



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 06:45 am BST

PDB ID : 5G1P
Title : Aspartate transcarbamoylase domain of human CAD bound to carbamoyl phosphate
Authors : Ruiz-Ramos, A.; Grande-Garcia, A.; Moreno-Morcillo, M.D.; Ramon-Maiques, S.
Deposited on : 2016-03-29
Resolution : 3.19 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	:	1.8.5 (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

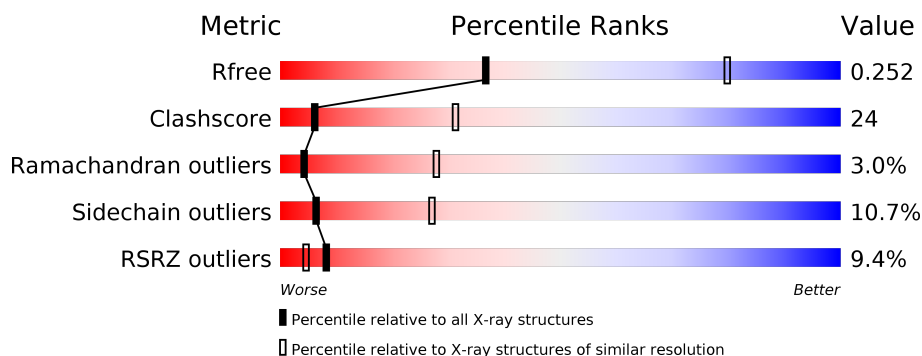
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.19 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 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 $\leq 5\%$. 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	314	<div> <div>9%</div> <div>49% 38% 6% 7%</div> </div>
1	B	314	<div> <div>8%</div> <div>50% 36% 5% 8%</div> </div>
1	C	314	<div> <div>10%</div> <div>49% 38% 5% 8%</div> </div>
1	D	314	<div> <div>11%</div> <div>51% 38% • 7%</div> </div>
1	E	314	<div> <div>6%</div> <div>54% 35% 6% • 5%</div> </div>
1	F	314	<div> <div>10%</div> <div>53% 37% 6% • •</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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	CP	D	3226	-	-	X	-
2	CP	F	3226	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 13193 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 CAD PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	292	Total	C	N	O	S	0	0	0
			2175	1363	385	407	20			
1	B	288	Total	C	N	O	S	0	0	0
			2128	1337	378	393	20			
1	C	288	Total	C	N	O	S	0	0	0
			2141	1344	384	394	19			
1	D	292	Total	C	N	O	S	0	0	0
			2171	1365	384	401	21			
1	E	299	Total	C	N	O	S	0	0	0
			2197	1380	396	401	20			
1	F	304	Total	C	N	O	S	0	0	0
			2238	1407	400	414	17			

There are 18 discrepancies between the modelled and reference sequences:

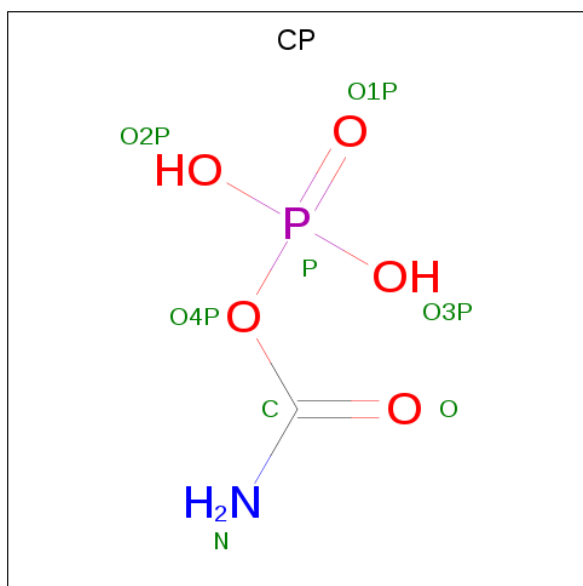
Chain	Residue	Modelled	Actual	Comment	Reference
A	1912	GLY	-	expression tag	UNP P27708
A	1913	PRO	-	expression tag	UNP P27708
A	1914	MET	-	expression tag	UNP P27708
B	1912	GLY	-	expression tag	UNP P27708
B	1913	PRO	-	expression tag	UNP P27708
B	1914	MET	-	expression tag	UNP P27708
C	1912	GLY	-	expression tag	UNP P27708
C	1913	PRO	-	expression tag	UNP P27708
C	1914	MET	-	expression tag	UNP P27708
D	1912	GLY	-	expression tag	UNP P27708
D	1913	PRO	-	expression tag	UNP P27708
D	1914	MET	-	expression tag	UNP P27708
E	1912	GLY	-	expression tag	UNP P27708
E	1913	PRO	-	expression tag	UNP P27708
E	1914	MET	-	expression tag	UNP P27708
F	1912	GLY	-	expression tag	UNP P27708
F	1913	PRO	-	expression tag	UNP P27708

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	1914	MET	-	expression tag	UNP P27708

- Molecule 2 is PHOSPHORIC ACID MONO(FORMAMIDE)ESTER (three-letter code: CP) (formula: $\text{CH}_4\text{NO}_5\text{P}$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			8	1	1	5	1		
2	B	1	Total	C	N	O	P	0	0
			8	1	1	5	1		
2	C	1	Total	C	N	O	P	0	0
			8	1	1	5	1		
2	D	1	Total	C	N	O	P	0	0
			8	1	1	5	1		
2	E	1	Total	C	N	O	P	0	0
			8	1	1	5	1		
2	F	1	Total	C	N	O	P	0	0
			8	1	1	5	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	16	Total	O	0	0
			16	16		
3	B	14	Total	O	0	0
			14	14		

Continued on next page...

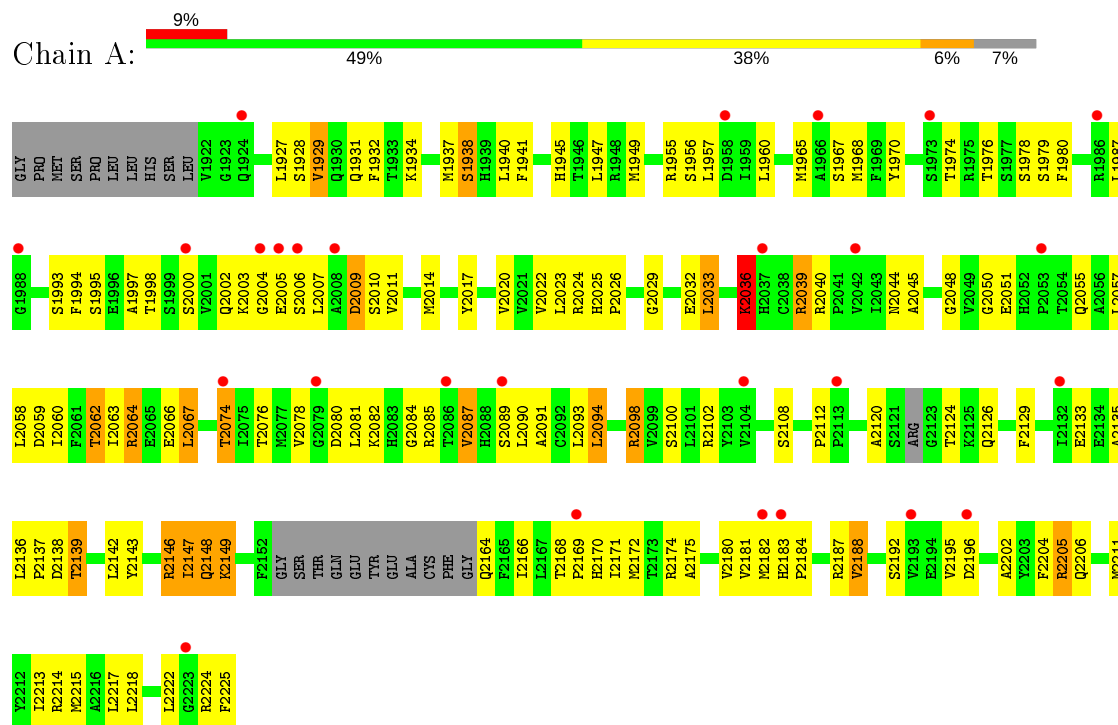
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	13	Total 13	O 13	0	0
3	D	15	Total 15	O 15	0	0
3	E	22	Total 22	O 22	0	0
3	F	15	Total 15	O 15	0	0

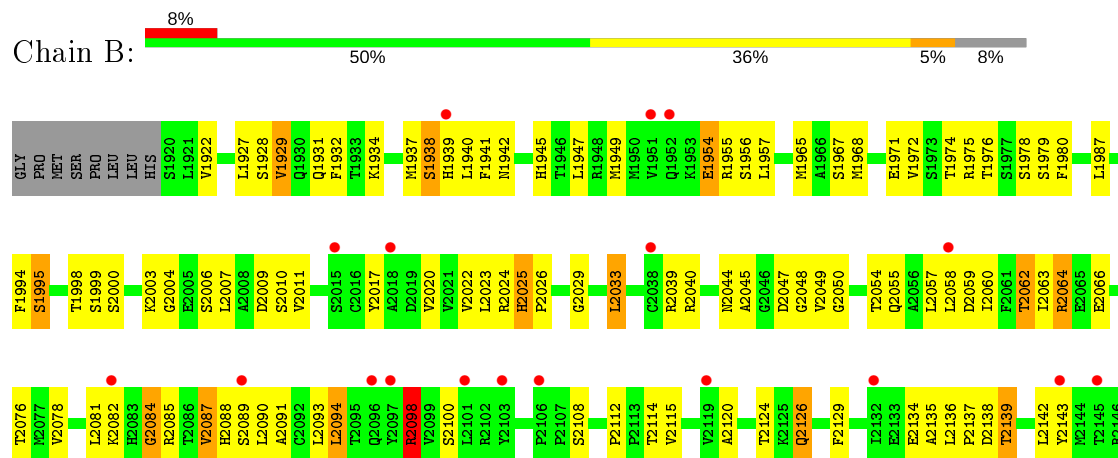
3 Residue-property plots [i](#)

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.

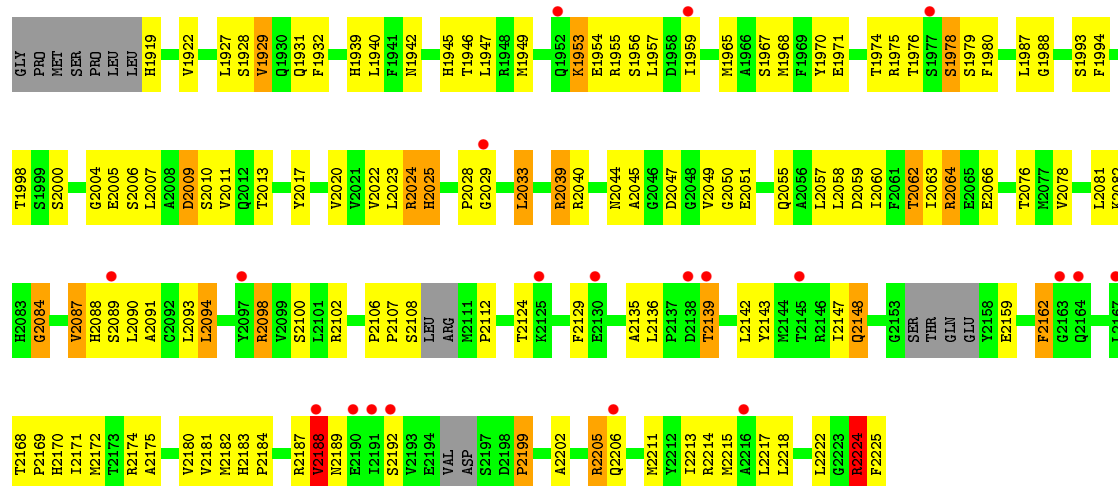
• Molecule 1: CAD PROTEIN



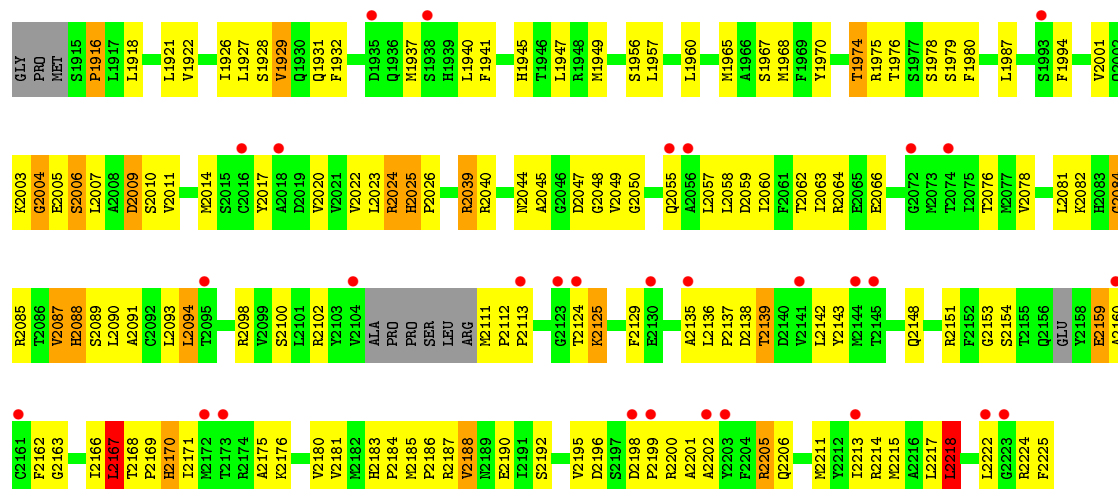
• Molecule 1: CAD PROTEIN







• Molecule 1: CAD PROTEIN



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	83.67Å 157.67Å 83.54Å 90.00° 120.08° 90.00°	Depositor
Resolution (Å)	42.54 – 3.19 42.53 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.3 (42.54-3.19) 99.3 (42.53-3.20)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.24 (at 3.19Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.221 , 0.251 0.222 , 0.252	Depositor DCC
R_{free} test set	1545 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	59.8	Xtriage
Anisotropy	0.431	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 38.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	0.349 for l,k,-h-l 0.349 for -h-l,k,h 0.388 for l,-k,h 0.338 for -h-l,-k,l 0.408 for h,-k,-h-l	Xtriage
Reported twinning fraction	0.284 for H, K, L 0.225 for H, -K, -H-L 0.188 for L, -K, H 0.119 for L, K, -H-L 0.099 for -H-L, K, H 0.085 for H+L, -K, -L	Depositor
Outliers	0 of 30893 reflections	Xtriage
F_o, F_c correlation	0.85	EDS
Total number of atoms	13193	wwPDB-VP
Average B, all atoms (Å ²)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.75% 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 ⓘ

5.1 Standard geometry ⓘ

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

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	0.60	0/2209	0.89	7/2992 (0.2%)
1	B	0.61	0/2164	0.90	10/2937 (0.3%)
1	C	0.60	0/2174	0.89	6/2943 (0.2%)
1	D	0.60	0/2206	0.94	12/2986 (0.4%)
1	E	0.61	0/2234	1.05	10/3028 (0.3%)
1	F	0.60	0/2274	0.92	11/3083 (0.4%)
All	All	0.60	0/13261	0.93	56/17969 (0.3%)

There are no bond length outliers.

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	2024	ARG	NE-CZ-NH1	22.40	131.50	120.30
1	E	2024	ARG	NE-CZ-NH2	-21.46	109.57	120.30
1	F	2064	ARG	NE-CZ-NH1	12.47	126.54	120.30
1	D	2064	ARG	NE-CZ-NH1	11.59	126.09	120.30
1	A	2205	ARG	NE-CZ-NH1	10.84	125.72	120.30

There are no chirality outliers.

There are no planarity outliers.

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	2175	0	2130	117	1
1	B	2128	0	2084	102	0
1	C	2141	0	2109	120	0
1	D	2171	0	2116	113	1
1	E	2197	0	2110	101	0
1	F	2238	0	2149	115	0
2	A	8	0	2	0	0
2	B	8	0	2	0	0
2	C	8	0	2	2	0
2	D	8	0	2	4	0
2	E	8	0	2	1	0
2	F	8	0	2	4	0
3	A	16	0	0	1	0
3	B	14	0	0	0	0
3	C	13	0	0	0	0
3	D	15	0	0	1	0
3	E	22	0	0	1	0
3	F	15	0	0	0	0
All	All	13193	0	12710	632	1

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 24.

The worst 5 of 632 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:2032:GLU:OE1	3:A:4009:HOH:O	1.59	1.16
1:F:2066:GLU:OE2	1:F:2205:ARG:NH1	1.82	1.13
1:E:2084:GLY:O	1:E:2088:HIS:NE2	1.85	1.10
1:A:2063:ILE:O	1:A:2067:LEU:HD12	1.52	1.08
1:E:2076:THR:OG1	1:E:2139:THR:OG1	1.74	1.04

All (1) 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:1945:HIS:NE2	1:D:1945:HIS:CE1[1_454]	1.95	0.25

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	286/314 (91%)	247 (86%)	32 (11%)	7 (2%)	6	34
1	B	284/314 (90%)	251 (88%)	28 (10%)	5 (2%)	8	41
1	C	282/314 (90%)	247 (88%)	27 (10%)	8 (3%)	5	29
1	D	282/314 (90%)	241 (86%)	32 (11%)	9 (3%)	4	26
1	E	291/314 (93%)	247 (85%)	34 (12%)	10 (3%)	3	24
1	F	298/314 (95%)	249 (84%)	37 (12%)	12 (4%)	3	21
All	All	1723/1884 (92%)	1482 (86%)	190 (11%)	51 (3%)	4	28

5 of 51 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2149	LYS
1	D	2147	ILE
1	D	2148	GLN
1	F	1916	PRO
1	F	2125	LYS

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	226/268 (84%)	201 (89%)	25 (11%)	6	25
1	B	220/268 (82%)	192 (87%)	28 (13%)	4	20
1	C	221/268 (82%)	199 (90%)	22 (10%)	7	30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	224/268 (84%)	202 (90%)	22 (10%)	8	31
1	E	218/268 (81%)	193 (88%)	25 (12%)	5	24
1	F	223/268 (83%)	202 (91%)	21 (9%)	8	33
All	All	1332/1608 (83%)	1189 (89%)	143 (11%)	6	27

5 of 143 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	2039	ARG
1	D	2006	SER
1	F	2088	HIS
1	C	2064	ARG
1	C	2174	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	1936	GLN
1	D	1942	ASN
1	F	1942	ASN
1	B	2183	HIS
1	C	2183	HIS

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

6 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	CP	F	3226	-	6,7,7	1.09	1 (16%)	7,10,10	1.43	2 (28%)
2	CP	D	3226	-	6,7,7	0.77	0	7,10,10	2.13	5 (71%)
2	CP	A	3226	-	6,7,7	1.40	1 (16%)	7,10,10	0.93	0
2	CP	E	3226	-	6,7,7	1.06	0	7,10,10	1.52	1 (14%)
2	CP	C	3226	-	6,7,7	1.29	1 (16%)	7,10,10	1.83	1 (14%)
2	CP	B	3226	-	6,7,7	1.12	0	7,10,10	1.62	2 (28%)

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	CP	F	3226	-	-	0/3/5/5	-
2	CP	D	3226	-	-	0/3/5/5	-
2	CP	A	3226	-	-	0/3/5/5	-
2	CP	E	3226	-	-	0/3/5/5	-
2	CP	C	3226	-	-	0/3/5/5	-
2	CP	B	3226	-	-	0/3/5/5	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	3226	CP	P-O4P	2.98	1.63	1.59
2	C	3226	CP	P-O4P	2.42	1.62	1.59
2	F	3226	CP	P-O4P	2.29	1.62	1.59

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	C	3226	CP	O-C-N	-3.95	119.00	125.51
2	B	3226	CP	O-C-N	-2.87	120.77	125.51
2	E	3226	CP	O3P-P-O2P	2.81	118.38	107.64
2	D	3226	CP	O3P-P-O1P	2.62	120.94	110.68
2	D	3226	CP	O3P-P-O2P	2.45	117.00	107.64

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	3226	CP	4	0
2	D	3226	CP	4	0
2	E	3226	CP	1	0
2	C	3226	CP	2	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	292/314 (92%)	0.37	27 (9%) 9 5	23, 51, 80, 94	0
1	B	288/314 (91%)	0.38	24 (8%) 11 6	24, 53, 83, 100	0
1	C	288/314 (91%)	0.45	30 (10%) 6 4	22, 47, 72, 88	0
1	D	292/314 (92%)	0.49	35 (11%) 4 2	17, 49, 82, 110	0
1	E	299/314 (95%)	0.29	20 (6%) 17 10	26, 49, 74, 90	0
1	F	304/314 (96%)	0.49	30 (9%) 7 4	19, 49, 75, 98	0
All	All	1763/1884 (93%)	0.41	166 (9%) 8 4	17, 49, 78, 110	0

The worst 5 of 166 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	2074	THR	9.0
1	F	2198	ASP	7.0
1	D	2089	SER	6.4
1	C	1974	THR	5.8
1	D	2121	SER	5.5

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	CP	A	3226	8/8	0.84	0.24	34,49,61,64	0
2	CP	F	3226	8/8	0.92	0.17	43,46,52,56	0
2	CP	C	3226	8/8	0.93	0.16	41,44,49,51	0
2	CP	E	3226	8/8	0.94	0.15	28,32,34,35	0
2	CP	D	3226	8/8	0.94	0.17	25,29,32,32	0
2	CP	B	3226	8/8	0.96	0.14	31,33,34,34	0

6.5 Other polymers [i](#)

There are no such residues in this entry.