W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Oct 23, 2021 – 09:34 AM EDT

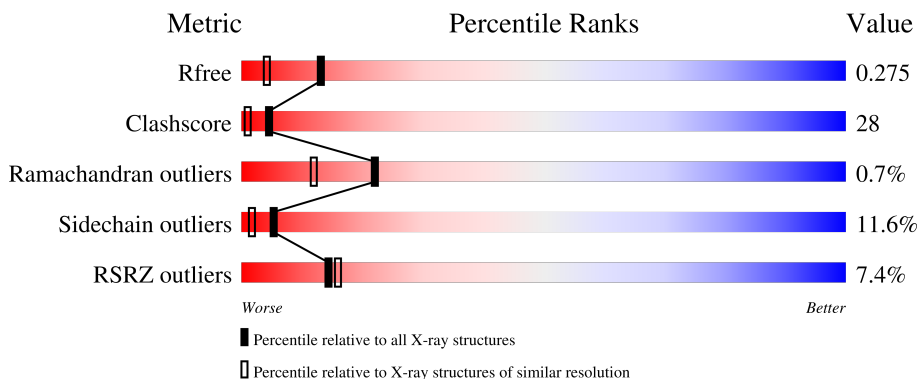|		
PDB ID	:	1QSJ
Title	:	N-TERMINALLY TRUNCATED C3DG FRAGMENT
Authors	:	Zanotti, G.; Bassetto, A.; Battistutta, R.; Stoppini, M.; Folli, C.; Berni, R.
Deposited on	:	1999-06-22
Resolution	:	1.90 Å(reported)

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 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	Percentile Ranks	Value
Rfree		0.275
Clashscore		28
Ramachandran outliers		0.7%
Sidechain outliers		11.6%
RSRZ outliers		7.4%

Worse *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

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)
RSRZ outliers	127900	6082 (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% 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	277	8% / 58% / 35% / 6%
1	B	277	7% / 61% / 33% / 5%
1	C	277	7% / 55% / 39% / 5%
1	D	277	8% / 59% / 35% / 5% •

2   Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 9162 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 C3 PRECURSOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total	C	N	O	S	0	0	0
			2189	1406	364	411	8			
1	B	276	Total	C	N	O	S	0	0	0
			2189	1406	364	411	8			
1	C	276	Total	C	N	O	S	0	0	0
			2189	1406	364	411	8			
1	D	276	Total	C	N	O	S	0	0	0
			2189	1406	364	411	8			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1081	TYR	MET	engineered mutation	UNP P01026
B	1081	TYR	MET	engineered mutation	UNP P01026
C	1081	TYR	MET	engineered mutation	UNP P01026
D	1081	TYR	MET	engineered mutation	UNP P01026
A	1082	VAL	TRP	engineered mutation	UNP P01026
B	1082	VAL	TRP	engineered mutation	UNP P01026
C	1082	VAL	TRP	engineered mutation	UNP P01026
D	1082	VAL	TRP	engineered mutation	UNP P01026
A	1085	VAL	SER	engineered mutation	UNP P01026
B	1085	VAL	SER	engineered mutation	UNP P01026
C	1085	VAL	SER	engineered mutation	UNP P01026
D	1085	VAL	SER	engineered mutation	UNP P01026

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	87	Total	O	0	0
			87	87		
2	B	115	Total	O	0	0
			115	115		

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	101	Total 101	O 101	0	0
2	D	103	Total 103	O 103	0	0

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: COMPLEMENT C3 PRECURSOR
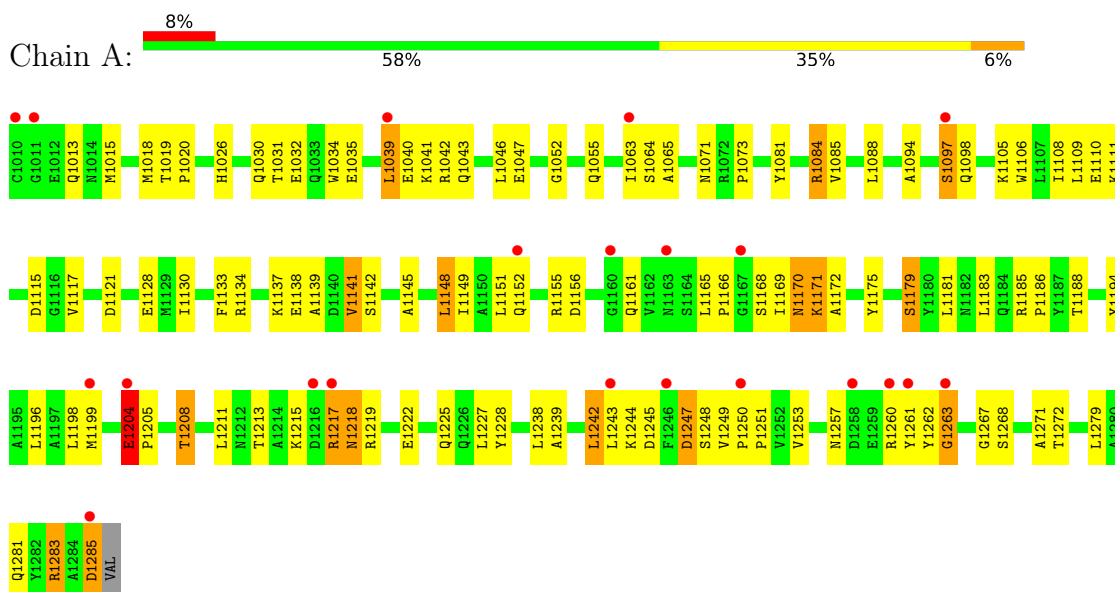
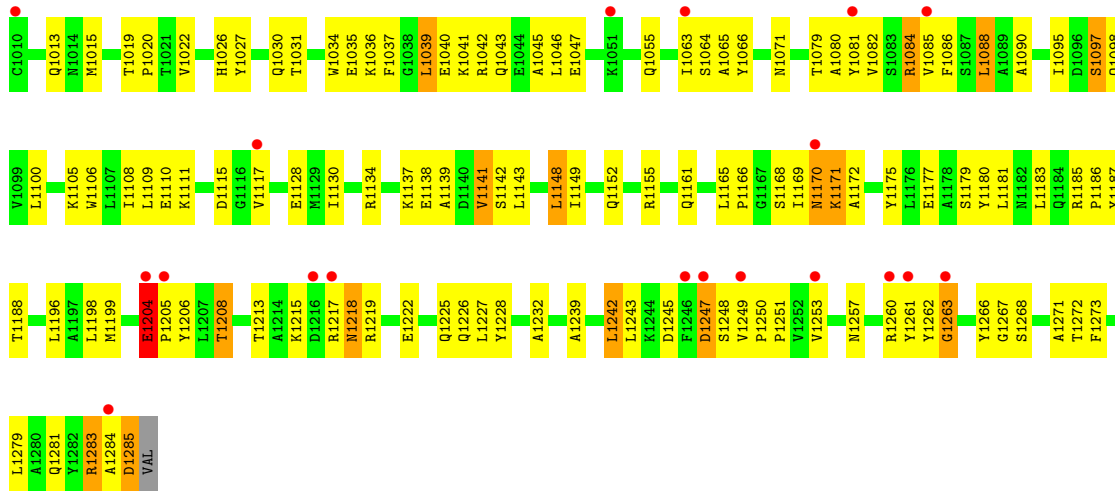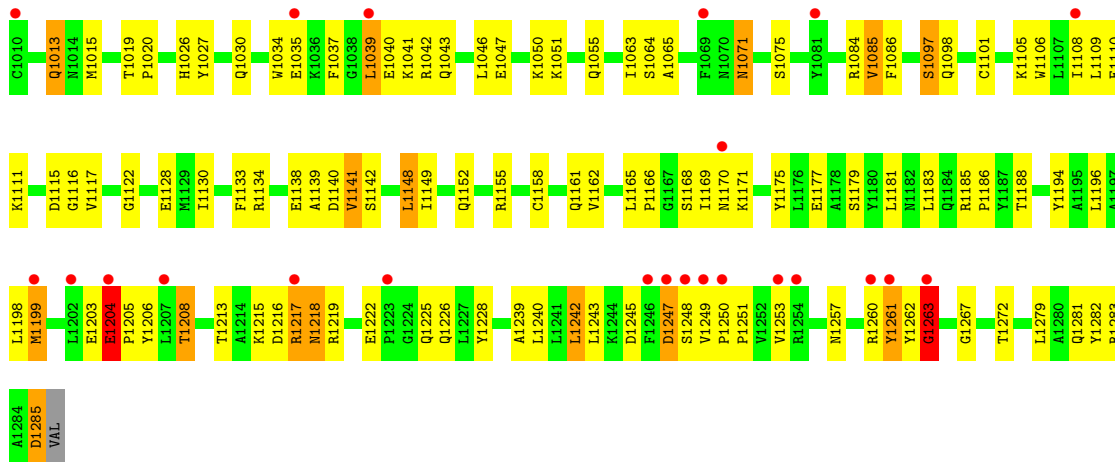Chain A:

• Molecule 1: COMPLEMENT C3 PRECURSOR

Chain B:

• Molecule 1: COMPLEMENT C3 PRECURSOR

Chain C:

• Molecule 1: COMPLEMENT C3 PRECURSOR

Chain D:
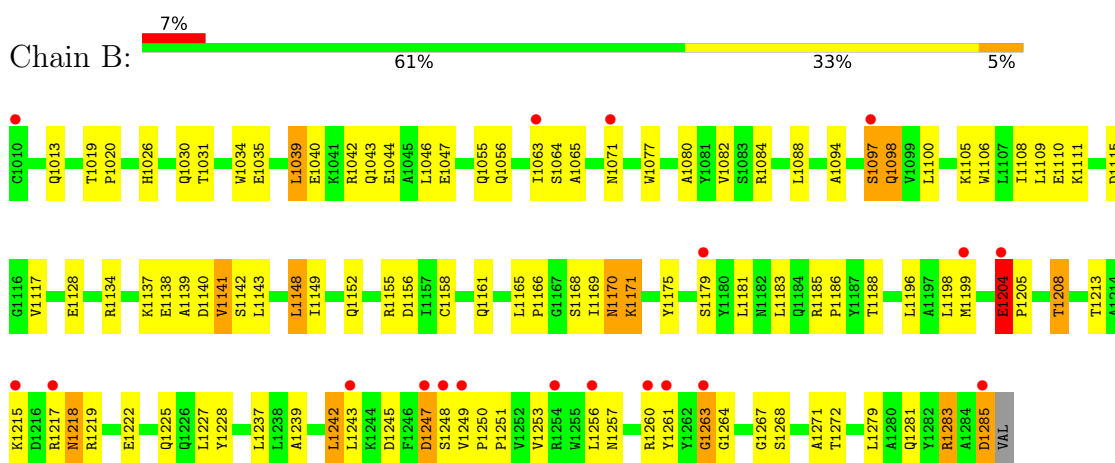
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	99.62Å 99.93Å 114.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	55.00 – 1.90 25.13 – 1.94	Depositor EDS
% Data completeness (in resolution range)	(Not available) (55.00-1.90) 75.0 (25.13-1.94)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	15.24 (at 1.95Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.212 , 0.253 0.230 , 0.275	Depositor DCC
R_{free} test set	6491 reflections (10.09%)	wwPDB-VP
Wilson B-factor (Å^2)	16.8	Xtriage
Anisotropy	0.307	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 45.9	EDS
L-test for twinning[2]	$< |L| > = 0.40, < L^2 > = 0.23$	Xtriage
Estimated twinning fraction	0.274 for k,h,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	9162	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.16% 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	0.57	2/2238 (0.1%)	0.59	1/3040 (0.0%)
1	B	0.50	2/2238 (0.1%)	0.65	3/3040 (0.1%)
1	C	0.43	1/2238 (0.0%)	0.68	2/3040 (0.1%)
1	D	0.44	1/2238 (0.0%)	0.67	3/3040 (0.1%)
All	All	0.49	6/8952 (0.1%)	0.65	9/12160 (0.1%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1285	ASP	C-O	-15.56	0.93	1.23
1	B	1285	ASP	C-O	-10.41	1.03	1.23
1	A	1285	ASP	CA-C	9.53	1.77	1.52
1	C	1285	ASP	C-O	-6.70	1.10	1.23
1	B	1285	ASP	CA-CB	6.69	1.68	1.53

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	1285	ASP	CA-C-O	18.25	158.43	120.10
1	D	1285	ASP	CA-C-O	17.38	156.59	120.10
1	B	1285	ASP	CA-C-O	10.79	142.77	120.10
1	B	1285	ASP	N-CA-C	6.82	129.42	111.00
1	D	1263	GLY	N-CA-C	5.33	126.41	113.10

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	2189	0	2163	128	0
1	B	2189	0	2165	113	0
1	C	2189	0	2165	141	0
1	D	2189	0	2165	125	0
2	A	87	0	0	28	0
2	B	115	0	0	28	0
2	C	101	0	0	31	0
2	D	103	0	0	18	0
All	All	9162	0	8658	491	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 28.

The worst 5 of 491 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:1285:ASP:C	1:A:1285:ASP:CA	1.77	1.51
1:A:1179:SER:HB2	2:A:1372:HOH:O	1.29	1.28
1:D:1162:VAL:HG22	2:D:1375:HOH:O	1.28	1.27
1:D:1261:TYR:OH	1:D:1263:GLY:HA2	1.11	1.27
1:D:1261:TYR:CZ	1:D:1263:GLY:HA2	1.78	1.19

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	274/277 (99%)	260 (95%)	12 (4%)	2 (1%)	22	12
1	B	274/277 (99%)	260 (95%)	12 (4%)	2 (1%)	22	12

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	274/277 (99%)	260 (95%)	12 (4%)	2 (1%)	22	12
1	D	274/277 (99%)	261 (95%)	11 (4%)	2 (1%)	22	12
All	All	1096/1108 (99%)	1041 (95%)	47 (4%)	8 (1%)	22	12

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1204	GLU
1	B	1204	GLU
1	C	1204	GLU
1	D	1204	GLU
1	D	1263	GLY

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	229/230 (100%)	202 (88%)	27 (12%)	5	2
1	B	229/230 (100%)	203 (89%)	26 (11%)	5	2
1	C	229/230 (100%)	202 (88%)	27 (12%)	5	2
1	D	229/230 (100%)	203 (89%)	26 (11%)	5	2
All	All	916/920 (100%)	810 (88%)	106 (12%)	5	2

5 of 106 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	1055	GLN
1	C	1204	GLU
1	D	1215	LYS
1	C	1084	ARG
1	C	1148	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 55 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1013	GLN
1	C	1170	ASN
1	D	1281	GLN
1	D	1163	ASN
1	C	1033	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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(Å²)	Q<0.9
1	A	276/277 (99%)	0.88	21 (7%)	13	15	9, 19, 40, 58	0
1	B	276/277 (99%)	0.76	19 (6%)	16	19	8, 17, 37, 57	0
1	C	276/277 (99%)	0.75	19 (6%)	16	19	7, 17, 39, 61	0
1	D	276/277 (99%)	0.86	23 (8%)	11	13	8, 18, 38, 52	0
All	All	1104/1108 (99%)	0.81	82 (7%)	14	16	7, 18, 39, 61	0

The worst 5 of 82 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	1261	TYR	10.3
1	D	1010	CYS	10.2
1	B	1261	TYR	9.3
1	A	1261	TYR	8.4
1	B	1010	CYS	8.4

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

6.5   Other polymers ⓘ

There are no such residues in this entry.