wwPDB X-ray Structure Validation Summary Report (i)

Mar 5, 2026 – 12:55 AM UTC

PDB ID	:	4XLP / pdb_00004xlp
Title	:	Crystal structure of T.aquaticus transcription initiation complex containing upstream fork promoter
Authors	:	Bae, B.; Darst, S.A.
Deposited on	:	2015-01-13
Resolution	:	4.00 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

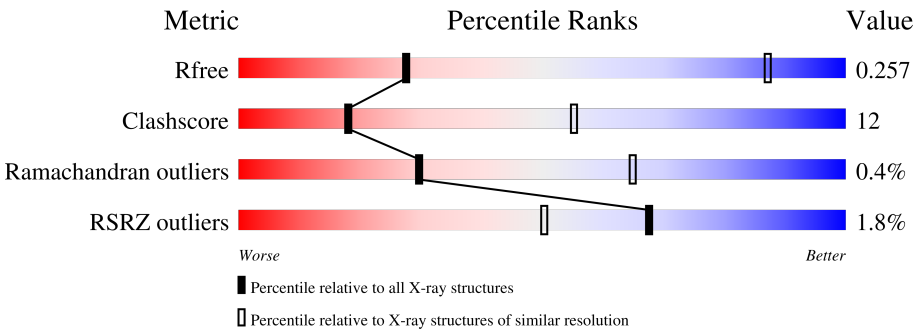
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 4.00 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}	180053	1082 (4.20-3.80)
Clashscore	190562	1129 (4.20-3.80)
Ramachandran outliers	187476	1064 (4.20-3.80)
RSRZ outliers	180081	1082 (4.20-3.80)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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	A	314	2% 55% 18% 28%
1	B	314	. % 51% 20% • 28%
1	G	314	2% 54% 18% • 28%
1	H	314	. % 50% 21% • 28%
2	C	1119	2% 68% 31% ••
2	I	1119	2% 67% 32% ••

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	1524	% 68% 29%
3	J	1524	% 63% 26% 10%
4	E	99	2% 69% 25% 6%
4	K	99	% 70% 24% 6%
5	F	347	4% 70% 29%
5	L	347	7% 68% 31%
6	O	30	73% 27%
6	R	30	77% 23%
7	P	25	56% 44%
7	S	25	52% 48%

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 56454 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 protein called DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	227	1770	1130	303	334	3	0	0	0
1	B	227	1770	1130	303	334	3	0	0	0
1	G	227	1770	1130	303	334	3	0	0	0
1	H	227	1770	1130	303	334	3	0	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	1112	8739	5531	1553	1632	23	0	0	0
2	I	1112	8739	5531	1553	1632	23	0	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	1490	11761	7439	2088	2196	38	0	0	0
3	J	1367	10779	6810	1923	2010	36	0	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	E	93	768	490	136	138	4	0	0	0
4	K	93	768	490	136	138	4	0	0	0

- Molecule 5 is a protein called RNA polymerase sigma factor SigA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	F	345	2787	1758	502	523	4	0	0	0
5	L	345	2787	1758	502	523	4	0	0	0

- Molecule 6 is a DNA chain called DNA (30-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
6	O	30	613	296	109	179	29	0	0	0
6	R	30	613	296	109	179	29	0	0	0

- Molecule 7 is a DNA chain called DNA (25-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
7	P	25	507	245	91	147	24	0	0	0
7	S	25	507	245	91	147	24	0	0	0

- Molecule 8 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
8	D	2	2	2	0	0
8	J	2	2	2	0	0

- Molecule 9 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

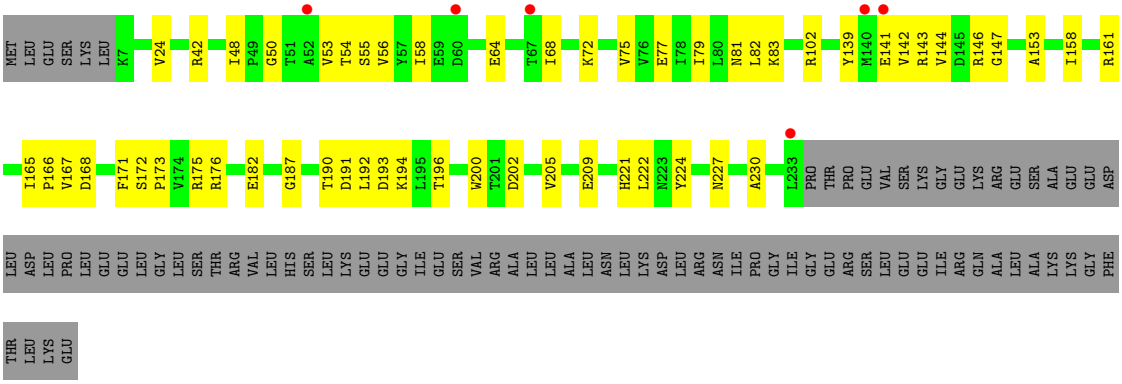
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
9	D	1	1	1	0	0
9	J	1	1	1	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

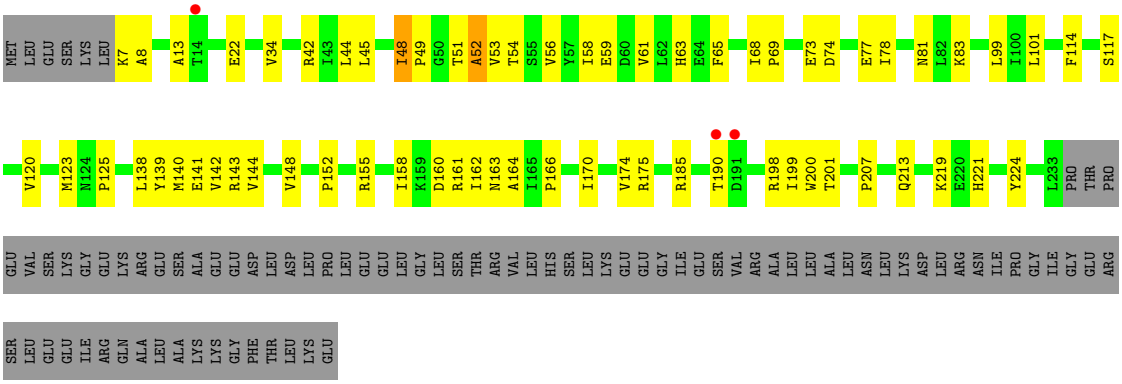
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A:



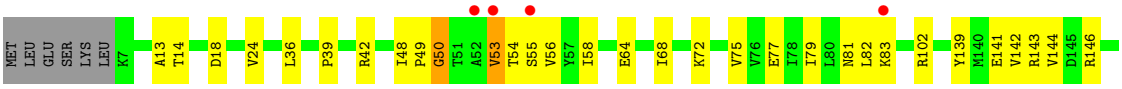
• Molecule 1: DNA-directed RNA polymerase subunit alpha

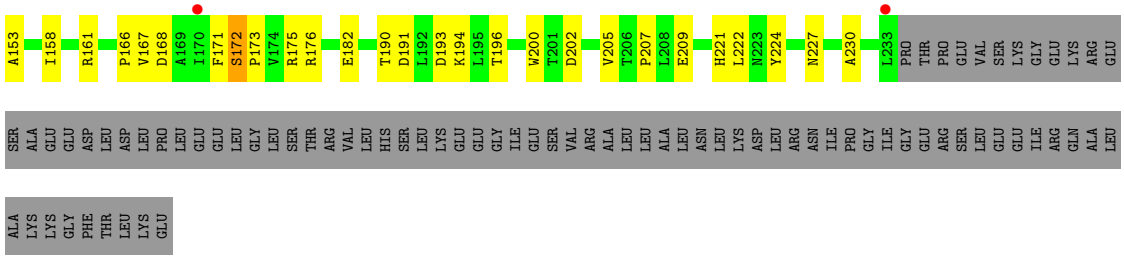
Chain B:



• Molecule 1: DNA-directed RNA polymerase subunit alpha

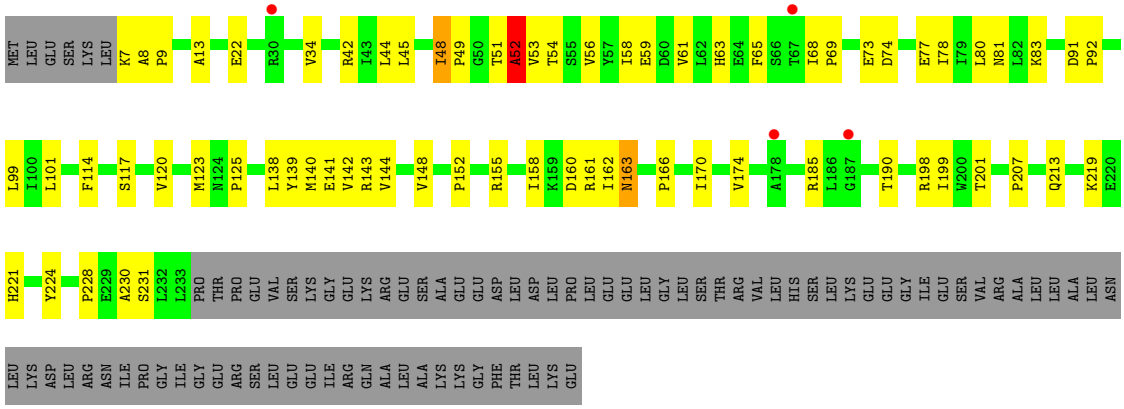
Chain G:





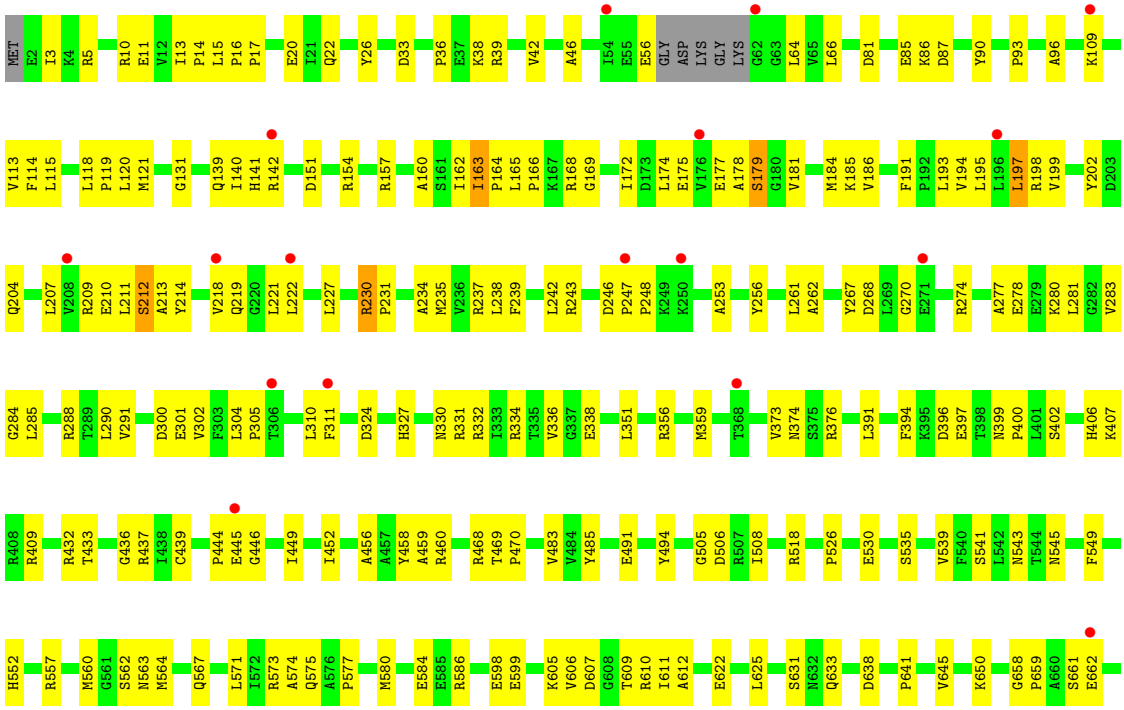
- Molecule 1: DNA-directed RNA polymerase subunit alpha

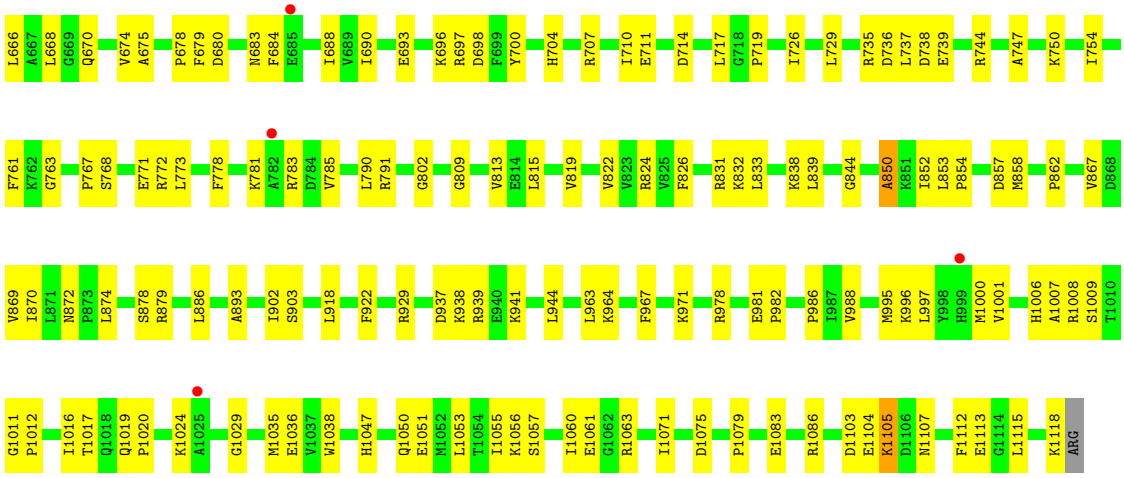
Chain H:



- Molecule 2: DNA-directed RNA polymerase subunit beta

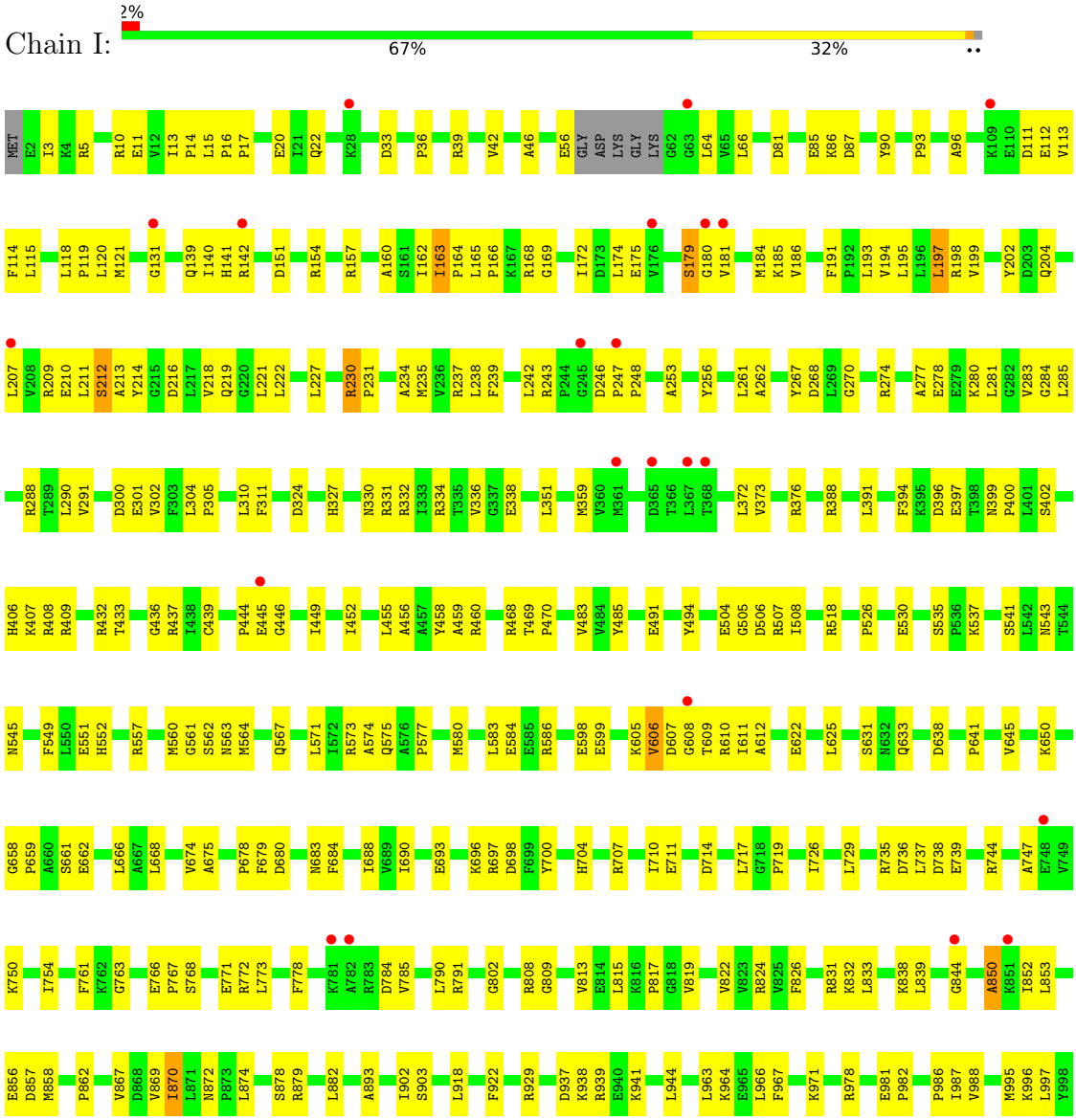
Chain C:

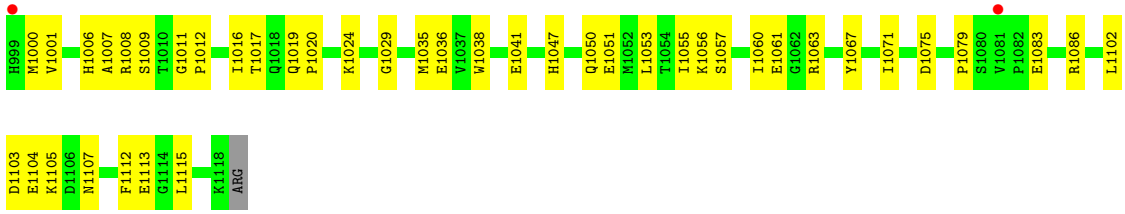




• Molecule 2: DNA-directed RNA polymerase subunit beta

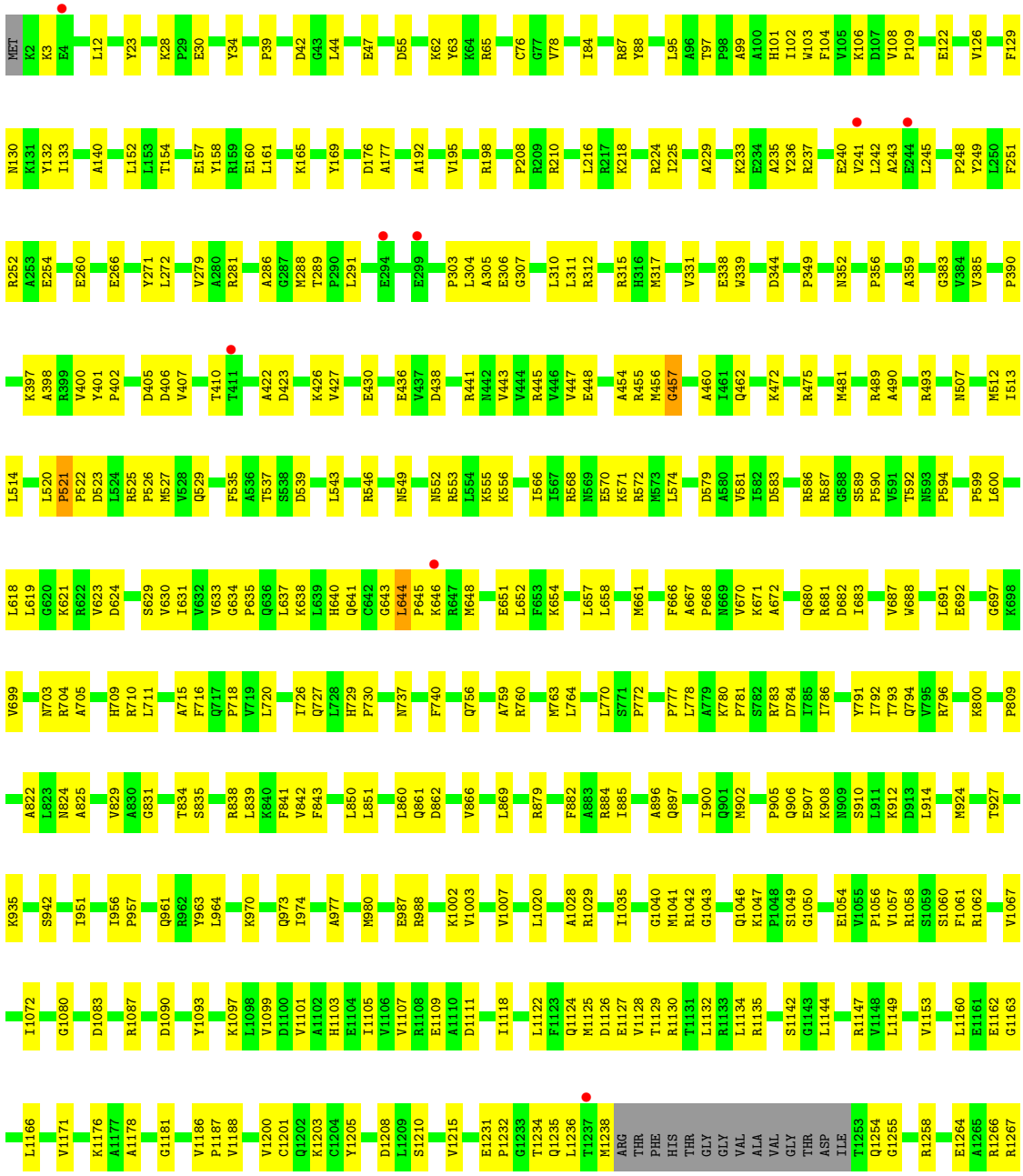
Chain I:

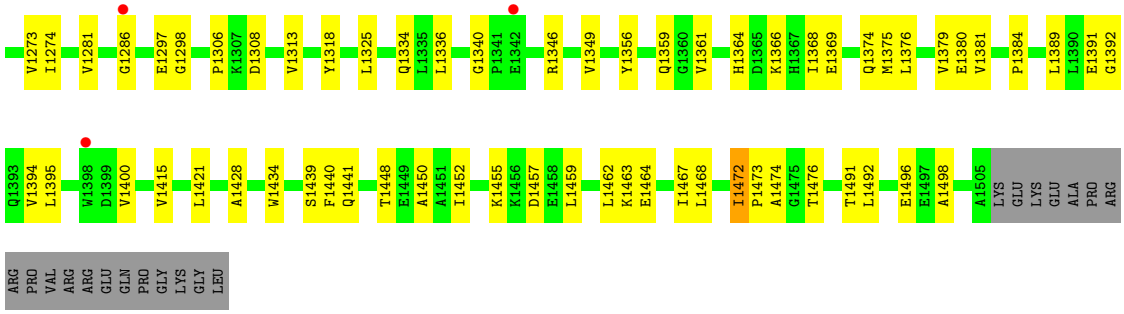




• Molecule 3: DNA-directed RNA polymerase subunit beta'

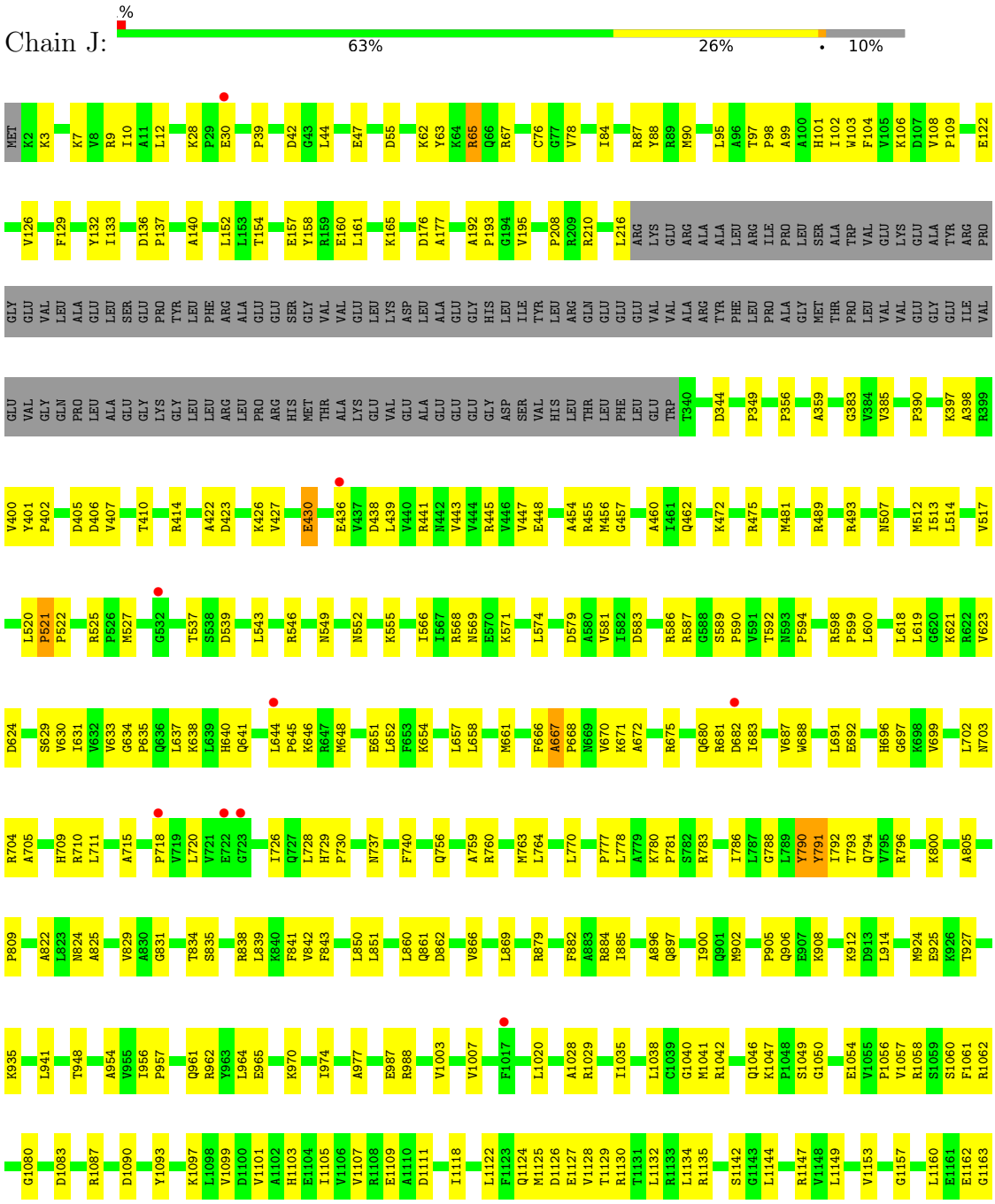
Chain D:

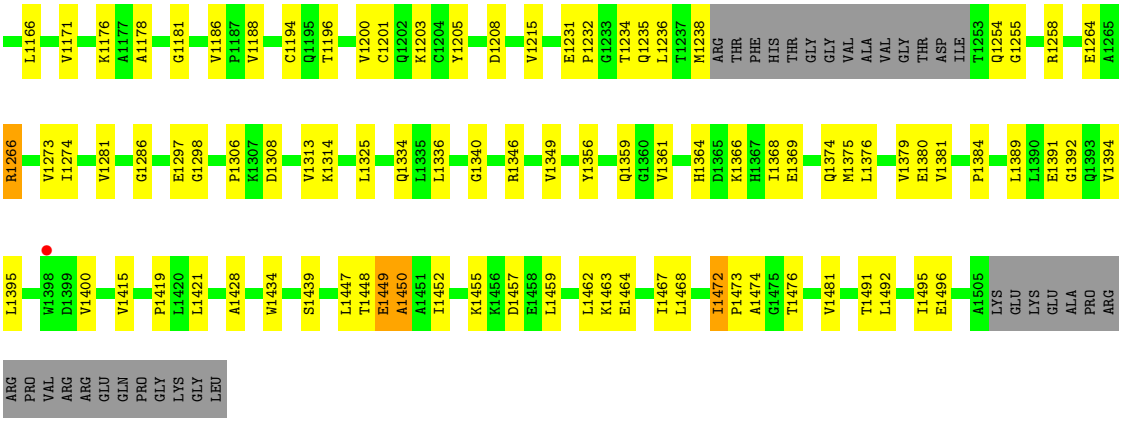




- Molecule 3: DNA-directed RNA polymerase subunit beta'

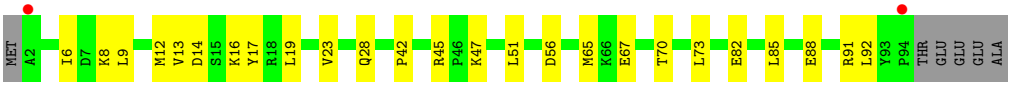
Chain J:





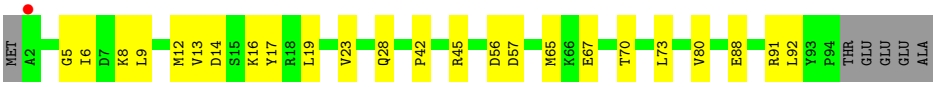
- Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E: 2% 69% 25% 6%



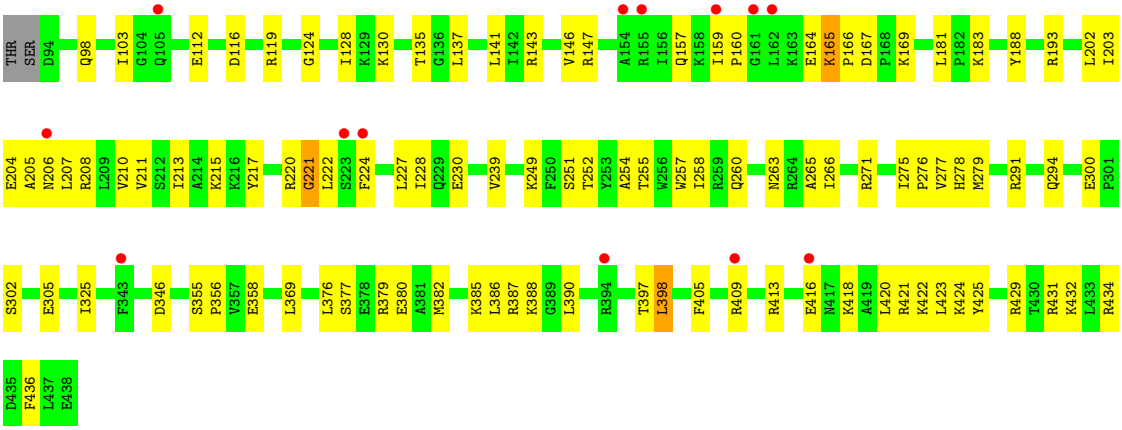
- Molecule 4: DNA-directed RNA polymerase subunit omega

Chain K: 2% 70% 24% 6%



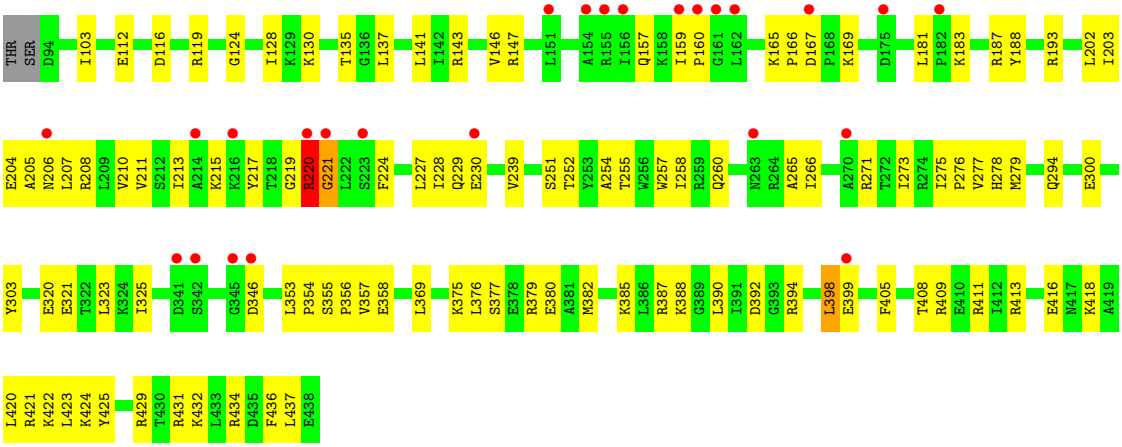
- Molecule 5: RNA polymerase sigma factor SigA

Chain F: 4% 70% 29% ••



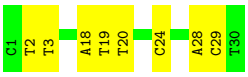
- Molecule 5: RNA polymerase sigma factor SigA

Chain L: 7% 68% 31% ••



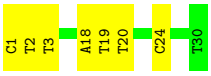
- Molecule 6: DNA (30-MER)

Chain O:



- Molecule 6: DNA (30-MER)

Chain R:

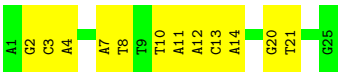


- Molecule 7: DNA (25-MER)

Chain P:

- Molecule 7: DNA (25-MER)

Chain S:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	289.87Å 289.87Å 537.36Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.03 – 4.00 50.03 – 4.00	Depositor EDS
% Data completeness (in resolution range)	96.3 (50.03-4.00) 96.3 (50.03-4.00)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.92 (at 3.77Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1839)	Depositor
R, R_{free}	0.257 , 0.297 (Not available) , 0.257	Depositor DCC
R_{free} test set	9509 reflections (4.26%)	wwPDB-VP
Wilson B-factor (Å^2)	135.6	Xtriage
Anisotropy	0.002	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 126.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.36$, $< L^2 > = 0.19$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	56454	wwPDB-VP
Average B, all atoms (Å^2)	141.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.82% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, MG

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	A	0.34	0/1804	0.86	2/2455 (0.1%)
1	B	0.32	0/1804	0.88	3/2455 (0.1%)
1	G	0.35	0/1804	0.88	3/2455 (0.1%)
1	H	0.33	0/1804	0.89	4/2455 (0.2%)
2	C	0.33	0/8905	0.86	16/12040 (0.1%)
2	I	0.33	0/8905	0.86	16/12040 (0.1%)
3	D	0.32	0/11963	0.83	15/16165 (0.1%)
3	J	0.32	0/10959	0.84	13/14802 (0.1%)
4	E	0.27	0/783	0.76	0/1054
4	K	0.27	0/783	0.74	0/1054
5	F	0.36	0/2829	0.87	6/3804 (0.2%)
5	L	0.37	0/2829	0.89	8/3804 (0.2%)
6	O	0.13	0/687	0.60	0/1059
6	R	0.12	0/687	0.60	0/1059
7	P	0.13	0/568	0.66	0/874
7	S	0.13	0/568	0.65	0/874
All	All	0.32	0/57682	0.84	86/78449 (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
2	C	0	6
2	I	0	5
3	J	0	3
5	F	0	2
5	L	0	3
All	All	0	19

There are no bond length outliers.

The worst 5 of 86 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	I	179	SER	N-CA-C	-11.49	97.90	113.18
3	J	1286	GLY	N-CA-C	-9.28	98.27	113.02
3	D	1286	GLY	N-CA-C	-9.27	98.28	113.02
2	C	213	ALA	N-CA-C	-8.98	101.86	113.17
2	I	213	ALA	N-CA-C	-8.96	101.88	113.17

There are no chirality outliers.

5 of 19 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	178	ALA	Peptide
2	C	197	LEU	Peptide
2	C	212	SER	Peptide
2	C	230	ARG	Peptide
2	C	737	LEU	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	A	1770	0	1799	37	0
1	B	1770	0	1799	46	0
1	G	1770	0	1799	41	0
1	H	1770	0	1799	50	0
2	C	8739	0	8841	248	0
2	I	8739	0	8841	257	0
3	D	11761	0	11976	320	0
3	J	10779	0	10993	298	0
4	E	768	0	784	18	0
4	K	768	0	784	18	0
5	F	2787	0	2866	74	0
5	L	2787	0	2866	82	0
6	O	613	0	343	8	0
6	R	613	0	343	7	0
7	P	507	0	285	7	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	S	507	0	285	9	0
8	D	2	0	0	0	0
8	J	2	0	0	0	0
9	D	1	0	0	0	0
9	J	1	0	0	0	0
All	All	56454	0	56403	1374	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 12.

The worst 5 of 1374 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)
5:L:398:LEU:HG	5:L:409:ARG:HB2	1.45	0.99
5:F:398:LEU:HG	5:F:409:ARG:HB2	1.48	0.91
1:H:53:VAL:HG23	1:H:144:VAL:HG22	1.55	0.88
1:B:53:VAL:HG23	1:B:144:VAL:HG22	1.55	0.86
1:A:42:ARG:HH12	2:C:857:ASP:HB3	1.41	0.86

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	
1	G	225/314 (72%)	209 (93%)	15 (7%)	1 (0%)	30	65
1	H	225/314 (72%)	208 (92%)	16 (7%)	1 (0%)	30	65
2	I	1108/1119 (99%)	1040 (94%)	62 (6%)	6 (0%)	24	60
3	D	694/1524 (46%)	663 (96%)	30 (4%)	1 (0%)	48	81
3	J	1361/1524 (89%)	1281 (94%)	74 (5%)	6 (0%)	30	65

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	E	91/99 (92%)	87 (96%)	4 (4%)	0	100	100
4	K	91/99 (92%)	88 (97%)	3 (3%)	0	100	100
5	F	343/347 (99%)	324 (94%)	17 (5%)	2 (1%)	21	57
5	L	343/347 (99%)	326 (95%)	15 (4%)	2 (1%)	21	57
All	All	4481/5687 (79%)	4226 (94%)	236 (5%)	19 (0%)	30	65

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	1128	VAL
1	G	53	VAL
3	J	1128	VAL
2	I	111	ASP
3	J	791	TYR

5.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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	A	227/314 (72%)	0.12	6 (2%)	57	42	87, 140, 176, 201	0
1	B	227/314 (72%)	0.03	3 (1%)	75	57	62, 117, 166, 206	0
1	G	227/314 (72%)	0.33	6 (2%)	57	42	124, 168, 194, 220	0
1	H	227/314 (72%)	0.23	4 (1%)	67	50	90, 147, 178, 196	0
2	C	1112/1119 (99%)	0.15	21 (1%)	66	49	59, 139, 200, 249	0
2	I	1112/1119 (99%)	0.23	24 (2%)	62	46	80, 159, 207, 252	0
3	D	1490/1524 (97%)	-0.02	11 (0%)	84	68	52, 113, 165, 216	0
3	J	1367/1524 (89%)	0.09	10 (0%)	84	68	66, 134, 184, 230	0
4	E	93/99 (93%)	0.07	2 (2%)	62	46	65, 117, 169, 188	0
4	K	93/99 (93%)	0.22	1 (1%)	78	60	92, 145, 192, 221	0
5	F	345/347 (99%)	0.36	13 (3%)	44	34	76, 146, 205, 235	0
5	L	345/347 (99%)	0.50	25 (7%)	21	20	100, 161, 214, 248	0
6	O	30/30 (100%)	-0.01	0	100	100	107, 178, 224, 227	0
6	R	30/30 (100%)	0.09	0	100	100	134, 169, 201, 211	0
7	P	25/25 (100%)	0.08	0	100	100	138, 174, 220, 243	0
7	S	25/25 (100%)	0.08	0	100	100	141, 177, 196, 203	0
All	All	6975/7544 (92%)	0.14	126 (1%)	67	50	52, 139, 196, 252	0

The worst 5 of 126 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	F	159	ILE	5.9
5	L	160	PRO	5.4
5	L	159	ILE	5.3
1	A	52	ALA	4.8
2	I	368	THR	4.6

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 oligosaccharides 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. The B-factors column lists the minimum, median, 95th 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	B-factors(Å^2)	Q<0.9
9	MG	J	2003	1/1	0.68	0.13	255,255,255,255	0
9	MG	D	2003	1/1	0.82	0.15	205,205,205,205	0
8	ZN	D	2002	1/1	0.98	0.04	118,118,118,118	0
8	ZN	J	2002	1/1	0.98	0.03	154,154,154,154	0
8	ZN	J	2001	1/1	1.00	0.05	151,151,151,151	0
8	ZN	D	2001	1/1	1.00	0.04	107,107,107,107	0

6.5 Other polymers (i)

There are no such residues in this entry.