wwPDB X-ray Structure Validation Summary Report (i)

Dec 2, 2023 – 01:18 pm GMT

PDB ID	:	2JG9
Title	:	Crystallographic structure of human C1q globular heads (P1)
Authors	:	Paidassi, H.; Tacnet-Delorme, P.; Garlatti, V.; Darnault, C.; Ghebrehiwet, B.; Gaboriaud, C.; Arlaud, G.J.; Frachet, P.
Deposited on	:	2007-02-09
Resolution	:	1.90 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:

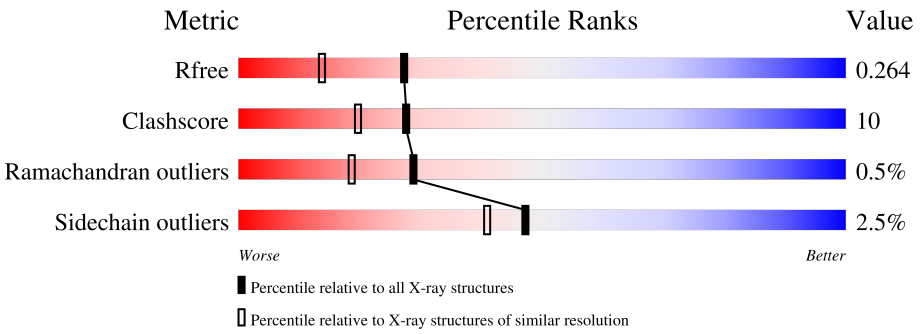
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)

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%

Mol	Chain	Length	Quality of chain
1	A	134	83% 16% ••
1	D	134	78% 20% •
2	B	136	76% 19% • •
2	E	136	70% 26% • •
3	C	131	81% 15% • • •
3	F	131	84% 12% • • •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6263 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 Complement C1q subcomponent subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	134	1042	666	178	193	5	0	0	1
1	D	134	1042	666	178	193	5	0	0	1

- Molecule 2 is a protein called Complement C1q subcomponent subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	133	1028	653	179	189	7	0	0	1
2	E	133	1028	653	179	189	7	0	0	1

- Molecule 3 is a protein called Complement C1q subcomponent subunit C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	129	1006	646	168	188	4	0	0	0
3	F	129	1006	646	168	188	4	0	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
4	B	1	1	1	0	0
4	E	1	1	1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	24	24	24	0	0
			Total	O		
5	B	10	10	10	0	0
			Total	O		
5	C	26	26	26	0	0
			Total	O		
5	D	25	25	25	0	0
			Total	O		
5	E	10	10	10	0	0
			Total	O		
5	F	14	14	14	0	0

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. 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. 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: Complement C1q subcomponent subunit A

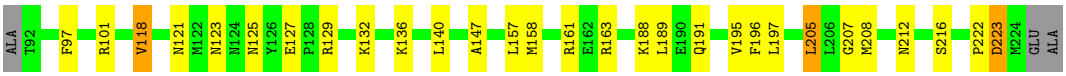
Chain A:

• Molecule 1: Complement C1q subcomponent subunit A

Chain D:

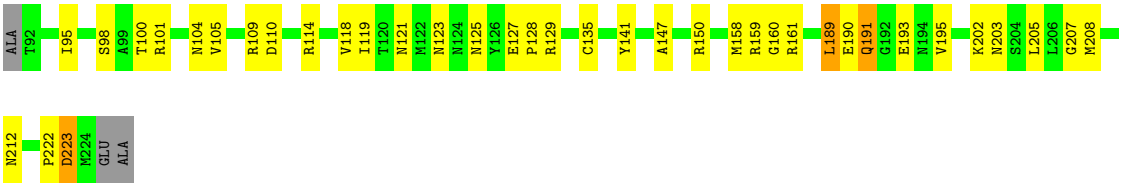
• Molecule 2: Complement C1q subcomponent subunit B

Chain B:



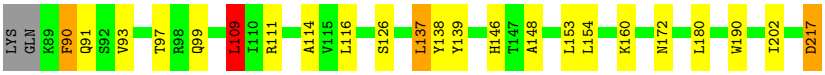
• Molecule 2: Complement C1q subcomponent subunit B

Chain E:



• Molecule 3: Complement C1q subcomponent subunit C

Chain C:



• Molecule 3: Complement C1q subcomponent subunit C

Chain F:

84%
12%

LYS
GLN
K89
F90
Q91
S92
V93
T97
R98
Q99
L109
I110
R111
L116
T117
N118
P119
S126
G136
Y139
H146
L153
K160
L180
W190
D217

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	48.15Å 48.24Å 87.79Å 92.45° 92.65° 113.53°	Depositor
Resolution (Å)	15.00 – 1.90 44.13 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (15.00-1.90) 92.2 (44.13-1.90)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.11 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.214 , 0.260 0.218 , 0.264	Depositor DCC
R_{free} test set	5293 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	11.0	Xtriage
Anisotropy	0.487	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 19.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.38$, $< L^2 > = 0.20$	Xtriage
Estimated twinning fraction	0.225 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6263	wwPDB-VP
Average B, all atoms (Å^2)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.85% 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)

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

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	0.74	0/1070	0.81	1/1453 (0.1%)
1	D	0.73	0/1070	0.77	0/1453
2	B	0.69	0/1049	0.76	0/1419
2	E	0.70	1/1049 (0.1%)	0.77	0/1419
3	C	0.75	0/1033	0.82	2/1408 (0.1%)
3	F	0.74	0/1033	0.83	1/1408 (0.1%)
All	All	0.73	1/6304 (0.0%)	0.80	4/8560 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	135	CYS	CB-SG	-5.09	1.73	1.81

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	109	LEU	CA-CB-CG	8.73	135.39	115.30
3	C	109	LEU	CA-CB-CG	8.35	134.51	115.30
1	A	99	ARG	NE-CZ-NH2	-6.50	117.05	120.30
3	C	137	LEU	CA-CB-CG	-5.10	103.57	115.30

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	1042	0	998	18	0
1	D	1042	0	998	21	0
2	B	1028	0	993	18	1
2	E	1028	0	993	27	1
3	C	1006	0	968	25	0
3	F	1006	0	968	23	0
4	B	1	0	0	0	0
4	E	1	0	0	0	0
5	A	24	0	0	4	0
5	B	10	0	0	0	0
5	C	26	0	0	1	0
5	D	25	0	0	2	0
5	E	10	0	0	1	0
5	F	14	0	0	1	0
All	All	6263	0	5918	121	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 10.

The worst 5 of 121 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)
3:F:217:ASP:OXT	3:F:217:ASP:OD1	1.53	1.25
3:C:99:GLN:OE1	3:F:97:THR:HG21	1.53	1.08
2:B:125:ASN:OD1	2:B:136:LYS:HE3	1.53	1.06
3:C:217:ASP:OXT	3:C:217:ASP:OD1	1.79	1.00
2:E:159:ARG:HG3	2:E:193:GLU:OE1	1.66	0.96

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)
2:B:163:ARG:O	2:E:190:GLU:OE2[1_556]	2.12	0.08

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	132/134 (98%)	124 (94%)	7 (5%)	1 (1%)	19	9
1	D	132/134 (98%)	124 (94%)	7 (5%)	1 (1%)	19	9
2	B	131/136 (96%)	122 (93%)	8 (6%)	1 (1%)	19	9
2	E	131/136 (96%)	122 (93%)	8 (6%)	1 (1%)	19	9
3	C	127/131 (97%)	123 (97%)	4 (3%)	0	100	100
3	F	127/131 (97%)	123 (97%)	4 (3%)	0	100	100
All	All	780/802 (97%)	738 (95%)	38 (5%)	4 (0%)	29	18

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	223	ASP
1	A	222	SER
1	D	222	SER
2	E	223	ASP

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	115/117 (98%)	114 (99%)	1 (1%)	78	79
1	D	115/117 (98%)	112 (97%)	3 (3%)	46	39
2	B	109/117 (93%)	106 (97%)	3 (3%)	43	36

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	E	109/117 (93%)	105 (96%)	4 (4%)	34	25
3	C	112/115 (97%)	109 (97%)	3 (3%)	44	38
3	F	112/115 (97%)	109 (97%)	3 (3%)	44	38
All	All	672/698 (96%)	655 (98%)	17 (2%)	47	41

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	90	PHE
3	F	217	ASP
1	D	119	GLU
1	D	157	SER
1	D	165	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	91	GLN
3	F	102	GLN
3	F	146	HIS
3	C	146	HIS
1	D	186	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)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.