Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 05:40 pm BST

PDB ID	:	3C2G
Title	:	Crystal complex of SYS-1/POP-1 at 2.5A resolution
Authors	:	Liu, J.; Phillips, B.T.; Amaya, M.F.; Kimble, J.; Xu, W.
Deposited on	:	2008-01-24
Resolution	:	2.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ symbol.

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

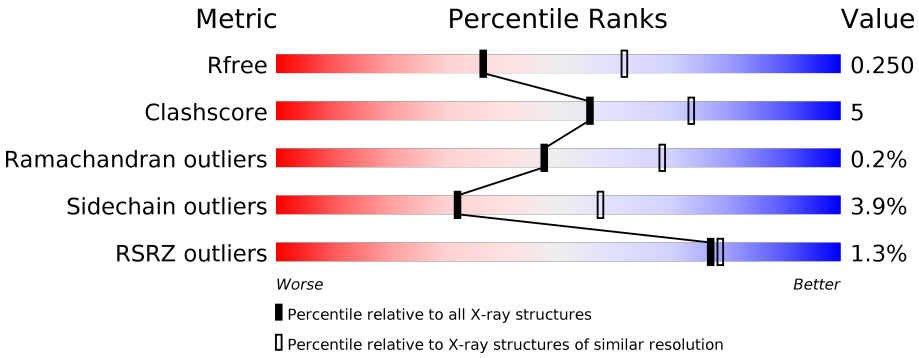
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	619	% 86% 12%
1	B	619	% 84% 13%
2	C	8	88% 13%
2	D	8	88% 13%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9957 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 Sys-1 protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	618	Total	C	N	O	S	Se	0	0	0
			4893	3150	853	858	17	15			
1	B	615	Total	C	N	O	S	Se	0	0	0
			4864	3132	848	852	17	15			

- Molecule 2 is a protein called Pop-1 8-residue peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	8	Total	C	N	O	0	0	0
			62	40	12	10			
2	D	8	Total	C	N	O	0	0	0
			66	42	12	12			

- Molecule 3 is water.

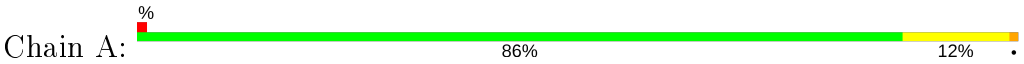
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	37	Total	O	0	0
			37	37		
3	B	35	Total	O	0	0
			35	35		

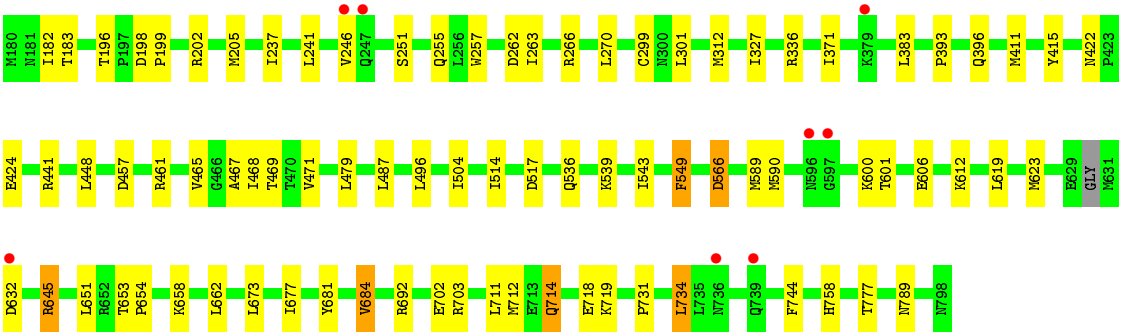
3 Residue-property plots (i)

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: Sys-1 protein

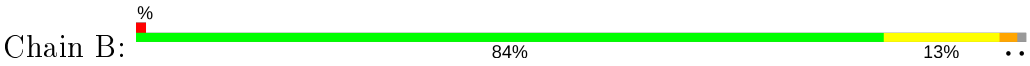
Chain A:

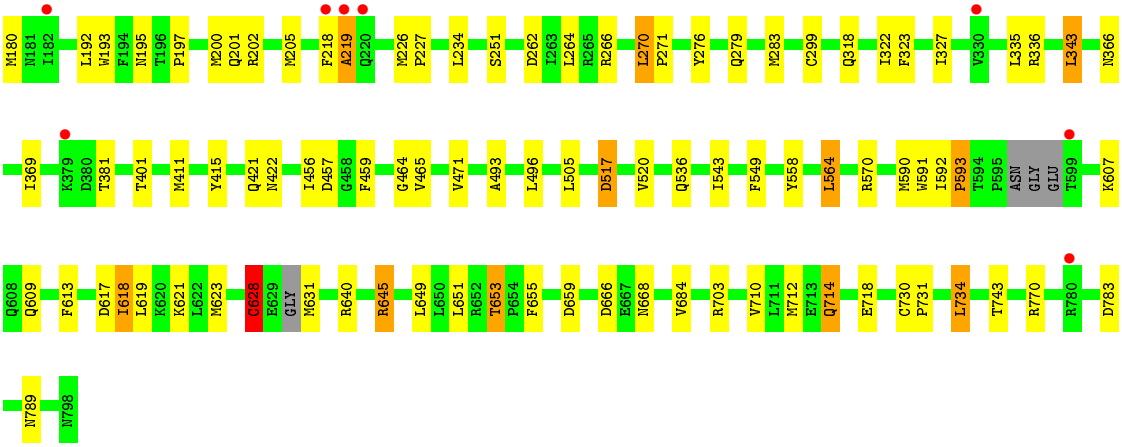




- Molecule 1: Sys-1 protein

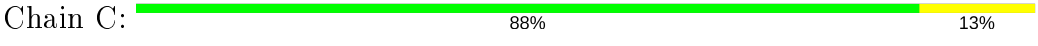
Chain B:





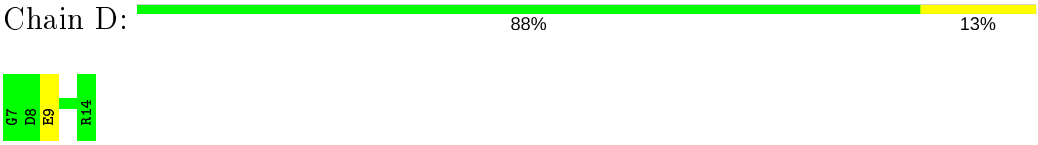
- Molecule 2: Pop-1 8-residue peptide

Chain C:



• Molecule 2: Pop-1 8-residue peptide

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	84.97Å 85.34Å 93.66Å 65.22° 78.08° 83.02°	Depositor
Resolution (Å)	48.00 – 2.50 47.97 – 2.32	Depositor EDS
% Data completeness (in resolution range)	93.7 (48.00-2.50) 81.2 (47.97-2.32)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.95 (at 2.32Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.216 , 0.261 0.209 , 0.250	Depositor DCC
R_{free} test set	4080 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	42.2	Xtriage
Anisotropy	0.356	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 39.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9957	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.34% 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	0.45	0/4992	0.55	0/6775
1	B	0.46	1/4961 (0.0%)	0.56	0/6733
2	C	0.64	0/62	0.50	0/81
2	D	0.60	0/66	0.42	0/86
All	All	0.46	1/10081 (0.0%)	0.56	0/13675

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	B	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	628	CYS	CB-SG	-5.74	1.72	1.81

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	200	MSE	Peptide

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	A	4893	0	4949	48	0
1	B	4864	0	4906	54	0
2	C	62	0	61	1	0
2	D	66	0	65	0	0
3	A	37	0	0	0	0
3	B	35	0	0	1	0
All	All	9957	0	9981	101	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 5.

All (101) 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:777:THR:HG23	1:A:789:ASN:HD22	1.38	0.89
1:A:411:MSE:HE3	1:A:471:VAL:HG21	1.59	0.82
1:A:301:LEU:HD23	1:A:312:MSE:HE1	1.61	0.82
1:B:649:LEU:O	1:B:653:THR:HG23	1.92	0.70
1:A:651:LEU:O	1:A:714:GLN:HG2	1.91	0.69
1:B:623:MSE:CE	1:B:666:ASP:HB3	2.24	0.68
1:A:411:MSE:HE3	1:A:471:VAL:CG2	2.25	0.65
1:B:234:LEU:HD13	1:B:283:MSE:HE3	1.79	0.65
1:A:327:ILE:O	1:A:336:ARG:NH2	2.30	0.64
1:A:712:MSE:HG2	1:A:718:GLU:HA	1.80	0.63
1:A:202:ARG:HA	1:A:205:MSE:HE3	1.81	0.62
1:B:651:LEU:O	1:B:714:GLN:HG2	2.02	0.60
1:B:712:MSE:HG2	1:B:718:GLU:HA	1.85	0.58
1:A:645:ARG:HG3	1:A:703:ARG:CZ	2.35	0.57
1:A:468:ILE:HG21	1:A:487:LEU:HG	1.87	0.56
1:B:543:ILE:HD11	1:B:590:MSE:HE1	1.86	0.56
1:B:564:LEU:HD23	1:B:570:ARG:HA	1.87	0.56
1:A:241:LEU:HD23	1:A:257:TRP:CD1	2.41	0.56
1:B:628:CYS:HB2	1:B:640:ARG:HH21	1.71	0.55
1:B:558:TYR:O	1:B:621:LYS:NZ	2.35	0.55
1:A:469:THR:HG21	1:A:504:ILE:HG23	1.88	0.55
1:B:202:ARG:HA	1:B:205:MSE:HE3	1.88	0.55
1:B:623:MSE:HE3	1:B:666:ASP:HB3	1.89	0.54
1:A:465:VAL:HG21	1:A:496:LEU:HD21	1.90	0.53
1:A:411:MSE:HE1	1:A:467:ALA:C	2.29	0.52
1:A:536:GLN:HE22	1:A:590:MSE:SE	2.42	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:628:CYS:HB2	1:B:640:ARG:NH2	2.24	0.52
1:A:719:LYS:HE2	1:A:758:HIS:HD2	1.74	0.51
1:A:673:LEU:O	1:A:677:ILE:HG12	2.10	0.51
1:A:196:THR:HB	1:A:205:MSE:HE1	1.94	0.50
1:A:651:LEU:O	1:A:658:LYS:HE2	2.11	0.50
1:B:366:ASN:HB3	1:B:369:ILE:HD12	1.93	0.50
1:A:393:PRO:HD2	1:A:396:GLN:HG3	1.94	0.50
1:A:255:GLN:HG3	1:B:422:ASN:ND2	2.27	0.50
1:B:262:ASP:HB2	1:B:318:GLN:HE22	1.77	0.49
1:B:195:ASN:O	1:B:197:PRO:HD3	2.13	0.49
1:A:731:PRO:O	1:A:734:LEU:HB2	2.13	0.49
1:B:465:VAL:HG21	1:B:496:LEU:HD21	1.95	0.49
1:A:262:ASP:OD2	1:A:266:ARG:NH1	2.46	0.48
1:A:301:LEU:CD2	1:A:312:MSE:HE1	2.39	0.48
1:A:422:ASN:HD21	1:A:424:GLU:HB2	1.77	0.48
1:A:415:TYR:HB2	1:A:471:VAL:HG22	1.94	0.48
1:B:456:ILE:HG22	1:B:493:ALA:HB2	1.96	0.47
1:A:457:ASP:O	1:A:461:ARG:HG3	2.14	0.47
1:B:609:GLN:NE2	3:B:68:HOH:O	2.45	0.47
1:A:539:LYS:NZ	2:C:8:ASP:OD1	2.47	0.47
1:A:411:MSE:HE1	1:A:467:ALA:HB1	1.97	0.47
1:B:327:ILE:O	1:B:336:ARG:NH2	2.48	0.46
1:B:543:ILE:CD1	1:B:590:MSE:HE1	2.45	0.46
1:B:415:TYR:HB2	1:B:471:VAL:HG22	1.96	0.46
1:B:649:LEU:O	1:B:653:THR:CG2	2.62	0.46
1:B:234:LEU:HD13	1:B:283:MSE:CE	2.44	0.45
1:A:198:ASP:HA	1:A:199:PRO:HD3	1.83	0.45
1:B:730:CYS:HA	1:B:731:PRO:HD3	1.87	0.45
1:B:323:PHE:HE1	1:B:343:LEU:HD11	1.82	0.45
1:B:645:ARG:HG3	1:B:703:ARG:CZ	2.47	0.45
1:A:182:ILE:HD12	1:A:183:THR:H	1.81	0.45
1:B:549:PHE:H	1:B:609:GLN:HE21	1.64	0.45
1:B:335:LEU:HD12	1:B:335:LEU:HA	1.81	0.45
1:A:600:LYS:HE3	1:A:600:LYS:HB2	1.86	0.44
1:B:234:LEU:HB3	1:B:283:MSE:HE3	1.99	0.44
1:B:592:ILE:HA	1:B:593:PRO:HD2	1.78	0.44
1:A:479:LEU:C	1:A:479:LEU:HD23	2.37	0.44
1:B:590:MSE:HE3	1:B:591:TRP:CZ2	2.53	0.44
1:B:264:LEU:HD23	1:B:322:ILE:HD13	1.99	0.44
1:A:371:ILE:HD11	1:A:441:ARG:HH12	1.83	0.44
1:B:613:PHE:HA	1:B:618:ILE:HG21	2.00	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:662:LEU:HD21	1:A:711:LEU:HD22	1.99	0.44
1:A:383:LEU:HA	1:A:383:LEU:HD12	1.92	0.43
1:A:566:ASP:OD1	1:A:566:ASP:N	2.40	0.43
1:B:226:MSE:HG2	1:B:276:TYR:CZ	2.53	0.43
1:A:543:ILE:HG21	1:A:606:GLU:HG3	1.99	0.43
1:B:226:MSE:HB3	1:B:227:PRO:HD3	2.00	0.43
1:A:702:GLU:HG2	1:A:744:PHE:CD1	2.53	0.43
1:B:270:LEU:HA	1:B:271:PRO:HD2	1.91	0.43
1:A:549:PHE:CZ	1:A:612:LYS:HD2	2.54	0.43
1:B:459:PHE:CZ	1:B:464:GLY:HA3	2.53	0.43
1:A:411:MSE:HB3	1:A:411:MSE:HE2	1.70	0.43
1:B:279:GLN:O	1:B:283:MSE:HE2	2.19	0.43
1:A:198:ASP:O	1:A:202:ARG:HG3	2.19	0.43
1:B:218:PHE:O	1:B:219:ALA:C	2.58	0.42
1:A:182:ILE:H	1:A:182:ILE:HG13	1.58	0.42
1:B:202:ARG:HA	1:B:205:MSE:CE	2.50	0.42
1:B:457:ASP:OD1	1:B:493:ALA:HB1	2.20	0.42
1:A:681:TYR:O	1:A:684:VAL:HG12	2.20	0.42
1:A:237:ILE:HD11	1:A:263:ILE:HG22	2.02	0.42
1:A:619:LEU:O	1:A:623:MSE:HG2	2.20	0.42
1:A:653:THR:HA	1:A:654:PRO:HD3	1.88	0.41
1:B:226:MSE:HG2	1:B:276:TYR:CE2	2.55	0.41
1:B:505:LEU:HD12	1:B:505:LEU:HA	1.87	0.41
1:B:607:LYS:HG2	1:B:655:PHE:CD1	2.56	0.41
1:B:731:PRO:O	1:B:734:LEU:HB2	2.20	0.41
1:B:266:ARG:O	1:B:270:LEU:HB2	2.20	0.41
1:B:192:LEU:O	1:B:193:TRP:C	2.58	0.41
1:A:266:ARG:O	1:A:270:LEU:HB2	2.21	0.41
1:B:270:LEU:HA	1:B:270:LEU:HD12	1.98	0.41
1:B:517:ASP:OD1	1:B:520:VAL:HG13	2.21	0.40
1:B:590:MSE:CE	1:B:591:TRP:CZ2	3.04	0.40
1:B:536:GLN:HE22	1:B:590:MSE:SE	2.55	0.40
1:B:411:MSE:HG2	1:B:471:VAL:HG21	2.04	0.40
1:B:710:VAL:O	1:B:714:GLN:HB2	2.21	0.40

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	A	614/619 (99%)	601 (98%)	12 (2%)	1 (0%)	47 68
1	B	609/619 (98%)	581 (95%)	26 (4%)	2 (0%)	41 61
2	C	6/8 (75%)	5 (83%)	1 (17%)	0	100 100
2	D	6/8 (75%)	5 (83%)	1 (17%)	0	100 100
All	All	1235/1254 (98%)	1192 (96%)	40 (3%)	3 (0%)	47 68

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	219	ALA
1	A	632	ASP
1	B	593	PRO

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
1	A	541/548 (99%)	526 (97%)	15 (3%)	43 70
1	B	536/548 (98%)	509 (95%)	27 (5%)	24 46
2	C	6/7 (86%)	6 (100%)	0	100 100
2	D	7/7 (100%)	6 (86%)	1 (14%)	3 6
All	All	1090/1110 (98%)	1047 (96%)	43 (4%)	32 57

All (43) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	246	VAL
1	A	251	SER
1	A	299	CYS
1	A	448	LEU
1	A	514	ILE
1	A	517	ASP
1	A	549	PHE
1	A	566	ASP
1	A	589	MSE
1	A	601	THR
1	A	645	ARG
1	A	684	VAL
1	A	692	ARG
1	A	714	GLN
1	A	734	LEU
1	B	180	MSE
1	B	201	GLN
1	B	251	SER
1	B	270	LEU
1	B	299	CYS
1	B	343	LEU
1	B	381	THR
1	B	401	THR
1	B	421	GLN
1	B	517	ASP
1	B	564	LEU
1	B	617	ASP
1	B	618	ILE
1	B	619	LEU
1	B	628	CYS
1	B	631	MSE
1	B	645	ARG
1	B	653	THR
1	B	659	ASP
1	B	668	ASN
1	B	684	VAL
1	B	714	GLN
1	B	734	LEU
1	B	743	THR
1	B	770	ARG
1	B	783	ASP
1	B	789	ASN
2	D	9	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	247	GLN
1	A	278	HIS
1	A	515	HIS
1	A	536	GLN
1	A	714	GLN
1	A	758	HIS
1	A	789	ASN
1	B	220	GLN
1	B	259	ASN
1	B	305	ASN
1	B	307	GLN
1	B	318	GLN
1	B	342	HIS
1	B	536	GLN
1	B	546	ASN
1	B	609	GLN
1	B	699	GLN

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

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	A	603/619 (97%)	-0.25	8 (1%) 77 79	23, 42, 70, 86	14 (2%)
1	B	600/619 (96%)	-0.18	8 (1%) 77 79	22, 43, 73, 91	19 (3%)
2	C	8/8 (100%)	-0.42	0 100 100	41, 43, 45, 55	0
2	D	8/8 (100%)	0.20	0 100 100	57, 59, 62, 65	0
All	All	1219/1254 (97%)	-0.22	16 (1%) 77 79	22, 43, 71, 91	33 (2%)

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	182	ILE	4.3
1	B	220	GLN	3.1
1	A	247	GLN	2.9
1	B	219	ALA	2.9
1	B	379	LYS	2.7
1	A	379	LYS	2.7
1	A	596	ASN	2.6
1	A	739	GLN	2.6
1	B	780	ARG	2.5
1	B	218	PHE	2.4
1	A	597	GLY	2.3
1	B	330	VAL	2.3
1	A	736	ASN	2.2
1	B	599	THR	2.1
1	A	632	ASP	2.0
1	A	246	VAL	2.0

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

There are no such residues in this entry.