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

wwPDB X-ray Structure Validation Summary Report (i)

Sep 11, 2023 – 10:34 AM EDT

|		
PDB ID	:	4MVL
Title	:	Crystal structure of an engineered lipocalin (Anticalin H1GA) in complex with the Alzheimer amyloid peptide Abeta1-40
Authors	:	Eichinger, A.; Skerra, A.
Deposited on	:	2013-09-24
Resolution	:	2.30 Å(reported)

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 2.30 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.
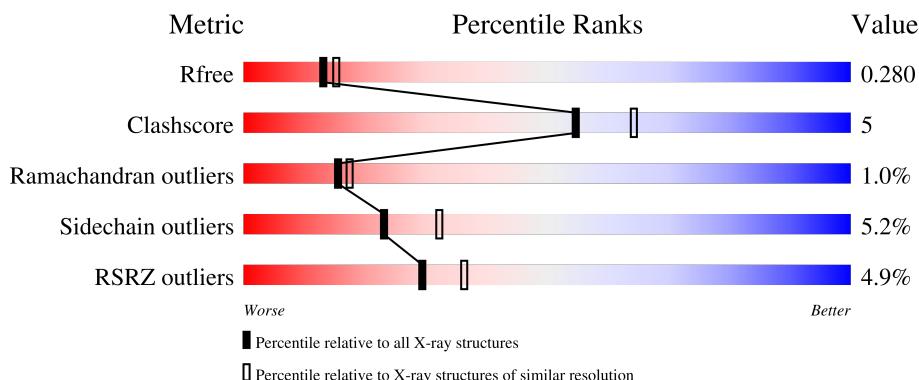
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 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	188	4% / 73% / 17% / · / 8%
1	B	188	4% / 77% / 13% / · / 9%
1	C	188	7% / 78% / 13% / · / 7%
1	D	188	5% / 73% / 16% / ·· / 9%
2	E	40	28% / 72%
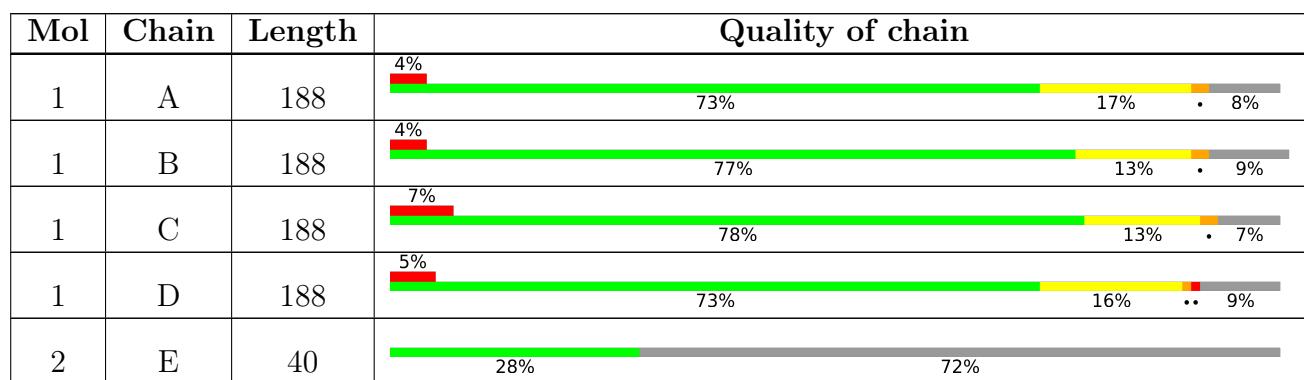
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	40	25% · 72%
2	G	40	20% · 78%
2	H	40	22% 78%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 5993 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 Neutrophil gelatinase-associated lipocalin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	173	Total	C	N	O	S	0	0	0
			1394	897	234	259	4			
1	B	172	Total	C	N	O	S	0	0	0
			1388	894	233	257	4			
1	C	174	Total	C	N	O	S	0	0	0
			1398	899	235	260	4			
1	D	171	Total	C	N	O	S	0	0	0
			1380	890	232	254	4			

There are 116 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	28	HIS	GLN	engineered mutation	UNP P80188
A	36	ALA	LEU	engineered mutation	UNP P80188
A	40	VAL	ALA	engineered mutation	UNP P80188
A	41	LEU	ILE	engineered mutation	UNP P80188
A	49	LEU	GLN	engineered mutation	UNP P80188
A	70	GLY	LEU	engineered mutation	UNP P80188
A	72	ASP	ARG	engineered mutation	UNP P80188
A	73	ASP	LYS	engineered mutation	UNP P80188
A	77	LEU	ASP	engineered mutation	UNP P80188
A	79	LYS	TRP	engineered mutation	UNP P80188
A	87	SER	CYS	engineered mutation	UNP P80188
A	96	ARG	ASN	engineered mutation	UNP P80188
A	100	GLU	TYR	engineered mutation	UNP P80188
A	103	GLY	LEU	engineered mutation	UNP P80188
A	106	TRP	TYR	engineered mutation	UNP P80188
A	125	GLU	LYS	engineered mutation	UNP P80188
A	127	ALA	SER	engineered mutation	UNP P80188
A	132	THR	TYR	engineered mutation	UNP P80188
A	134	ASN	LYS	engineered mutation	UNP P80188
A	179	SER	-	expression tag	UNP P80188
A	180	ALA	-	expression tag	UNP P80188

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	181	TRP	-	expression tag	UNP P80188
A	182	SER	-	expression tag	UNP P80188
A	183	HIS	-	expression tag	UNP P80188
A	184	PRO	-	expression tag	UNP P80188
A	185	GLN	-	expression tag	UNP P80188
A	186	PHE	-	expression tag	UNP P80188
A	187	GLU	-	expression tag	UNP P80188
A	188	LYS	-	expression tag	UNP P80188
B	28	HIS	GLN	engineered mutation	UNP P80188
B	36	ALA	LEU	engineered mutation	UNP P80188
B	40	VAL	ALA	engineered mutation	UNP P80188
B	41	LEU	ILE	engineered mutation	UNP P80188
B	49	LEU	GLN	engineered mutation	UNP P80188
B	70	GLY	LEU	engineered mutation	UNP P80188
B	72	ASP	ARG	engineered mutation	UNP P80188
B	73	ASP	LYS	engineered mutation	UNP P80188
B	77	LEU	ASP	engineered mutation	UNP P80188
B	79	LYS	TRP	engineered mutation	UNP P80188
B	87	SER	CYS	engineered mutation	UNP P80188
B	96	ARG	ASN	engineered mutation	UNP P80188
B	100	GLU	TYR	engineered mutation	UNP P80188
B	103	GLY	LEU	engineered mutation	UNP P80188
B	106	TRP	TYR	engineered mutation	UNP P80188
B	125	GLU	LYS	engineered mutation	UNP P80188
B	127	ALA	SER	engineered mutation	UNP P80188
B	132	THR	TYR	engineered mutation	UNP P80188
B	134	ASN	LYS	engineered mutation	UNP P80188
B	179	SER	-	expression tag	UNP P80188
B	180	ALA	-	expression tag	UNP P80188
B	181	TRP	-	expression tag	UNP P80188
B	182	SER	-	expression tag	UNP P80188
B	183	HIS	-	expression tag	UNP P80188
B	184	PRO	-	expression tag	UNP P80188
B	185	GLN	-	expression tag	UNP P80188
B	186	PHE	-	expression tag	UNP P80188
B	187	GLU	-	expression tag	UNP P80188
B	188	LYS	-	expression tag	UNP P80188
C	28	HIS	GLN	engineered mutation	UNP P80188
C	36	ALA	LEU	engineered mutation	UNP P80188
C	40	VAL	ALA	engineered mutation	UNP P80188
C	41	LEU	ILE	engineered mutation	UNP P80188
C	49	LEU	GLN	engineered mutation	UNP P80188

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	70	GLY	LEU	engineered mutation	UNP P80188
C	72	ASP	ARG	engineered mutation	UNP P80188
C	73	ASP	LYS	engineered mutation	UNP P80188
C	77	LEU	ASP	engineered mutation	UNP P80188
C	79	LYS	TRP	engineered mutation	UNP P80188
C	87	SER	CYS	engineered mutation	UNP P80188
C	96	ARG	ASN	engineered mutation	UNP P80188
C	100	GLU	TYR	engineered mutation	UNP P80188
C	103	GLY	LEU	engineered mutation	UNP P80188
C	106	TRP	TYR	engineered mutation	UNP P80188
C	125	GLU	LYS	engineered mutation	UNP P80188
C	127	ALA	SER	engineered mutation	UNP P80188
C	132	THR	TYR	engineered mutation	UNP P80188
C	134	ASN	LYS	engineered mutation	UNP P80188
C	179	SER	-	expression tag	UNP P80188
C	180	ALA	-	expression tag	UNP P80188
C	181	TRP	-	expression tag	UNP P80188
C	182	SER	-	expression tag	UNP P80188
C	183	HIS	-	expression tag	UNP P80188
C	184	PRO	-	expression tag	UNP P80188
C	185	GLN	-	expression tag	UNP P80188
C	186	PHE	-	expression tag	UNP P80188
C	187	GLU	-	expression tag	UNP P80188
C	188	LYS	-	expression tag	UNP P80188
D	28	HIS	GLN	engineered mutation	UNP P80188
D	36	ALA	LEU	engineered mutation	UNP P80188
D	40	VAL	ALA	engineered mutation	UNP P80188
D	41	LEU	ILE	engineered mutation	UNP P80188
D	49	LEU	GLN	engineered mutation	UNP P80188
D	70	GLY	LEU	engineered mutation	UNP P80188
D	72	ASP	ARG	engineered mutation	UNP P80188
D	73	ASP	LYS	engineered mutation	UNP P80188
D	77	LEU	ASP	engineered mutation	UNP P80188
D	79	LYS	TRP	engineered mutation	UNP P80188
D	87	SER	CYS	engineered mutation	UNP P80188
D	96	ARG	ASN	engineered mutation	UNP P80188
D	100	GLU	TYR	engineered mutation	UNP P80188
D	103	GLY	LEU	engineered mutation	UNP P80188
D	106	TRP	TYR	engineered mutation	UNP P80188
D	125	GLU	LYS	engineered mutation	UNP P80188
D	127	ALA	SER	engineered mutation	UNP P80188
D	132	THR	TYR	engineered mutation	UNP P80188

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	134	ASN	LYS	engineered mutation	UNP P80188
D	179	SER	-	expression tag	UNP P80188
D	180	ALA	-	expression tag	UNP P80188
D	181	TRP	-	expression tag	UNP P80188
D	182	SER	-	expression tag	UNP P80188
D	183	HIS	-	expression tag	UNP P80188
D	184	PRO	-	expression tag	UNP P80188
D	185	GLN	-	expression tag	UNP P80188
D	186	PHE	-	expression tag	UNP P80188
D	187	GLU	-	expression tag	UNP P80188
D	188	LYS	-	expression tag	UNP P80188

- Molecule 2 is a protein called Beta-amyloid protein 40.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	11	Total 84	C 55	N 12	O 17	0	0	0
2	F	11	Total 84	C 55	N 12	O 17	0	0	0
2	G	9	Total 68	C 45	N 9	O 14	0	0	0
2	H	9	Total 68	C 45	N 9	O 14	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	33	Total 33	O 33	0	0
3	B	26	Total 26	O 26	0	0
3	C	28	Total 28	O 28	0	0
3	D	34	Total 34	O 34	0	0
3	G	1	Total 1	O 1	0	0
3	H	7	Total 7	O 7	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: Neutrophil gelatinase-associated lipocalin
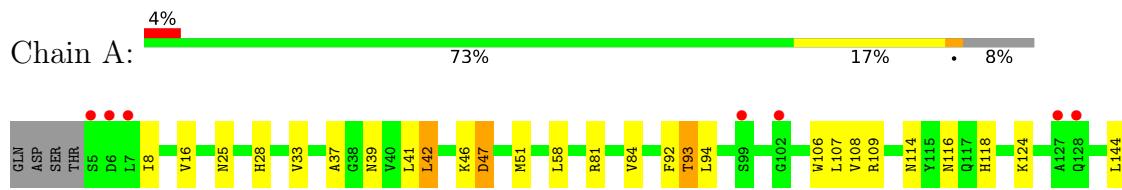
• Molecule 1: Neutrophil gelatinase-associated lipocalin

• Molecule 1: Neutrophil gelatinase-associated lipocalin

• Molecule 1: Neutrophil gelatinase-associated lipocalin

`V126 A127 Q128 Y138 G139 R140 R154 V167 I172 D173 Q174 C175 I176 D177 GLY SER ALA TRP SER HIS PRO GLN PHE GLU LYS`
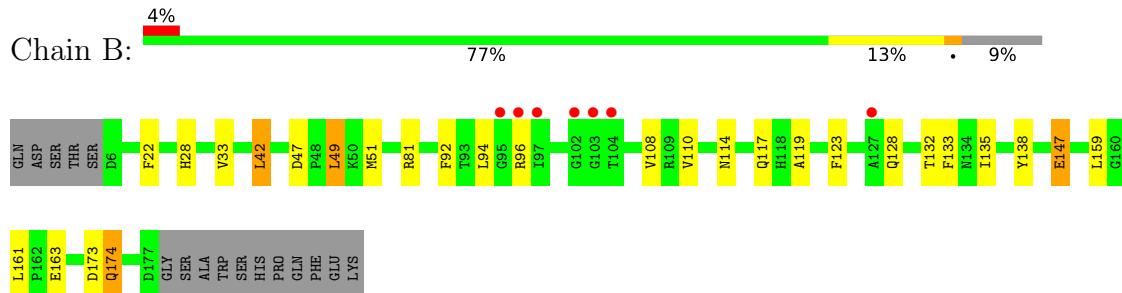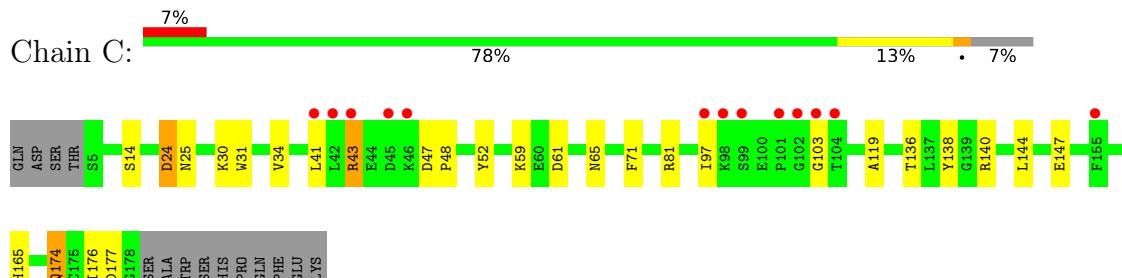
• Molecule 2: Beta-amyloid protein 40

Chain E: 28%   72%

`ASP ALA GLU PHE ARG HIS ASP SER GLY TYR GLU VAL HIS HIS GLN LYS L17 N27 LYS GLY ALA ILE ILE GLY LEU MET VAL GLY GLY VAL VAL`

• Molecule 2: Beta-amyloid protein 40

Chain F: 25%   ·   72%

`ASP ALA GLU PHE ARG HIS ASP SER GLY TYR GLU VAL HIS HIS GLN LYS L17 E22 N27 LYS GLY ALA ILE ILE GLY LEU MET VAL GLY GLY VAL VAL`

• Molecule 2: Beta-amyloid protein 40

Chain G: 20%   ·   78%

`ASP ALA GLU PHE ARG HIS ASP SER GLY TYR GLU VAL HIS HIS GLN LYS LEU V18 A21 S26 ASN LYS GLY ALA ILE ILE GLY LEU MET VAL GLY GLY VAL VAL`

• Molecule 2: Beta-amyloid protein 40

Chain H: 22%   78%

`ASP ALA GLU PHE ARG HIS ASP SER GLY TYR GLU VAL HIS HIS GLN LYS LEU V18 S26 ASN LYS GLY ALA ILE ILE GLY LEU MET VAL GLY GLY VAL VAL`
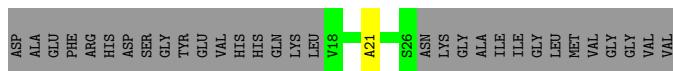
4   Data and refinement statistics   (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	79.30Å 79.18Å 142.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	71.30 – 2.30 39.65 – 2.30	Depositor EDS
% Data completeness (in resolution range)	97.6 (71.30-2.30) 97.7 (39.65-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	4.88 (at 2.29Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.219 , 0.279 0.219 , 0.280	Depositor DCC
R_{free} test set	1998 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	26.3	Xtriage
Anisotropy	0.077	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 13.2	EDS
L-test for twinning[2]	$< |L| > = 0.37, < L^2 > = 0.20$	Xtriage
Estimated twinning fraction	0.457 for k,h,-l	Xtriage
Reported twinning fraction	0.516 for H, K, L 0.484 for K, H, -L	Depositor
Outliers	0 of 39910 reflections	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	5993	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.44% 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 ⓘ

5.1 Standard geometry ⓘ

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	0.44	1/1429 (0.1%)	0.53	0/1936
1	B	0.44	0/1423	0.57	0/1928
1	C	0.42	0/1433	0.52	0/1941
1	D	0.44	1/1415 (0.1%)	0.53	0/1917
2	E	0.39	0/85	0.61	0/114
2	F	0.41	0/85	0.59	0/114
2	G	0.44	0/69	0.51	0/92
2	H	0.40	0/69	0.54	0/92
All	All	0.43	2/6008 (0.0%)	0.54	0/8134

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	31	TRP	CD2-CE2	5.16	1.47	1.41
1	A	106	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 ⓘ

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	1394	0	1382	16	0
1	B	1388	0	1377	15	0
1	C	1398	0	1385	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1380	0	1373	15	0
2	E	84	0	75	0	0
2	F	84	0	75	1	0
2	G	68	0	58	1	0
2	H	68	0	58	0	0
3	A	33	0	0	0	0
3	B	26	0	0	1	0
3	C	28	0	0	0	0
3	D	34	0	0	0	0
3	G	1	0	0	0	0
3	H	7	0	0	0	0
All	All	5993	0	5783	61	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 5.

The worst 5 of 61 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:D:42:LEU:HD21	1:D:167:VAL:HG21	1.72	0.71
1:C:97:ILE:HG21	1:C:103:GLY:HA3	1.76	0.68
1:B:173:ASP:O	1:B:174:GLN:HB2	1.96	0.66
1:B:114:ASN:HD21	1:B:117:GLN:HG2	1.65	0.61
1:D:57:GLU:HB3	1:D:65:ASN: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	171/188 (91%)	153 (90%)	17 (10%)	1 (1%)	25	31

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	170/188 (90%)	156 (92%)	12 (7%)	2 (1%)	13	14
1	C	172/188 (92%)	157 (91%)	13 (8%)	2 (1%)	13	14
1	D	169/188 (90%)	154 (91%)	13 (8%)	2 (1%)	13	14
2	E	9/40 (22%)	6 (67%)	3 (33%)	0	100	100
2	F	9/40 (22%)	6 (67%)	3 (33%)	0	100	100
2	G	7/40 (18%)	5 (71%)	2 (29%)	0	100	100
2	H	7/40 (18%)	5 (71%)	2 (29%)	0	100	100
All	All	714/912 (78%)	642 (90%)	65 (9%)	7 (1%)	15	17

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	47	ASP
1	D	97	ILE
1	C	174	GLN
1	D	128	GLN
1	B	128	GLN

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	156/169 (92%)	148 (95%)	8 (5%)	24	33
1	B	155/169 (92%)	146 (94%)	9 (6%)	20	27
1	C	156/169 (92%)	151 (97%)	5 (3%)	39	54
1	D	154/169 (91%)	142 (92%)	12 (8%)	12	16
2	E	9/31 (29%)	9 (100%)	0	100	100
2	F	9/31 (29%)	9 (100%)	0	100	100
2	G	7/31 (23%)	7 (100%)	0	100	100
2	H	7/31 (23%)	7 (100%)	0	100	100
All	All	653/800 (82%)	619 (95%)	34 (5%)	23	32

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	94	LEU
1	D	97	ILE
1	D	154	ARG
1	B	81	ARG
1	B	49	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	26	GLN
1	A	116	ASN
1	A	174	GLN
1	C	88	GLN
1	D	26	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 (i)

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(Å^2)	Q<0.9
1	A	173/188 (92%)	0.01	7 (4%)	38	45	15, 20, 43, 51	0
1	B	172/188 (91%)	0.13	7 (4%)	37	44	15, 24, 52, 67	0
1	C	174/188 (92%)	0.30	13 (7%)	14	19	17, 26, 52, 59	0
1	D	171/188 (90%)	0.09	9 (5%)	26	33	16, 23, 51, 61	0
2	E	11/40 (27%)	-0.15	0	100	100	18, 19, 25, 29	0
2	F	11/40 (27%)	0.06	0	100	100	23, 25, 37, 40	0
2	G	9/40 (22%)	0.00	0	100	100	27, 28, 34, 34	0
2	H	9/40 (22%)	-0.31	0	100	100	20, 21, 22, 22	0
All	All	730/912 (80%)	0.12	36 (4%)	29	36	15, 23, 51, 67	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	42	LEU	6.7
1	A	99	SER	6.5
1	D	97	ILE	5.2
1	D	101	PRO	5.0
1	C	103	GLY	4.9

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.

6.5 Other polymers ⓘ

There are no such residues in this entry.