wwPDB X-ray Structure Validation Summary Report (i)

May 23, 2020 – 09:08 am BST

PDB ID	:	3QMZ
Title	:	Crystal structure of the cytoplasmic dynein heavy chain motor domain
Authors	:	Cho, C.; Carter, A.P.; Jin, L.; Vale, R.D.
Deposited on	:	2011-02-07
Resolution	:	6.00 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 following versions of software and data (see references (1)) were used in the production of this report:

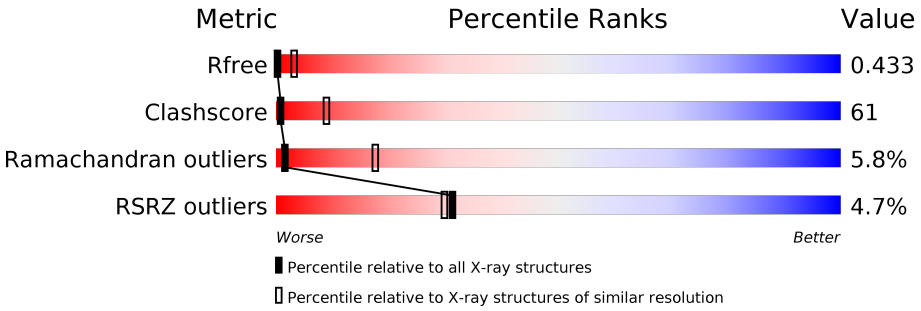
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 6.00 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}	130704	1000 (8.00-3.88)
Clashscore	141614	1049 (8.00-3.90)
Ramachandran outliers	138981	1016 (8.00-3.86)
RSRZ outliers	127900	1015 (8.20-3.78)

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 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	2486	3% 53% 19% 10% • 14%
1	B	2486	3% 53% 19% 10% • 14%
2	S	219	18% 92% 5% ••
2	T	219	12% 91% 6% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 23302 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 Cytoplasmic dynein heavy chain.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	2136	Total	C	N	O	0	0	0
			10585	6313	2136	2136			
1	B	2136	Total	C	N	O	0	0	0
			10586	6314	2136	2136			

- Molecule 2 is a protein called Glutathione-S-transferase.

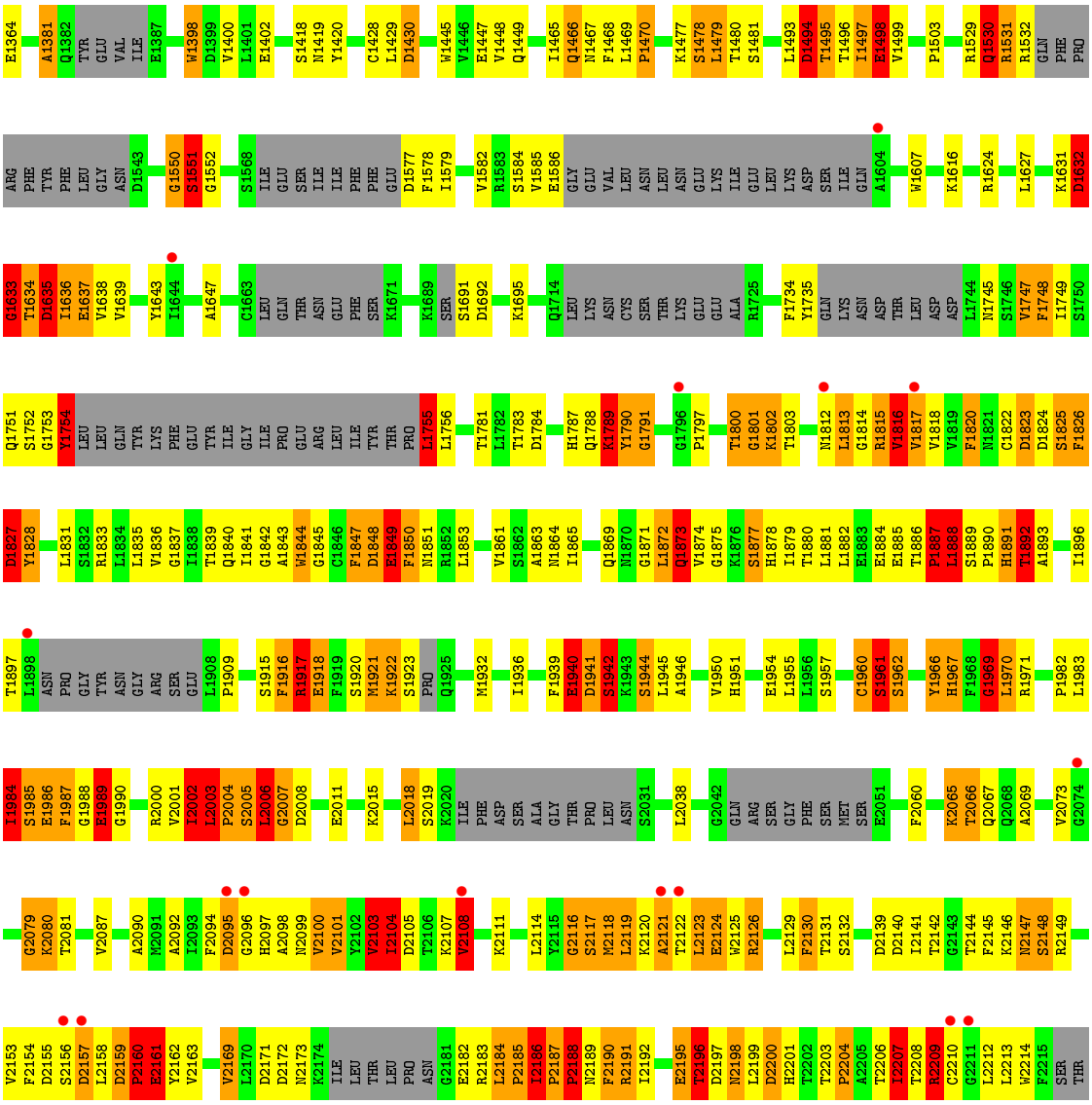
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	T	216	Total	C	N	O	0	0	0
			1066	634	216	216			
2	S	216	Total	C	N	O	0	0	0
			1065	633	216	216			

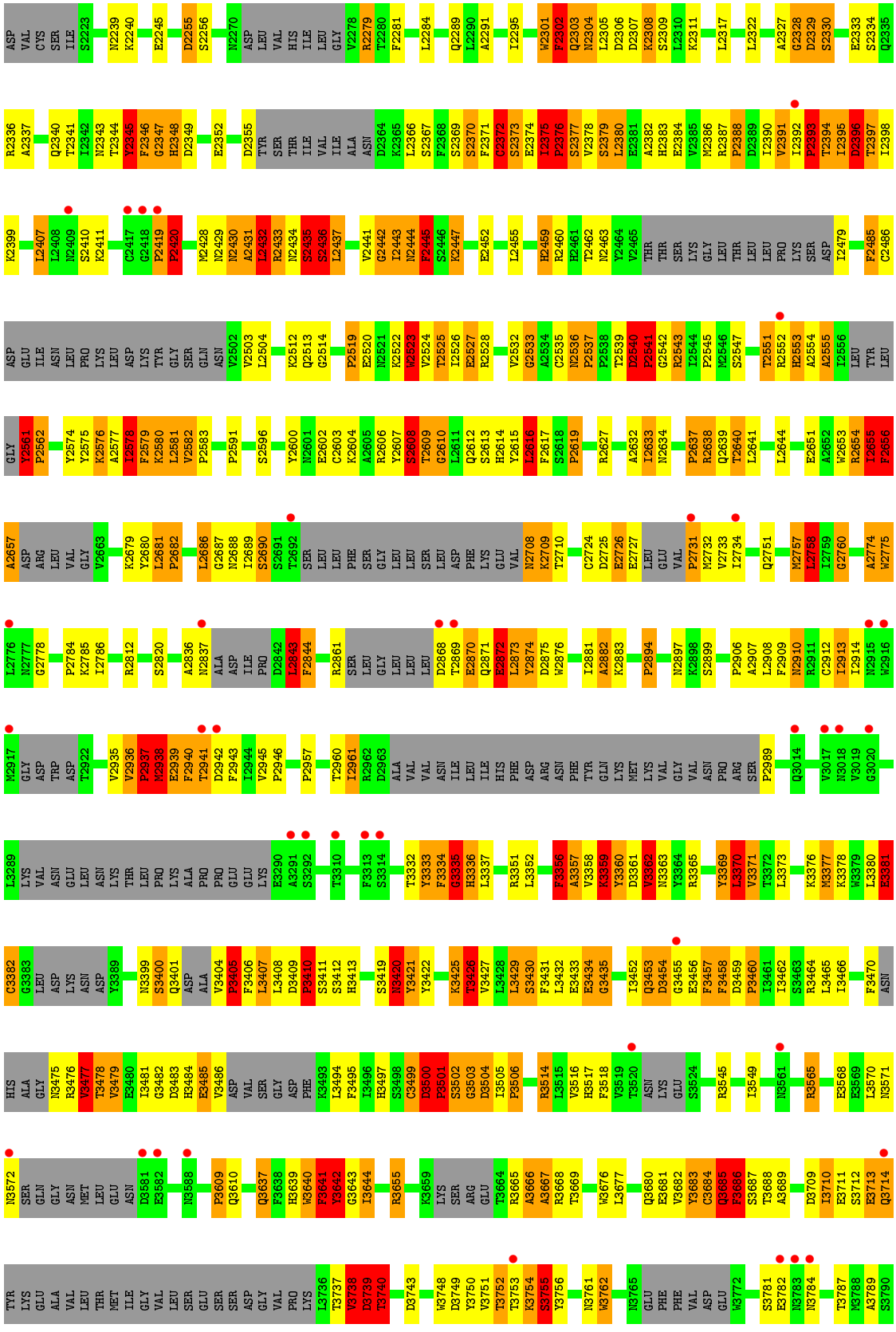
3 Residue-property plots (i)

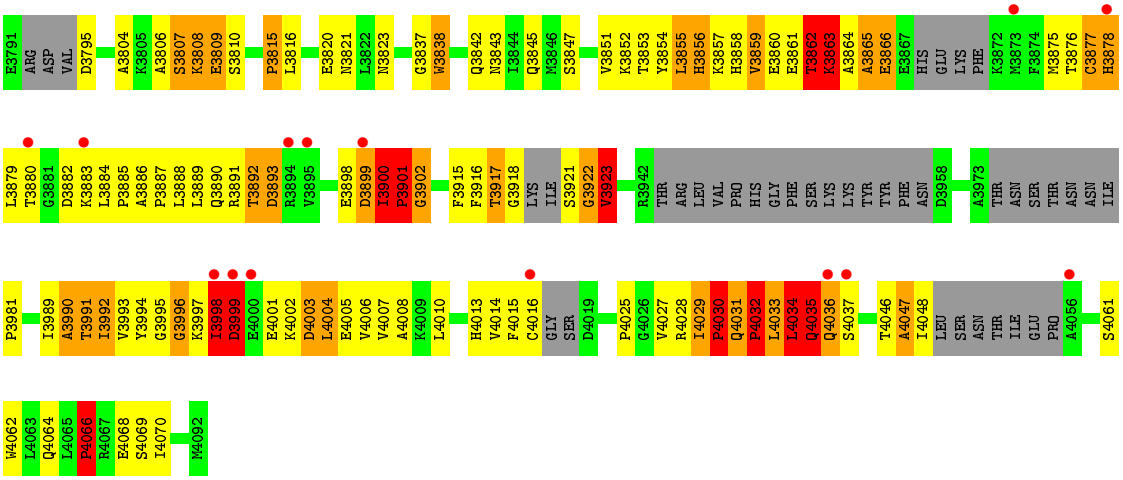
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Cytoplasmic dynein heavy chain

Chain A:

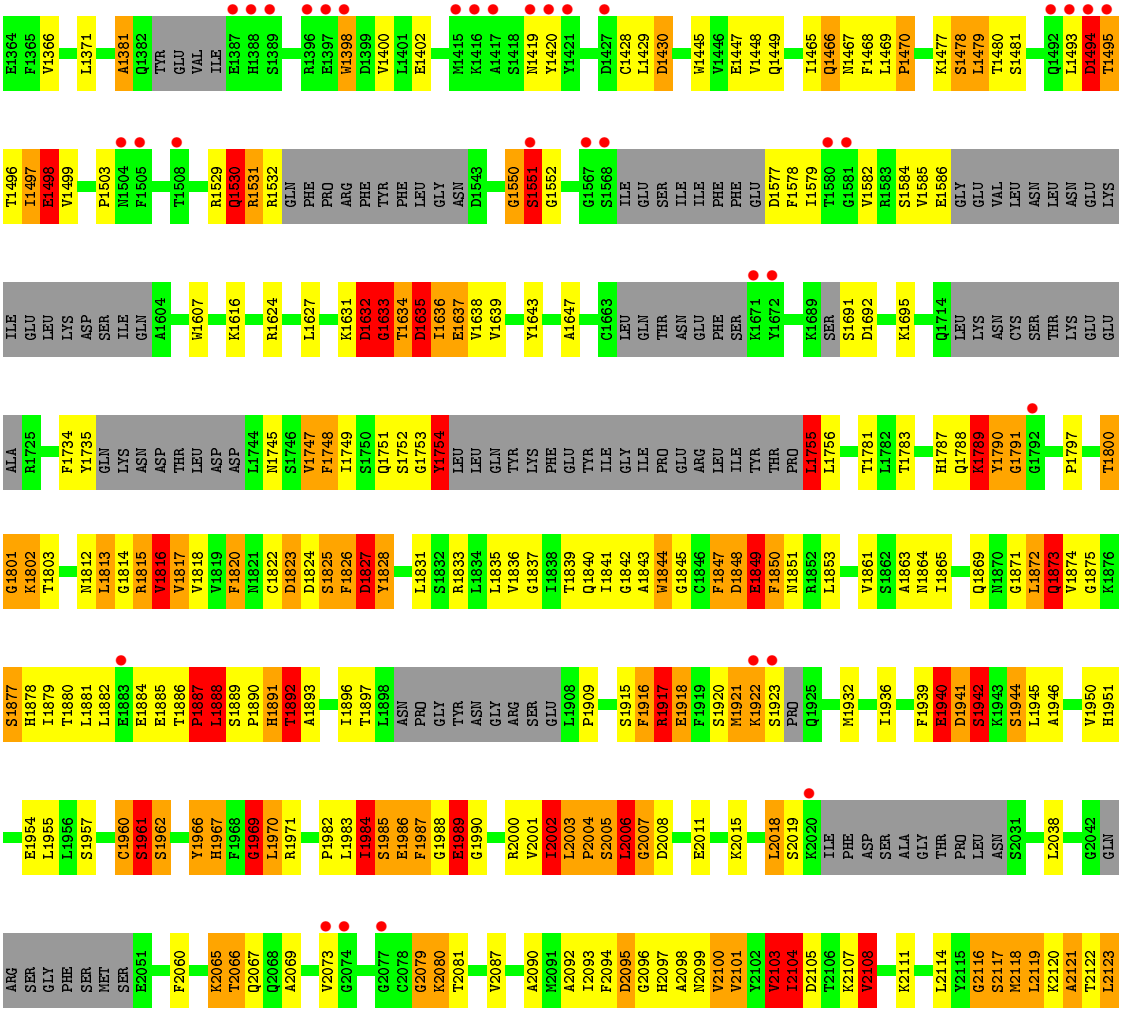


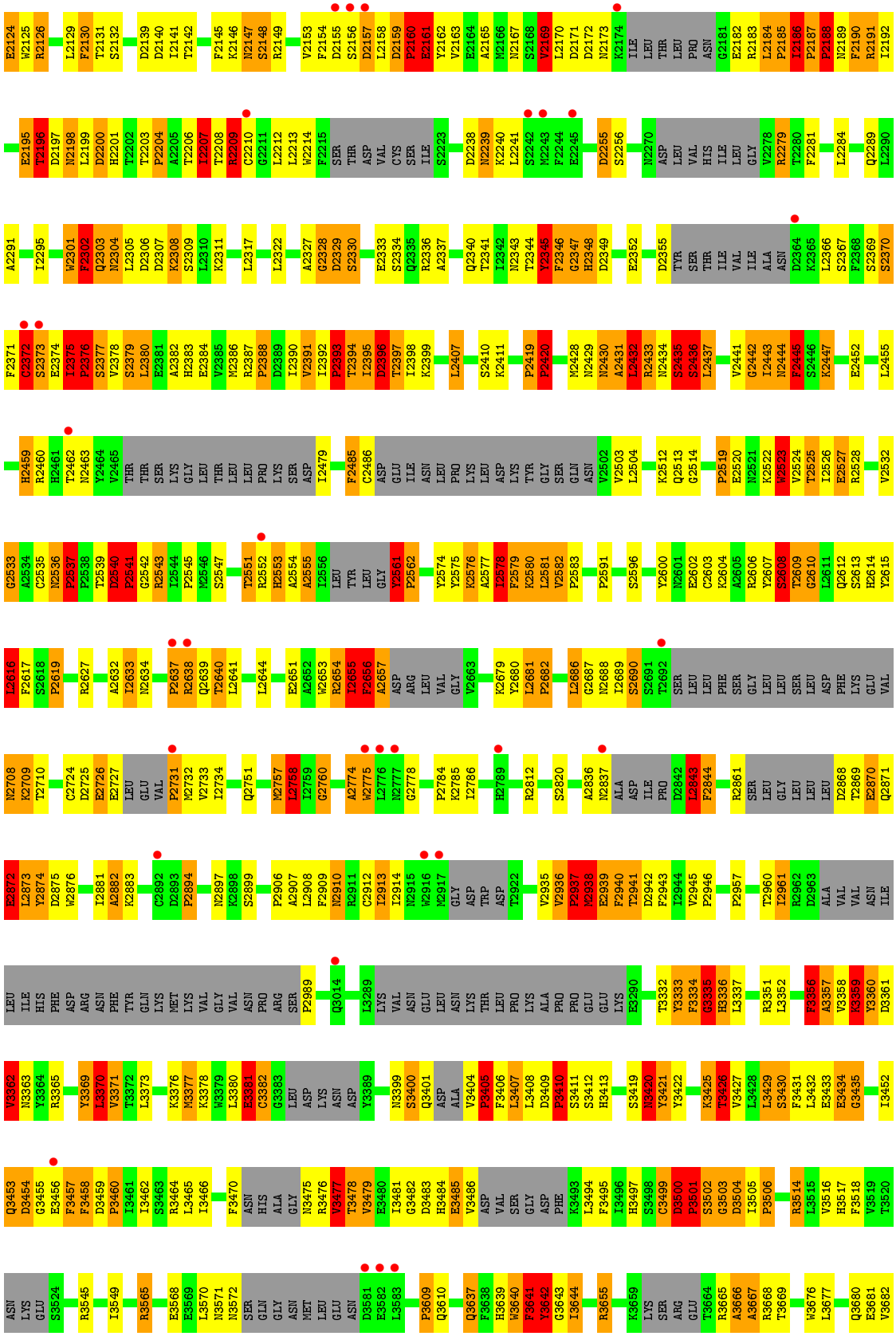


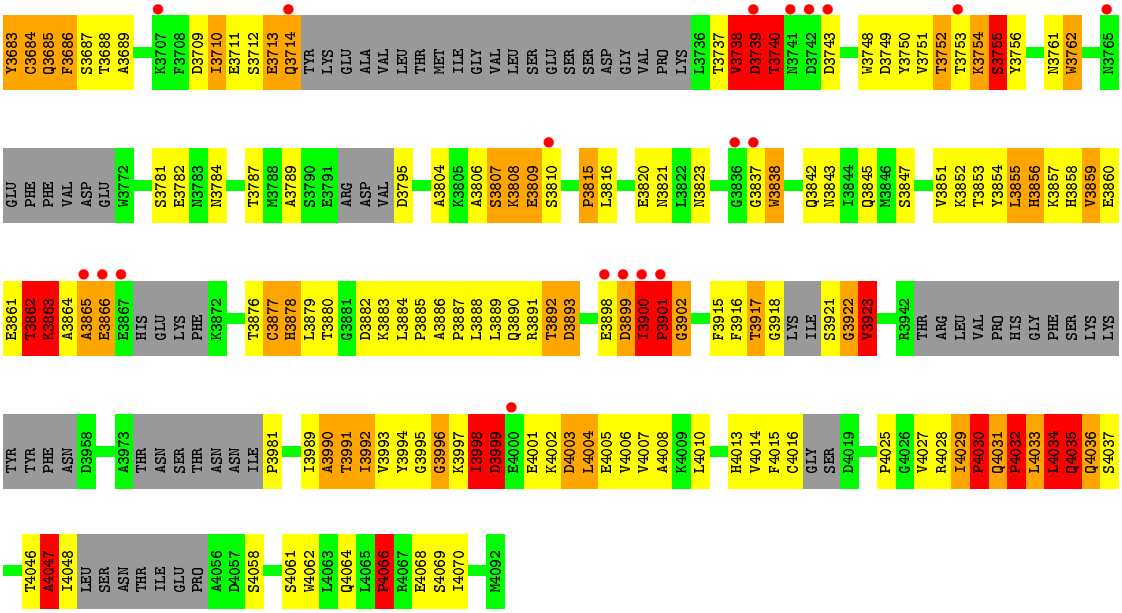


• Molecule 1: Cytoplasmic dynein heavy chain

Chain B:

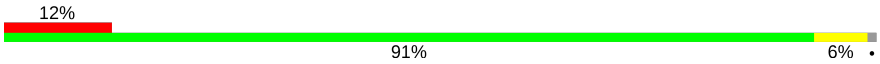


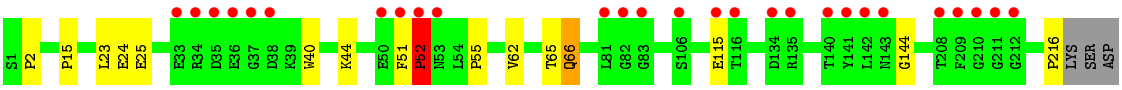




- Molecule 2: Glutathione-S-transferase

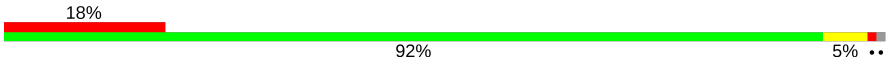
Chain T:

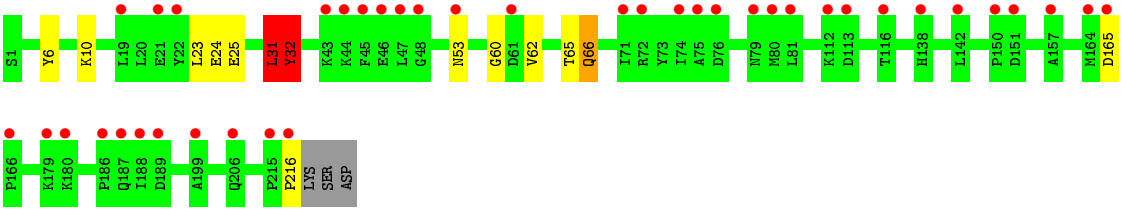




- Molecule 2: Glutathione-S-transferase

Chain S:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	174.08Å 118.92Å 200.51Å 90.00° 90.27° 90.00°	Depositor
Resolution (Å)	50.00 – 6.00 50.13 – 6.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-6.00) 99.9 (50.13-6.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	9.38 (at 6.15Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.430 , 0.430 0.440 , 0.433	Depositor DCC
R_{free} test set	1079 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å^2)	259.8	Xtriage
Anisotropy	0.257	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.24 , 756.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.014 for h,-k,-l	Xtriage
F_o,F_c correlation	0.72	EDS
Total number of atoms	23302	wwPDB-VP
Average B, all atoms (Å^2)	222.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.18% 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	A	1.63	201/10536 (1.9%)	2.24	571/14615 (3.9%)
1	B	1.63	201/10537 (1.9%)	2.24	572/14617 (3.9%)
2	S	1.94	2/1064 (0.2%)	0.86	3/1479 (0.2%)
2	T	0.39	0/1065	0.78	6/1481 (0.4%)
All	All	1.61	404/23202 (1.7%)	2.15	1152/32192 (3.6%)

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
1	A	3	136
1	B	3	136
2	S	0	3
All	All	6	275

The worst 5 of 404 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	S	32	TYR	N-CA	49.28	2.44	1.46
2	S	32	TYR	CA-C	37.84	2.51	1.52
1	B	4047	ALA	C-N	-30.16	0.64	1.34
1	A	4047	ALA	C-N	-30.14	0.64	1.34
1	A	3426	THR	C-N	-25.27	0.76	1.34

The worst 5 of 1152 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1466	GLN	O-C-N	-31.44	72.39	122.70
1	A	1466	GLN	O-C-N	-31.43	72.42	122.70
1	B	2436	SER	O-C-N	-28.51	77.09	122.70
1	A	2436	SER	O-C-N	-28.49	77.12	122.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1466	GLN	CA-C-N	-28.26	55.02	117.20

5 of 6 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	1888	LEU	CA
1	A	2375	ILE	CA
1	A	3641	PHE	CA
1	B	1888	LEU	CA
1	B	2375	ILE	CA

5 of 275 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1381	ALA	Mainchain
1	A	1430	ASP	Mainchain
1	A	1479	LEU	Mainchain
1	A	1494	ASP	Mainchain,Peptide
1	A	1530	GLN	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	10585	0	4454	1006	3
1	B	10586	0	4455	994	3
2	S	1065	0	464	9	0
2	T	1066	0	465	22	0
All	All	23302	0	9838	2013	3

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 61.

The worst 5 of 2013 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)
1:A:1364:GLU:CA	2:T:216:PRO:HA	1.19	1.65
1:A:1932:MET:CB	1:A:1946:ALA:HB1	1.24	1.64
1:B:2060:PHE:CB	1:B:2087:VAL:CB	1.78	1.62
1:A:2060:PHE:CB	1:A:2087:VAL:CB	1.78	1.57
1:A:3666:ALA:HB2	1:A:3668:ARG:CB	1.32	1.57

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:2245:GLU:N	1:B:2241:LEU:CB[1_554]	1.43	0.77
1:A:2245:GLU:CA	1:B:2241:LEU:CB[1_554]	1.81	0.39
1:A:2245:GLU:CB	1:B:2241:LEU:CB[1_554]	2.05	0.15

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	A	2039/2486 (82%)	1605 (79%)	307 (15%)	127 (6%)	1 16
1	B	2039/2486 (82%)	1605 (79%)	308 (15%)	126 (6%)	1 16
2	S	214/219 (98%)	194 (91%)	16 (8%)	4 (2%)	8 38
2	T	214/219 (98%)	190 (89%)	20 (9%)	4 (2%)	8 38
All	All	4506/5410 (83%)	3594 (80%)	651 (14%)	261 (6%)	1 17

5 of 261 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1470	PRO
1	A	1494	ASP
1	A	1498	GLU
1	A	1635	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1637	GLU

5.3.2 Protein sidechains

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

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	109
1	A	109

The worst 5 of 218 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1756:LEU	C	1776:LEU	N	12.66
1	B	1756:LEU	C	1776:LEU	N	12.66
1	A	3028:VAL	C	3289:LEU	N	11.45
1	B	3028:VAL	C	3289:LEU	N	11.45
1	A	3981:PRO	C	3982:TRP	N	5.84

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	A	2136/2486 (85%)	-0.03	69 (3%)	47	41	183, 210, 384, 480	0
1	B	2136/2486 (85%)	0.07	84 (3%)	39	34	183, 210, 384, 480	0
2	S	216/219 (98%)	0.96	40 (18%)	1	3	209, 233, 253, 268	0
2	T	216/219 (98%)	0.66	27 (12%)	3	7	201, 219, 248, 273	0
All	All	4704/5410 (86%)	0.09	220 (4%)	31	30	183, 210, 376, 480	0

The worst 5 of 220 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	S	80	MET	7.1
1	B	2156	SER	6.0
1	B	1397	GLU	5.5
1	B	3899	ASP	5.3
2	T	52	PRO	5.2

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

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.