



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 17, 2021 – 12:09 AM EDT

PDB ID : 1U15
Title : Crystal structure of a duck-delta-crystallin-1 double loop mutant (DLM)
Authors : Tsai, M.; Sampaleanu, L.M.; Greene, C.; Creagh, L.; Haynes, C.; Howell, P.L.
Deposited on : 2004-07-14
Resolution : 2.50 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
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)	:	2.23.2

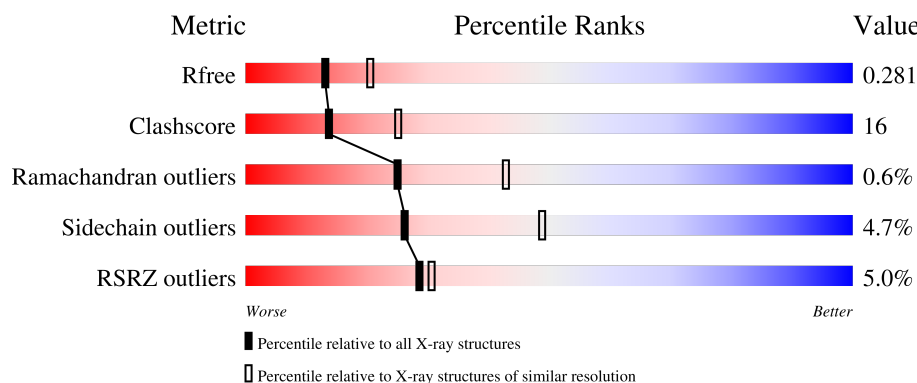
1 Overall quality at a glance

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 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 $\leq 5\%$. 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	472	<div> <div>5%</div> <div> <div></div> <div>66%</div> <div>26%</div> <div>• 5%</div> </div> </div>
1	B	472	<div> <div>6%</div> <div> <div></div> <div>65%</div> <div>27%</div> <div>• 5%</div> </div> </div>
1	C	472	<div> <div>4%</div> <div> <div></div> <div>67%</div> <div>25%</div> <div>• 5%</div> </div> </div>
1	D	472	<div> <div>4%</div> <div> <div></div> <div>66%</div> <div>27%</div> <div>• 5%</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14536 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 Delta crystallin I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	449	Total	C	N	O	S	0	0	0
			3475	2201	586	677	11			
1	B	450	Total	C	N	O	S	0	0	0
			3486	2207	590	678	11			
1	C	448	Total	C	N	O	S	0	0	0
			3472	2199	587	675	11			
1	D	450	Total	C	N	O	S	0	0	0
			3486	2207	590	678	11			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	GLU	GLN	engineered mutation	UNP P24057
A	23	LYS	MET	engineered mutation	UNP P24057
A	25	ASN	SER	engineered mutation	UNP P24057
A	26	SER	THR	engineered mutation	UNP P24057
A	29	ALA	SER	engineered mutation	UNP P24057
A	30	TYR	THR	engineered mutation	UNP P24057
A	31	ASP	GLU	engineered mutation	UNP P24057
A	74	TRP	LEU	engineered mutation	UNP P24057
A	79	PHE	ILE	engineered mutation	UNP P24057
A	82	LYS	THR	engineered mutation	UNP P24057
A	89	HIS	GLN	engineered mutation	UNP P24057
A	467	HIS	-	expression tag	UNP P24057
A	468	HIS	-	expression tag	UNP P24057
A	469	HIS	-	expression tag	UNP P24057
A	470	HIS	-	expression tag	UNP P24057
A	471	HIS	-	expression tag	UNP P24057
A	472	HIS	-	expression tag	UNP P24057
B	22	GLU	GLN	engineered mutation	UNP P24057
B	23	LYS	MET	engineered mutation	UNP P24057
B	25	ASN	SER	engineered mutation	UNP P24057
B	26	SER	THR	engineered mutation	UNP P24057

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	29	ALA	SER	engineered mutation	UNP P24057
B	30	TYR	THR	engineered mutation	UNP P24057
B	31	ASP	GLU	engineered mutation	UNP P24057
B	74	TRP	LEU	engineered mutation	UNP P24057
B	79	PHE	ILE	engineered mutation	UNP P24057
B	82	LYS	THR	engineered mutation	UNP P24057
B	89	HIS	GLN	engineered mutation	UNP P24057
B	467	HIS	-	expression tag	UNP P24057
B	468	HIS	-	expression tag	UNP P24057
B	469	HIS	-	expression tag	UNP P24057
B	470	HIS	-	expression tag	UNP P24057
B	471	HIS	-	expression tag	UNP P24057
B	472	HIS	-	expression tag	UNP P24057
C	22	GLU	GLN	engineered mutation	UNP P24057
C	23	LYS	MET	engineered mutation	UNP P24057
C	25	ASN	SER	engineered mutation	UNP P24057
C	26	SER	THR	engineered mutation	UNP P24057
C	29	ALA	SER	engineered mutation	UNP P24057
C	30	TYR	THR	engineered mutation	UNP P24057
C	31	ASP	GLU	engineered mutation	UNP P24057
C	74	TRP	LEU	engineered mutation	UNP P24057
C	79	PHE	ILE	engineered mutation	UNP P24057
C	82	LYS	THR	engineered mutation	UNP P24057
C	89	HIS	GLN	engineered mutation	UNP P24057
C	467	HIS	-	expression tag	UNP P24057
C	468	HIS	-	expression tag	UNP P24057
C	469	HIS	-	expression tag	UNP P24057
C	470	HIS	-	expression tag	UNP P24057
C	471	HIS	-	expression tag	UNP P24057
C	472	HIS	-	expression tag	UNP P24057
D	22	GLU	GLN	engineered mutation	UNP P24057
D	23	LYS	MET	engineered mutation	UNP P24057
D	25	ASN	SER	engineered mutation	UNP P24057
D	26	SER	THR	engineered mutation	UNP P24057
D	29	ALA	SER	engineered mutation	UNP P24057
D	30	TYR	THR	engineered mutation	UNP P24057
D	31	ASP	GLU	engineered mutation	UNP P24057
D	74	TRP	LEU	engineered mutation	UNP P24057
D	79	PHE	ILE	engineered mutation	UNP P24057
D	82	LYS	THR	engineered mutation	UNP P24057
D	89	HIS	GLN	engineered mutation	UNP P24057
D	467	HIS	-	expression tag	UNP P24057

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	468	HIS	-	expression tag	UNP P24057
D	469	HIS	-	expression tag	UNP P24057
D	470	HIS	-	expression tag	UNP P24057
D	471	HIS	-	expression tag	UNP P24057
D	472	HIS	-	expression tag	UNP P24057

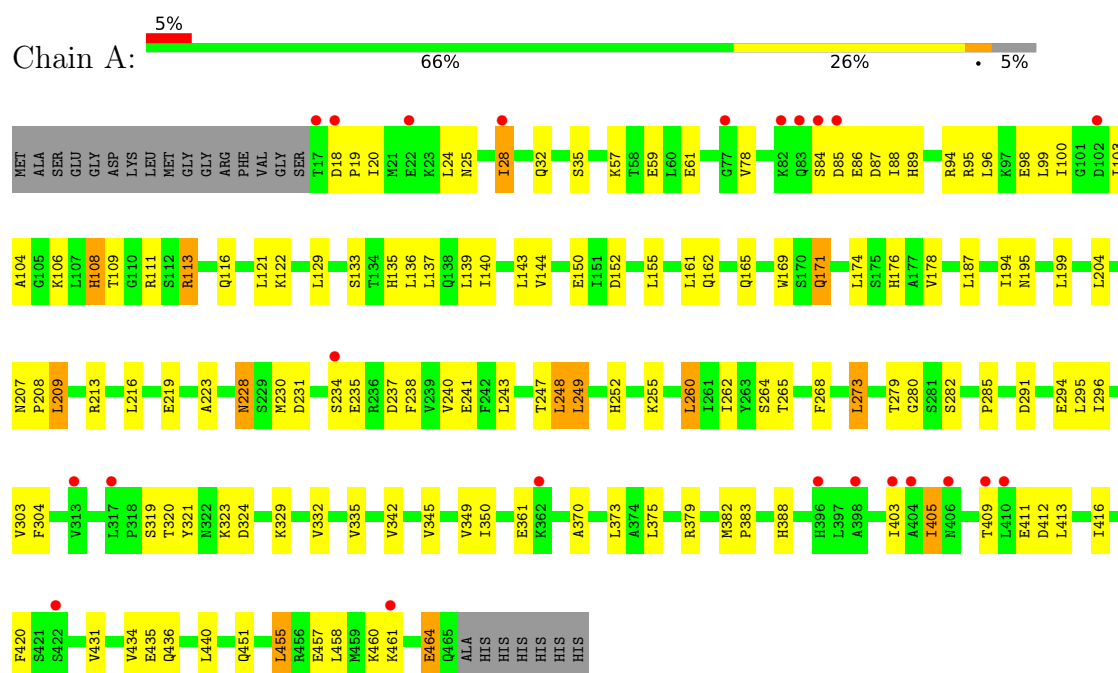
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	169	Total	O	0	0
			169	169		
2	B	128	Total	O	0	0
			128	128		
2	C	173	Total	O	0	0
			173	173		
2	D	147	Total	O	0	0
			147	147		

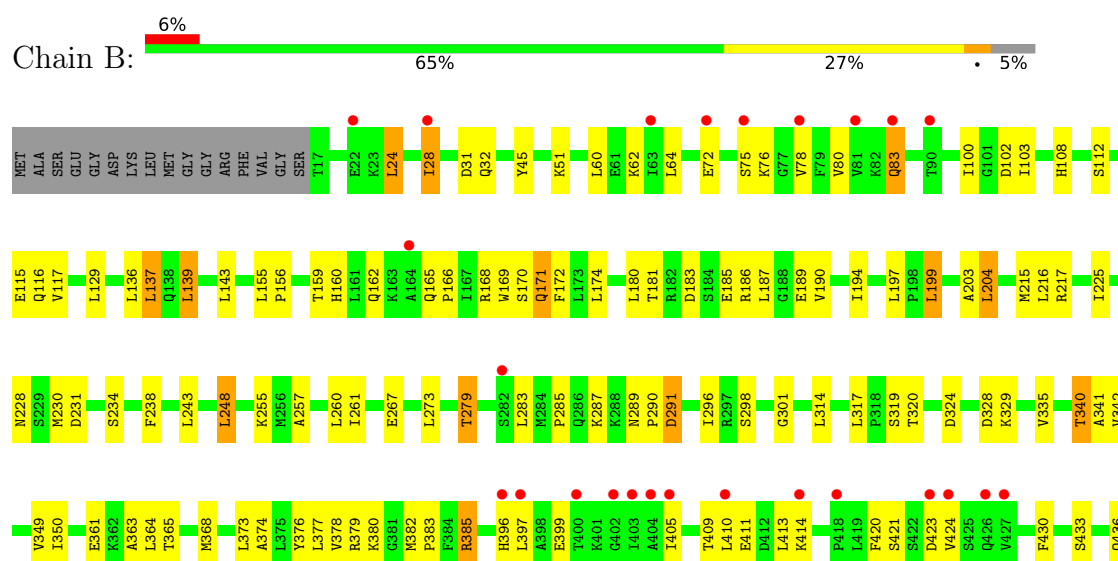
3 Residue-property plots

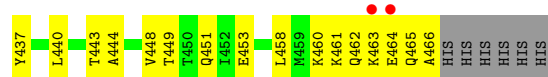
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

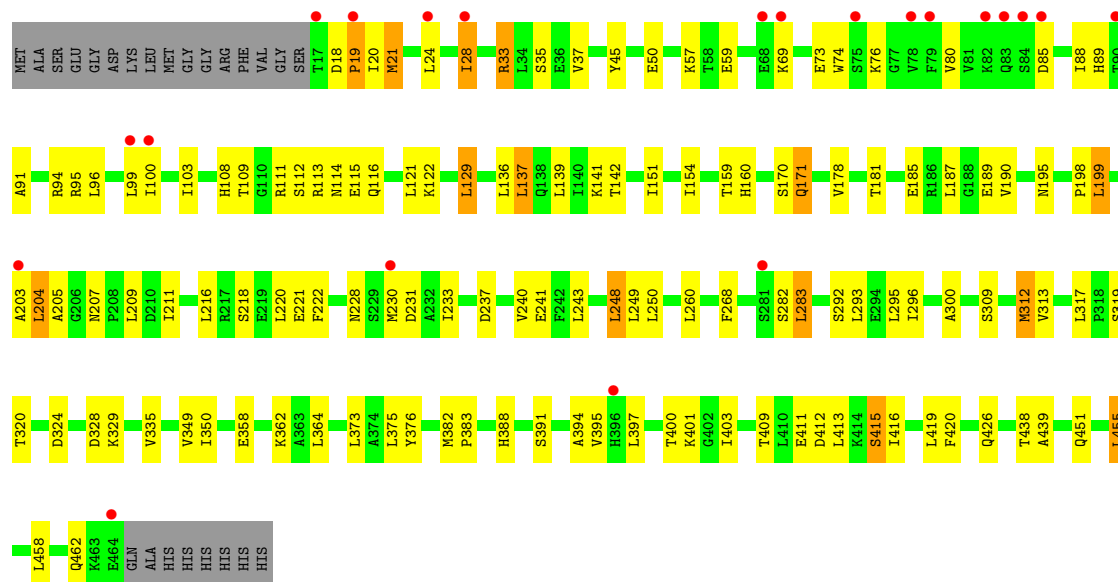


• Molecule 1: Delta crystallin I

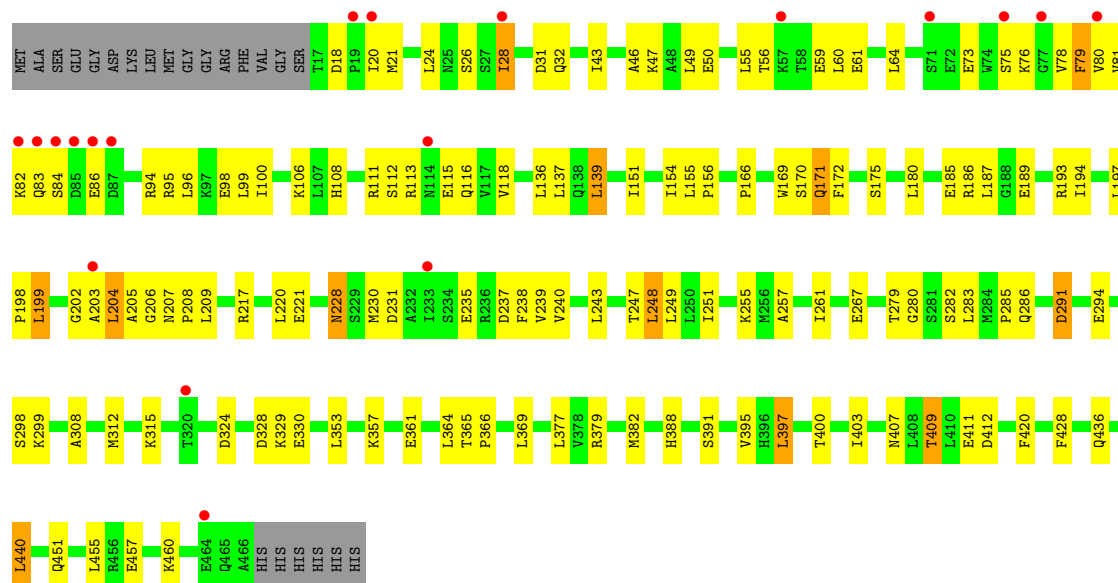




• Molecule 1: Delta crystallin I



• Molecule 1: Delta crystallin I



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	93.70Å 98.80Å 107.40Å 90.00° 101.50° 90.00°	Depositor
Resolution (Å)	35.19 – 2.50 35.19 – 2.50	Depositor EDS
% Data completeness (in resolution range)	98.3 (35.19-2.50) 98.2 (35.19-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.73 (at 2.51Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.221 , 0.281 0.220 , 0.281	Depositor DCC
R_{free} test set	6627 reflections (10.13%)	wwPDB-VP
Wilson B-factor (Å ²)	40.9	Xtriage
Anisotropy	0.211	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 49.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	14536	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.99% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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](#)

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	0.41	0/3518	0.64	0/4750
1	B	0.42	0/3529	0.63	0/4764
1	C	0.42	0/3515	0.63	0/4745
1	D	0.43	0/3529	0.62	0/4764
All	All	0.42	0/14091	0.63	0/19023

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	3475	0	3595	119	0
1	B	3486	0	3611	132	0
1	C	3472	0	3598	122	0
1	D	3486	0	3611	119	0
2	A	169	0	0	12	0
2	B	128	0	0	9	0
2	C	173	0	0	13	0
2	D	147	0	0	10	0
All	All	14536	0	14415	451	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 451 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:C:295:LEU:HB3	2:C:636:HOH:O	1.25	1.24
1:B:24:LEU:HD13	2:C:636:HOH:O	1.32	1.22
1:B:24:LEU:CD1	2:C:636:HOH:O	1.86	1.14
1:D:409:THR:HG22	1:D:412:ASP:H	0.91	1.05
1:D:409:THR:HG22	1:D:412:ASP:N	1.70	1.05

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	447/472 (95%)	426 (95%)	18 (4%)	3 (1%)	22	39
1	B	448/472 (95%)	432 (96%)	14 (3%)	2 (0%)	34	54
1	C	446/472 (94%)	429 (96%)	15 (3%)	2 (0%)	34	54
1	D	448/472 (95%)	425 (95%)	19 (4%)	4 (1%)	17	31
All	All	1789/1888 (95%)	1712 (96%)	66 (4%)	11 (1%)	25	43

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	204	LEU
1	D	204	LEU
1	A	85	ASP
1	A	204	LEU
1	C	204	LEU

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	390/409 (95%)	370 (95%)	20 (5%)	24	45
1	B	391/409 (96%)	374 (96%)	17 (4%)	29	53
1	C	390/409 (95%)	371 (95%)	19 (5%)	25	47
1	D	391/409 (96%)	373 (95%)	18 (5%)	27	50
All	All	1562/1636 (96%)	1488 (95%)	74 (5%)	26	49

5 of 74 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	118	VAL
1	D	409	THR
1	D	137	LEU
1	D	249	LEU
1	B	137	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 51 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	114	ASN
1	C	228	ASN
1	D	407	ASN
1	C	116	GLN
1	C	165	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 monosaccharides 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 [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	449/472 (95%)	0.34	23 (5%) 28 29	20, 37, 60, 82	0
1	B	450/472 (95%)	0.48	27 (6%) 21 22	24, 42, 67, 82	0
1	C	448/472 (94%)	0.33	21 (4%) 31 33	19, 35, 59, 79	0
1	D	450/472 (95%)	0.33	19 (4%) 36 39	20, 40, 68, 79	0
All	All	1797/1888 (95%)	0.37	90 (5%) 28 30	19, 39, 65, 82	0

The worst 5 of 90 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	84	SER	5.1
1	C	100	ILE	4.6
1	A	77	GLY	4.5
1	D	84	SER	4.3
1	B	404	ALA	4.3

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 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.