Electronic Supplementary Material (ESI) for MedChemComm.

Electronic Supplementary Information for

Identification of an auxiliary druggable pocket in DNA gyrase ATPase domain using fragment probes

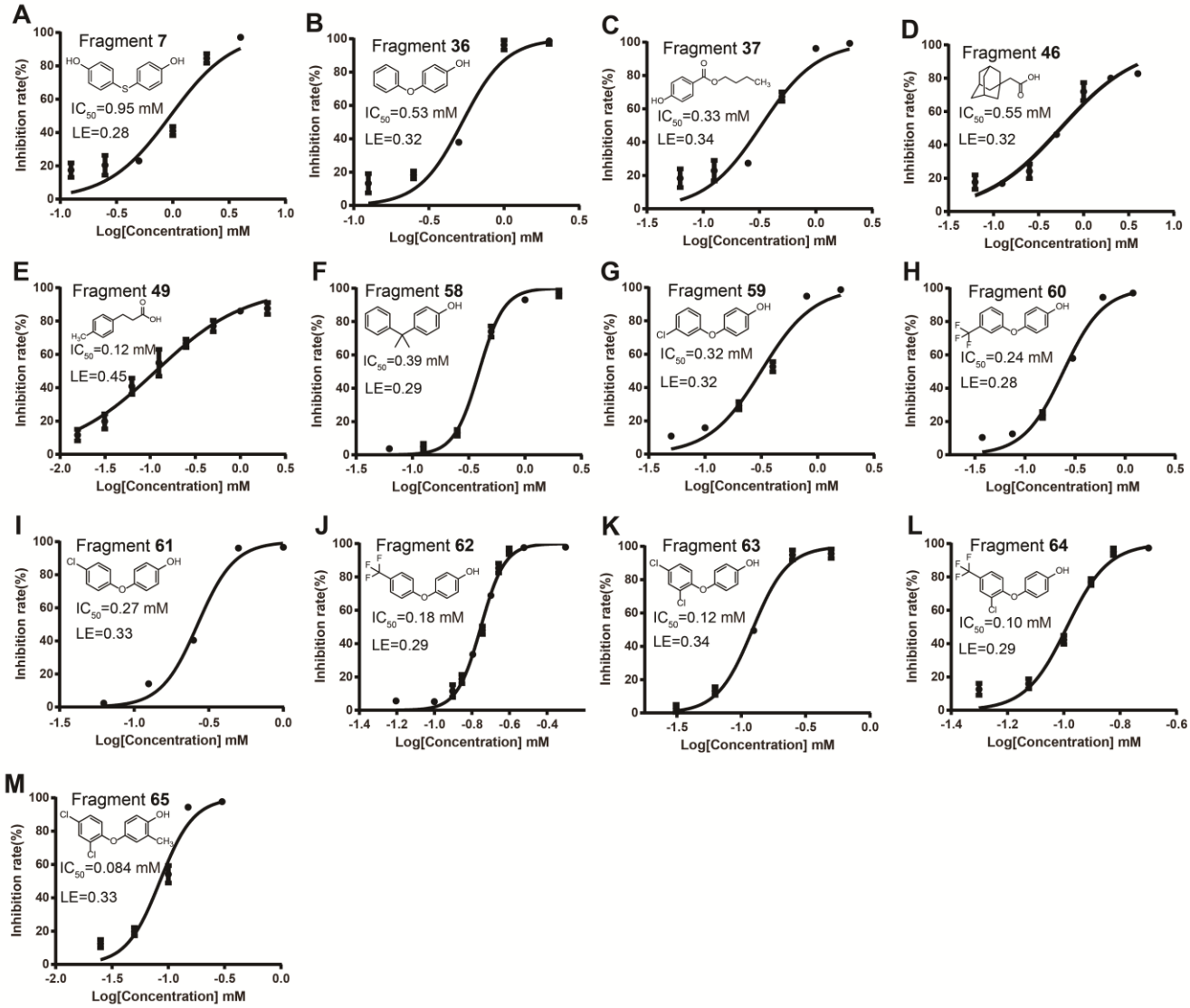
Xiaojie Huang, Junsong Guo, Qi Liu, Qiong Gu, Jun Xu, and Huihao Zhou*

Research Center for Drug Discovery, School of Pharmaceutical Sciences, Sun Yat-Sen University, Guangzhou 510006, China

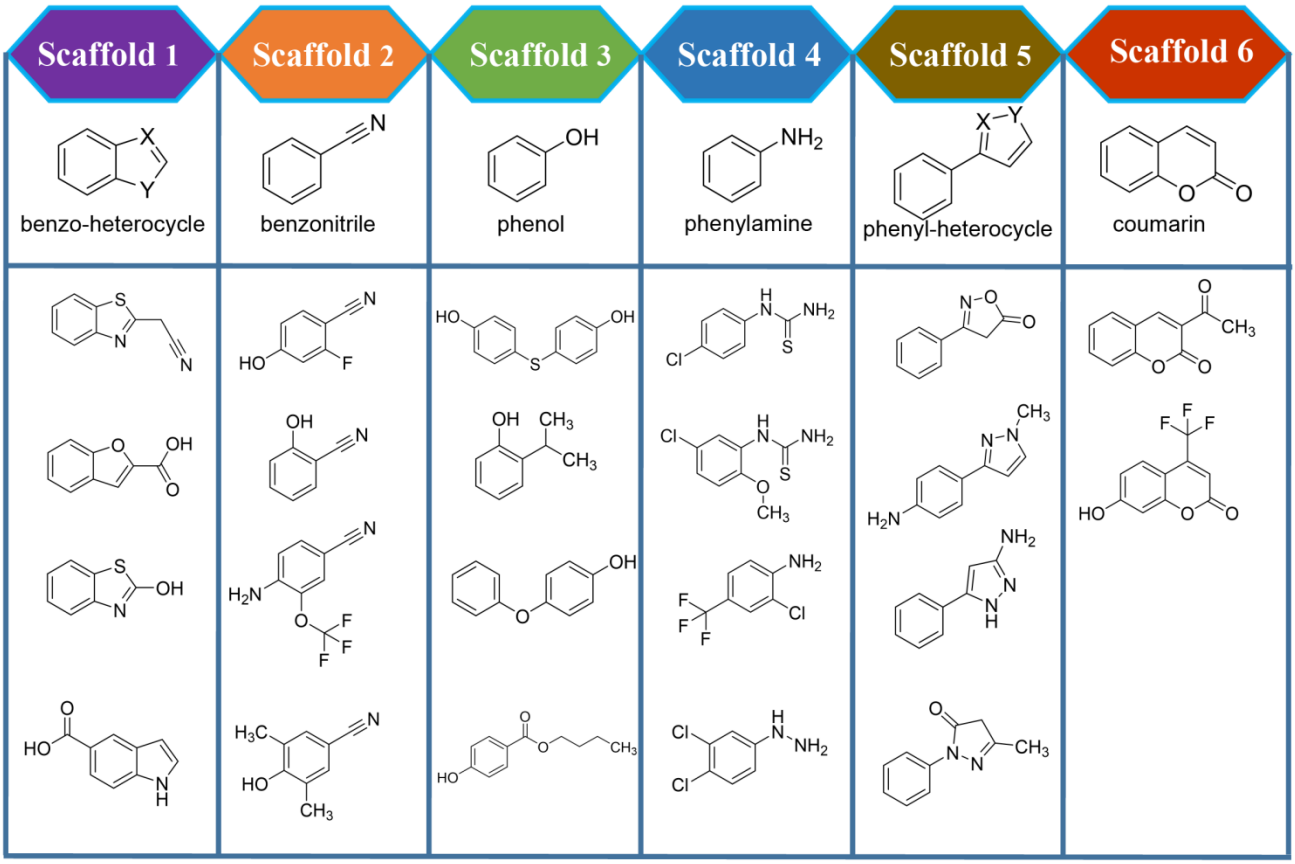
*Correspondence and requests for materials should be addressed to H.Z. (zhuihao@mail.sysu.edu.cn)

Contents

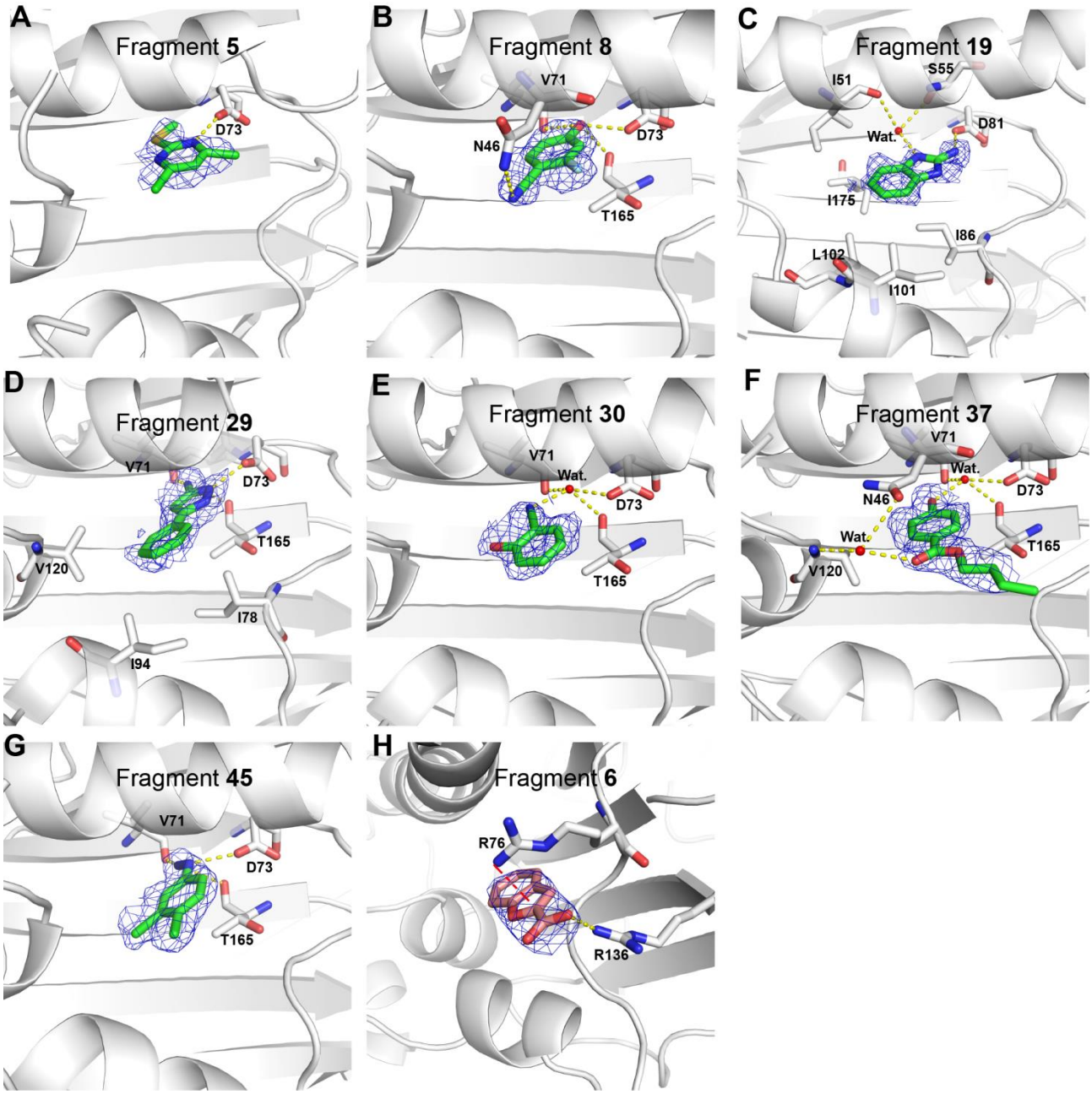
SUPPLEMENTARY FIGURES



Supplementary Fig. S1 The inhibition curves of the ATPase activity of *E. coli* DNA gyrase by fragments **7**, **36**, **37**, **46**, **49**, **58**-**65**. The IC_{50} values were calculated by fitting the curves of inhibitory rates *versus* fragment concentrations. The ligand efficiency (LE) for each fragment was calculated from the IC_{50} values using the equation described in "Methods".

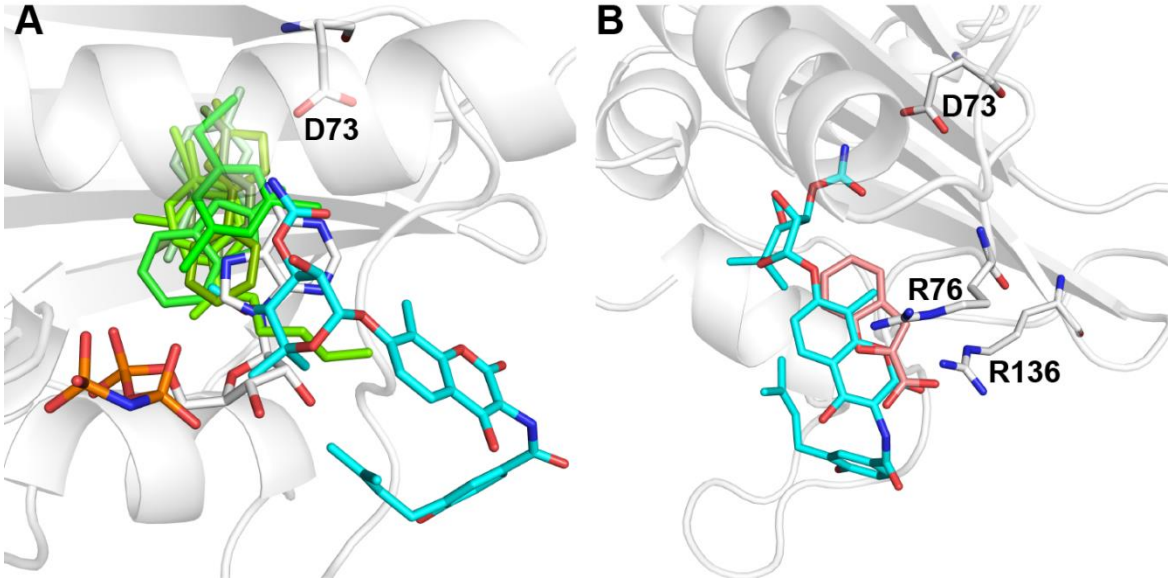


Supplementary Fig. S2 Six scaffolds of the fragment hits. Most fragment hits belong to the six scaffolds according to their chemical structures: benzo-heterocyclic fragments, benzonitrile analogs, phenol fragments, aniline fragments, phenyl-heterocyclic fragments and coumarins. The representative fragments are listed.

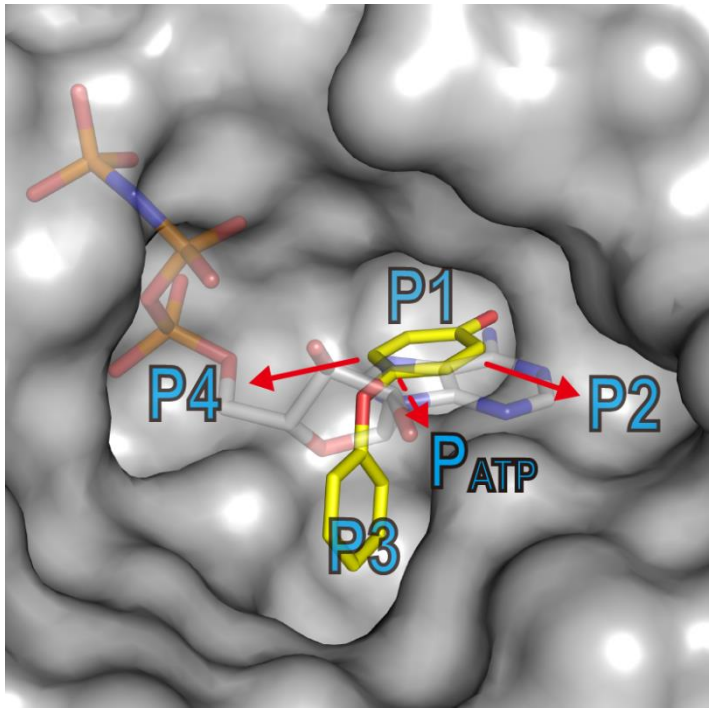


Supplementary Fig. S3 The GyrB-binding mechanisms of fragments **5**, **6**, **8**, **19**, **29**, **30**, **37** and **45** revealed by X-ray crystallography. (A) Fragment **5** (4,6-dichloro-2-(methylthio)pyrimidine) mainly interacts with residue Asp73 of *Ec*GyrB_AD, a key residue in coordinating adenosine ring of ATP, to competes with the substrate ATP. (B) The phenolic hydroxy group of fragment **8** (2-fluoro-4-hydroxybenzonitrile) forms H-bonding interactions with the carbonyl oxygen of residue

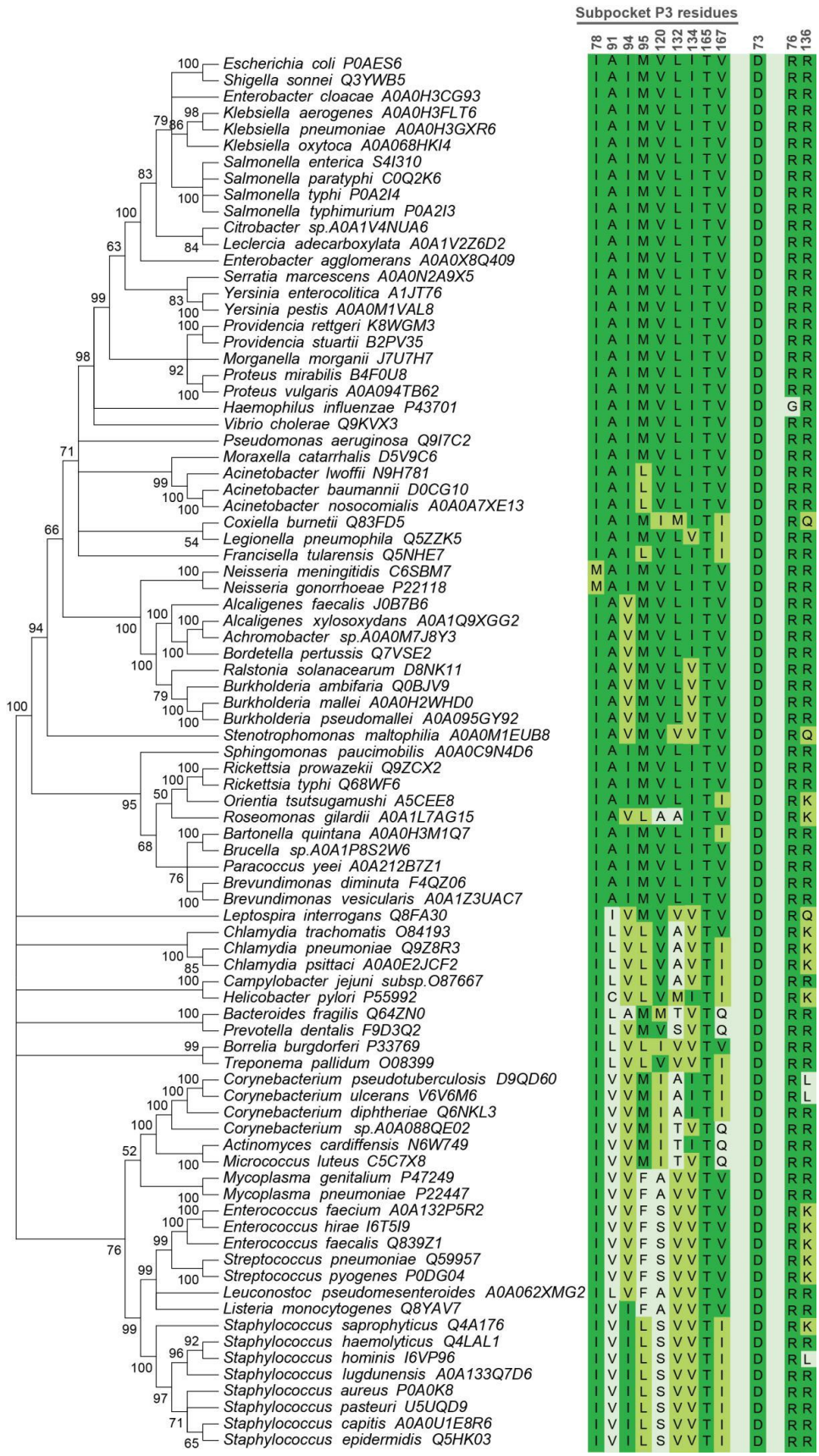
Asp73 as well as the backbone oxygen of Val71 and Thr165. The cyano group forms another H-bond with the side chain nitrogen of Asn46. (C) The X-ray cocrystal structure of fragment **19** (2-aminobenzimidazole) was solved using a *Sa*GyrB_AD crystal. The amino group forms an H-bond with the side chain of residue Asp81 (corresponding to Asp73 in *E. coli* GyrB). The N1 atom of the benzimidazole is involved in water-mediated H-bond network with residues Ile55 and Ser55. The benzene ring forms hydrophobic contacts with Ile51, Ile86, Ile102, Leu103 and Ile175. (D) Fragment **29** (3-amino-5-phenylpyrazole) uses its pyrazole nitrogen atom to form two H-bonds with residues Asp73 and Thr165. Another H-bond is formed between the amino group of fragment **29** and the backbone oxygen of residue Val71. In addition, the benzene ring of fragment **29** is in hydrophobic contacts with residues Ile78, Ile94 and Val120. (E) A water molecule in the active site forms a water-bridged H-bonding network between the cyano group of fragment **30** (2-cyanophenol) and the oxygen atoms of residues Val71, Asp73, and Thr165. (F) Fragment **37** (butyl 4-hydroxybenzoate) is also with a phenol scaffold. Interestingly, its phenolic hydroxy group does not directly interact with the key residue Asp73 but bridged by a water molecule. Its carbonyl oxygen forms water-bridged H-bonding interactions with the side chains of residues Asn46 and Val120. (G) The nitrogen atom N1 of fragment **45** ((3,4-dichlorophenyl)hydrazine) forms H-bonds with residues Asp73 and Thr165. Its nitrogen atom N2 makes another H-bond interaction with the backbone oxygen of Val71. (H) Fragment **6** is unique to all the fragments described above as it binds to a different site about 7 Å away from the key residue Asp73. Fragment **6** binds to GyrB mainly through a H-bond interaction with the side chain of Arg136 as well as a cation-π interaction with Arg76. The H-bonds were shown as yellow dotted lines, and the cation-π interaction was shown as red dotted line.



Supplementary Fig. S4 The binding mechanisms of the class I and II fragments in comparison with ADPNP and novobiocin. (A) Cocrystal structure superimposition of the seven fragments in class I with ADPNP and novobiocin. Similar to ADPNP and novobiocin, all the class I fragments form H-bonding interactions with the key residue Asp73, resulting in the partially overlap of class I fragments with the ADPNP adenosine group or the novobiocin noviose sugar group. Interestingly, all the class I fragments bind deeper than ADPNP and novobiocin in the pocket. (B) Fragment **6** (class II fragment) partially overlaps with the coumarin ring of novobiocin. Both fragment **6** and the cumarin ring interact with Arg76 and Arg136. Class I fragments are drawn in green sticks, class II in pink, ADPNP in white and novobiocin in cyan.



Supplementary Fig. S5 The multiple subpockets in GyrB ATPase site. Subpocket P1 is around the key residue Asp73, and it is occupied by phenol group of fragments **7** and **36** as well as by most GyrB inhibitors reported. Subpocket P2 is a tunnel-like pocket playing an important role for novobiocin binding. Subpocket P3 is a hydrophobic pocket induced by fragment **7** and **36**. Subpocket P4 is another hydrophobic pocket mainly formed by Val43, Val120, Val123, Leu130, Val167 and Phe169, and it was rarely touched by the inhibitors reported. Substrate ATP binds to a pocket in the front of the screen (P_{ATP}). These five subpockets connect to each other. Fragments **7** and **36** are mainly occupying P1 and P2, and they have plenty of room to grow in different directions to achieve high affinity. The protein is shown as gray surface. The residues of GyrB staying in the front of the screen and covering the active pocket are omitted for clarity.



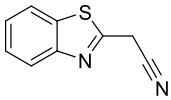
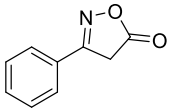
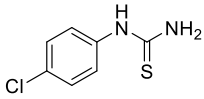
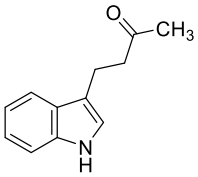
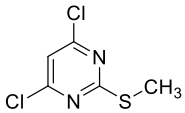
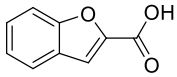
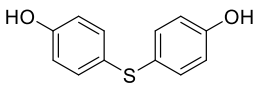
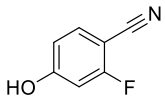
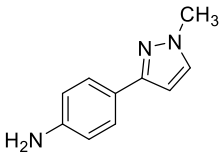
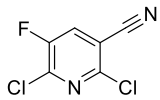
Supplementary Fig. S6

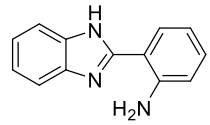
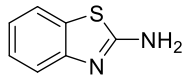
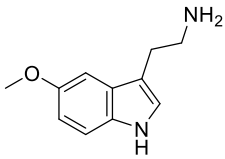
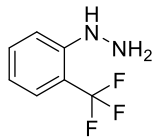
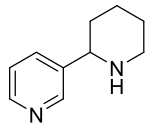
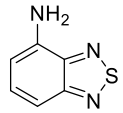
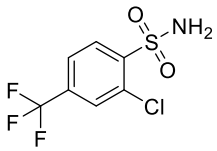
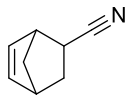
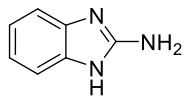
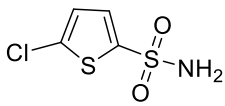
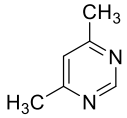
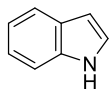
Supplementary Fig. S6 The phylogenetic analysis of 85 GyrBs (Bode Science Center. http://www.bode-science-center.com/center/relevant-pathogens-from-a-z.html) from pathogenic bacteria. Residues surrounding the new identified hydrophobic subpocket (P3) are highly conserved among the pathogenic bacteria in *Proteobacteria* phylum (Gram-negative). In GyrBs of Gram-positive bacteria, these residues are still similar (most are hydrophobic residues), although not absolutely conserved. The key residue Asp73 (*E. coli* numbered) for ATP binding is absolute conserved among all GyrBs analyzed. And, the novobiocin-binding residue Arg76 is also highly conserved among bacterial GyrBs with only one exception in *Haemophilus influenzae*. In contrast, another novobiocin-binding key residue Arg136 is less conserved in the GyrBs from Gram-positive bacteria in *Staphylococcus*, *Corynebacterium* and *Enterococcus*, which is consistent with that mutations of Arg136 were frequently observed in most novobiocin-resistant bacterial strains. Sequence alignments and phylogenetic tree construction were performed using program Mega5[1], and the uniprot codes of each GyrB were provided after the names of the bacteria.

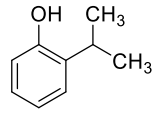
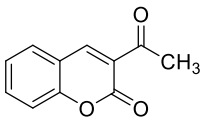
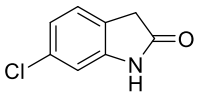
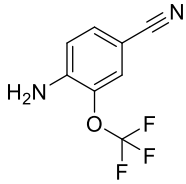
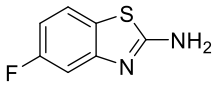
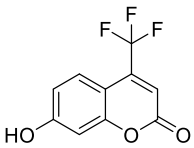
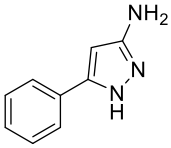
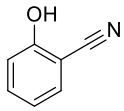
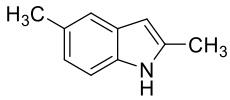
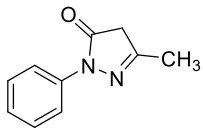
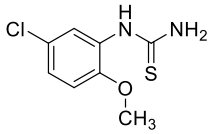
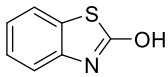
1. K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei and S. Kumar, *Mol Biol Evol*, 2011, **28**, 2731-2739.

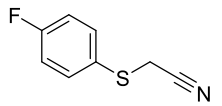
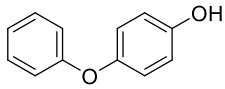
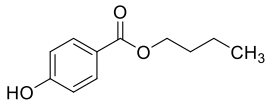
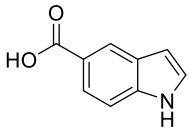
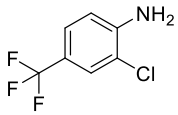
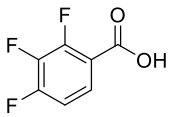
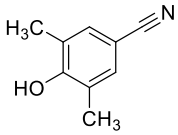
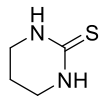
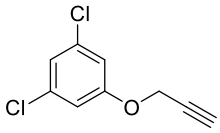
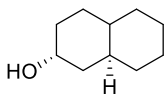
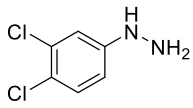
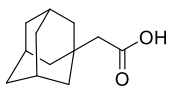
SUPPLEMENTARY TABLES

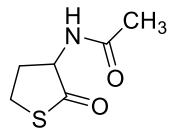
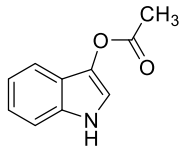
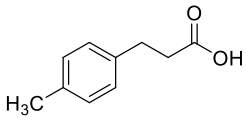
Table S1. The list of the fragment hits.

No.	Structure	ΔT_m [a] (°C)	ATPase inhibition [b] (%)	IC_{50} (mM)	PDB code
1		3.1	< 30	nd	-
2		2.0	< 30	nd	-
3		3.0	< 30	nd	-
4		2.9	< 30	nd	-
5		1.8	< 30	nd	5Z9N
6		2.3	< 30	nd	5Z9M
7		2.0	60.2	0.95	5Z4H
8		1.3	31.1	nd	5Z9L
9		1.5	< 30	nd	-
10		1.4	< 30	nd	-

11		1.6	< 30	nd	-
12		1.6	30.6	nd	-
13		1.5	< 30	nd	-
14		1.8	41.6	nd	-
15		1.0	< 30	nd	-
16		0.9	< 30	nd	-
17		1.2	< 30	nd	-
18		0.8	< 30	nd	-
19		1.4	< 30	nd	5Z9P
20		0.6	< 30	nd	-
21		0.6	< 30	nd	-
22		1.0	< 30	nd	-

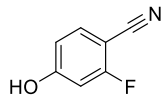
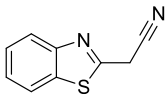
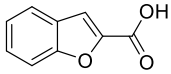
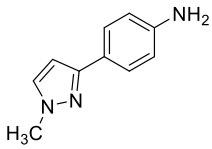
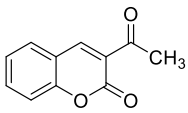
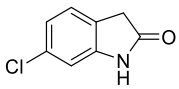
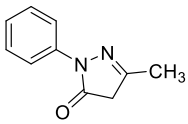
23		0.7	31.2	nd	-
24		0.8	45.9	nd	-
25		2.0	37.4	nd	-
26		0.6	< 30	nd	-
27		0.6	< 30	nd	-
28		0.6	< 30	nd	-
29		0.8	< 30	nd	5Z9Q
30		0.8	30.6	nd	5Z9F
31		1.4	40.5	nd	-
32		0.5	< 30	nd	-
33		1.0	< 30	nd	-
34		< 0.5	41.3	nd	-

35		< 0.5	33.8	nd	-
36		< 0.5	96.0	0.53	5Z4O
37		< 0.5	96.6	0.33	5Z9E
38		< 0.5	35.2	nd	-
39		< 0.5	45.8	nd	-
40		< 0.5	50.2	nd	-
41		< 0.5	44.4	nd	-
42		< 0.5	33.8	nd	-
43		< 0.5	50.7	nd	-
44		< 0.5	37.9	nd	-
45		< 0.5	50.4	nd	5Z9B
46		< 0.5	93.4	0.55	-

47		< 0.5	35.5	nd	-
48		< 0.5	58.7	nd	-
49		< 0.5	97.4	0.12	-

[a,b]Values are measured at 1 mM.

Table S2. TSA results of the fragments that potentially bind to an alternative site.

Fragment	Structure	Fragments in the complex	T_m (°C)	ΔT_m [a] (°C)	$\Delta\Delta T_m$ [b] (°C)
		Native protein	55.3	-	-
8		Fragment **8**	56.6	1.3	-
1		Fragment **8** + **1**	58.9	3.1	1.5
6		Fragment **8** + **6**	58.3	2.3	1.7
9		Fragment **8** + **9**	57.4	1.5	0.8
24		Fragment **8** + **24**	58.1	0.8	1.5
25		Fragment **8** + **25**	57.3	2.0	0.7
32		Fragment **8** + **32**	57.6	0.5	1.0

[a] ΔT_m is calculated by subtracting the thermal melting temperature (T_m) of apo *Ec*GyrB_AD from the T_m of *Ec*GyrB supplemented with the fragments.

[b] The $\Delta\Delta T_m$ is calculated by subtracting the ΔTm of *Ec*GyrB_AD caused by fragment **8** from that by two fragments together.

Table S3. Data collection and refinement statistics.

	*Ec*GyrB:fragment **5**	*Ec*GyrB:fragment **6**	*Ec*GyrB:fragment **7**	*Ec*GyrB:fragment **8**	*Sa*GyrB:fragment **19**
PDB code	5Z9N	5Z9M	5Z4H	5Z9L	5Z9P
Data collection					
Space group	$P4_22_12$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	C121
Cell dimensions					
a, *b*, *c* (Å)	*a*=97.68, *b*= 97.68, *c*=54.40	*a*=97.38, *b*= 101.79, *c*=104.86	*a*=61.66, *b*= 67.53, *c*=102.99	*a*=61.07, *b*= 69.07, *c*=102.12	*a*=143.10, *b*= 55.66, *c*=51.05
α, *β*, *γ* (°)	*α*=90, *β*=90, *γ*=90	*α*=90, *β*=90, *γ*=90	*α*=90, *β*=90, *γ*=90	*α*=90, *β*=90, *γ*=90	*α*=90, *β*=100.2, *γ*=90
Resolution (Å)	97.69-2.55 (2.64-2.55)[a]	73.04-2.74 (2.88-2.74)	50.00-2.00 (2.07-2.00)	57.22-1.60 (1.68-1.60)	70.41-1.45 (1.50-1.45)
R_{merge}[b] (%)	12.3 (46.7)	7.1 (39.9)	9.9 (42.8)	7.1 (37.9)	3.9 (7.3)
I/σI	42.5 (11.6)	17.4 (4.5)	34.7 (7.7)	10.8 (2.7)	41.9 (25.1)
Completeness (%)	99.9 (99.8)	93.1 (95.1)	99.8 (100.0)	96.9 (89.6)	96.1 (88.6)
Redundancy	11.6	5.5	5.0	5.2	3.0
Refinement					
Resolution (Å)	69.17-2.54	73.04-2.74	50.00-2.00	57.21-1.60	70.40-1.45
No. reflections	8609	24332	29857	53016	63865
R_{work}[c] / R_{free}[d] (%)	20.5 (23.1)	19.9 (21.6)	18.6 (22.1)	22.8 (24.1)	19.3 (20.9)
No. non-hydrogen atoms					
Protein	1428	2830	2884	2878	2937
Ligand/ion	29	79	60	50	22
Water	32	53	180	176	371
Mean B-factors	35.2	57.3	26.0	23.0	11.1
RMSD bonds (Å)	0.008	0.007	0.008	0.008	0.007
RMSD angles (º)	1.285	1.225	1.316	1.329	1.191
Ramachandran plot (%)					
Favoured	97.8	99.2	98.4	98.6	97.6
Allowed	99.5	99.7	99.7	99.5	100.0
Outliers	0.5	0.3	0.3	0.5	0

Table S3. *continued.*

	*Ec*GyrB:fragment **29**	*Ec*GyrB:fragment **30**	*Ec*GyrB:fragment **36**	*Ec*GyrB:fragment **37**	*Ec*GyrB:fragment **45**
PDB code	5Z9Q	5Z9F	5Z4O	5Z9E	5Z9B
Data collection					
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions					
a, *b*, *c* (Å)	*a*=60.92，*b*=68.72，*c*=102.35	*a*=61.59,，*b*=68.10，*c*=102.12	*a*=61.88，*b*=67.75，*c*=102.96	*a*=61.38，*b*=68.56，*c*=102.70	*a*=61.78，*b*=68.2，*c*=102.53
α, *β*, *γ* (°)	*α*=90, *β*=90, *γ*=90	*α*=90, *β*=90, *γ*=90	*α*=90, *β*=90, *γ*=90	*α*=90, *β*=90, *γ*=90	*α*=90, *β*=90, *γ*=90
Resolution (Å)	57.06-1.75 (1.81-1.75)	56.66-1.76 (1.82-1.76)	67.75-1.73 (1.77-1.73)	57.03-1.80 (1.86-1.80)	56.81-2.10 (2.18-2.10)
R_{merge}[b](%)	6.0 (30.8)	6.0 (19.4)	4.4 (49.6)	6.3 (37.4)	7.8 (23.5)
$I/\sigma I$	47.3 (7.2)	52.0 (14.1)	16.5 (2.4)	40.4 (5.4)	31.6 (8.2)
Completeness (%)	89.5 (81.3)	91.5 (100.0)	95.3 (97.0)	97.9 (99.3)	97.6 (100.0)
Redundancy	5.4	5.5	4.5	4.9	4.5
Refinement					
Resolution (Å)	57.06-1.80	56.66-1.76	56.66-1.73	57.02-1.80	56.80-2.10
No. reflections	34845	37573	41311	37995	25391
R_{work}[c]/R_{free}[d] (%)	20.8 (22.9)	19.8 (22.4)	19.6 (22.7)	19.6 (22.2)	20.0 (22.6)
No. non-hydrogen atoms					
Protein	2817	2875	2875	2867	2900
Ligand/ion	54	48	58	58	50
Water	196	185	214	148	153
Mean B-factors	26.6	31.4	31.0	25.9	26.1
RMSD bonds (Å)	0.007	0.007	0.007	0.008	0.008
RMSD angles (º)	1.281	1.254	1.187	1.307	1.338
Ramachandran plot (%)					
Favoured	98.3	98.4	98.5	98.4	98.1
Allowed	99.4	99.5	99.7	99.7	99.7
Outliers	0.6	0.5	0.3	0.3	0.3

[a]Values in parentheses are for the highest resolution shell.

[b] R_{merge} = $\Sigma_h \Sigma_I$ | $I(h)_I$ - <$I(h)$>|/ $\Sigma_h \Sigma_I I(h)_I$, where $I(h)_I$ is the Ith observation of the reflection h and <$I(h)$> is the weighted average intensity for all observations I of reflection h.

[c] $R_{work}=\sum_h||F_{obs}(h)|-|F_{cal}(h)||/\sum_h|F_{obs}(h)|$, where $F_{obs}(h)$ and $F_{cal}(h)$ are the observed and calculated structure factors for reflection h respectively.

[d] R_{free} was calculated as R_{work} using 5% of the reflections which were selected randomly and omitted from refinement.