGTATTCAGTA	This study
qPCR-dinI-R	GAATTTCGCTAATGCGCTGT	This study
qPCR-soxR-F	GCATTAAAGCGCTGCTAACC	This study
qPCR-soxR-R	CGTTTATATCGCCGCTGATT	This study
qPCR-sodA-F	GCCTGTTCTGGAAAGGTCTG	This study
qPCR-sodA-R	GCCAGTTTATCGCCTTTCAG	This study
BamHI-recA-F	CTAGGATCCATGGCTATCGACGAAAACAAAC	This study
XhoI-recA-R	CGTCGCTCGAGTTAAAAATCTTCGTTAGTTTCTG	This study
recA-Y103A-F	CACGCGCTGGACCCAATCGCCGCACGTAAACTGGGCGTC	This study
recA-Y103A-R	GACGCCCAGTTTACGTGCGGCGATTGGGTCCAGCGCGTG	This study

Table S5. Worksheet for RecA (or RadA) from 20 species of probiotics or archaea[a], Related to Figure 1.

Entry	Name	NCBI accession	G$^+$ or G$^-$	Relationship to host
B1	*Aggregatibacter actinomycetemcomitans*	PHO21993	-	Probiotics
B2	*Bifidobacterium bifidum*	BBA55515	+	Probiotics
B3	*Bifidobacterium breve*	AAD20261	+	Probiotics
B4	*Bifidobacterium dentium*	OQM55798	+	Probiotics
B5	*Bifidobacterium infantis*	CEE99551	+	Probiotics
B6	*Bifidobacterium longum*	AAN25214	+	Probiotics
B7	*Eikenella corrodens*	SNW09484	-	Probiotics
B8	*Lactobacillus acidophilus*	ASN46574	+	Probiotics
B9	*Lactobacillus brevis*	PBQ23313	+	Probiotics
B10	*Lactobacillus casei*	PCL43130	+	Probiotics
B11	*Lactobacillus plantarum*	CDN27496	+	Probiotics
B12	*Lactobacillus rhamnosus*	PIN33973	+	Probiotics
B13	*Rothia dentocariosa*	PEN16378	+	Commensal
B14	*Rothia mucilaginosa*	EHB87709	+	Commensal
B15	*Akkermansia muciniphila*	PND15691	-	Probiotics
B16	*Elizabethkingia anophelis*	ATC48335	-	Probiotics
Archaea species				
B17[b]	*Haloquadratum walsbyi*	WP_011572623	-	Archaea
B18[b]	*Methanosarcina acetivorans*	WP_011023460	-	Archaea
B19[b]	*Nanoarchaeum equitans*	Q74MX9	-	Archaea
B20[b]	*Sulfolobus solfataricus*	AAC44123	-	Archaea

[a] References for Table S4 to Table S9 are available at request at zhouziyuan@pku.edu.cn or wongnksz@163.com

[b] RadA proteins.

Table S6. Worksheet for RecA (or Rad51) from 20 prokaryotic or eukaryotic species. Related to Figure 1.

Entry	Species	NCBI accession	G$^+$ or G$^-$
C1	*Aeromonas caviae*	AKE49279	-
C2	*Aeromonas hydrophila*	AKE49297	-
C3	*Elizabethkingia meningoseptica*	OOH98005	-
C4	*Plesiomonas shigelloides*	PVU67488	-
C5	*Planktothrix agardhii*	KEI68811	-
C6	*Arthrospira platensis*	AAA75282	-
Eukaryotic species			
C7	*Microcystis aeruginosa*	ARI83212	N[a]
C8	*Dictyostelium discoideum*	FAA00018	N
C9	*Cyanidioschyzon merolae*	XP_005538367	N
C10	*Chlamydomonas reinhardtii*	XP_001692733	N
C11	*Arabidopsis thaliana*	OAO95923	N
C12	*Zea mays*	NP_001104918	N
C13	*Oryza sativa*	BAB85492	N
C14	*Caenorhabditis elegans*	AAD10194	N
C15	*Candida albican*	EEQ43332	N
C16	*Drosophila melanogaster* (fruit fly)	BAA04580	N
C17	*Gallus gallus domesticus* (chicken)	NP_990504	N
C18	*homo sapien* (human)	CAG38796	N
C19	*mus musculus* (house mouse)	NP_035364	N
C20	*Saccharomyces cerevisiae* (brewer's yeast; budding yeast)	CAA45563	N

[a] N, not applicable

Table S7. Molecular properties for compounds analyzed in the training set. Related to Figure 3 and Scheme 1.

Entry	Molecular properties		
	Net Charge	Molecular weight	Molecular span
cpd01	-6	1289	30
cpd02	-2	594	18
cpd03	0	652	26
cpd04	-2	891	18
cpd05	-1	298	11
cpd06	0	285	12
cpd07	0	452	16
cpd08	0	454	17
cpd09	1	371	20
cpd10	0	372	18
cpd11	0	373	22
cpd12	-3	550	21
cpd13	-3	514	20
cpd14	-3	528	23

Table S8. Predicted binding sites and key amino acid residues involved in RecA-inhibitor interactions. Related to Figure 2.

Related pockets	Key residues	SA^a area (Å^2)	SA volume (Å^3)	Reported functions
A	Gly66, Pro67, **Glu68**^{b}, Ser69, Ser70, Gly71, <u>**Lys72**</u>^{c}, <u>Thr73</u>, Asp94, <u>Glu96</u>, Ala98, Lys99, <u>Tyr103</u>, **Asp144**, Ser145, Gln194, Ile195, Arg196	90.6	48.4	Protein-Protein interaction, ATPase function
B	Glu38, Thr39, Ile40, Ser41, Gly43, Ala53, Gly54, Gln78, Ala81, Ala82, Ala83, Arg85, Glu86, Leu107, Glu226, Glu266, Gly267, Ile268, Asn269, Phe270, Thr271, Lys310, Glu320, Lys321, Arg324, Glu325, Leu328, Pro331, Asn332, Ser333	497.9	499.6	Unknown
C	Ser46, Ile49, Ala50, Glu241, Gln257, Ala258, Glu259, Phe260, Gln261, Asn269, Gly272, Gly273, Asp276, Leu277, Lys280, Lys302, Leu327	142.6	121.2	dsDNA binding
D	His163, Gly165, Leu166, Ala167, Ala168, Arg196, Met197, Lys198, Ile199, Thr209, Thr210, Gly211, Gly212	91.7	143.4	ssDNA binding

^{a} SA, solvent-accessible.

^{b} Amino acid residues in bold are key residues in the highly conserved domains Walker A/B.

^{c} Underlined amino acid residues are reported to be implicated in key interactions with ADP or ATP analogues in crystal structures.

Table S9. Typical resistance attributes of major pathogens in Figure 1. Related to Figure 1.

| Bacteria species and AMR	Carbapenems	Cephalosporins	Vancomycin	Penicillins	Quinolones	Clarithromycin
Acinetobacter baumannii	R (resistant)	R	R	R	R	R
Acinetobacter seifertii	R	unknown	unknown	unknown	unknown	unknown
Bacillus anthracis	unknown	R	unknown	unknown	unknown	unknown
Burkholderia cepacia	R	R	unknown	unknown	R	unknown
Campylobacter coli	unknown	unknown	unknown	unknown	R	unknown
Campylobacter jejuni	unknown	unknown	unknown	unknown	R	unknown
Enterobacter aerogenes	R	R	unknown	unknown	unknown	unknown
Enterobacter cloacae	R	R	unknown	unknown	unknown	unknown
Enterococcus faecalis	R	R	R	unknown	unknown	unknown
Enterococcus faecium	R	R	R	unknown	unknown	unknown
Enterococcus gallinarum	R	unknown	unknown	unknown	unknown	unknown
Escherichia coli	R	R	unknown	unknown	unknown	unknown
Escherichia fergusonii	R	R	unknown	unknown	unknown	unknown
Escherichia vulneris	R	unknown	unknown	unknown	unknown	unknown
Francisella novicida	R	R	unknown	unknown	unknown	unknown
Francisella tularensis	R	R	unknown	unknown	unknown	unknown
Haemophilus aegyptius	R	unknown	unknown	unknown	unknown	unknown
Haemophilus influenzae	R	R	R	R	unknown	unknown
Helicobacter canis	unknown	unknown	unknown	R	unknown	unknown
Helicobacter pylori	R	unknown	R	R	R	R
Klebsiella pneumoniae	R	R	R	R	R	R
Mycobacterium tuberculosis	R	R	R	R	R	R
Neisseria gonorrhoeae	R	R	R	R	R	R
Pseudomonas aeruginosa	R	R	R	R	R	R
Salmonella bongori	unknown	unknown	unknown	unknown	R	unknown
Salmonella enteritidis	unknown	unknown	unknown	unknown	R	unknown
Salmonella Typhi	unknown	unknown	unknown	unknown	R	unknown
Salmonella Typhimurium	unknown	unknown	unknown	unknown	R	unknown
Shigella boydii	unknown	unknown	unknown	unknown	R	unknown
Shigella dysenteriae	unknown	unknown	unknown	unknown	R	unknown
Shigella flexneri	unknown	unknown	unknown	unknown	R	unknown
Shigella sonnei	unknown	unknown	unknown	unknown	R	unknown
Staphylococcus aureus	R	R	R	R	R	R
Staphylococcus epidermidis	R	R	R	R	R	R
Staphylococcus haemolyticus	R	R	R	R	R	R
Streptococcus pneumoniae	R	R	R	R	unknown	R
Streptococcus anginosus	R	R	R	R	unknown	R
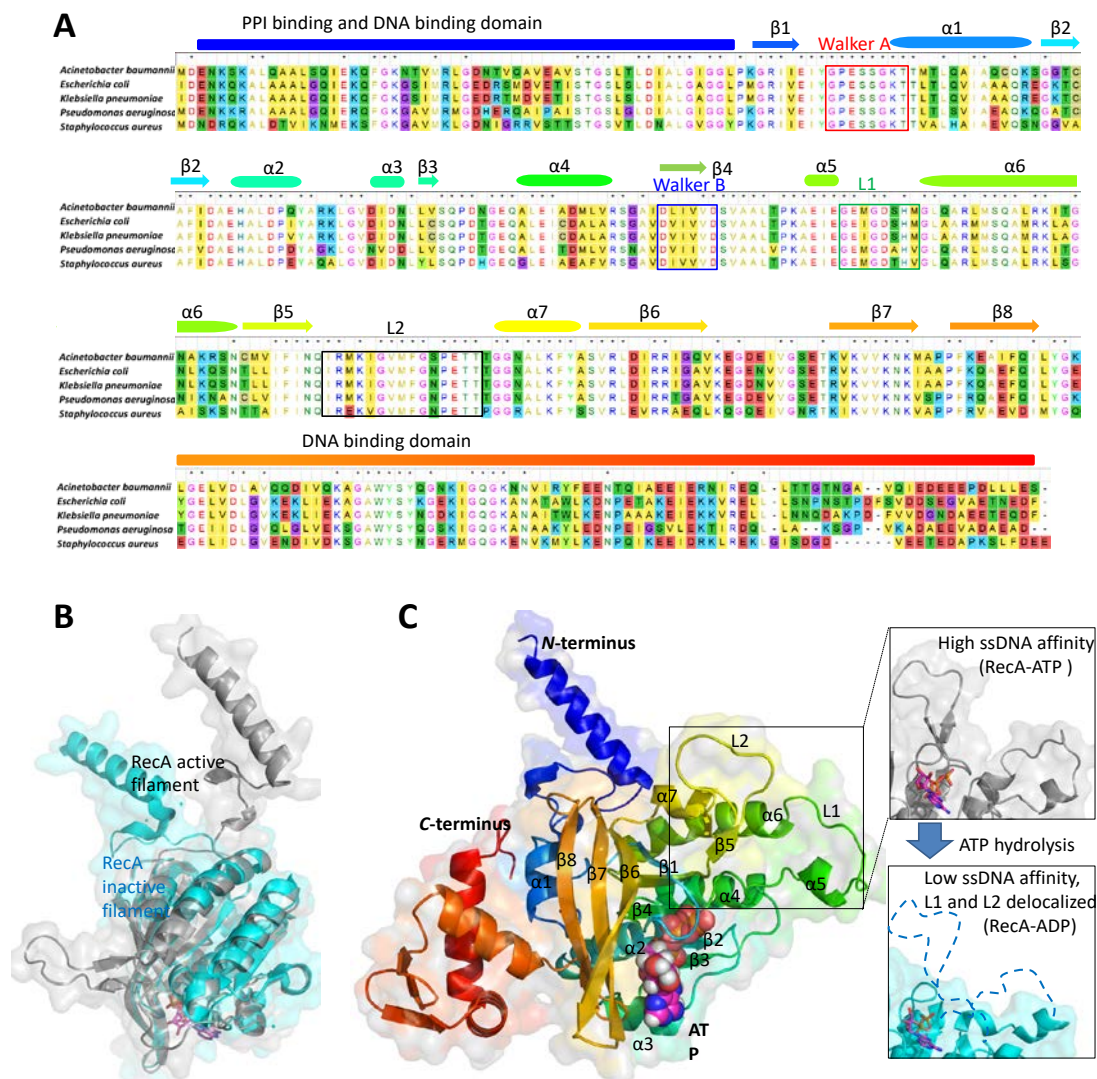
Figure S1. Bacterial RecA structural attributes and functional implications. Related to Figure 1 and Figure 2. **(A)** Comparison of *A. baummannii* (ASF77594), *E. coli* (ANK02966), *K. penumoniae* (CDO12858), *P. aeruginosa* (KFL09345) and *S. aureus* (AAK15276) RecA proteins. Secondary structures are indicaed on top of the alignment from blue to red. Colors used corresponds to the rainbow color below. In addition, the Walker A and Walker B regions implicated in nucleotide binding are depicted, along with the L1 and L2 regions, which are putatively important to ssDNA binding. **(B)** Crystal structures of *E. coli* RecA with the ATP minic ADP-AlF$_4$ (PDB 3CTM) and with ADP (PDB 1XMV) have been reported. **(C)** A close-up view of the nucleotide-bound form of RecA (colored in gray in the left and upper right inset) shows that when *E. coli* RecA protein is bound with ssDNA and ADP-AlF$_4$ (modified to ATP), it will adopt a high-affinity ssDNA binding state. In contrast, when bound to ADP, *E. coli* RecA protein will assume a low-affinity binding state with ssDNA (colored cyan in the left and lower right inset).

The active and inactive states are two distinct conformations of RecA filament (**Figure S1B**) critical to its biological functions. An inactive filament can form in the absence of DNA and ATP or stable ATP analogues. In contrast, a filament formed in the presence of a non-hydrolysable ATP analog and DNA is considered to be active and capable of catalyzing DNA strand exchange. ATPase activity of RecA protein is DNA-dependent, whose regulation in eubacteria species is allosteric. RecA was found to have a lower affinity for DNA binding in its

inactive state (ADP-bound) than that in its active state (ATP-bound). Majority of the reported crystallographic filaments of RecA are in their inactive states without the loop regions, with the ssDNA-binding L1 and L2 regions being in a disordered state not readily resolvable in structural analysis. In this sense, while most of their X-ray structures are considered inactive, crystallographic filaments representing the active state of RecA, RadA, and Rad51 do exist, which include the crystal structures of archaeal RadA reported by Wu *et al.*, and that of yeast Rad51 reported by Rice *et al*. Unfortunately, due to the absence of DNA, these structures are later considered altered conformations somewhat different from their authentic inactive states, rather than the active DNA-binding conformation.

Interfaces critical to RecA protein interactions display different conformational variations associated with ATP- or ADP-binding status, which gives rise to high-affinity or low-affinity ssDNA binding potencies at pocket D. A detailed structural view of RecA protien and an intuitionistic exhibition of the differences between the active and inactive states are as shown in **Figure S1**.
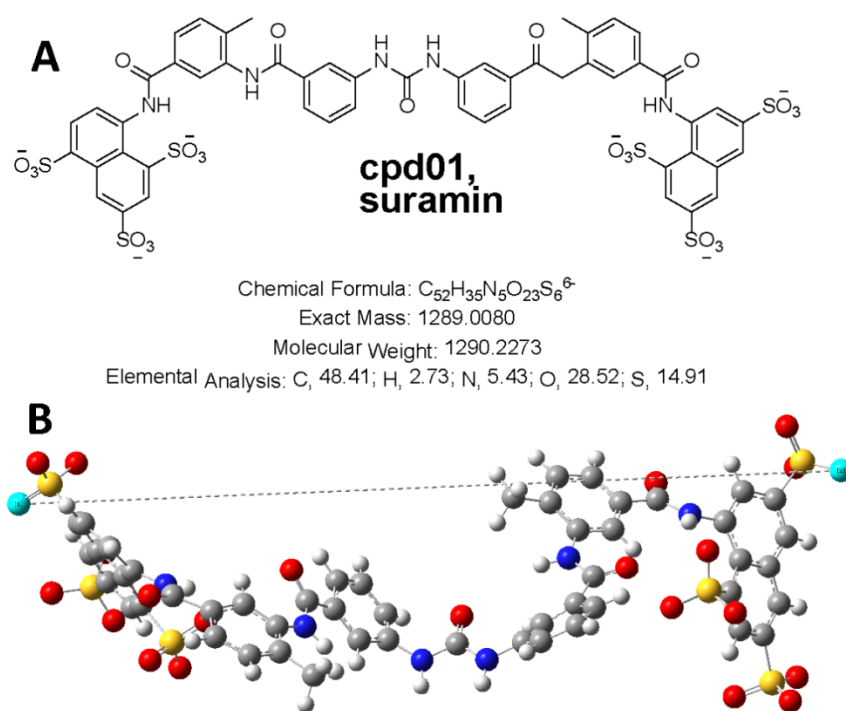

Figure S2. Molecular properties of suramin. Related to Figure 3. **(A)** Net charge, molecular weight. **(B)** Molecular span.
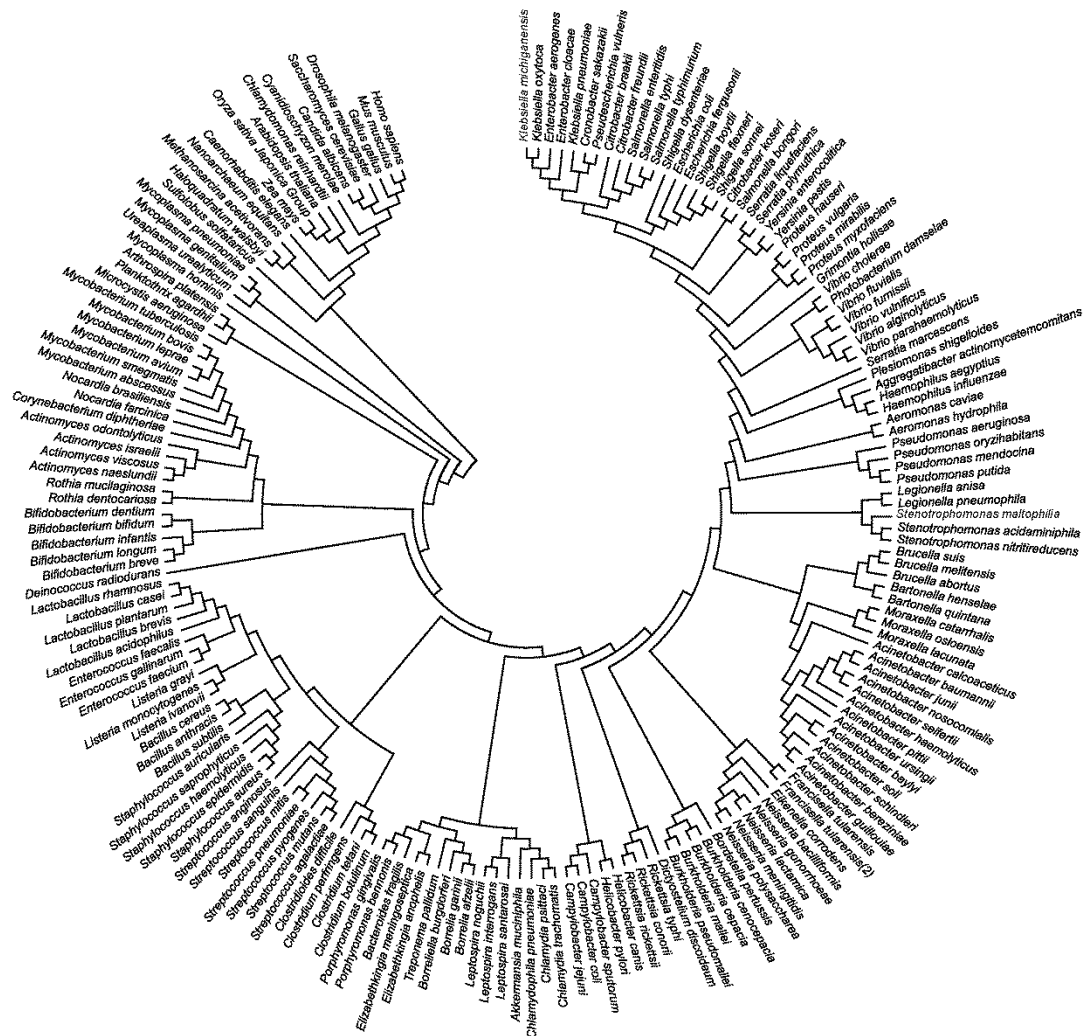
Figure S3. Molecular phylogenetic analysis on RecA by Maximum Likelihood method. Related to Figure 1. Evolutionary relationships were inferred by using the Maximum Likelihood method based on the JTT matrix-based model. A tree with the highest log likelihood (-17446.98) is as shown. Initial tree(s) for heuristic searches were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated by using a JTT model, and then selecting the topology with superior log likelihood values. The analysis involved 185 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 217 positions in the final dataset. Evolutionary analyses were conducted with MEGA7. The bacterial RecA protein and its eukaryotic homolog Rad51 are crucial for the initiation of homologous recombination, DNA damage repair and maintenance of genomic stability. While only one RecA gene has been reported for each eubacteria species, some of the archaeal species reportedly contain two RecA-like genes, namely RadA and RadB, and multiple Rad51-like genes exist in vertebrates and in plants.
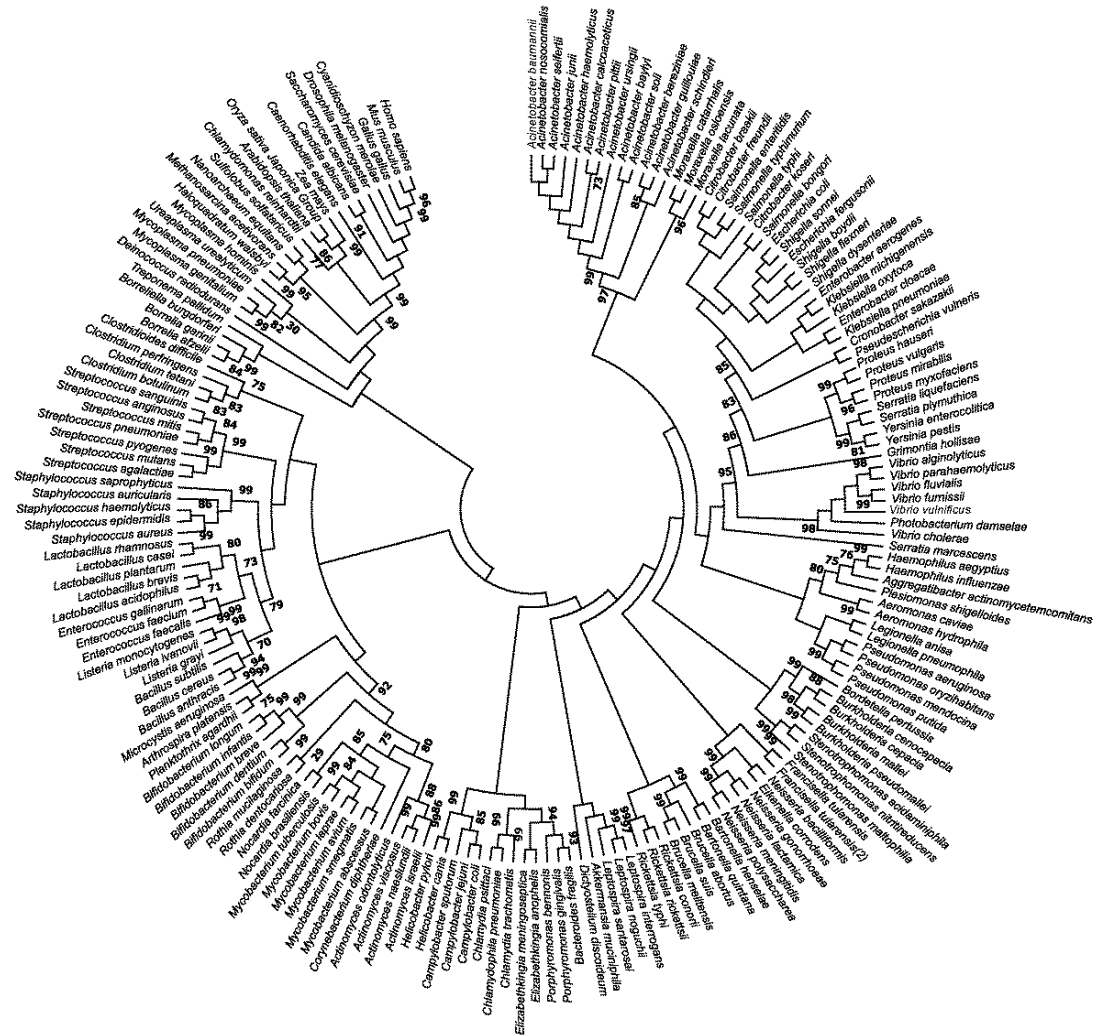
Figure S4. Phylogenetic analysis on RecA by Minimum Evolution method. Related to Figure 1. Evolutionary relationships were inferred by using the Minimum Evolution method. Bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 0% bootstrap replicates are collapsed. Evolutionary distances were computed by using the Poisson correction method and are presented in terms of the number of amino acid substitutions per site. Neighbor-joining algorithm was used to generate an initial tree. The analysis involved 185 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 217 positions in the final dataset. Evolutionary analyses were conducted with MEGA7.
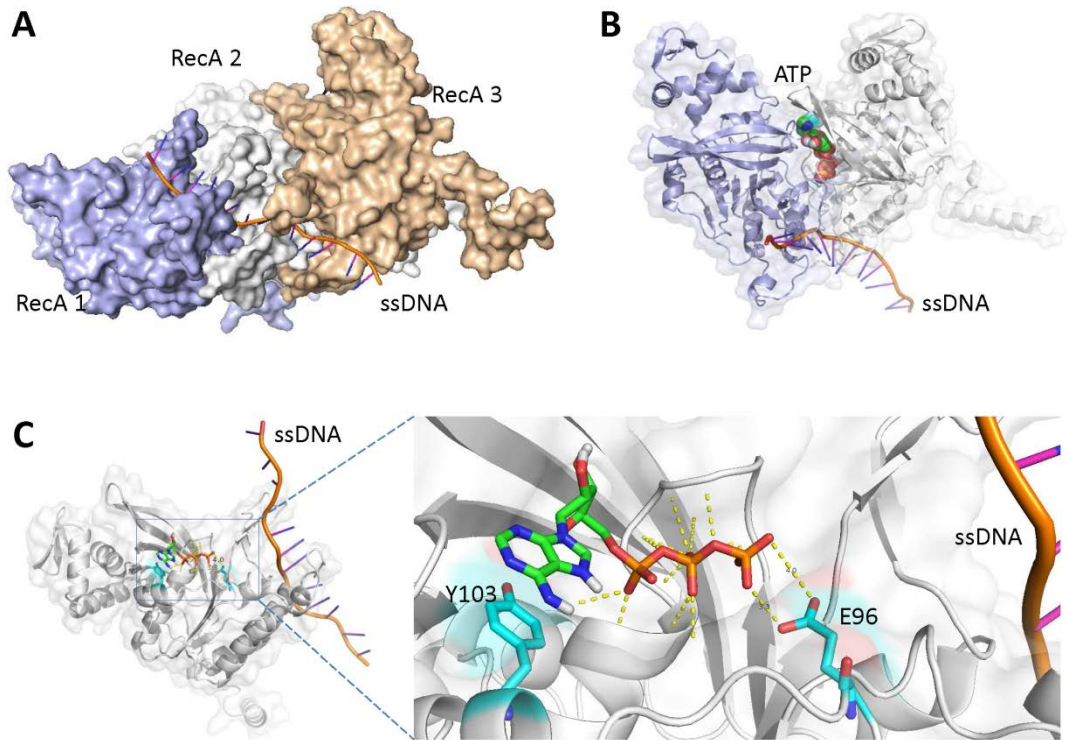
Figure S5. The ATP binding site of RecA plays a key role in the filamentous status of the RecA-ssDNA complex (PDB ID, 3CMT). Related to Figure 2. **(A)** Structure of RecA polymer. **(B)** ATP is sandwiched between two RecA monomers. **(C)** Y103 and E96 are crucial for ATP binding.
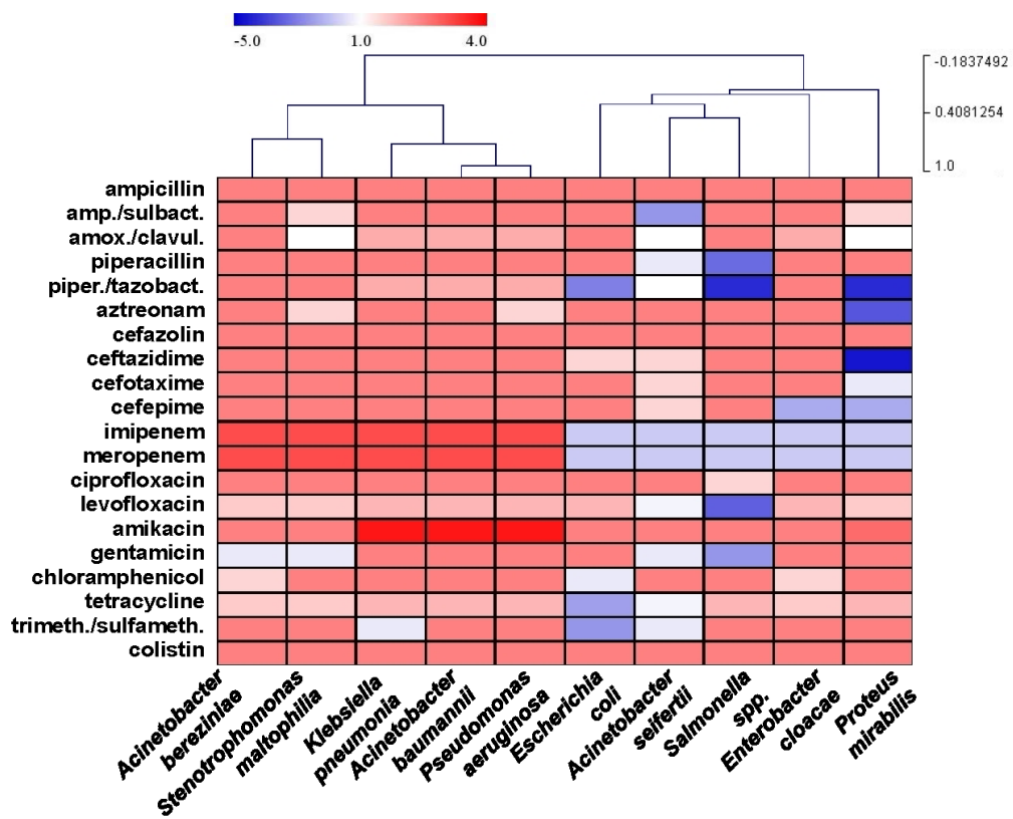
Figure S6. Heat map illustrating representative drug resistance profiles of clinical Gram-negative bacteria. Related to Figure 1. Laboratory-determined relative antibiotic susceptibility of the Gram-negative bacterial species is shown here by heat map. Columns represent antibiotics and rows represent bacterial strains. Red blocks indicate resistance, whereas blue blocks indicate antibiotic susceptibility and white blocks indicate intermediate degree of resistance.
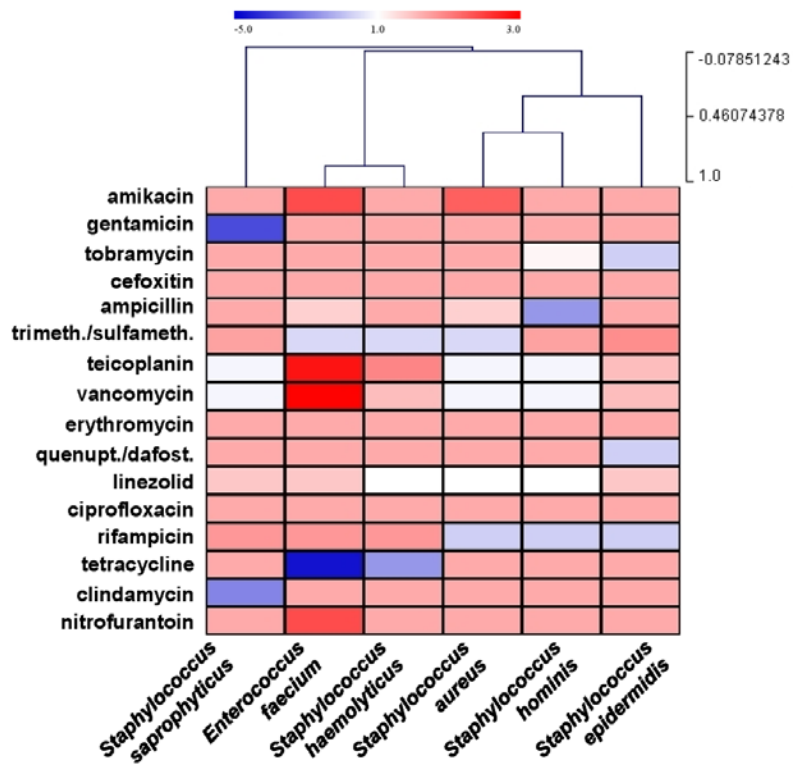
Figure S7. Heat map illustrating representative drug resistance profiles of clinical Gram-positive bacteria. Related to Figure 1. Laboratory-determined relative antibiotic susceptibilities of Gram-positive bacterial species are as shown in the heat map. Columns represent antibiotics and rows represent bacterial strains. Red blocks indicate resistance, whereas blue blocks indicate antibiotic susceptibility and white blocks indicate intermediate degree of resistance.

Transparent Methods

COMPUTATIONAL DETAILS

Data retrieval of protein sequences

Protein sequences of the RecA for *E. coli* and other bacteria species were retrieved form NCBI (National Center for Biotechnology Information) databases and were used as queries. 185 reported amino acid sequences from representative species of bacteria, eukaryotes, and archaea were selected. Their accession numbers and other details are listed in **Tables S4**-**S6**.

Sequence alignment and phylogenetic analyses

Preliminary multiple sequence alignment analysis was performed by using MEGA 7.0 (Kumar et al., 2016) with default parameter or settings. A comprehensive NJ tree (**Figure 1**), in addition to the preliminary ML tree and ME tree (**Figure S3** and **Figure S4**), was generated by using MEGA 7.0 to calculate the number and composition of subgroups. Reliability of interior branches was assessed with 1,000 bootstrap resampling by using ''pairwise deletion option'' of the amino acid sequences with gamma parameters. Phylogenetic tree files were viewed by using MEGA 7.0 and refined in PowerPoint.

Molecular docking of RecA and inhibitors

In order to clarify the binding mode of RecA with its inhibitors, molecular docking was performed by using the AutoDock Vina program.(Morris et al., 2009) (Trott et al, 2010) Molecular dynamics (MD) simulations were carried out by using the PMEMD module of the AMBER 16 package(D.A. Case et al. 2017).

Modeling of RecA-inhibitor complexes

Initial X-ray crystal structures of the RecA complex (inactive state of *E. coli* RecA) was downloaded from the RCSB database (PDB entry: 3CMT). A constructed system was subjected to multiple steps of energy minimization by using the AMBER 16 package(D.A. Case et al. 2017).

Energy minimization

Energy minimization was initially performed for 100,000 steps by using the steepest-descent method and then another 100,000 steps by using the conjugated gradient method implemented with the PMEMD module of the AMBER 16 package (D.A. Case et al. 2017). In order to effectively avoid steric hindrance between protein atoms and atoms of the protein environment, additional forces with a force constant of 50 kcal/mol were first applied as constraints to all atoms of protein, such that the rest of the solvent environment molecules became energy-minimized. Here, for convenience, the protein environment was defined as all solvent water molecules, Na^+ ions, Cl^- ions, and lipid molecules. Constraints were then applied only to the atoms of the protein, while the atomic positions of the protein environment were taken to be energy-minimized. In the third step, constraints were applied to atoms of the protein backbone, while atoms of the protein side-chains and atoms of the protein environment were taken to be energy-minimized. Finally, the whole system was energy-minimized without any constraints.

Molecular dynamics (MD) simulations

Each simulation system was gradually heated to a temperature (T) of 298 K by applying Langevin dynamics,

and was allowed to be equilibrated for 50 ns. As long-time simulations were needed for equilibration of lipids, the backbone atoms of the protein were fixed in the first 5 ns of the equilibration stage. Then, all of the restraints were removed in the remaining 10 ns of equilibration. During MD simulations, a 10.0 Å non-bonded interaction cutoff and a 2.0 Å non-bonded list updating cutoff were used. Motion from the mass center of the system was removed every 1,000 steps. The particle-mesh Ewald (PME) method was used to treat long-range electrostatic interactions. Lengths of covalent bonds involving a hydrogen atom were fixed with the SHAKE algorithm, enabling the use of a 2 fs time step to numerically integrate equations of motion. The production of MD simulations was allowed to proceed for 10 ns with a periodic boundary condition (PBC) at T = 298 K.

BIOLOGICAL METHODS

Bacterial strains

    E. coli K-12 (sub-strain MG1655) supplied the template of *recA* cloning. The *E. coli* strain DH5α was used in cloning and mutagenesis. The *E. coli* strain BL21(DE3) was used for recombinant protein overexpression. The *E. coli* type strain ATCC 25922 was used in drug treatments and RT-qPCR.

Cell growth and pharmacological treatments

    E. coli ATCC 25922 was cultured in LB (Luria-Bertani broth) at 37 °C, 250 rpm to log phase ($OD_{600\,nm}$ = 0.4). Ten mL of the cells were treated with 50 μM suramin, 0.25 mg/mL ciprofloxacin, or 0.25 mg/mL ciprofloxacin plus 50 μM suramin, respectively, followed by incubation with shaking for 60 min at 37 °C. Seven mL of the cell culture was pelleted by centrifugation at 8,000×g for 5 min and stored in -80°C until use.

RNA extraction

    Cells were lysed in 100 μL lysis buffer [TE with 20 U RNase inhibitor (Applied Biosystems) and 4 mg lysozyme (Sigma)] for 3 min at RT. One mL RNAiso plus (Takara) was added and mixed well. Subsequently, RNA was extracted by mixing with 400 μL chloroform and centrifugation at 12,000×g twice. Supernatant was recovered and mixed with 4-fold volume of 100% ethanol. RNA was then bound, washed and eluted following the instructions of PureLink miRNA Isolation Kit (Invitrogen). Total RNA was quantified and qualified by NanoDrop 2000 (Thermo) and agarose electrophoresis.

Reverse transcription and quantitative real-time PCR (RT-qPCR)

    One μg total RNA was reverse transcribed into cDNA by using PrimeScript RT reagent Kit with gDNA Eraser (Takara) following its instructions. Primers for real-time PCR were designed with Primer3 (http://primer3.ut.ee/) and listed in **Table S3**. Gene expression of glyceraldehyde 3-phosphate dehydrogenase A (*gapA*) was set as the endogenous control. The qPCR assay was run on a qTower3 real-time PCR System (Analytik Jena AG) with a default program as 95°C 30 s for 1 cycle, and 95°C 5 s, 60°C 30 min for 40 cycles. The relative quantity (RQ) was calculated by Ct values from all PCR reactions in triplicate using the ΔΔCt method (**Figure 5**). Statistical significance was determined by Student's t-test at $*p < 0.05$, $**p < 0.01$ and $***p < 0.001$.

Construction of pSumo-recA and pSumo-recA-Y103A

    Full-length *recA* from *E. coli* K-12 MG1655 was PCR-amplified and cloned into *Bam*H I and *Xho* I sites of a modified pET-Sumo vector (pSumo-Duet) to generate pSumo-recA, such that the target RecA protein fuses with a 6×His-tag and a yeast Sumo-tag at the *N*-terminus. The $recA^{Y103A}$ mutant was generated by *Dpn* I mediated site-directed mutagenesis by using pSumo-recA as a template. The sequencing-verified plasmids pSumo-recA and pSumo-recA-Y103A were used to transform *E. coli* BL21(DE3) for recombinant protein overexpression, respectively.

Induction and purification of RecA and RecA^{Y103A}

    E. coli BL21(DE3) harboring pSumo-recA or pSumo-recA-Y103A was grown in LB broth to $OD_{600\,nm}$ ~0.6. Expression of the recombinant proteins RecA and RecA^{Y103A} was induced by adding IPTG to a final concentration of 0.2 mM, followed by further culture at 16°C overnight. Bactdrial cells were harvest by centrifugation, and re-suspended in a buffer containing 25 mM Tris (pH 8.0), 500 mM NaCl, 5% glycerol, 0.2 mM PMSF, and 15 mM

imidazole. After sonication and centrifugation, the supernatant was used to recover recombinant RecA and RecA^{Y103A} proteins by a sequential purification procedure with Ni-NTA (GE Healthcare), Q-Sepharose FF (GE Healthcare), and Superdex G200 columns (GE Healthcare) at 4°C. The 6×His-tag and Sumo-tag were subsequently removed by SUMO protease-mediated cleavage. Purified proteins were dialyzed against a buffer containing 100 mM NaCl, and 20 mM HEPES (pH 7.0) at 4°C overnight, and concentrated with Amicon Ultra-15 (MWCO 10k) (Millipore). Concentrations and purity of proteins were determined by NanoDrop 2000 (Thermo) and SDS-PAGE, respectively.

ITC-based RecA-NPS binding assay

Recombinant RecA and RecA^{Y103A} proteins were diluted to 0.03-2.0 mM with a dialysis buffer (100 mM NaCl, 20 mM HEPES, at pH 7.0). Test compounds were also dissolved in the same buffer with concentrations adjusted to 0.05-0.3 mM. ITC-based binding assay was performed by using a standard protocol on MicroCal PEAQ-ITC instrument (Malvern) at 25°C. Data were fitted into a curve by using Origin 7.0 (https://www.originlab.com).

Supplemental References

Kumar, S.; Stecher, G.; Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol. 33*, 1870-1874.

Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem 30*, 2785-91.

Trott, O.; Olson, A. J. (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem. 31*, 455-461.

D.A. Case, D. S. C., T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman (2017) AMBER 16. *University of California, San Francisco.*