wwPDB X-ray Structure Validation Summary Report (i)

Sep 28, 2024 – 04:21 PM EDT

PDB ID	:	1PV4
Title	:	X-ray crystal structure of the Rho transcription termination factor in complex with single stranded DNA
Authors	:	Skordalakes, E.; Berger, J.M.
Deposited on	:	2003-06-26
Resolution	:	3.00 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:

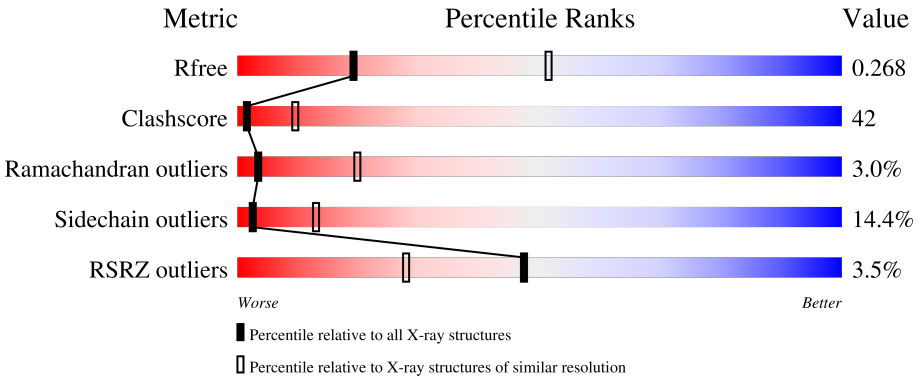
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 3.00 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.



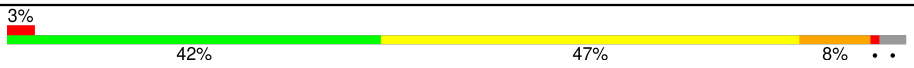
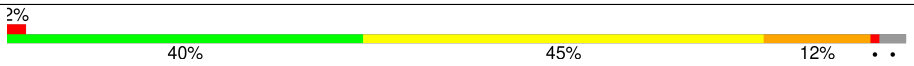
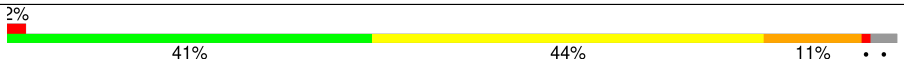
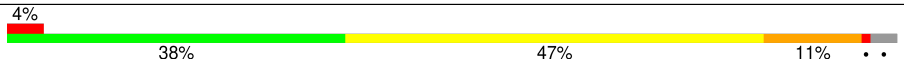
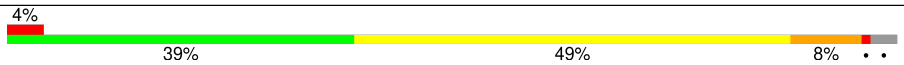
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	2511 (3.00-3.00)
Clashscore	180529	2866 (3.00-3.00)
Ramachandran outliers	177936	2778 (3.00-3.00)
Sidechain outliers	177891	2781 (3.00-3.00)
RSRZ outliers	164620	2523 (3.00-3.00)

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	G	2	100%
1	H	2	50% 50% 50%
1	J	2	50% 50% 50%
1	K	2	50% 50%
1	L	2	50% 100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	A	419	
2	B	419	
2	C	419	
2	D	419	
2	E	419	
2	F	419	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 19077 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*CP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	2	Total	C	N	O	P	0	0	0
			38	18	6	12	2			
1	H	2	Total	C	N	O	P	0	0	0
			38	18	6	12	2			
1	J	2	Total	C	N	O	P	0	0	0
			38	18	6	12	2			
1	K	2	Total	C	N	O	P	0	0	0
			38	18	6	12	2			
1	L	2	Total	C	N	O	P	0	0	0
			38	18	6	12	2			

- Molecule 2 is a protein called Transcription termination factor rho.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	A	408	Total	C	N	O	S	Se	0	0	0
			3208	2025	562	604	1	16			
2	B	358	Total	C	N	O	S	Se	0	0	0
			2813	1776	494	529	1	13			
2	C	408	Total	C	N	O	S	Se	0	0	0
			3208	2025	562	604	1	16			
2	D	408	Total	C	N	O	S	Se	0	0	0
			3208	2025	562	604	1	16			
2	E	407	Total	C	N	O	S	Se	0	0	0
			3201	2020	561	603	1	16			
2	F	408	Total	C	N	O	S	Se	0	0	0
			3208	2025	562	604	1	16			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	modified residue	UNP P0AG30
A	21	MSE	MET	modified residue	UNP P0AG30
A	29	MSE	MET	modified residue	UNP P0AG30

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	147	MSE	MET	modified residue	UNP P0AG30
A	186	MSE	MET	modified residue	UNP P0AG30
A	205	MSE	MET	modified residue	UNP P0AG30
A	219	MSE	MET	modified residue	UNP P0AG30
A	245	MSE	MET	modified residue	UNP P0AG30
A	327	MSE	MET	modified residue	UNP P0AG30
A	341	MSE	MET	modified residue	UNP P0AG30
A	380	MSE	MET	modified residue	UNP P0AG30
A	390	MSE	MET	modified residue	UNP P0AG30
A	396	MSE	MET	modified residue	UNP P0AG30
A	405	MSE	MET	modified residue	UNP P0AG30
A	415	MSE	MET	modified residue	UNP P0AG30
A	416	MSE	MET	modified residue	UNP P0AG30
B	1	MSE	MET	modified residue	UNP P0AG30
B	21	MSE	MET	modified residue	UNP P0AG30
B	29	MSE	MET	modified residue	UNP P0AG30
B	147	MSE	MET	modified residue	UNP P0AG30
B	186	MSE	MET	modified residue	UNP P0AG30
B	205	MSE	MET	modified residue	UNP P0AG30
B	219	MSE	MET	modified residue	UNP P0AG30
B	245	MSE	MET	modified residue	UNP P0AG30
B	327	MSE	MET	modified residue	UNP P0AG30
B	341	MSE	MET	modified residue	UNP P0AG30
B	380	MSE	MET	modified residue	UNP P0AG30
B	390	MSE	MET	modified residue	UNP P0AG30
B	396	MSE	MET	modified residue	UNP P0AG30
B	405	MSE	MET	modified residue	UNP P0AG30
B	415	MSE	MET	modified residue	UNP P0AG30
B	416	MSE	MET	modified residue	UNP P0AG30
C	1	MSE	MET	modified residue	UNP P0AG30
C	21	MSE	MET	modified residue	UNP P0AG30
C	29	MSE	MET	modified residue	UNP P0AG30
C	147	MSE	MET	modified residue	UNP P0AG30
C	186	MSE	MET	modified residue	UNP P0AG30
C	205	MSE	MET	modified residue	UNP P0AG30
C	219	MSE	MET	modified residue	UNP P0AG30
C	245	MSE	MET	modified residue	UNP P0AG30
C	327	MSE	MET	modified residue	UNP P0AG30
C	341	MSE	MET	modified residue	UNP P0AG30
C	380	MSE	MET	modified residue	UNP P0AG30
C	390	MSE	MET	modified residue	UNP P0AG30
C	396	MSE	MET	modified residue	UNP P0AG30

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	405	MSE	MET	modified residue	UNP P0AG30
C	415	MSE	MET	modified residue	UNP P0AG30
C	416	MSE	MET	modified residue	UNP P0AG30
D	1	MSE	MET	modified residue	UNP P0AG30
D	21	MSE	MET	modified residue	UNP P0AG30
D	29	MSE	MET	modified residue	UNP P0AG30
D	147	MSE	MET	modified residue	UNP P0AG30
D	186	MSE	MET	modified residue	UNP P0AG30
D	205	MSE	MET	modified residue	UNP P0AG30
D	219	MSE	MET	modified residue	UNP P0AG30
D	245	MSE	MET	modified residue	UNP P0AG30
D	327	MSE	MET	modified residue	UNP P0AG30
D	341	MSE	MET	modified residue	UNP P0AG30
D	380	MSE	MET	modified residue	UNP P0AG30
D	390	MSE	MET	modified residue	UNP P0AG30
D	396	MSE	MET	modified residue	UNP P0AG30
D	405	MSE	MET	modified residue	UNP P0AG30
D	415	MSE	MET	modified residue	UNP P0AG30
D	416	MSE	MET	modified residue	UNP P0AG30
E	1	MSE	MET	modified residue	UNP P0AG30
E	21	MSE	MET	modified residue	UNP P0AG30
E	29	MSE	MET	modified residue	UNP P0AG30
E	147	MSE	MET	modified residue	UNP P0AG30
E	186	MSE	MET	modified residue	UNP P0AG30
E	205	MSE	MET	modified residue	UNP P0AG30
E	219	MSE	MET	modified residue	UNP P0AG30
E	245	MSE	MET	modified residue	UNP P0AG30
E	327	MSE	MET	modified residue	UNP P0AG30
E	341	MSE	MET	modified residue	UNP P0AG30
E	380	MSE	MET	modified residue	UNP P0AG30
E	390	MSE	MET	modified residue	UNP P0AG30
E	396	MSE	MET	modified residue	UNP P0AG30
E	405	MSE	MET	modified residue	UNP P0AG30
E	415	MSE	MET	modified residue	UNP P0AG30
E	416	MSE	MET	modified residue	UNP P0AG30
F	1	MSE	MET	modified residue	UNP P0AG30
F	21	MSE	MET	modified residue	UNP P0AG30
F	29	MSE	MET	modified residue	UNP P0AG30
F	147	MSE	MET	modified residue	UNP P0AG30
F	186	MSE	MET	modified residue	UNP P0AG30
F	205	MSE	MET	modified residue	UNP P0AG30
F	219	MSE	MET	modified residue	UNP P0AG30

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	245	MSE	MET	modified residue	UNP P0AG30
F	327	MSE	MET	modified residue	UNP P0AG30
F	341	MSE	MET	modified residue	UNP P0AG30
F	380	MSE	MET	modified residue	UNP P0AG30
F	390	MSE	MET	modified residue	UNP P0AG30
F	396	MSE	MET	modified residue	UNP P0AG30
F	405	MSE	MET	modified residue	UNP P0AG30
F	415	MSE	MET	modified residue	UNP P0AG30
F	416	MSE	MET	modified residue	UNP P0AG30

- Molecule 3 is water.

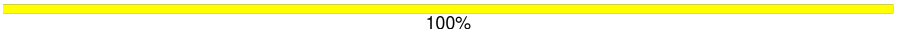
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	G	1	Total O 1 1	0	0
3	J	1	Total O 1 1	0	0
3	A	3	Total O 3 3	0	0
3	B	2	Total O 2 2	0	0
3	C	7	Total O 7 7	0	0
3	D	7	Total O 7 7	0	0
3	E	14	Total O 14 14	0	0
3	F	6	Total O 6 6	0	0

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: 5'-D(P*CP*C)-3'

Chain G:



- Molecule 1: 5'-D(P*CP*C)-3'

Chain H:

- Molecule 1: 5'-D(P*CP*C)-3'

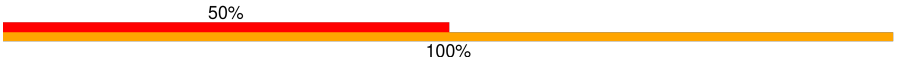
Chain J:

- Molecule 1: 5'-D(P*CP*C)-3'

Chain K:

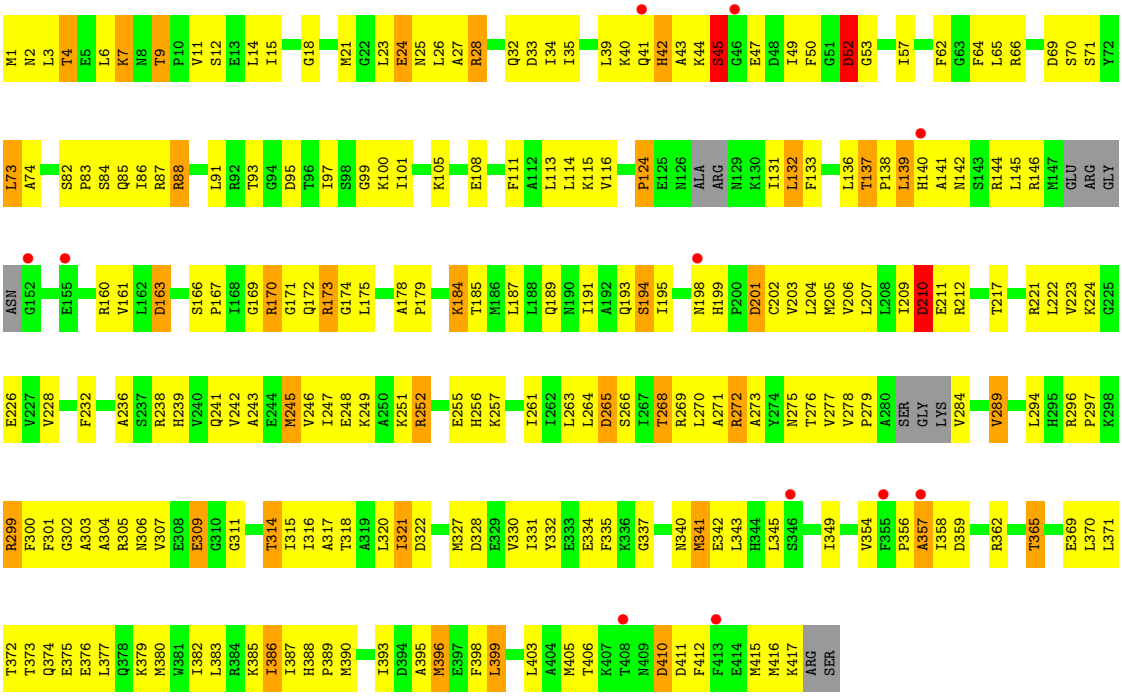
- Molecule 1: 5'-D(P*CP*C)-3'

Chain L:



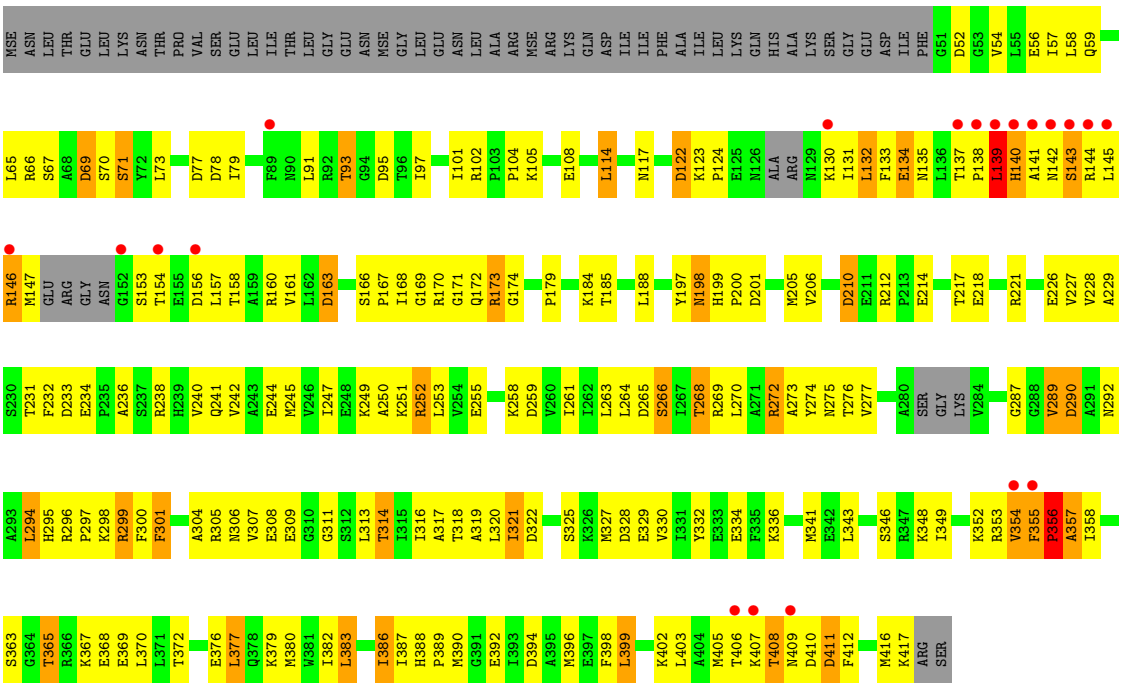
- Molecule 2: Transcription termination factor rho

Chain A:



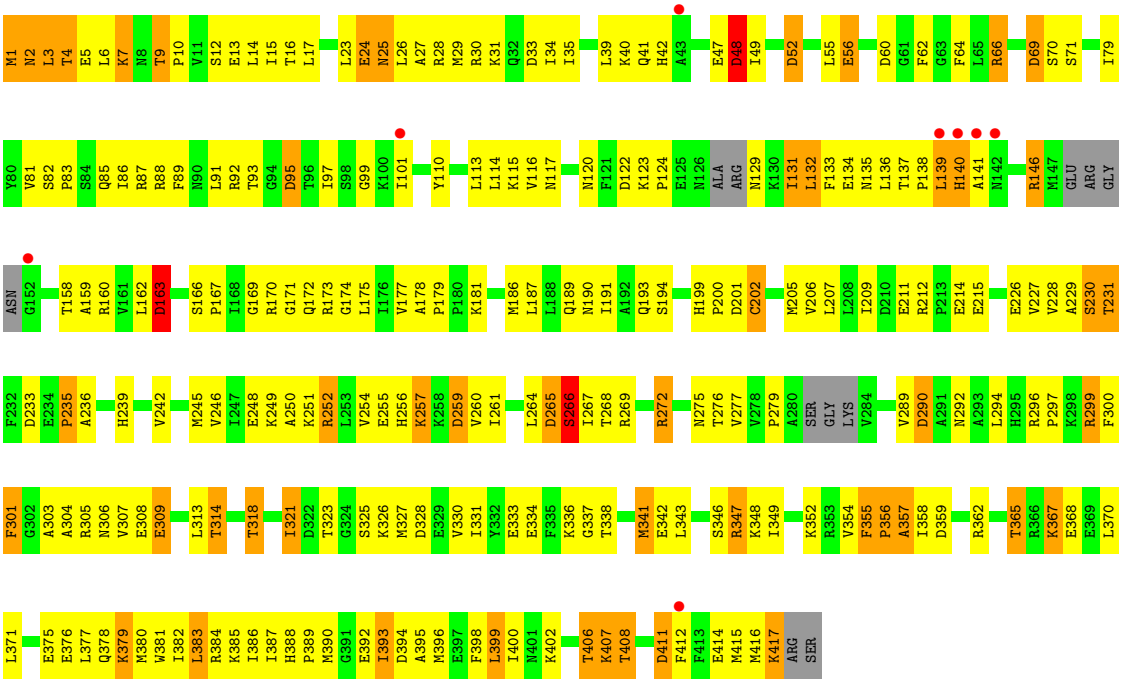
• Molecule 2: Transcription termination factor rho

Chain B:



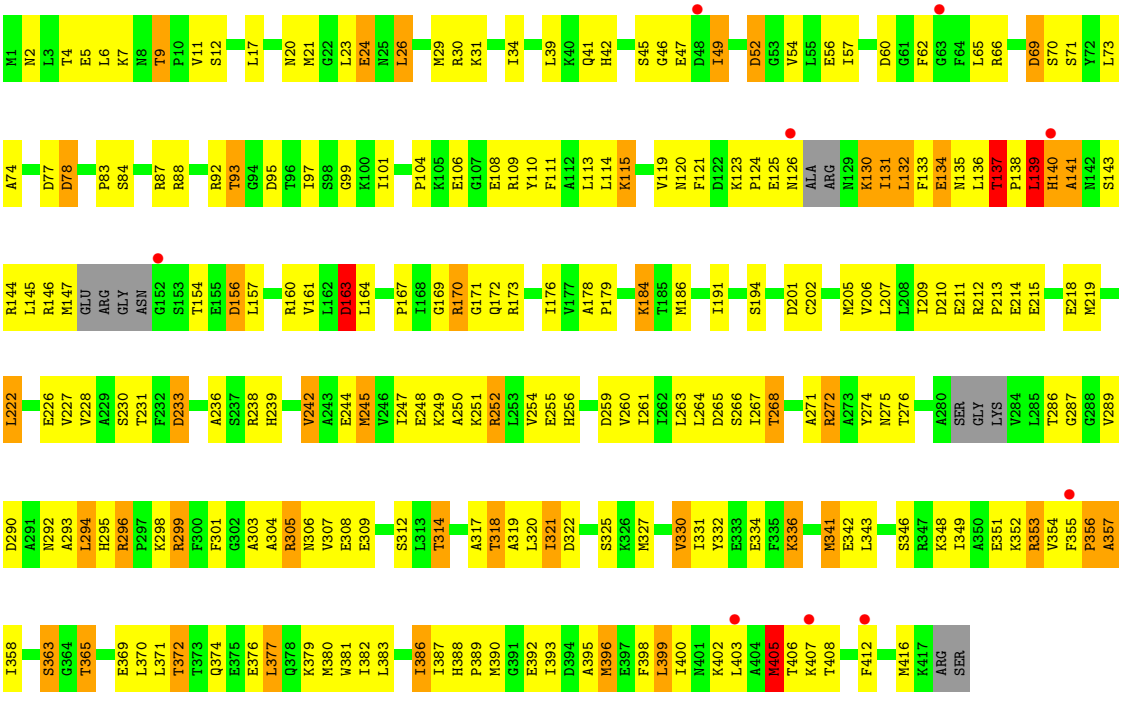
• Molecule 2: Transcription termination factor rho

Chain C:



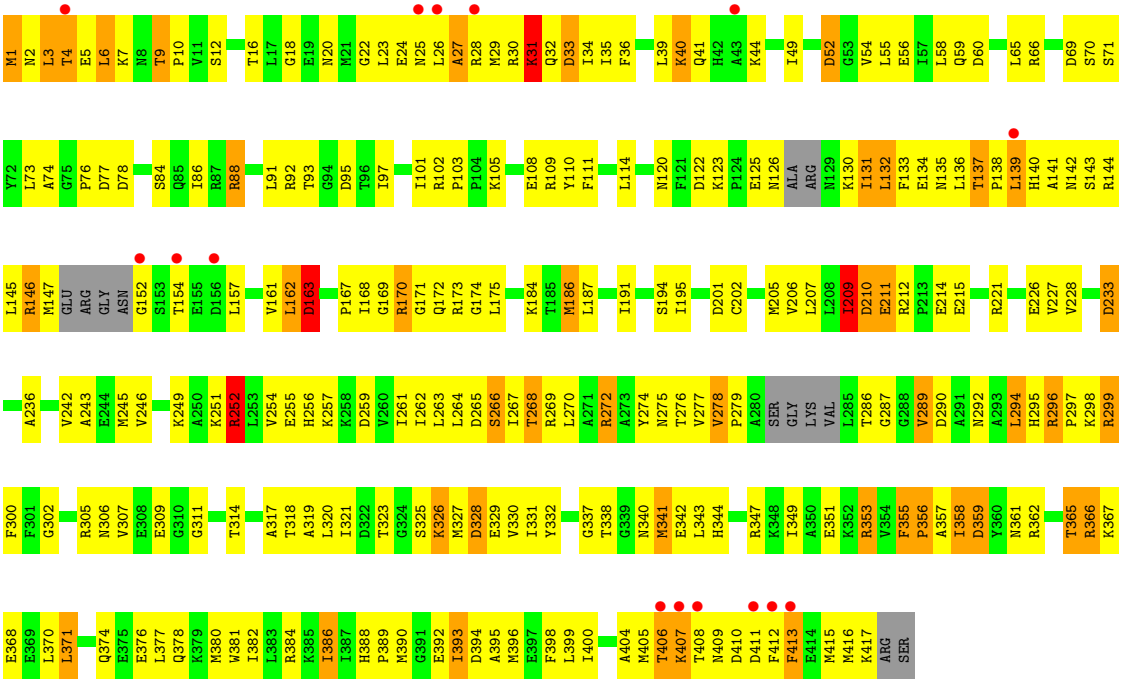
• Molecule 2: Transcription termination factor rho

Chain D:



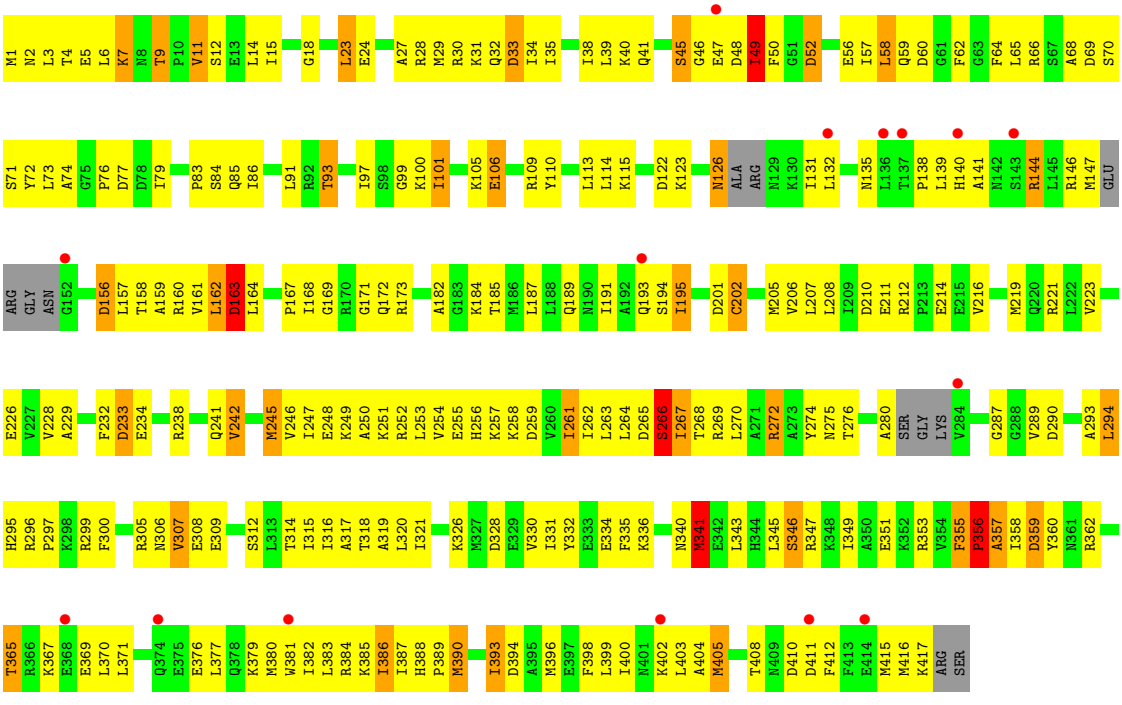
• Molecule 2: Transcription termination factor rho

Chain E:



• Molecule 2: Transcription termination factor rho

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	118.56Å 204.33Å 147.37Å 90.00° 95.86° 90.00°	Depositor
Resolution (Å)	20.00 – 3.00 20.00 – 3.00	Depositor EDS
% Data completeness (in resolution range)	97.9 (20.00-3.00) 96.7 (20.00-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	1.97 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.270 , 0.296 0.244 , 0.268	Depositor DCC
R_{free} test set	3379 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	86.7	Xtriage
Anisotropy	0.162	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 89.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	19077	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.08% 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

5.1 Standard geometry

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	G	0.67	0/41	1.16	0/60
1	H	0.89	0/41	1.34	1/60 (1.7%)
1	J	0.99	0/41	1.58	1/60 (1.7%)
1	K	0.95	0/41	1.59	1/60 (1.7%)
1	L	0.93	0/41	1.79	3/60 (5.0%)
2	A	0.57	0/3238	0.77	9/4334 (0.2%)
2	B	0.57	0/2841	0.80	13/3805 (0.3%)
2	C	0.76	0/3238	0.88	11/4334 (0.3%)
2	D	0.74	0/3238	0.90	11/4334 (0.3%)
2	E	0.70	0/3231	0.89	11/4324 (0.3%)
2	F	0.60	0/3238	0.80	11/4334 (0.3%)
All	All	0.67	0/19229	0.85	72/25765 (0.3%)

There are no bond length outliers.

The worst 5 of 72 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	233	ASP	CB-CG-OD2	10.35	127.61	118.30
2	E	163	ASP	CB-CG-OD2	7.97	125.47	118.30
2	D	233	ASP	CB-CG-OD1	-7.79	111.29	118.30
2	D	77	ASP	CB-CG-OD2	7.54	125.08	118.30
2	A	265	ASP	CB-CG-OD2	7.30	124.87	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	G	38	0	23	2	0
1	H	38	0	23	3	0
1	J	38	0	23	5	0
1	K	38	0	23	5	0
1	L	38	0	23	8	0
2	A	3208	0	3284	278	0
2	B	2813	0	2870	229	0
2	C	3208	0	3284	308	0
2	D	3208	0	3284	273	0
2	E	3201	0	3275	308	0
2	F	3208	0	3284	259	0
3	A	3	0	0	2	0
3	B	2	0	0	2	0
3	C	7	0	0	1	0
3	D	7	0	0	3	0
3	E	14	0	0	6	0
3	F	6	0	0	3	0
3	G	1	0	0	0	0
3	J	1	0	0	0	0
All	All	19077	0	19396	1597	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 42.

The worst 5 of 1597 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)
2:F:355:PHE:HB3	2:F:356:PRO:CD	1.59	1.33
2:C:379:LYS:HG2	2:C:412:PHE:CD2	1.72	1.23
2:C:177:VAL:CG1	2:C:321:ILE:HD12	1.69	1.22
2:A:265:ASP:O	2:A:318:THR:HB	1.36	1.22
2:C:141:ALA:CB	2:C:370:LEU:HB2	1.72	1.19

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
2	A	400/419 (96%)	363 (91%)	30 (8%)	7 (2%)	7 32
2	B	350/419 (84%)	306 (87%)	31 (9%)	13 (4%)	2 15
2	C	400/419 (96%)	354 (88%)	37 (9%)	9 (2%)	5 26
2	D	400/419 (96%)	352 (88%)	33 (8%)	15 (4%)	2 15
2	E	399/419 (95%)	346 (87%)	39 (10%)	14 (4%)	3 16
2	F	400/419 (96%)	352 (88%)	35 (9%)	13 (3%)	3 18
All	All	2349/2514 (93%)	2073 (88%)	205 (9%)	71 (3%)	3 20

5 of 71 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	357	ALA
2	B	140	HIS
2	B	355	PHE
2	C	140	HIS
2	C	408	THR

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
2	A	350/343 (102%)	308 (88%)	42 (12%)	4 18
2	B	306/343 (89%)	273 (89%)	33 (11%)	5 22
2	C	350/343 (102%)	288 (82%)	62 (18%)	1 8
2	D	350/343 (102%)	291 (83%)	59 (17%)	1 9
2	E	349/343 (102%)	291 (83%)	58 (17%)	2 9
2	F	350/343 (102%)	309 (88%)	41 (12%)	4 19
All	All	2055/2058 (100%)	1760 (86%)	295 (14%)	2 13

5 of 295 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	286	THR
2	F	341	MSE
2	E	341	MSE
2	F	40	LYS
2	C	146	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	129	ASN
2	E	295	HIS
2	E	306	ASN
2	E	275	ASN
2	D	41	GLN

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)

There are no ligands in this entry.

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	G	2/2 (100%)	1.42	0 100 100	150, 150, 150, 150	0
1	H	2/2 (100%)	1.85	1 (50%) 0 0	150, 150, 150, 150	0
1	J	2/2 (100%)	1.91	1 (50%) 0 0	150, 150, 150, 150	0
1	K	2/2 (100%)	1.95	0 100 100	150, 150, 150, 150	0
1	L	2/2 (100%)	1.98	1 (50%) 0 0	150, 150, 150, 150	0
2	A	392/419 (93%)	0.03	11 (2%) 55 33	35, 69, 121, 127	0
2	B	345/419 (82%)	0.08	20 (5%) 30 17	30, 69, 114, 127	0
2	C	392/419 (93%)	-0.30	8 (2%) 64 43	18, 44, 96, 113	0
2	D	392/419 (93%)	-0.34	9 (2%) 61 39	14, 45, 77, 126	0
2	E	391/419 (93%)	-0.25	15 (3%) 44 26	18, 47, 87, 126	0
2	F	392/419 (93%)	0.00	15 (3%) 44 26	36, 67, 114, 124	0
All	All	2314/2524 (91%)	-0.12	81 (3%) 47 28	14, 57, 110, 150	0

The worst 5 of 81 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	139	LEU	5.3
2	D	355	PHE	5.2
2	F	140	HIS	4.5
2	A	152	GLY	4.4
2	B	141	ALA	4.3

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.