wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 04:28 PM GMT

PDB ID	:	4ESV
Title	:	A New Twist on the Translocation Mechanism of Helicases from the Structure of DnaB with its Substrates
Authors	:	Itsathitphaisarn, O.; Wing, R.A.; Eliason, W.K.; Wang, J.; Steitz, T.A.
Deposited on	:	2012-04-23
Resolution	:	3.20 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

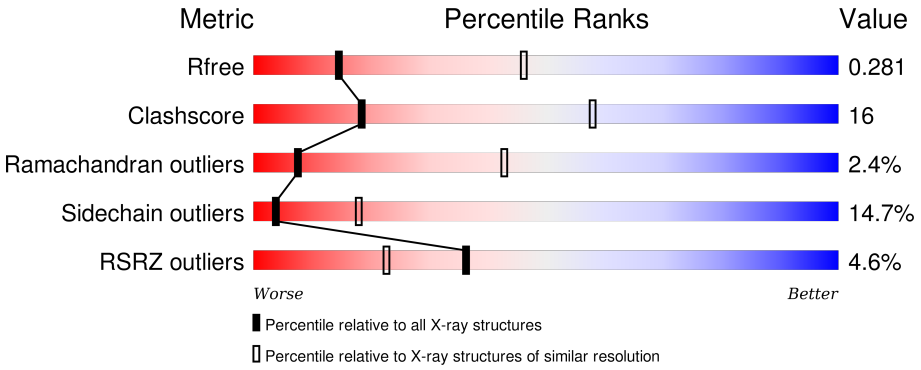
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



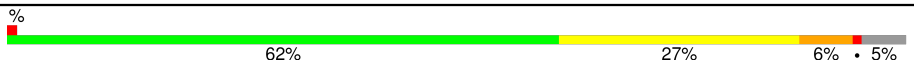
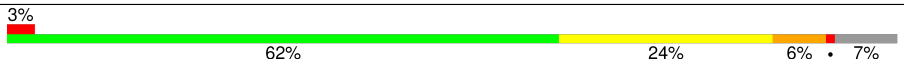
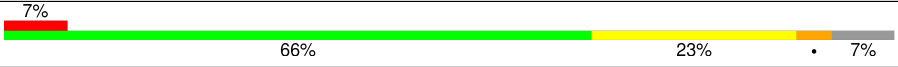
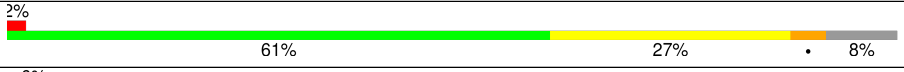
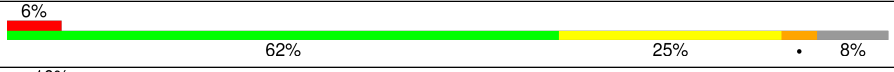
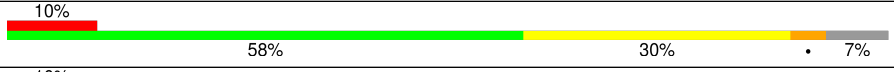
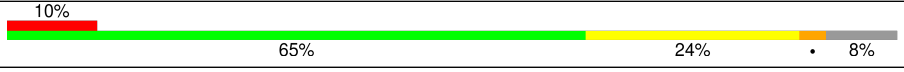
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1124 (3.24-3.16)
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	V	14	7%, 14%, 36%, 50%
2	W	13	8%, 23%, 69%
3	A	454	2%, 57%, 26%, 8%, 8%
3	B	454	2%, 55%, 27%, 11%, 5%
3	C	454	%, 60%, 28%, 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	454	
3	E	454	
3	F	454	
3	G	454	
3	H	454	
3	I	454	
3	J	454	
3	K	454	
3	L	454	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MES	C	505	-	-	-	X
4	MES	W	101	-	-	-	X
5	CA	A	503	-	-	-	X
5	CA	A	506	-	-	-	X
5	CA	C	502	-	-	-	X
5	CA	E	502	-	-	-	X
6	GDP	I	501	-	-	X	-
7	ALF	A	504	-	-	X	-
7	ALF	C	504	-	-	-	X
7	ALF	D	503	-	-	-	X
7	ALF	E	503	-	-	-	X
7	ALF	J	502	-	-	-	X

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 40296 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a DNA chain called 5'-D(P*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	V	14	Total	C	N	O	P	0	0	0
			280	140	28	98	14			

- Molecule 2 is a DNA chain called 5'-D(P*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	W	13	Total	C	N	O	P	0	0	0
			260	130	26	91	13			

- Molecule 3 is a protein called Replicative helicase.

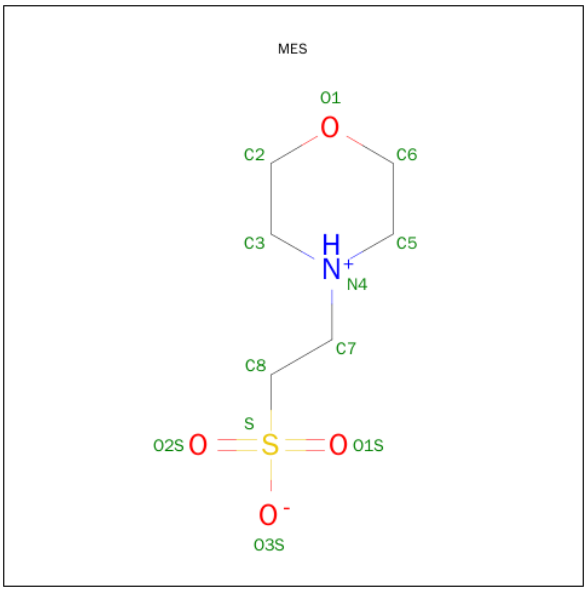
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	418	Total	C	N	O	S	8	1	0
			3255	2036	570	636	13			
3	B	430	Total	C	N	O	S	0	0	0
			3350	2093	591	652	14			
3	C	434	Total	C	N	O	S	0	0	0
			3367	2104	591	658	14			
3	D	431	Total	C	N	O	S	0	0	0
			3333	2084	583	652	14			
3	E	418	Total	C	N	O	S	0	0	0
			3243	2028	565	637	13			
3	F	421	Total	C	N	O	S	0	0	0
			3274	2049	569	643	13			
3	G	421	Total	C	N	O	S	0	0	0
			3259	2037	568	641	13			
3	H	425	Total	C	N	O	S	0	0	0
			3283	2052	572	645	14			
3	I	418	Total	C	N	O	S	0	0	0
			3245	2034	561	637	13			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	416	Total	C	N	O	S	0	0	0
			3237	2025	563	636	13			
3	K	421	Total	C	N	O	S	0	0	0
			3267	2044	568	642	13			
3	L	419	Total	C	N	O	S	0	0	0
			3246	2031	565	637	13			

- Molecule 4 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	W	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
4	C	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
4	E	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
4	I	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

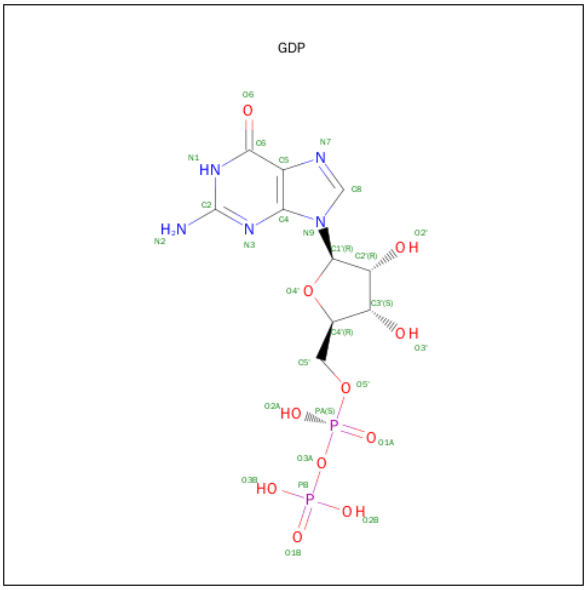
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	G	1	Total	Ca	0	0
			1	1		
5	D	2	Total	Ca	0	0
			2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	1	Total Ca 1 1	0	0
5	B	1	Total Ca 1 1	0	0
5	I	1	Total Ca 1 1	0	0
5	C	1	Total Ca 1 1	0	0
5	A	5	Total Ca 5 5	0	0
5	L	1	Total Ca 1 1	0	0
5	F	1	Total Ca 1 1	0	0

- Molecule 6 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).



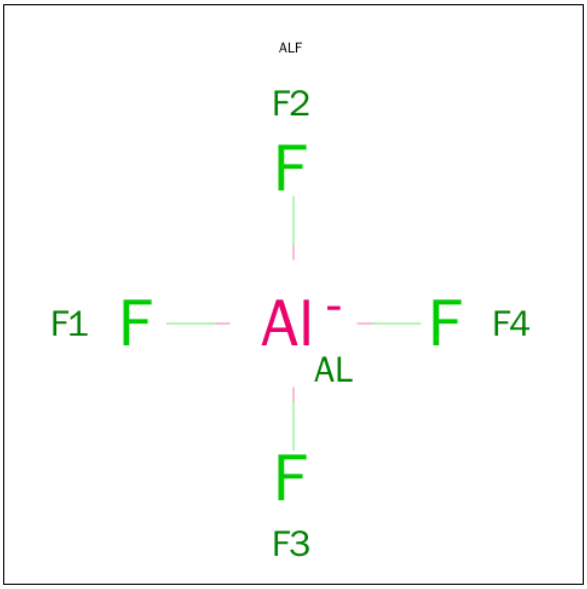
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C N O P 28 10 5 11 2	0	0
6	C	1	Total C N O P 28 10 5 11 2	0	0
6	D	1	Total C N O P 28 10 5 11 2	0	0
6	E	1	Total C N O P 28 10 5 11 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	F	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
6	G	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
6	I	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
6	J	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
6	K	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
6	L	1	Total	C	N	O	P	0	0
			28	10	5	11	2		

- Molecule 7 is TETRAFLUOROALUMINATE ION (three-letter code: ALF) (formula: AlF_4).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	Al	F	0	0
			5	1	4		
7	C	1	Total	Al	F	0	0
			5	1	4		
7	C	1	Total	Al	F	0	0
			5	1	4		
7	D	1	Total	Al	F	0	0
			5	1	4		
7	E	1	Total	Al	F	0	0
			5	1	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	G	1	Total Al F 5 1 4	0	0
7	I	1	Total Al F 5 1 4	0	0
7	J	1	Total Al F 5 1 4	0	0
7	K	1	Total Al F 5 1 4	0	0
7	L	1	Total Al F 5 1 4	0	0

- Molecule 8 is water.

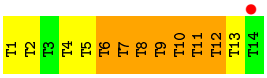
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total O 1 1	0	0
8	B	1	Total O 1 1	0	0
8	D	1	Total O 1 1	0	0
8	E	1	Total O 1 1	0	0
8	K	1	Total O 1 1	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: 5'-D(P*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*T)-3'

Chain V:

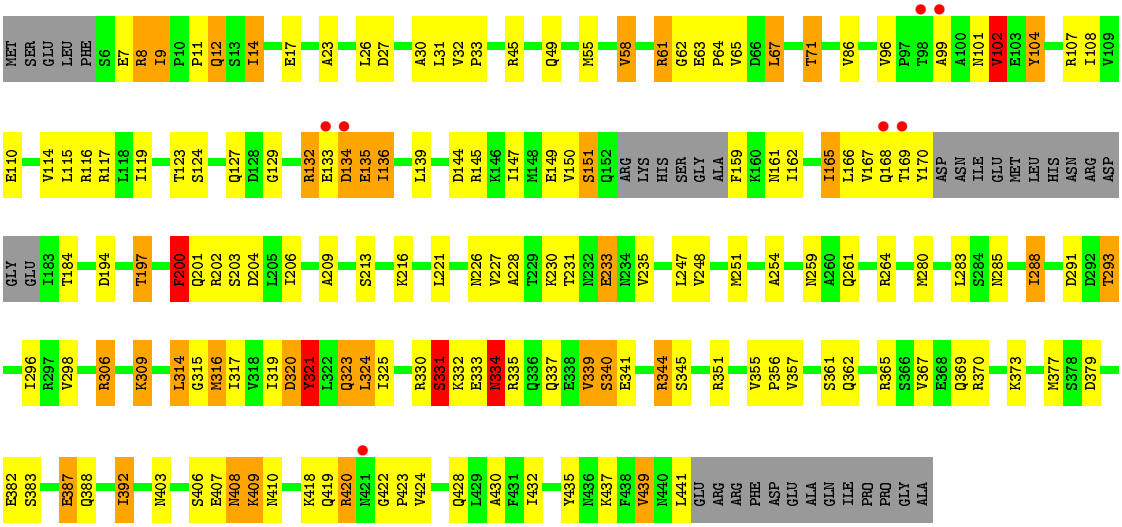


- Molecule 2: 5'-D(P*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*TP*T)-3'

Chain W:

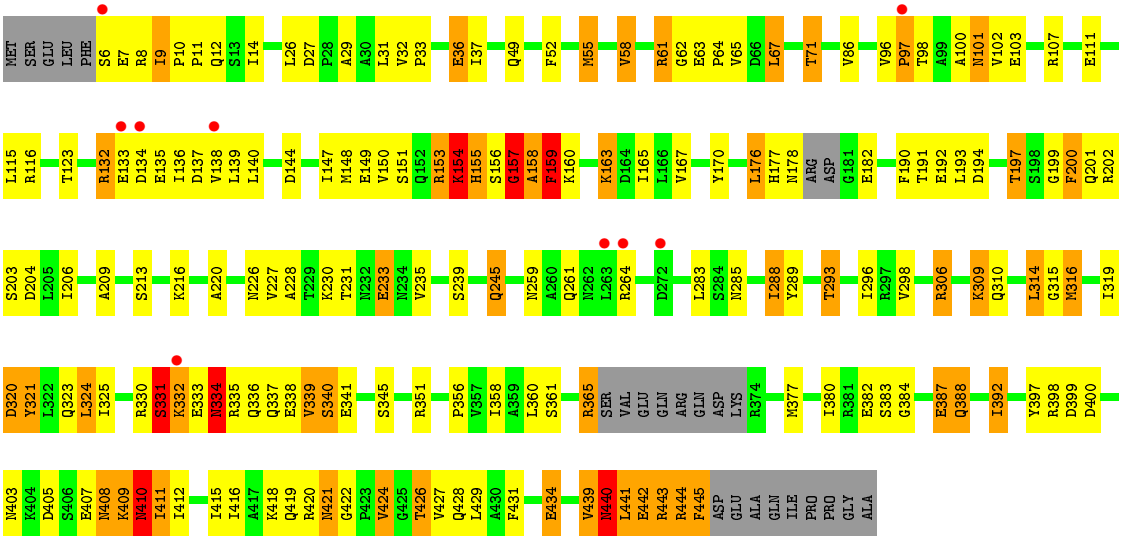
- Molecule 3: Replicative helicase

Chain A:



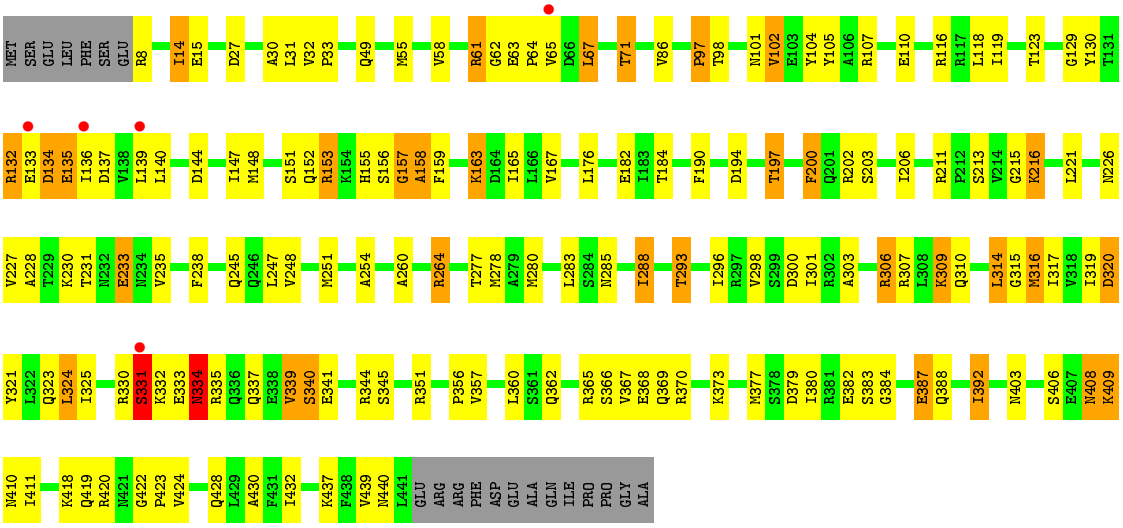
- Molecule 3: Replicative helicase

Chain B:



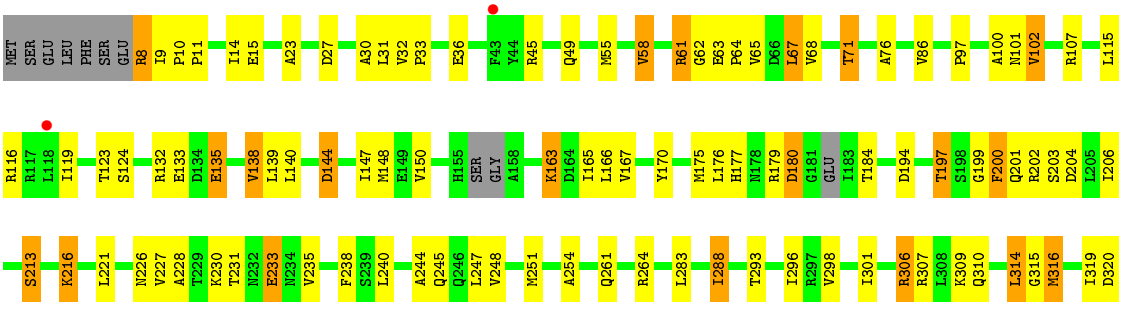
• Molecule 3: Replicative helicase

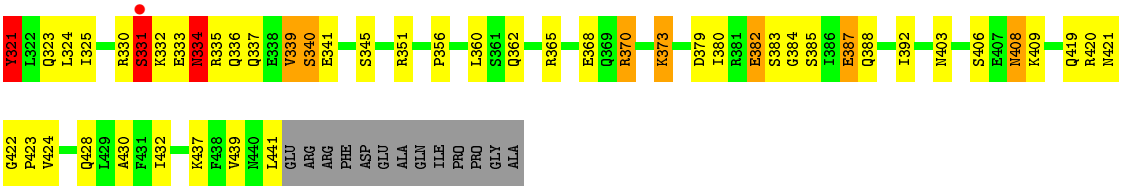
Chain C:



• Molecule 3: Replicative helicase

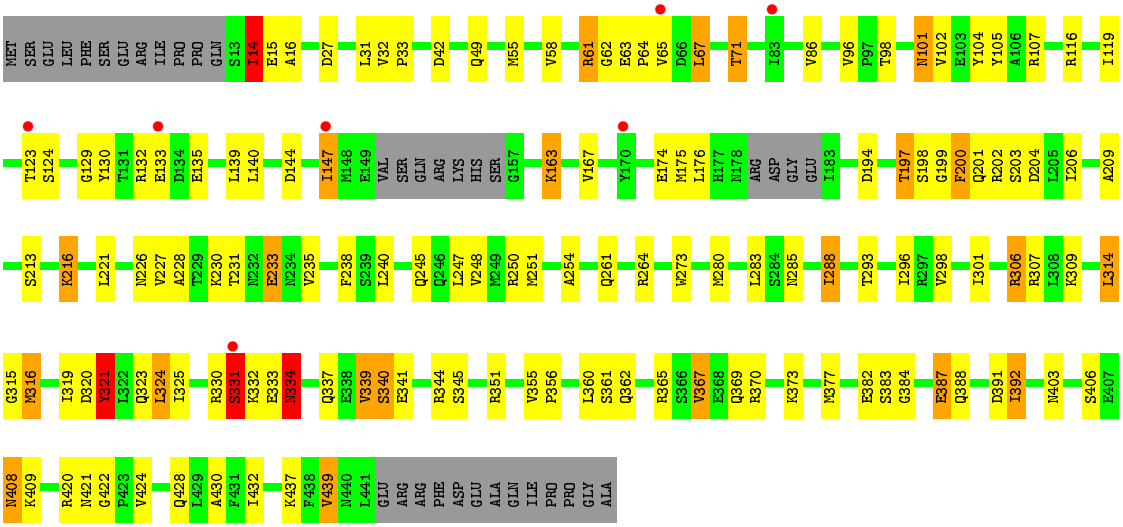
Chain D:





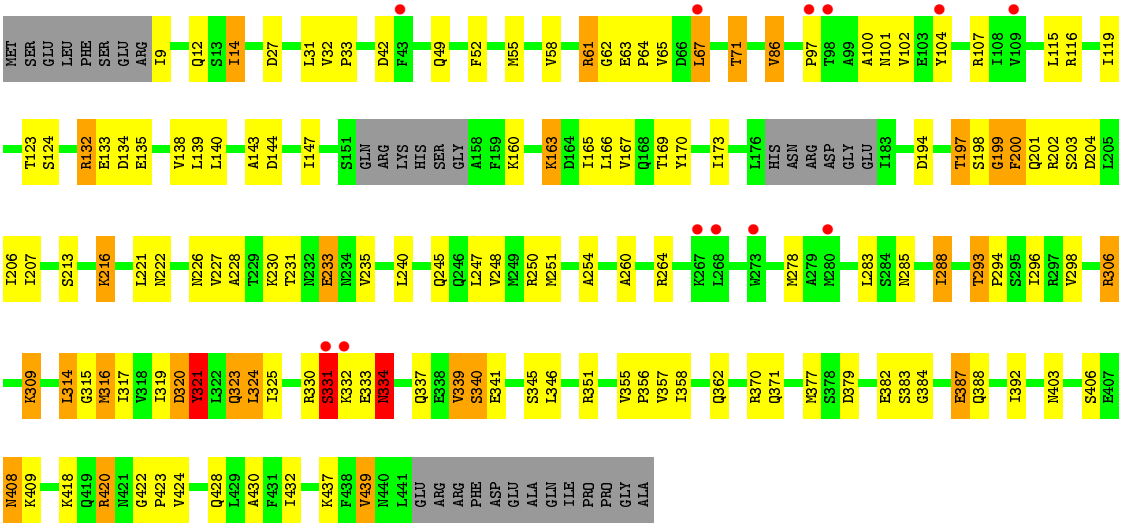
• Molecule 3: Replicative helicase

Chain E:



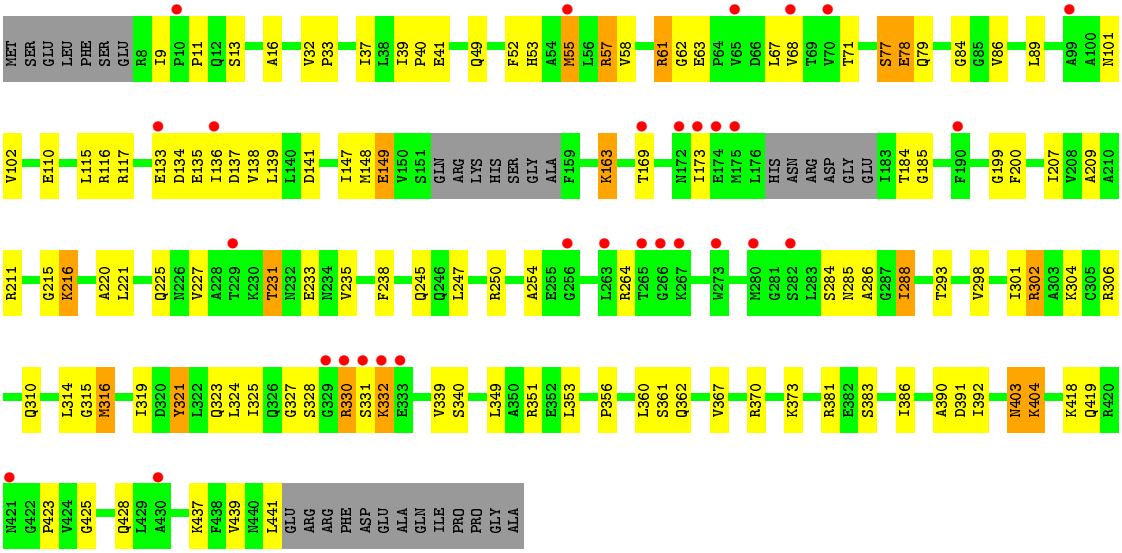
• Molecule 3: Replicative helicase

Chain F:



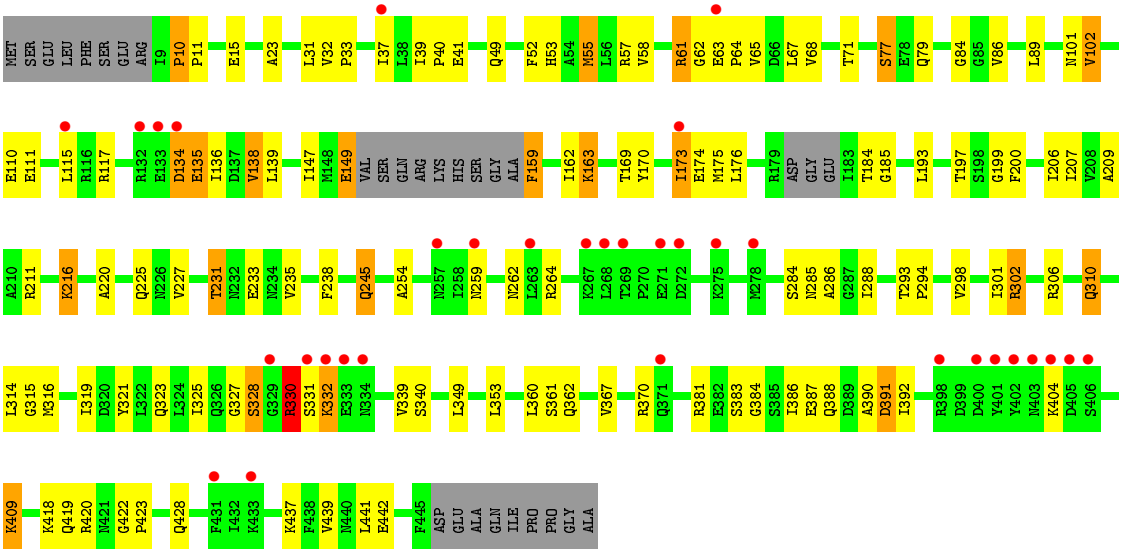
• Molecule 3: Replicative helicase

Chain G:



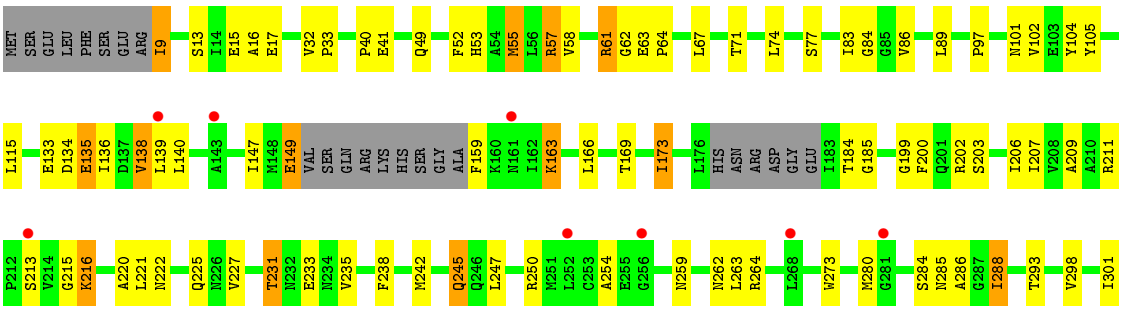
• Molecule 3: Replicative helicase

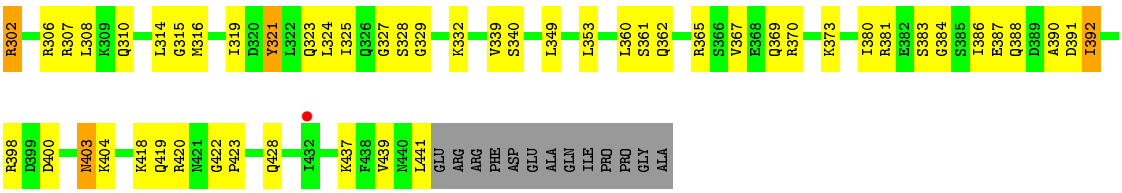
Chain H:



• Molecule 3: Replicative helicase

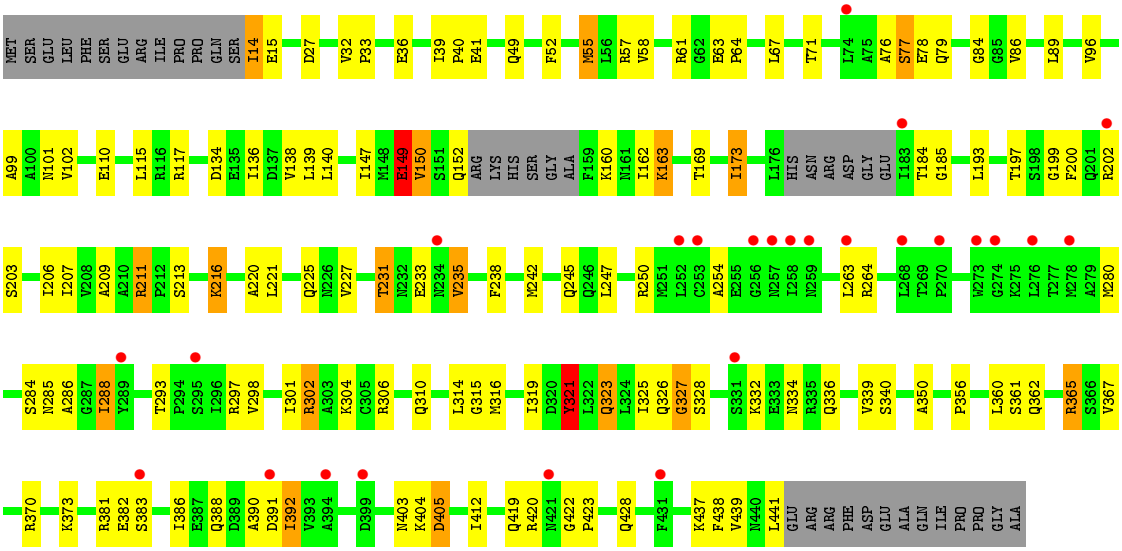
Chain I:





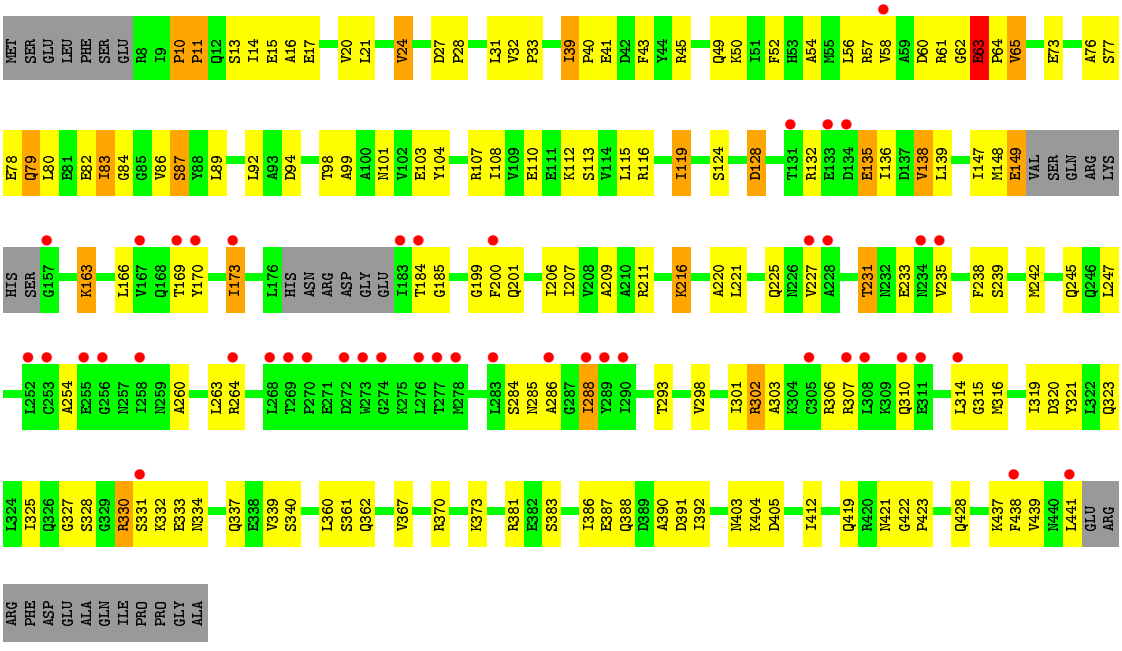
• Molecule 3: Replicative helicase

Chain J:



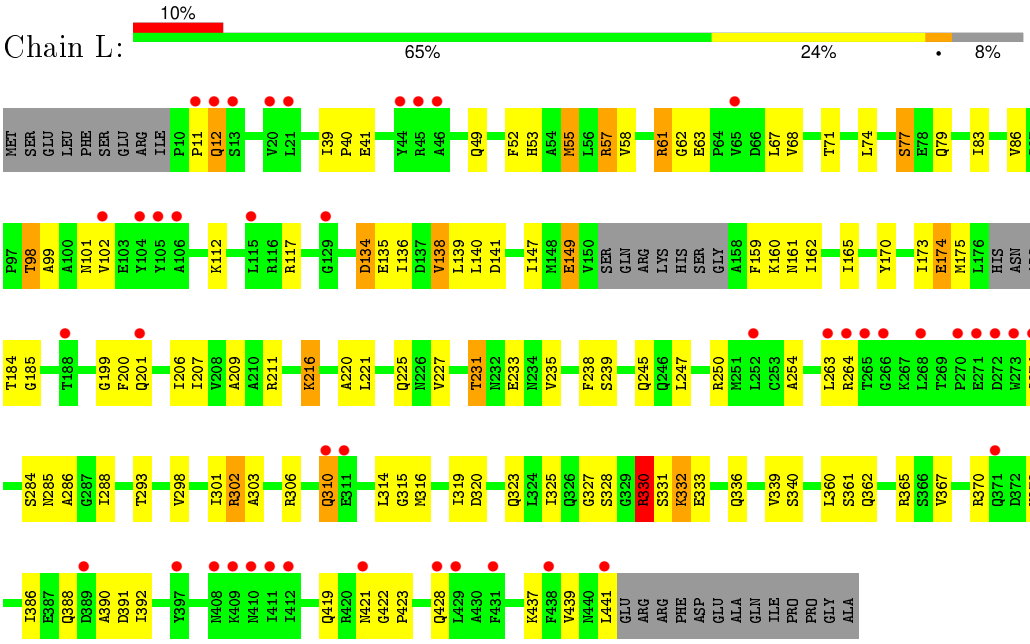
• Molecule 3: Replicative helicase

Chain K:



• Molecule 3: Replicative helicase

Chain L:


10%
65%
24%
8%
MET SER GLU LEU PHE SER GLU ARG ILE P10 P11 Q12 S13 V20 L21 I39 P40 E41 Y44 R45 A46 Q49 F52 H53 A54 M55 L56 R57 V58 R61 G62 E63 P64 V65 D66 L67 V68 T71 L74 S77 E78 Q79 I83 V86 S87 Y88 L89 V96
P97 T98 A99 A100 N101 V102 E103 Y104 Y105 A106 K112 L115 R116 R117 G129 D134 E135 I136 D137 V138 L139 L140 D141 I147 M148 E149 V150 SER GLN ARG LYS HIS SER GLY A158 F159 K160 N161 I162 I165 Y170 I173 E174 M175 L176 HIS ASN ARG ASP GLY GLU I183
T184 G185 T188 G199 F200 Q201 I206 I207 V208 A209 A210 R211 K216 A220 L221 Q225 N226 V227 T231 N232 E233 N234 V235 F238 S239 Q245 Q246 L247 R250 M251 L252 C253 A254 L263 R264 T265 G266 K267 L268 T269 P270 E271 D272 W273 G274 M278 A279 M280
S284 N285 A286 G287 I288 T293 V298 I301 R302 A303 R306 Q310 E311 I314 G315 M316 I319 D320 Q323 L324 I325 Q326 G327 S328 G329 R330 S331 K332 E333 Q336 V339 S340 L360 S361 Q362 R365 S366 V367 R370 Q371 D372 K373 R381 E382 S383
I386 E387 Q388 D389 A390 D391 I392 Y397 N408 K409 N410 I411 I412 Q419 R420 N421 G422 P423 Q428 L429 A430 F431 K437 F438 V439 N440 L441 GLU ARG ARG PHE ASP GLU ALA GLN ILE PRO PRO GLY ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	149.12Å 180.32Å 279.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.92 – 3.20 47.92 – 3.20	Depositor EDS
% Data completeness (in resolution range)	87.8 (47.92-3.20) 87.8 (47.92-3.20)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.32 (at 3.19Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.242 , 0.289 0.239 , 0.281	Depositor DCC
R_{free} test set	6214 reflections (6.01%)	DCC
Wilson B-factor (Å^2)	94.6	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 83.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	0 of 108863 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	40296	wwPDB-VP
Average B, all atoms (Å^2)	140.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.00% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GDP, ALF, CA, MES

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	V	1.30	2/307 (0.7%)	2.43	29/472 (6.1%)
2	W	1.61	2/285 (0.7%)	2.37	18/438 (4.1%)
3	A	0.61	0/3295	0.77	3/4450 (0.1%)
3	B	0.61	1/3390 (0.0%)	0.82	2/4576 (0.0%)
3	C	0.58	0/3408	0.75	2/4607 (0.0%)
3	D	0.52	0/3372	0.72	0/4557
3	E	0.54	0/3279	0.73	1/4428 (0.0%)
3	F	0.54	0/3312	0.74	2/4475 (0.0%)
3	G	0.40	0/3296	0.61	1/4454 (0.0%)
3	H	0.44	0/3319	0.64	3/4482 (0.1%)
3	I	0.43	0/3283	0.62	1/4437 (0.0%)
3	J	0.43	0/3273	0.65	2/4420 (0.0%)
3	K	0.46	0/3305	0.68	3/4467 (0.1%)
3	L	0.41	0/3284	0.62	0/4437
All	All	0.53	5/40408 (0.0%)	0.76	67/54700 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	A	0	4
3	B	0	5
3	C	0	1
3	D	0	2
3	E	0	1
3	F	0	1
3	I	0	1
3	J	0	1
3	K	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	0	17

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	421	ASN	CB-CG	5.99	1.64	1.51
2	W	6	DT	C1'-N1	5.61	1.56	1.49
1	V	12	DT	C1'-N1	5.56	1.56	1.49
2	W	11	DT	N1-C2	5.10	1.42	1.38
1	V	8	DT	C3'-O3'	5.03	1.50	1.44

The worst 5 of 67 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	V	8	DT	O4'-C1'-N1	12.56	116.79	108.00
2	W	7	DT	O4'-C1'-N1	10.63	115.44	108.00
2	W	4	DT	C4-C5-C7	10.36	125.22	119.00
2	W	11	DT	O4'-C4'-C3'	-9.32	100.41	106.00
2	W	10	DT	C1'-O4'-C4'	-9.01	101.09	110.10

There are no chirality outliers.

5 of 17 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	166	LEU	Peptide
3	A	167	VAL	Peptide
3	A	184	THR	Peptide
3	A	320	ASP	Peptide
3	B	151	SER	Peptide

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	V	280	0	169	11	0
2	W	260	0	157	24	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	3255	0	3296	128	0
3	B	3350	0	3382	177	0
3	C	3367	0	3384	118	0
3	D	3333	0	3337	119	0
3	E	3243	0	3262	110	0
3	F	3274	0	3305	105	0
3	G	3259	0	3274	87	0
3	H	3283	0	3282	110	0
3	I	3245	0	3268	118	0
3	J	3237	0	3265	108	0
3	K	3267	0	3287	115	0
3	L	3246	0	3263	89	0
4	C	12	0	12	0	0
4	E	12	0	12	0	0
4	I	12	0	12	0	0
4	W	12	0	12	4	0
5	A	5	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
5	D	2	0	0	0	0
5	E	1	0	0	0	0
5	F	1	0	0	0	0
5	G	1	0	0	0	0
5	I	1	0	0	0	0
5	L	1	0	0	0	0
6	A	28	0	12	4	0
6	C	28	0	12	4	0
6	D	28	0	12	3	0
6	E	28	0	12	4	0
6	F	28	0	12	4	0
6	G	28	0	12	5	0
6	I	28	0	12	10	0
6	J	28	0	12	6	0
6	K	28	0	12	3	0
6	L	28	0	12	5	0
7	A	5	0	0	2	0
7	C	10	0	0	0	0
7	D	5	0	0	0	0
7	E	5	0	0	0	0
7	G	5	0	0	0	0
7	I	5	0	0	0	0
7	J	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	K	5	0	0	1	0
7	L	5	0	0	0	0
8	A	1	0	0	0	0
8	B	1	0	0	0	0
8	D	1	0	0	0	0
8	E	1	0	0	0	0
8	K	1	0	0	0	0
All	All	40296	0	40099	1249	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 16.

The worst 5 of 1249 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:J:403:ASN:HB3	3:J:404:LYS:CA	1.63	1.28
3:A:145:ARG:HD2	3:B:310:GLN:NE2	1.58	1.18
3:E:14:ILE:HD13	3:E:15:GLU:H	1.10	1.16
2:W:6:DT:H2"	2:W:7:DT:H5'	1.29	1.14
3:J:403:ASN:CB	3:J:404:LYS:HA	1.75	1.14

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	A	413/454 (91%)	369 (89%)	31 (8%)	13 (3%)	5 34
3	B	424/454 (93%)	365 (86%)	38 (9%)	21 (5%)	3 21
3	C	432/454 (95%)	385 (89%)	29 (7%)	18 (4%)	3 26
3	D	425/454 (94%)	376 (88%)	37 (9%)	12 (3%)	6 37

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	E	412/454 (91%)	373 (90%)	27 (7%)	12 (3%)	6 36
3	F	415/454 (91%)	377 (91%)	27 (6%)	11 (3%)	6 39
3	G	415/454 (91%)	374 (90%)	37 (9%)	4 (1%)	19 65
3	H	419/454 (92%)	375 (90%)	39 (9%)	5 (1%)	16 60
3	I	412/454 (91%)	375 (91%)	34 (8%)	3 (1%)	26 72
3	J	410/454 (90%)	369 (90%)	37 (9%)	4 (1%)	19 65
3	K	415/454 (91%)	354 (85%)	51 (12%)	10 (2%)	7 43
3	L	413/454 (91%)	372 (90%)	36 (9%)	5 (1%)	16 60
All	All	5005/5448 (92%)	4464 (89%)	423 (8%)	118 (2%)	7 43

5 of 118 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	97	PRO
3	B	154	LYS
3	B	331	SER
3	B	409	LYS
3	B	410	ASN

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	A	353/386 (92%)	291 (82%)	62 (18%)	2 11
3	B	362/386 (94%)	293 (81%)	69 (19%)	2 10
3	C	361/386 (94%)	305 (84%)	56 (16%)	3 15
3	D	356/386 (92%)	303 (85%)	53 (15%)	4 17
3	E	348/386 (90%)	293 (84%)	55 (16%)	3 15
3	F	355/386 (92%)	298 (84%)	57 (16%)	3 14
3	G	351/386 (91%)	307 (88%)	44 (12%)	6 26
3	H	350/386 (91%)	308 (88%)	42 (12%)	6 28

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	I	350/386 (91%)	307 (88%)	43 (12%)	6 27
3	J	350/386 (91%)	303 (87%)	47 (13%)	5 22
3	K	352/386 (91%)	300 (85%)	52 (15%)	4 18
3	L	349/386 (90%)	307 (88%)	42 (12%)	6 28
All	All	4237/4632 (92%)	3615 (85%)	622 (15%)	4 18

5 of 622 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	E	340	SER
3	F	420	ARG
3	K	373	LYS
3	E	408	ASN
3	F	147	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 124 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	323	GLN
3	G	362	GLN
3	K	419	GLN
3	F	334	ASN
3	G	161	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 38 ligands modelled in this entry, 14 are monoatomic - leaving 24 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
6	GDP	A	502	5	23,30,30	1.39	3 (13%)	30,47,47	2.26	11 (36%)
7	ALF	A	504	-	0,4,4	0.00	-	0,6,6	0.00	-
6	GDP	C	501	5	23,30,30	1.63	3 (13%)	30,47,47	2.16	12 (40%)
7	ALF	C	503	-	0,4,4	0.00	-	0,6,6	0.00	-
7	ALF	C	504	-	0,4,4	0.00	-	0,6,6	0.00	-
4	MES	C	505	-	11,12,12	0.61	0	14,16,16	2.65	6 (42%)
6	GDP	D	501	5	23,30,30	1.29	3 (13%)	30,47,47	1.76	7 (23%)
7	ALF	D	503	-	0,4,4	0.00	-	0,6,6	0.00	-
6	GDP	E	501	5	23,30,30	1.35	3 (13%)	30,47,47	2.33	11 (36%)
7	ALF	E	503	-	0,4,4	0.00	-	0,6,6	0.00	-
4	MES	E	504	-	11,12,12	0.86	0	14,16,16	8.83	7 (50%)
6	GDP	F	501	5	23,30,30	1.44	3 (13%)	30,47,47	2.23	10 (33%)
6	GDP	G	501	-	23,30,30	1.34	3 (13%)	30,47,47	1.82	6 (20%)
7	ALF	G	502	-	0,4,4	0.00	-	0,6,6	0.00	-
6	GDP	I	501	5	23,30,30	1.19	2 (8%)	30,47,47	2.28	13 (43%)
7	ALF	I	503	-	0,4,4	0.00	-	0,6,6	0.00	-
4	MES	I	504	-	11,12,12	0.78	0	14,16,16	2.80	6 (42%)
6	GDP	J	501	-	23,30,30	1.19	3 (13%)	30,47,47	1.81	7 (23%)
7	ALF	J	502	-	0,4,4	0.00	-	0,6,6	0.00	-
6	GDP	K	501	-	23,30,30	1.27	2 (8%)	30,47,47	2.23	9 (30%)
7	ALF	K	502	-	0,4,4	0.00	-	0,6,6	0.00	-
6	GDP	L	501	5	23,30,30	1.35	2 (8%)	30,47,47	2.47	12 (40%)
7	ALF	L	502	-	0,4,4	0.00	-	0,6,6	0.00	-
4	MES	W	101	-	11,12,12	0.55	0	14,16,16	2.56	7 (50%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	GDP	A	502	5	-	0/12/32/32	0/3/3/3
7	ALF	A	504	-	-	0/0/0/0	0/0/0/0
6	GDP	C	501	5	-	0/12/32/32	0/3/3/3
7	ALF	C	503	-	-	0/0/0/0	0/0/0/0
7	ALF	C	504	-	-	0/0/0/0	0/0/0/0
4	MES	C	505	-	-	0/6/14/14	0/1/1/1
6	GDP	D	501	5	-	0/12/32/32	0/3/3/3
7	ALF	D	503	-	-	0/0/0/0	0/0/0/0
6	GDP	E	501	5	-	0/12/32/32	0/3/3/3
7	ALF	E	503	-	-	0/0/0/0	0/0/0/0
4	MES	E	504	-	-	0/6/14/14	0/1/1/1
6	GDP	F	501	5	-	0/12/32/32	0/3/3/3
6	GDP	G	501	-	-	0/12/32/32	0/3/3/3
7	ALF	G	502	-	-	0/0/0/0	0/0/0/0
6	GDP	I	501	5	-	0/12/32/32	0/3/3/3
7	ALF	I	503	-	-	0/0/0/0	0/0/0/0
4	MES	I	504	-	-	0/6/14/14	0/1/1/1
6	GDP	J	501	-	-	0/12/32/32	0/3/3/3
7	ALF	J	502	-	-	0/0/0/0	0/0/0/0
6	GDP	K	501	-	-	0/12/32/32	0/3/3/3
7	ALF	K	502	-	-	0/0/0/0	0/0/0/0
6	GDP	L	501	5	-	0/12/32/32	0/3/3/3
7	ALF	L	502	-	-	0/0/0/0	0/0/0/0
4	MES	W	101	-	-	0/6/14/14	0/1/1/1

The worst 5 of 27 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	G	501	GDP	O4'-C1'	2.16	1.43	1.41
6	J	501	GDP	O4'-C1'	2.34	1.44	1.41
6	E	501	GDP	O4'-C1'	2.35	1.44	1.41
6	F	501	GDP	O4'-C1'	2.46	1.44	1.41
6	I	501	GDP	C5-C4	2.52	1.46	1.40

The worst 5 of 124 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	E	504	MES	O2S-S-C8	-22.70	87.53	106.91
4	E	504	MES	O1S-S-C8	-21.89	88.23	106.91
6	L	501	GDP	C4'-O4'-C1'	-5.89	103.25	109.72
6	K	501	GDP	C2'-C1'-N9	-5.57	105.78	114.29
6	E	501	GDP	C2'-C1'-N9	-5.51	105.88	114.29

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

13 monomers are involved in 54 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	502	GDP	4	0
7	A	504	ALF	2	0
6	C	501	GDP	4	0
6	D	501	GDP	3	0
6	E	501	GDP	4	0
6	F	501	GDP	4	0
6	G	501	GDP	5	0
6	I	501	GDP	10	0
6	J	501	GDP	6	0
6	K	501	GDP	3	0
7	K	502	ALF	1	0
6	L	501	GDP	5	0
4	W	101	MES	4	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	V	14/14 (100%)	-0.16	1 (7%) 19 10	82, 91, 181, 202	0
2	W	13/13 (100%)	-0.55	0 100 100	92, 107, 167, 208	0
3	A	418/454 (92%)	-0.16	7 (1%) 73 60	69, 100, 162, 214	3 (0%)
3	B	430/454 (94%)	-0.04	9 (2%) 67 52	66, 111, 181, 236	0
3	C	434/454 (95%)	-0.14	5 (1%) 81 69	66, 102, 157, 213	0
3	D	431/454 (94%)	-0.10	3 (0%) 89 83	67, 109, 171, 229	0
3	E	418/454 (92%)	-0.06	7 (1%) 73 60	68, 108, 164, 221	0
3	F	421/454 (92%)	0.01	12 (2%) 55 41	66, 108, 167, 212	0
3	G	421/454 (92%)	0.45	30 (7%) 19 10	100, 165, 252, 322	0
3	H	425/454 (93%)	0.42	33 (7%) 16 9	95, 149, 253, 284	0
3	I	418/454 (92%)	0.23	9 (2%) 65 50	92, 158, 248, 326	0
3	J	416/454 (91%)	0.41	26 (6%) 23 13	89, 177, 285, 314	1 (0%)
3	K	421/454 (92%)	0.45	45 (10%) 8 4	77, 169, 283, 310	0
3	L	419/454 (92%)	0.59	46 (10%) 7 4	112, 174, 247, 301	0
All	All	5099/5475 (93%)	0.17	233 (4%) 36 23	66, 134, 228, 326	4 (0%)

The worst 5 of 233 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	H	405	ASP	12.1
3	H	400	ASP	7.6
3	L	270	PRO	7.4
3	H	404	LYS	6.9
3	K	256	GLY	6.8

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
4	MES	W	101	12/12	0.68	0.66	11.13	247,250,253,253	0
4	MES	C	505	12/12	0.64	0.85	10.12	232,233,234,235	0
7	ALF	E	503	5/5	0.97	0.36	4.90	88,88,100,101	0
5	CA	A	506	1/1	0.90	0.35	4.85	140,140,140,140	0
5	CA	E	502	1/1	0.99	0.34	3.77	110,110,110,110	0
7	ALF	D	503	5/5	0.98	0.37	3.71	80,88,97,109	0
5	CA	A	503	1/1	0.98	0.28	3.62	103,103,103,103	0
5	CA	C	502	1/1	0.98	0.40	3.24	101,101,101,101	0
7	ALF	C	504	5/5	0.99	0.32	3.20	67,74,92,94	0
7	ALF	J	502	5/5	0.93	0.38	3.16	119,120,129,130	0
7	ALF	C	503	5/5	0.99	0.31	1.73	66,70,86,88	0
7	ALF	G	502	5/5	0.92	0.40	1.55	120,127,139,141	0
7	ALF	A	504	5/5	0.96	0.21	0.67	84,89,97,97	0
6	GDP	F	501	28/28	0.92	0.24	0.61	71,90,105,111	0
5	CA	L	503	1/1	0.99	0.30	0.52	155,155,155,155	0
7	ALF	K	502	5/5	0.96	0.31	0.48	121,126,137,144	0
6	GDP	C	501	28/28	0.95	0.24	0.47	66,69,83,87	0
6	GDP	E	501	28/28	0.90	0.23	0.46	67,89,110,118	0
6	GDP	A	502	28/28	0.95	0.20	0.08	67,75,87,90	0
5	CA	A	501	1/1	0.92	0.23	-0.21	109,109,109,109	0
6	GDP	K	501	28/28	0.82	0.24	-0.40	165,178,204,208	0
6	GDP	D	501	28/28	0.95	0.18	-0.50	65,81,90,97	0
6	GDP	L	501	28/28	0.86	0.23	-0.55	129,162,173,177	0
7	ALF	I	503	5/5	0.98	0.28	-0.61	89,90,95,97	0
6	GDP	J	501	28/28	0.89	0.22	-0.71	140,157,182,186	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
6	GDP	G	501	28/28	0.92	0.23	-0.89	118,157,181,186	0
7	ALF	L	502	5/5	0.98	0.21	-1.49	118,121,124,126	0
4	MES	E	504	12/12	0.91	0.16	-1.63	122,127,144,148	0
6	GDP	I	501	28/28	0.93	0.23	-1.73	71,84,103,109	0
4	MES	I	504	12/12	0.88	0.16	-1.74	111,121,151,151	0
5	CA	B	501	1/1	0.81	0.58	-	147,147,147,147	0
5	CA	G	503	1/1	0.80	0.48	-	141,141,141,141	0
5	CA	A	507	1/1	0.84	0.27	-	134,134,134,134	0
5	CA	D	504	1/1	0.92	0.16	-	131,131,131,131	0
5	CA	D	502	1/1	0.98	0.26	-	90,90,90,90	0
5	CA	I	502	1/1	0.99	0.32	-	117,117,117,117	0
5	CA	F	502	1/1	0.97	0.37	-	92,92,92,92	0
5	CA	A	505	1/1	0.85	0.26	-	117,117,117,117	0

6.5 Other polymers

There are no such residues in this entry.