



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 05:14 am GMT

PDB ID : 4A2M
Title : Structure of the periplasmic domain of the heparin and heparan sulphate sensing hybrid two component system BT4663 in apo and ligand bound forms
Authors : Lowe, E.C.; Basle, A.; Czjzek, M.; Firbank, S.J.; Bolam, D.N.
Deposited on : 2011-09-27
Resolution : 3.40 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
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

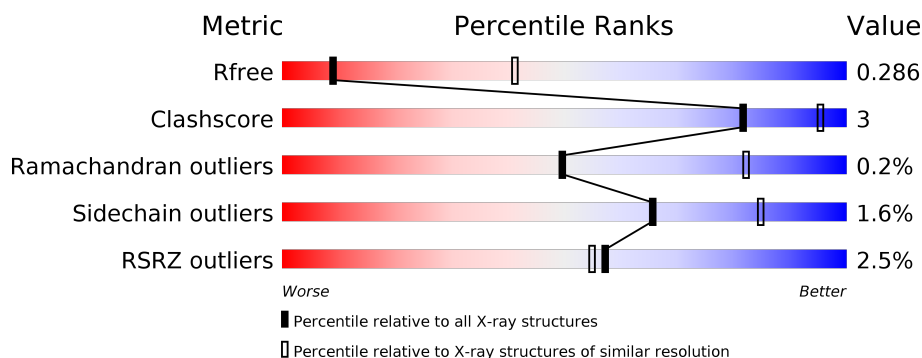
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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}	111664	1928 (3.50-3.30)
Clashscore	122126	2051 (3.50-3.30)
Ramachandran outliers	120053	2006 (3.50-3.30)
Sidechain outliers	120020	2006 (3.50-3.30)
RSRZ outliers	108989	1827 (3.50-3.30)

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. 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	795	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>8%</div> <div>7%</div> </div> </div>
1	B	795	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>9%</div> <div>6%</div> </div> </div>
1	C	795	<div> <div>3%</div> <div> <div></div> <div>85%</div> <div>8%</div> <div>7%</div> </div> </div>
1	D	795	<div> <div>3%</div> <div> <div></div> <div>85%</div> <div>9%</div> <div>6%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NGS	B	1785	X	-	-	-
3	NGS	C	1785	X	-	-	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 23370 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 TWO-COMPONENT SYSTEM SENSOR HISTIDINE KINASE/RESPONSE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	741	Total	C	N	O	S	0	0	0
			5791	3684	956	1139	12			
1	B	750	Total	C	N	O	S	0	0	0
			5893	3741	983	1157	12			
1	C	743	Total	C	N	O	S	0	0	0
			5794	3683	960	1139	12			
1	D	745	Total	C	N	O	S	0	0	0
			5772	3671	951	1138	12			

There are 32 discrepancies between the modelled and reference sequences:

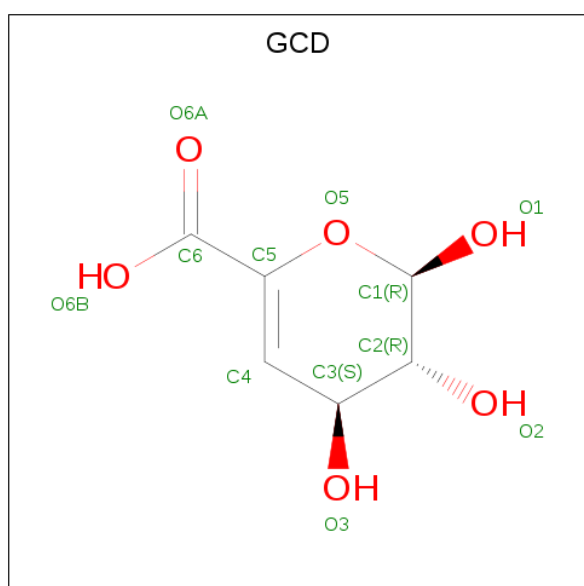
Chain	Residue	Modelled	Actual	Comment	Reference
A	788	LEU	-	expression tag	UNP Q89YR8
A	789	GLU	-	expression tag	UNP Q89YR8
A	790	HIS	-	expression tag	UNP Q89YR8
A	791	HIS	-	expression tag	UNP Q89YR8
A	792	HIS	-	expression tag	UNP Q89YR8
A	793	HIS	-	expression tag	UNP Q89YR8
A	794	HIS	-	expression tag	UNP Q89YR8
A	795	HIS	-	expression tag	UNP Q89YR8
B	788	LEU	-	expression tag	UNP Q89YR8
B	789	GLU	-	expression tag	UNP Q89YR8
B	790	HIS	-	expression tag	UNP Q89YR8
B	791	HIS	-	expression tag	UNP Q89YR8
B	792	HIS	-	expression tag	UNP Q89YR8
B	793	HIS	-	expression tag	UNP Q89YR8
B	794	HIS	-	expression tag	UNP Q89YR8
B	795	HIS	-	expression tag	UNP Q89YR8
C	788	LEU	-	expression tag	UNP Q89YR8
C	789	GLU	-	expression tag	UNP Q89YR8
C	790	HIS	-	expression tag	UNP Q89YR8
C	791	HIS	-	expression tag	UNP Q89YR8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	792	HIS	-	expression tag	UNP Q89YR8
C	793	HIS	-	expression tag	UNP Q89YR8
C	794	HIS	-	expression tag	UNP Q89YR8
C	795	HIS	-	expression tag	UNP Q89YR8
D	788	LEU	-	expression tag	UNP Q89YR8
D	789	GLU	-	expression tag	UNP Q89YR8
D	790	HIS	-	expression tag	UNP Q89YR8
D	791	HIS	-	expression tag	UNP Q89YR8
D	792	HIS	-	expression tag	UNP Q89YR8
D	793	HIS	-	expression tag	UNP Q89YR8
D	794	HIS	-	expression tag	UNP Q89YR8
D	795	HIS	-	expression tag	UNP Q89YR8

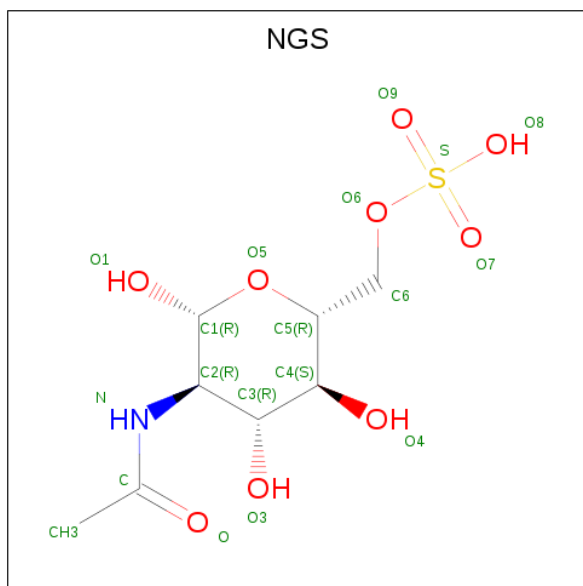
- Molecule 2 is 4,5-DEHYDRO-D-GLUCURONIC ACID (three-letter code: GCD) (formula: $C_6H_8O_6$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	6	6		
2	B	1	Total	C	O	0	0
			12	6	6		
2	C	1	Total	C	O	0	0
			12	6	6		
2	D	1	Total	C	O	0	0
			12	6	6		

- Molecule 3 is 2-(acetylamino)-2-deoxy-6-O-sulfo-beta-D-glucopyranose (three-letter code:

NGS) (formula: C₈H₁₅NO₉S).

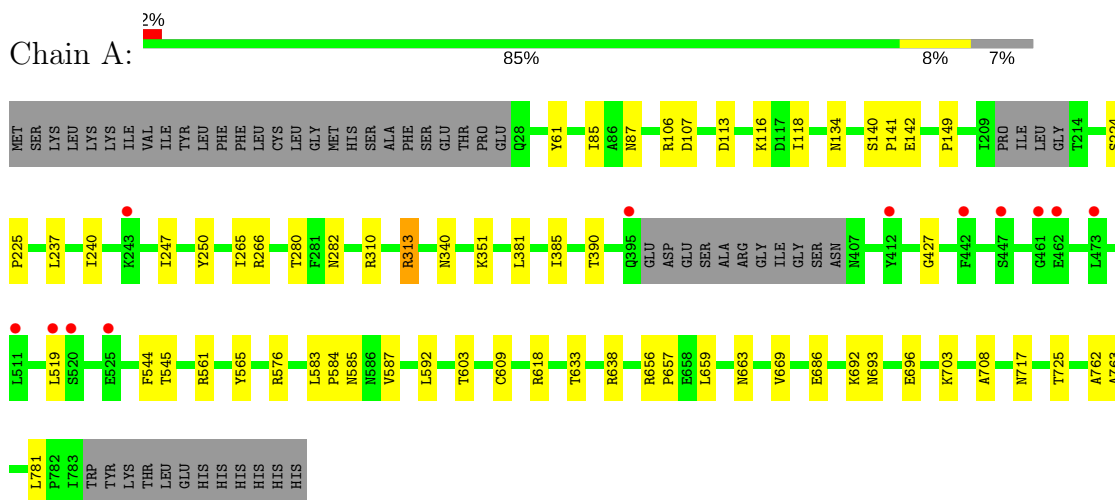


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			18	8	1	8	1		
3	B	1	Total	C	N	O	S	0	0
			18	8	1	8	1		
3	C	1	Total	C	N	O	S	0	0
			18	8	1	8	1		
3	D	1	Total	C	N	O	S	0	0
			18	8	1	8	1		

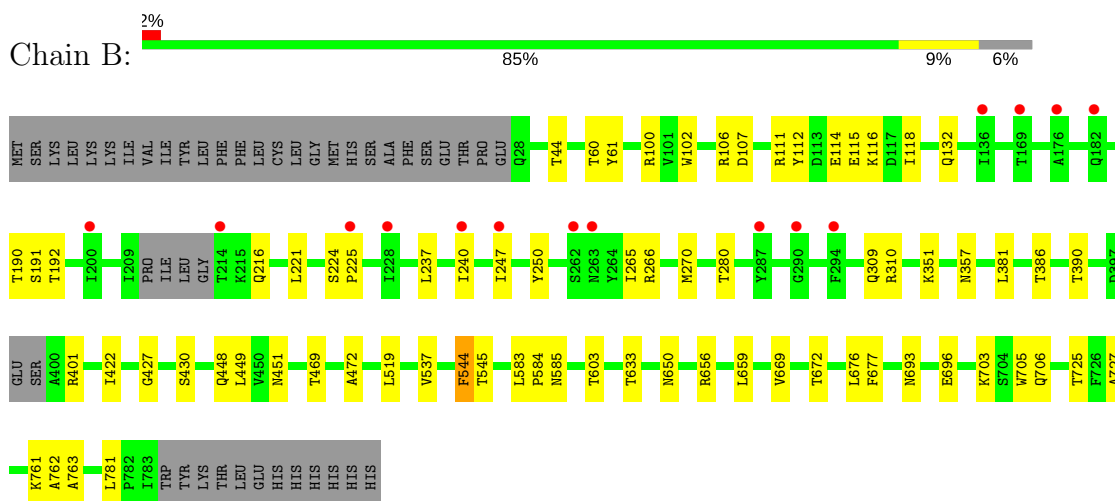
3 Residue-property plots [i](#)

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.

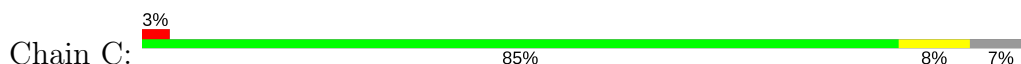
• Molecule 1: TWO-COMPONENT SYSTEM SENSOR HISTIDINE KINASE/RESPONSE

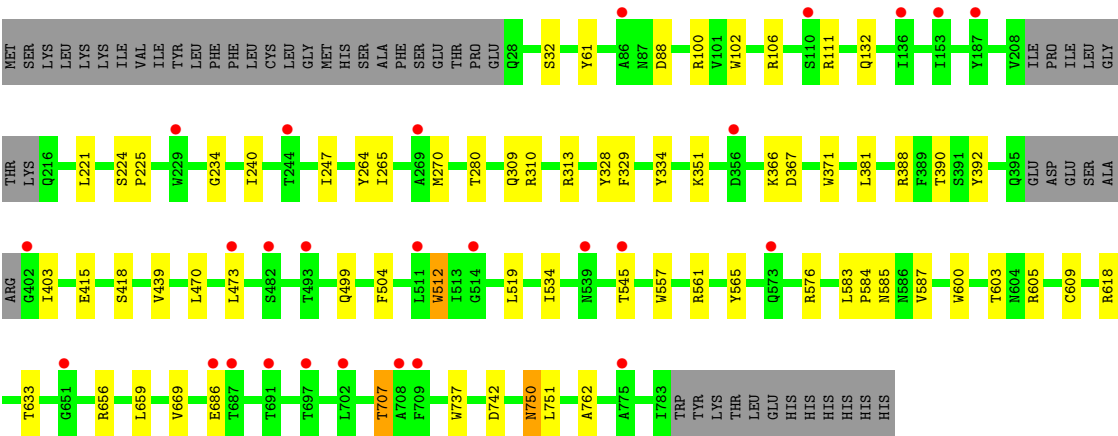


• Molecule 1: TWO-COMPONENT SYSTEM SENSOR HISTIDINE KINASE/RESPONSE

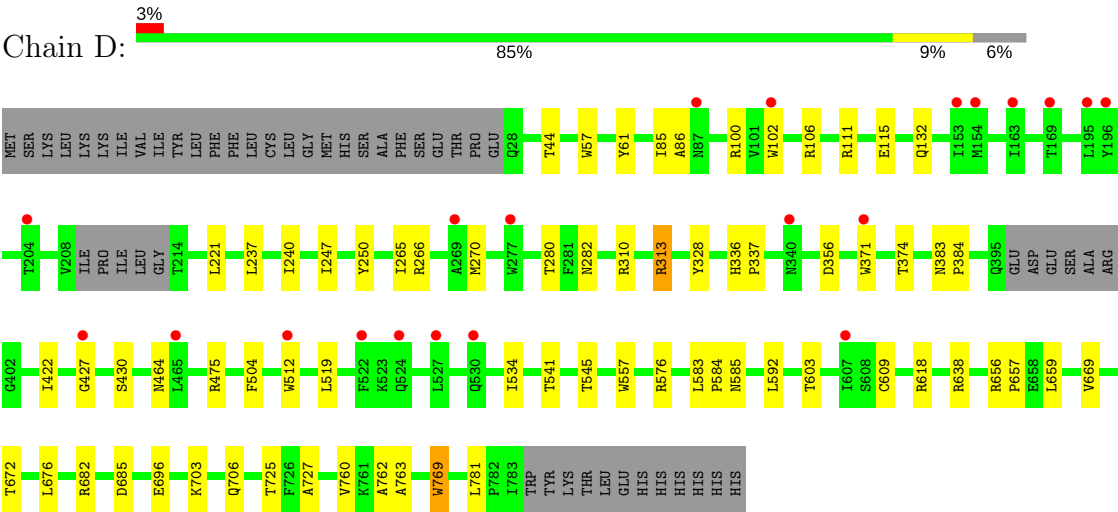


• Molecule 1: TWO-COMPONENT SYSTEM SENSOR HISTIDINE KINASE/RESPONSE





● Molecule 1: TWO-COMPONENT SYSTEM SENSOR HISTIDINE KINASE/RESPONSE



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	111.12Å 80.83Å 228.94Å 90.00° 93.96° 90.00°	Depositor
Resolution (Å)	55.73 – 3.40 55.73 – 3.40	Depositor EDS
% Data completeness (in resolution range)	97.5 (55.73-3.40) 97.6 (55.73-3.40)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.54 (at 3.40Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.268 , 0.286 0.267 , 0.286	Depositor DCC
R_{free} test set	2544 reflections (4.62%)	wwPDB-VP
Wilson B-factor (Å ²)	58.9	Xtriage
Anisotropy	0.457	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 52.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.85	EDS
Total number of atoms	23370	wwPDB-VP
Average B, all atoms (Å ²)	131.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.78% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GCD, NGS

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	0.43	0/5926	0.46	0/8076
1	B	0.43	2/6028 (0.0%)	0.47	0/8205
1	C	0.43	6/5929 (0.1%)	0.46	0/8081
1	D	0.43	5/5907 (0.1%)	0.46	0/8060
All	All	0.43	13/23790 (0.1%)	0.46	0/32422

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	371	TRP	CD2-CE2	5.11	1.47	1.41
1	D	557	TRP	CD2-CE2	5.08	1.47	1.41
1	D	102	TRP	CD2-CE2	5.05	1.47	1.41
1	C	102	TRP	CD2-CE2	5.05	1.47	1.41
1	C	512	TRP	CD2-CE2	5.04	1.47	1.41

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	5791	0	5420	30	0
1	B	5893	0	5554	31	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	5794	0	5419	29	0
1	D	5772	0	5349	33	0
2	A	12	0	6	0	0
2	B	12	0	6	1	0
2	C	12	0	6	2	0
2	D	12	0	6	0	0
3	A	18	0	13	0	0
3	B	18	0	13	1	0
3	C	18	0	13	4	0
3	D	18	0	13	0	0
All	All	23370	0	21818	126	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 3.

The worst 5 of 126 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:707:THR:HG22	1:C:750:ASN:HA	1.72	0.72
1:C:470:LEU:HD22	1:C:499:GLN:HE21	1.62	0.62
1:A:265:ILE:HG12	1:A:280:THR:HG22	1.84	0.59
1:C:609:CYS:HB3	1:C:618:ARG:HB3	1.84	0.59
1:D:609:CYS:HB3	1:D:618:ARG:HB3	1.85	0.59

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	735/795 (92%)	703 (96%)	30 (4%)	2 (0%)	43	77
1	B	744/795 (94%)	707 (95%)	35 (5%)	2 (0%)	43	77

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	737/795 (93%)	701 (95%)	34 (5%)	2 (0%)	43	77
1	D	739/795 (93%)	707 (96%)	31 (4%)	1 (0%)	53	85
All	All	2955/3180 (93%)	2818 (95%)	130 (4%)	7 (0%)	49	82

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	427	GLY
1	B	427	GLY
1	D	427	GLY
1	C	750	ASN
1	C	751	LEU

5.3.2 Protein sidechains ⓘ

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	620/712 (87%)	610 (98%)	10 (2%)	65	85
1	B	635/712 (89%)	621 (98%)	14 (2%)	55	81
1	C	620/712 (87%)	609 (98%)	11 (2%)	62	83
1	D	611/712 (86%)	606 (99%)	5 (1%)	83	92
All	All	2486/2848 (87%)	2446 (98%)	40 (2%)	65	85

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	386	THR
1	B	633	THR
1	D	313	ARG
1	B	537	VAL
1	B	672	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	241	ASN
1	B	448	GLN
1	C	678	ASN
1	B	132	GLN
1	C	499	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

8 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	GCD	A	1784	3	9,12,12	1.89	2 (22%)	10,17,17	0.80	0
3	NGS	A	1785	2	18,18,19	0.62	0	24,26,28	1.50	5 (20%)
2	GCD	B	1784	3	9,12,12	1.95	2 (22%)	10,17,17	0.97	1 (10%)
3	NGS	B	1785	2	18,18,19	0.76	0	24,26,28	1.30	2 (8%)
2	GCD	C	1784	3	9,12,12	1.84	1 (11%)	10,17,17	1.19	1 (10%)
3	NGS	C	1785	2	18,18,19	0.92	0	24,26,28	1.54	4 (16%)
2	GCD	D	1784	3	9,12,12	1.84	1 (11%)	10,17,17	1.01	1 (10%)
3	NGS	D	1785	2	18,18,19	0.65	0	24,26,28	1.48	5 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GCD	A	1784	3	-	0/0/20/20	0/1/1/1
3	NGS	A	1785	2	-	0/10/26/30	0/1/1/1
2	GCD	B	1784	3	-	0/0/20/20	0/1/1/1
3	NGS	B	1785	2	1/1/6/8	0/10/26/30	0/1/1/1
2	GCD	C	1784	3	-	0/0/20/20	0/1/1/1
3	NGS	C	1785	2	1/1/6/8	0/10/26/30	1/1/1/1
2	GCD	D	1784	3	-	0/0/20/20	0/1/1/1
3	NGS	D	1785	2	-	0/10/26/30	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1784	GCD	C4-C5	2.09	1.35	1.32
2	B	1784	GCD	C4-C5	2.13	1.35	1.32
2	C	1784	GCD	O5-C5	4.74	1.46	1.37
2	B	1784	GCD	O5-C5	4.82	1.46	1.37
2	A	1784	GCD	O5-C5	4.94	1.46	1.37

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1785	NGS	C1-C2-N	-4.39	105.64	110.73
3	D	1785	NGS	C1-C2-N	-3.83	106.29	110.73
3	A	1785	NGS	C1-C2-N	-3.47	106.70	110.73
3	A	1785	NGS	O5-C1-C2	-3.22	106.28	109.52
3	D	1785	NGS	O5-C1-C2	-3.15	106.35	109.52

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	B	1785	NGS	C4
3	C	1785	NGS	C4

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	1785	NGS	C1-C2-C3-C4-C5-O5

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1784	GCD	1	0
3	B	1785	NGS	1	0
2	C	1784	GCD	2	0
3	C	1785	NGS	4	0

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	741/795 (93%)	0.21	12 (1%) 72 69	76, 109, 211, 243	0
1	B	750/795 (94%)	0.20	15 (2%) 65 63	72, 105, 208, 257	0
1	C	743/795 (93%)	0.49	27 (3%) 42 41	129, 144, 167, 184	0
1	D	745/795 (93%)	0.43	21 (2%) 53 51	124, 139, 163, 179	0
All	All	2979/3180 (93%)	0.33	75 (2%) 57 55	72, 135, 191, 257	0

The worst 5 of 75 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	102	TRP	5.1
1	D	204	THR	4.5
1	A	520	SER	3.8
1	B	263	ASN	3.7
1	C	709	PHE	3.3

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 carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	GCD	C	1784	12/12	0.87	0.27	57,57,57,57	0
3	NGS	B	1785	18/19	0.87	0.24	57,57,57,57	0
2	GCD	B	1784	12/12	0.88	0.29	57,57,57,57	0
3	NGS	D	1785	18/19	0.90	0.23	57,57,57,57	0
2	GCD	D	1784	12/12	0.91	0.25	57,57,57,57	0
3	NGS	C	1785	18/19	0.91	0.20	57,57,57,57	0
3	NGS	A	1785	18/19	0.93	0.18	57,57,57,57	0
2	GCD	A	1784	12/12	0.94	0.19	57,57,57,57	0

6.5 Other polymers [i](#)

There are no such residues in this entry.