wwPDB X-ray Structure Validation Summary Report (i)

Jun 22, 2024 – 09:12 PM EDT

PDB ID	:	6D1S
Title	:	Crystal structure of an apo chimeric human alpha1GABAA receptor
Authors	:	Chen, Q.; Arjunan, P.; Cohen, A.E.; Xu, Y.; Tang, P.
Deposited on	:	2018-04-12
Resolution	:	3.20 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:

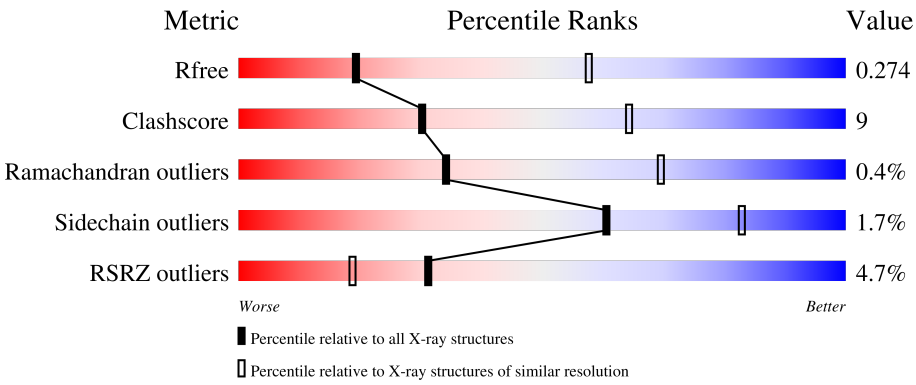
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	324	7% 75% 22% • •
1	B	324	4% 79% 17% • •
1	C	324	2% 74% 21% • •
1	D	324	4% 77% 19% • •
1	E	324	2% 75% 21% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	324	4% 75% 22%
1	G	324	4% 75% 21%
1	H	324	5% 77% 19%
1	I	324	6% 74% 22%
1	J	324	8% 73% 22%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 25690 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 chimeric alpha1GABAA receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	314	2569	1672	428	462	7	0	0	0
1	B	314	2569	1672	428	462	7	0	0	0
1	C	314	2569	1672	428	462	7	0	0	0
1	D	314	2569	1672	428	462	7	0	0	0
1	E	314	2569	1672	428	462	7	0	0	0
1	F	314	2569	1672	428	462	7	0	0	0
1	G	314	2569	1672	428	462	7	0	0	0
1	H	314	2569	1672	428	462	7	0	0	0
1	I	314	2569	1672	428	462	7	0	0	0
1	J	314	2569	1672	428	462	7	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	314	GLY	-	linker	UNP P14867
A	315	VAL	-	linker	UNP P14867
A	316	GLU	-	linker	UNP P14867
B	314	GLY	-	linker	UNP P14867
B	315	VAL	-	linker	UNP P14867
B	316	GLU	-	linker	UNP P14867
C	314	GLY	-	linker	UNP P14867
C	315	VAL	-	linker	UNP P14867
C	316	GLU	-	linker	UNP P14867

Continued on next page...

Continued from previous page...

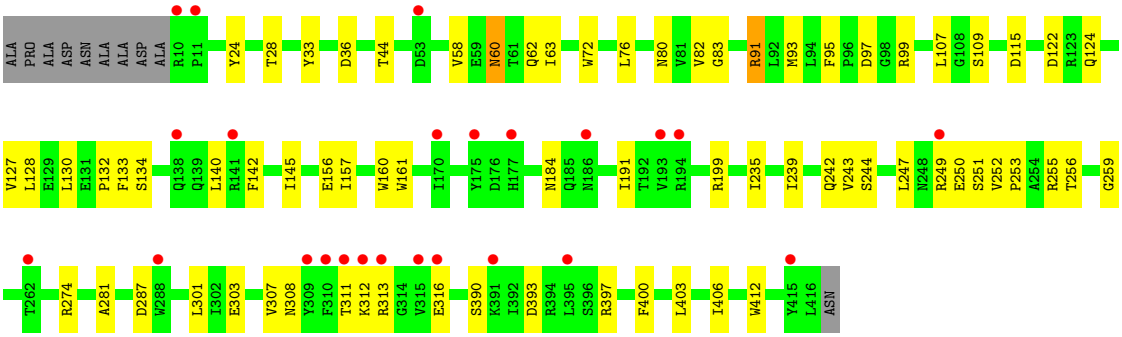
Chain	Residue	Modelled	Actual	Comment	Reference
D	314	GLY	-	linker	UNP P14867
D	315	VAL	-	linker	UNP P14867
D	316	GLU	-	linker	UNP P14867
E	314	GLY	-	linker	UNP P14867
E	315	VAL	-	linker	UNP P14867
E	316	GLU	-	linker	UNP P14867
F	314	GLY	-	linker	UNP P14867
F	315	VAL	-	linker	UNP P14867
F	316	GLU	-	linker	UNP P14867
G	314	GLY	-	linker	UNP P14867
G	315	VAL	-	linker	UNP P14867
G	316	GLU	-	linker	UNP P14867
H	314	GLY	-	linker	UNP P14867
H	315	VAL	-	linker	UNP P14867
H	316	GLU	-	linker	UNP P14867
I	314	GLY	-	linker	UNP P14867
I	315	VAL	-	linker	UNP P14867
I	316	GLU	-	linker	UNP P14867
J	314	GLY	-	linker	UNP P14867
J	315	VAL	-	linker	UNP P14867
J	316	GLU	-	linker	UNP P14867

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.

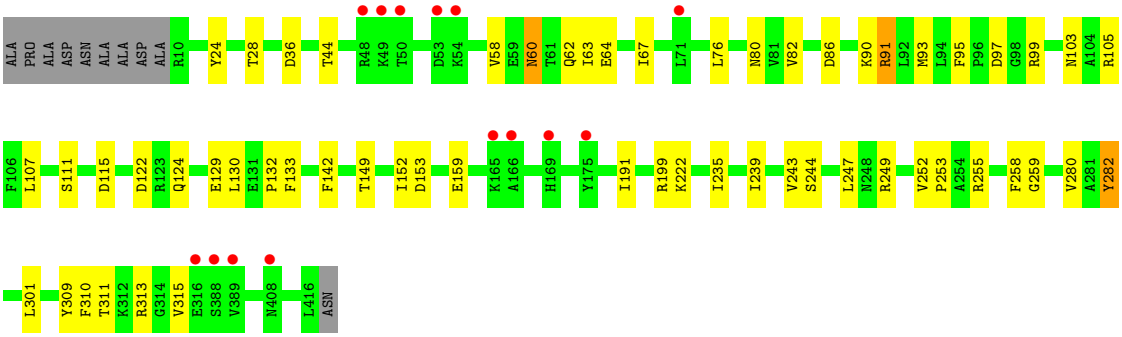
• Molecule 1: chimeric alpha1GABAA receptor

Chain A:



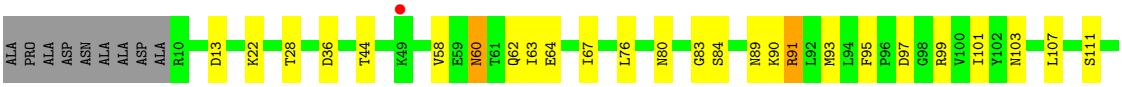
• Molecule 1: chimeric alpha1GABAA receptor

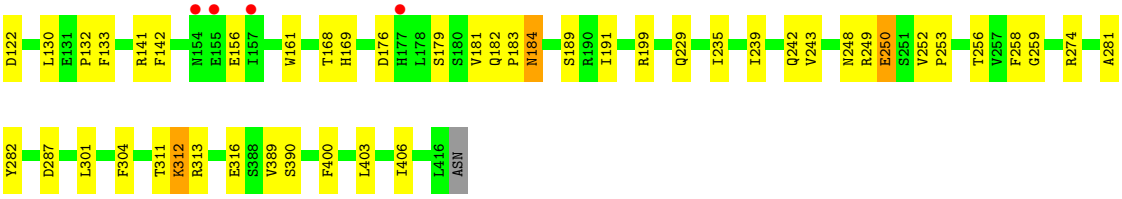
Chain B:



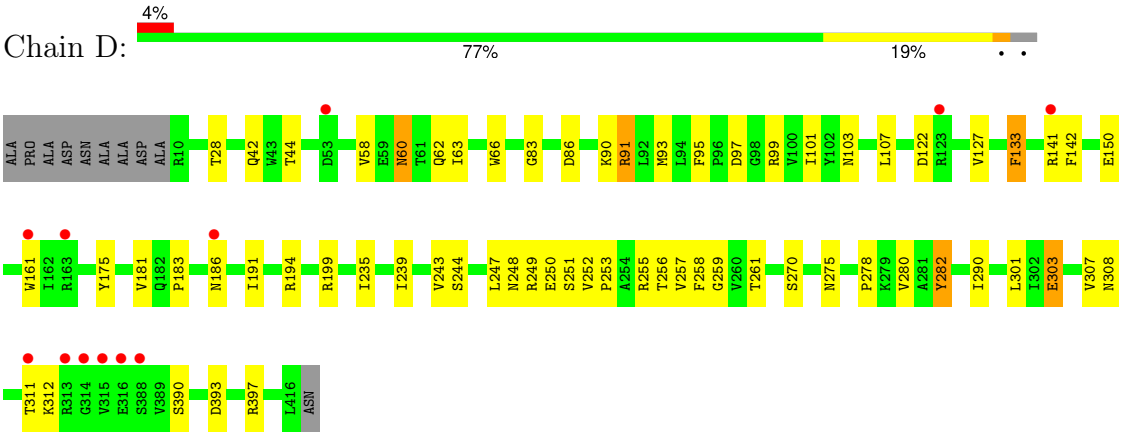
• Molecule 1: chimeric alpha1GABAA receptor

Chain C:

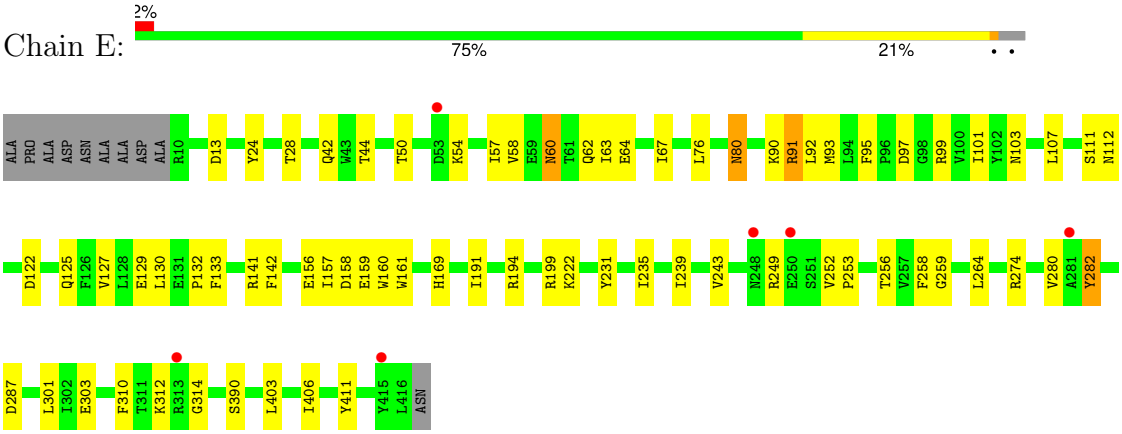




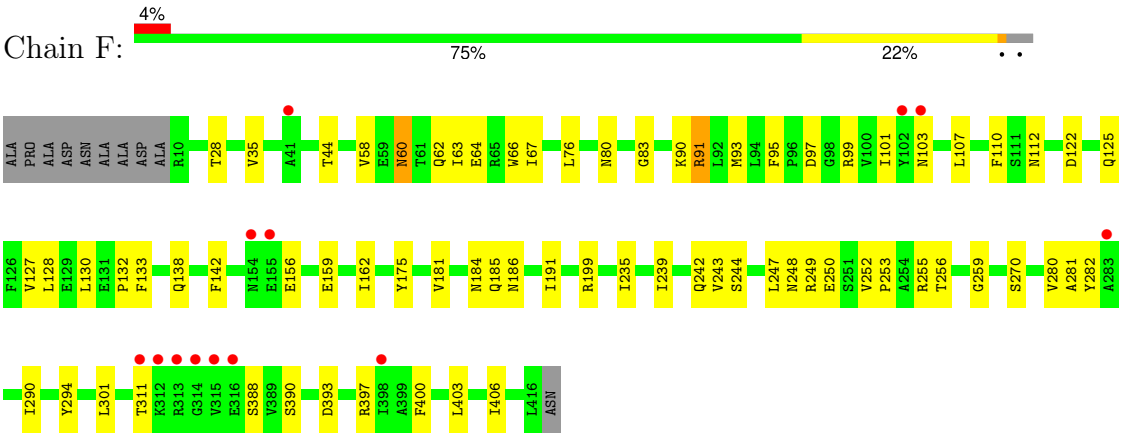
• Molecule 1: chimeric alpha1GABAA receptor



• Molecule 1: chimeric alpha1GABAA receptor

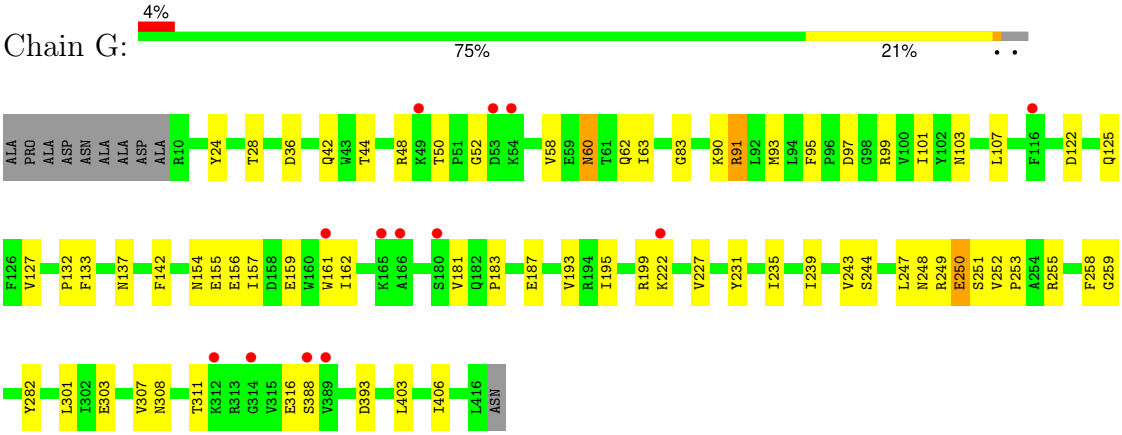


• Molecule 1: chimeric alpha1GABAA receptor



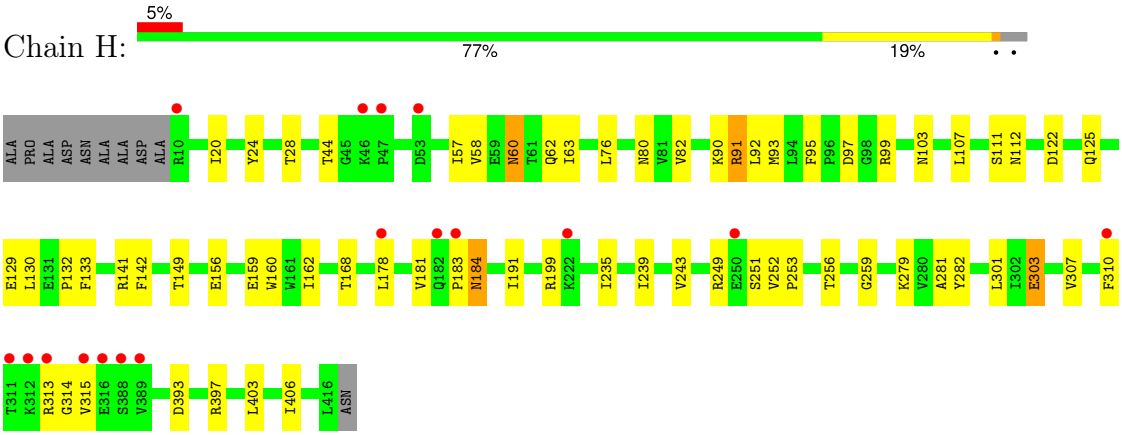
• Molecule 1: chimeric alpha1GABAA receptor

Chain G:



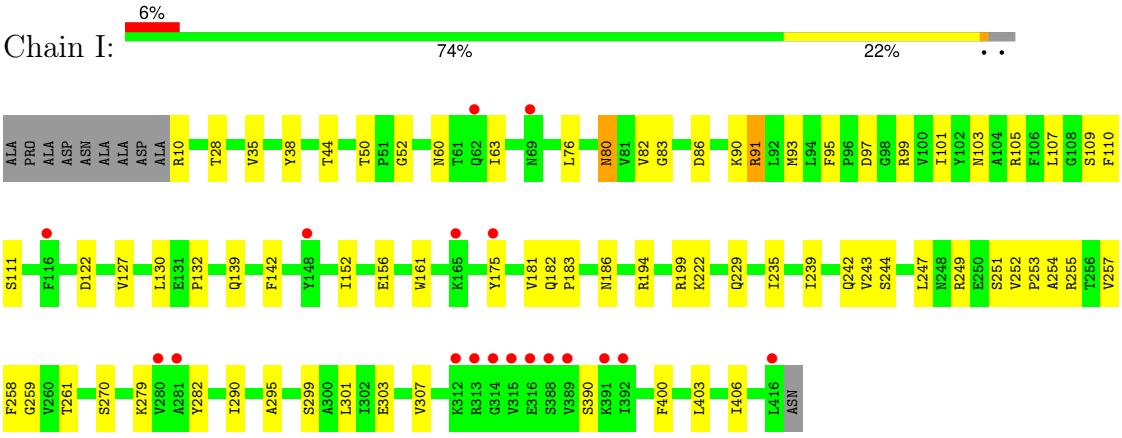
• Molecule 1: chimeric alpha1GABAA receptor

Chain H:



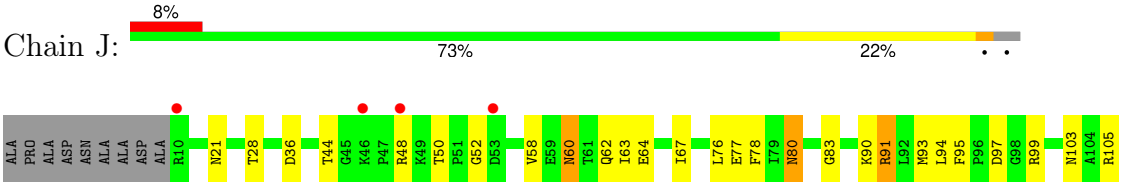
• Molecule 1: chimeric alpha1GABAA receptor

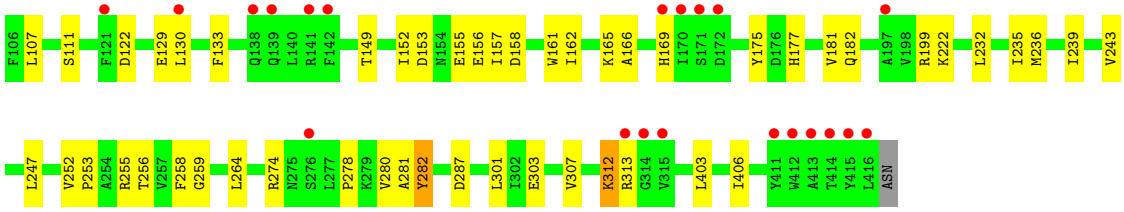
Chain I:



• Molecule 1: chimeric alpha1GABAA receptor

Chain J:




F106
L107
S111
F121
D122
E129
L130
F133
Q138
Q139
L140
R141
F142
T149
I152
D153
N154
E155
E156
I157
D158
W161
I162
K165
A166
H169
I170
S171
D172
Y175
D176
H177
V181
Q182
A197
V198
R199
K222
L232
I235
M236
I239
V243
L247
V252
P253
A254
R255
T256
V257
F258
G259
L264
R274
M275
S276
L277
P278
K279
V280
A281
Y282
D287
L301
I302
E303
V307
K312
R313
G314
V315
L403
I406
Y411
W412
A413
T414
Y415
L416
ASN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	108.52Å 264.85Å 109.15Å 90.00° 110.46° 90.00°	Depositor
Resolution (Å)	37.77 – 3.20 37.77 – 3.20	Depositor EDS
% Data completeness (in resolution range)	98.8 (37.77-3.20) 98.8 (37.77-3.20)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.35 (at 3.18Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.220 , 0.269 0.223 , 0.274	Depositor DCC
R_{free} test set	4686 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	126.9	Xtriage
Anisotropy	0.110	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 88.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.015 for l,-k,h	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	25690	wwPDB-VP
Average B, all atoms (Å^2)	142.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.98% 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)

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	0.29	0/2637	0.47	0/3592
1	B	0.28	0/2637	0.48	0/3592
1	C	0.29	0/2637	0.48	0/3592
1	D	0.30	0/2637	0.49	0/3592
1	E	0.29	0/2637	0.48	0/3592
1	F	0.29	0/2637	0.48	0/3592
1	G	0.29	0/2637	0.48	0/3592
1	H	0.29	0/2637	0.48	0/3592
1	I	0.29	0/2637	0.47	0/3592
1	J	0.29	0/2637	0.47	0/3592
All	All	0.29	0/26370	0.48	0/35920

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	2569	0	2550	48	0
1	B	2569	0	2550	42	0
1	C	2569	0	2550	55	0
1	D	2569	0	2550	50	0
1	E	2569	0	2550	53	0
1	F	2569	0	2550	54	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	2569	0	2550	57	0
1	H	2569	0	2550	54	0
1	I	2569	0	2550	55	0
1	J	2569	0	2550	59	0
All	All	25690	0	25500	476	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 9.

The worst 5 of 476 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:H:156:GLU:HG3	1:H:162:ILE:HD13	1.55	0.88
1:F:107:LEU:HD22	1:G:83:GLY:H	1.40	0.86
1:F:156:GLU:HG3	1:F:162:ILE:H	1.42	0.84
1:H:28:THR:HG21	1:H:199:ARG:HH12	1.40	0.83
1:C:58:VAL:HG13	1:C:62:GLN:HB2	1.61	0.83

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	312/324 (96%)	290 (93%)	21 (7%)	1 (0%)	41	74
1	B	312/324 (96%)	291 (93%)	20 (6%)	1 (0%)	41	74
1	C	312/324 (96%)	292 (94%)	18 (6%)	2 (1%)	25	64
1	D	312/324 (96%)	285 (91%)	26 (8%)	1 (0%)	41	74
1	E	312/324 (96%)	291 (93%)	19 (6%)	2 (1%)	25	64
1	F	312/324 (96%)	286 (92%)	25 (8%)	1 (0%)	41	74

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	G	312/324 (96%)	292 (94%)	19 (6%)	1 (0%)	41	74
1	H	312/324 (96%)	286 (92%)	24 (8%)	2 (1%)	25	64
1	I	312/324 (96%)	288 (92%)	24 (8%)	0	100	100
1	J	312/324 (96%)	289 (93%)	21 (7%)	2 (1%)	25	64
All	All	3120/3240 (96%)	2890 (93%)	217 (7%)	13 (0%)	34	69

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	312	LYS
1	E	314	GLY
1	J	312	LYS
1	D	60	ASN
1	J	60	ASN

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	284/289 (98%)	281 (99%)	3 (1%)	73	88
1	B	284/289 (98%)	281 (99%)	3 (1%)	73	88
1	C	284/289 (98%)	280 (99%)	4 (1%)	67	86
1	D	284/289 (98%)	277 (98%)	7 (2%)	47	77
1	E	284/289 (98%)	278 (98%)	6 (2%)	53	79
1	F	284/289 (98%)	279 (98%)	5 (2%)	59	82
1	G	284/289 (98%)	279 (98%)	5 (2%)	59	82
1	H	284/289 (98%)	278 (98%)	6 (2%)	53	79
1	I	284/289 (98%)	280 (99%)	4 (1%)	67	86
1	J	284/289 (98%)	278 (98%)	6 (2%)	53	79
All	All	2840/2890 (98%)	2791 (98%)	49 (2%)	60	83

5 of 49 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	91	ARG
1	H	141	ARG
1	G	133	PHE
1	G	303	GLU
1	H	282	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	31	GLN
1	G	177	HIS

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	314/324 (96%)	0.18	24 (7%)	13	7	102, 143, 224, 264	0
1	B	314/324 (96%)	0.18	14 (4%)	33	21	100, 144, 233, 284	0
1	C	314/324 (96%)	-0.13	5 (1%)	72	59	82, 118, 206, 313	0
1	D	314/324 (96%)	-0.03	12 (3%)	40	26	89, 120, 200, 338	0
1	E	314/324 (96%)	-0.09	6 (1%)	66	53	88, 124, 202, 273	0
1	F	314/324 (96%)	0.02	13 (4%)	37	24	95, 134, 229, 313	0
1	G	314/324 (96%)	-0.07	13 (4%)	37	24	97, 129, 206, 281	0
1	H	314/324 (96%)	0.03	17 (5%)	25	14	82, 124, 211, 287	0
1	I	314/324 (96%)	-0.03	18 (5%)	23	13	107, 142, 221, 272	0
1	J	314/324 (96%)	0.16	25 (7%)	12	6	99, 147, 210, 291	0
All	All	3140/3240 (96%)	0.02	147 (4%)	31	19	82, 133, 219, 338	0

The worst 5 of 147 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	155	GLU	10.4
1	J	415	TYR	7.7
1	D	315	VAL	7.6
1	G	388	SER	7.1
1	H	311	THR	6.7

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.