



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 11:19 PM GMT

PDB ID : 1HC1
Title : CRYSTAL STRUCTURE OF HEXAMERIC HAEMOCYANIN FROM PAN-
ULIRUS INTERRUPTUS REFINED AT 3.2 ANGSTROMS RESOLUTION
Authors : Volbeda, A.; Hol, W.G.J.
Deposited on : 1991-05-15
Resolution : 3.20 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

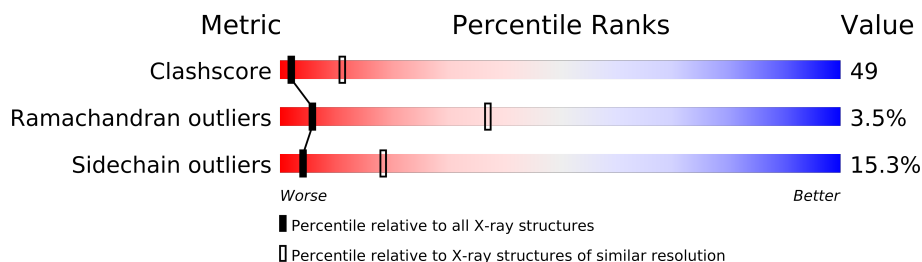
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	21963
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 3.20 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))
Clashscore	79885	1078 (3.26-3.14)
Ramachandran outliers	78287	1059 (3.26-3.14)
Sidechain outliers	78261	1058 (3.26-3.14)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	657	
1	B	657	
1	C	657	
1	D	657	
1	E	657	
1	F	657	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 32166 atoms, of which 0 are hydrogen and 0 are deuterium.

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 ARTHROPODAN HEMOCYANIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	634	Total	C	N	O	S	0	0	0
			5173	3283	892	977	21			
1	B	634	Total	C	N	O	S	0	0	0
			5173	3283	892	977	21			
1	C	634	Total	C	N	O	S	0	0	0
			5173	3283	892	977	21			
1	D	634	Total	C	N	O	S	0	0	0
			5173	3283	892	977	21			
1	E	634	Total	C	N	O	S	0	0	0
			5173	3283	892	977	21			
1	F	634	Total	C	N	O	S	0	0	0
			5173	3283	892	977	21			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	32	ASP	GLU	CONFLICT	UNP P04254
A	163	PRO	GLN	CONFLICT	UNP P04254
A	458	ASN	LYS	CONFLICT	UNP P04254
A	514	SER	LYS	CONFLICT	UNP P04254
B	32	ASP	GLU	CONFLICT	UNP P04254
B	163	PRO	GLN	CONFLICT	UNP P04254
B	458	ASN	LYS	CONFLICT	UNP P04254
B	514	SER	LYS	CONFLICT	UNP P04254
C	32	ASP	GLU	CONFLICT	UNP P04254
C	163	PRO	GLN	CONFLICT	UNP P04254
C	458	ASN	LYS	CONFLICT	UNP P04254
C	514	SER	LYS	CONFLICT	UNP P04254
D	32	ASP	GLU	CONFLICT	UNP P04254
D	163	PRO	GLN	CONFLICT	UNP P04254
D	458	ASN	LYS	CONFLICT	UNP P04254
D	514	SER	LYS	CONFLICT	UNP P04254
E	32	ASP	GLU	CONFLICT	UNP P04254

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	163	PRO	GLN	CONFLICT	UNP P04254
E	458	ASN	LYS	CONFLICT	UNP P04254
E	514	SER	LYS	CONFLICT	UNP P04254
F	32	ASP	GLU	CONFLICT	UNP P04254
F	163	PRO	GLN	CONFLICT	UNP P04254
F	458	ASN	LYS	CONFLICT	UNP P04254
F	514	SER	LYS	CONFLICT	UNP P04254

- Molecule 2 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	D	2	Total Cu 2 2	0	0
2	E	2	Total Cu 2 2	0	0
2	B	2	Total Cu 2 2	0	0
2	C	2	Total Cu 2 2	0	0
2	A	2	Total Cu 2 2	0	0
2	F	2	Total Cu 2 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	F	186	Total O 186 186	0	0
3	E	186	Total O 186 186	0	0
3	D	186	Total O 186 186	0	0
3	C	186	Total O 186 186	0	0
3	B	186	Total O 186 186	0	0
3	A	186	Total O 186 186	0	0

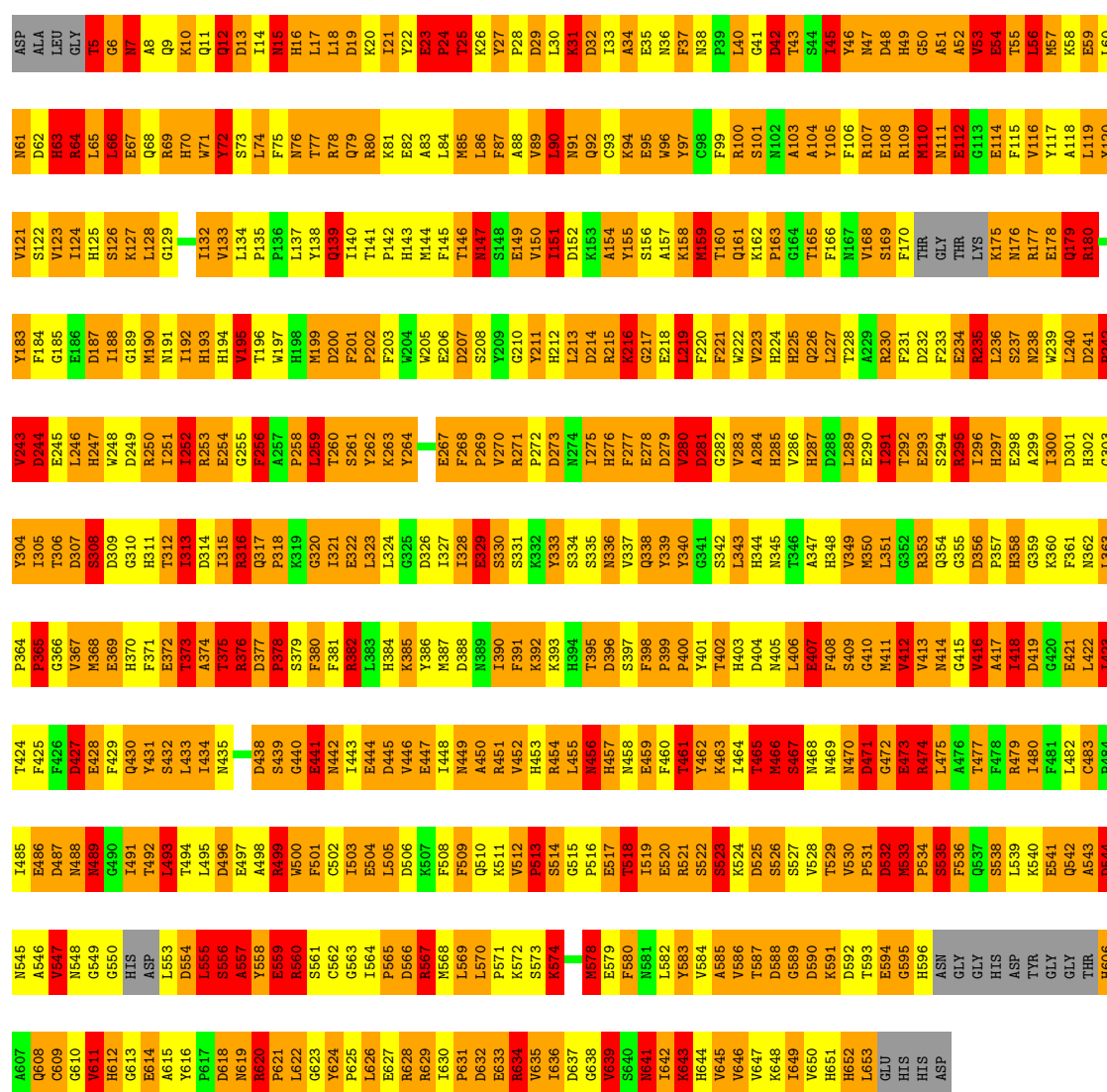
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Note EDS was not executed.

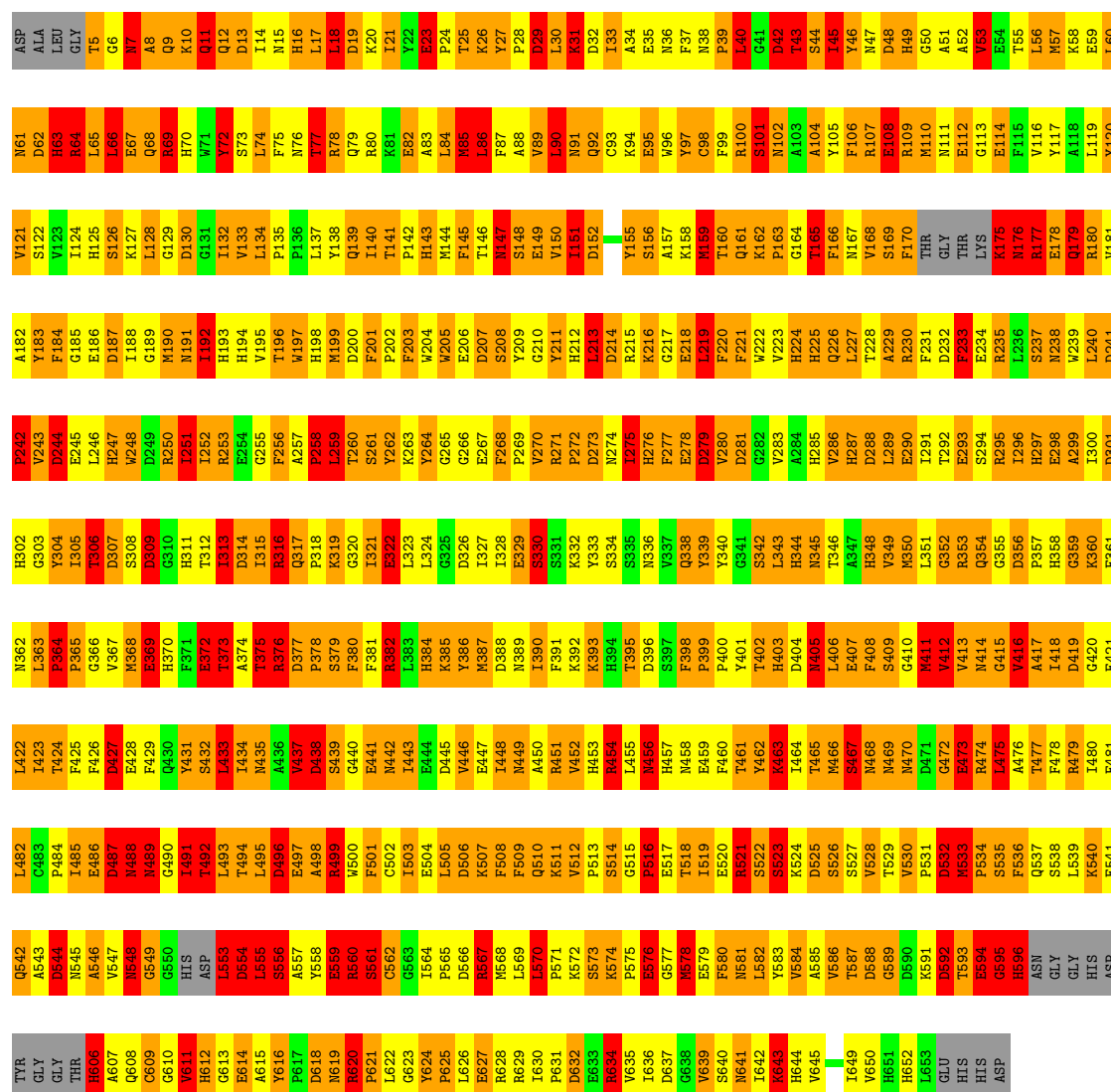
• Molecule 1: ARTHROPODAN HEMOCYANIN

Chain A:



• Molecule 1: ARTHROPODAN HEMOCYANIN

Chain B:



Chain E:

M619	HIS	H470	N405	G341	R271	F203	P135	R69	ASP
R620	ASP	D471	L406	S342	P272	W204	Y138	Y72	ALA
P621	L553	R474	E407	L343	D273	W205	Q139	L74	LEU
L622	D554	L475	F408	H344	H274	E206	I140	S73	GLY
G623	L555	A476	S409	N345	I275	D207		L74	
Y624	S556							F75	G6
P625	A557	T477	G410	H348	E278	G210	H143	N76	R7
R626	Y558	F478	M411	V349	D279	Y211	T77	T77	A8
E627	Y559	R479	V412	H350	V280	H212	F145	R78	Q9
R628	R560	L480	V413	L351	D281	L213		Q79	K10
R629		F481	N414	G352	G282	D214	S148	R80	Q11
L630	P565	L482	G415	R353	V283	R215	E149	K81	
P631	D566		V416	Q354	A284	R216	V150	E82	M15
D632	R567	L485		G355	H285	G217	L151	A83	H16
E633	M568	E486		D356		E218	D152	L84	L17
R634	L569	D487		P357	D288	L219	K153	M85	L18
V635	L570	N488	L422	R358	E290	F220	A154	L86	
L636		N489	I423	G359	E290	F221	Y155	F87	Y22
D637	S573	G490	I424			W222	S156	A88	
G638	K574	L491	F425	F361	E293	V223	A157	V89	T25
V639	P575	T492	F426	N362	S294		K158	L90	K26
S640	E576	L493	D427	L363	R295	Q226	M159	N91	Y27
M641	G577	T494	E428	P364		L227	T160	Q92	P28
L642	M578	L495	F429	P365	I300		Q161		D29
R643	E579	D496	Q430	G366	D301	R230	K162	E95	L30
H644	F580	A497	Y431	V367	H302	F232	T165	W96	K31
		R499	S432	M368	G303	D231		Y97	
L649	Y583	W500	L433	E369	Y304	F233	V168	C98	I33
		F501	D438	H370	T305	E234	S169	F99	A34
H652	V586	F504	S439	F371	T306	R235	F170	R100	E35
L653	T587	E504	N442	E372	D307	L236	THR	S101	N36
HIS	L505		I443	T373	H311	S237	GLY	N102	F37
HIS			E444	T375	T312	W239	THR	A103	P39
ASP	K591	F508	D445	R376	I313	L240	LVS	A104	P39
	F509	F509	D445	D377	D314	D241	K175	Y105	L40
			V446	P378	L315	P242	N176	F106	G41
	V512		E447	S379	R316	R242	R177	R107	T43
	G595		L448	F380	Q317	D244	E178	E108	S44
	H596		N449	F381	P318		Q179	R109	
	ASN	P516	A450	R382	K319	H247		M110	I45
	GLY		R451	L383	G320	W248	Y183	M111	Y46
	GLY	L519	V452	H384	I321	D249	F184	E112	N47
	HIS	E520	H453	K385	E322	R250		V116	D48
	ASP	R521	R453	Y386	L323	I251	E186		H49
	TTR	S522	R454	L324	L324	D252	D187	L119	
	GLY	S523	L455	M387	L324	G255	I188	Y120	V53
	GLY	K524	N456	D388	G325		G189	V121	
	THR	D525	H457	N389	D326	G256			L56
	H606		N458	I390	I327	F256	M190		M57
	A607	D532	E459	F391	I328	L259	N191	I124	K58
	Q608	M533	F460	K392	E329	T260	I192	H125	E59
	C609	P534	T461	K393	S330	S261		S126	L60
	G610		Y462	H394			V195	K127	N61
	V611		K463	T395	Y333		T196	L128	D62
	H612		L464	D396		Y264	I197	G129	H63
	G613	A543	T465	S397	N336		H198	D130	R64
	E614		M466	F398	V337	E267	M199	G131	L65
		V547	S467	P399	Q338	F268	D200	I132	L66
	P617		N468	P399	Y339	F269	V203	E67	
	D618		N469	T402	Y240	V620	F205	I134	O68

• Molecule 1: ARTHROPODAN HEMOCYANIN

Chain F:



ASP	ALA	LEU	GLY	T5	G6	R7	A8	Q9	K10	Q11	Q12		M15	H16	L17	L18	D19	K20	I21	Y22	E23	P24	T25	K26	Y27	P28	D29	L30	K31	D32	I33	A34	E35	N36	F37	N38	P39	L40	L40	G41	D42	T43	S44	L45	Y46	M47	D48	H49		V53	L56	M57	K58	E59	L60	N61	D62	R63	H64		
L65	L66	E67	D68	R69		Y72	S73	L74	F75	N76	T77		R78	Q79	R80	K81	E82	K83	L84	H85	L86	F87	A88	V89	L90	N91	Q92		E95	W96		F99	R100	S101	N102	A103	A104	Y105	F106	R107	E108	R109	M110	N111	E112		Y116		L119	Y120	V121	I124	H125	K126	L127	L128	N129	D130	L132	V133	
L134	P135		Y138	Q139	I140	T141	F142	H143	M144	F145		S148	E149	V150	I151	D152	K153	A154	Y155	S156	A157	K158	M159	T160	Q161		T165		V168	S169	F170	THR	GLY	THR	K175	M176	R177	E178	Q179	Y183	F184	G185	E186	D187	I188	G189	M190	N191	I192	V196	L196	K197	D200	F201	P202						
F203	W204	W205	E206	D207		G210	Y211	H212	L213	D214	K215	K216	G217	E218	L219	F220	F221	W222	V223		Q226	L227	R230	F231	D232	F233	E234	R235	L236	S237	N238	W239	L240	D241	P242	V243	D244		H247	W248	D249	R250	I251	I252	G255	M256	L259	T260	S261	Y264	E267	F268	P269								
V270	R271	P272	D273	I275		E278	D279	V280	D281	G282	A284	H285		D288	L289	E290		E293	S294	R295	I296	L300	D301	H302	G303	Y304	T306	D307		H311	T312	I313	D314	I315	R316	Q317	P318	K319	G320	I321	E322	L323	L324	G325	D326	I327	I328	E329	S330	Y333	N336	V337									
Q338	Y339	Y340	G341	S342	L343	H344	N345		H348	V349	H350	L351	G352	R353	K354	G355	D356		H358	P357	G359	K360		L363	P364	G365	G366	V367	E368	E369	H370	F371	E372	T373	A374	T375	R376	D377	P378	S379	F380	F381	R382	L383	H384	K385	Y386	M387	D388	N389	I390	F391	K392	K393	H394	T395	L396	D396	S397	F398	P399



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	119.80Å 193.10Å 122.20Å 90.00° 118.10° 90.00°	Depositor
Resolution (Å)	8.00 – 3.20	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-3.20)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	0.201 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	32166	wwPDB-VP
Average B, all atoms (Å ²)	17.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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

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	1.41	11/5316 (0.2%)	4.13	1053/7205 (14.6%)
1	B	1.40	14/5316 (0.3%)	3.74	1016/7205 (14.1%)
1	C	0.75	0/5316	1.64	87/7205 (1.2%)
1	D	0.77	1/5316 (0.0%)	1.66	94/7205 (1.3%)
1	E	0.76	2/5316 (0.0%)	1.64	87/7205 (1.2%)
1	F	0.75	0/5316	1.65	94/7205 (1.3%)
All	All	1.02	28/31896 (0.1%)	2.64	2431/43230 (5.6%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	2
All	All	0	3

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	550	GLY	C-O	7.84	1.36	1.23
1	B	208	SER	CB-OG	7.79	1.52	1.42
1	A	441	GLU	CB-CG	7.42	1.66	1.52
1	A	526	SER	CB-OG	6.86	1.51	1.42
1	B	267	GLU	CD-OE2	-6.23	1.18	1.25

The worst 5 of 2431 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	64	ARG	NE-CZ-NH2	89.27	164.93	120.30
1	A	207	ASP	CB-CG-OD1	43.98	157.88	118.30
1	B	271	ARG	NE-CZ-NH1	43.24	141.92	120.30
1	A	273	ASP	CB-CG-OD1	40.24	154.52	118.30
1	B	273	ASP	CB-CG-OD1	39.23	153.61	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	295	ARG	Sidechain
1	B	177	ARG	Sidechain
1	B	521	ARG	Sidechain

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5173	0	4880	687	0
1	B	5173	0	4883	651	1
1	C	5173	0	4888	416	3
1	D	5173	0	4888	472	1
1	E	5173	0	4888	428	0
1	F	5173	0	4888	416	1
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
2	E	2	0	0	0	0
2	F	2	0	0	0	0
3	A	186	0	0	39	0
3	B	186	0	0	15	0
3	C	186	0	0	11	0
3	D	186	0	0	11	0
3	E	186	0	0	10	0
3	F	186	0	0	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	32166	0	29315	2986	3

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 49.

The worst 5 of 2986 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:161:GLN:OE1	1:B:443:ILE:HD13	1.28	1.29
1:A:422:LEU:CD2	1:A:570:LEU:HD21	1.66	1.23
1:A:316:ARG:HD3	3:A:829:HOH:O	1.41	1.19
1:B:456:ASN:HD22	1:B:457:HIS:N	1.42	1.17
1:A:165:THR:CG2	1:A:449:ASN:HB2	1.73	1.17

All (3) 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	Distance(Å)	Clash(Å)
1:B:309:ASP:OD1	1:C:49:HIS:CD2[2.647]	1.85	0.35
1:C:594:GLU:OE1	1:F:471:ASP:CB[2.657]	2.13	0.07
1:C:474:ARG:NH2	1:D:41:GLY:O[2.656]	2.17	0.03

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	626/657 (95%)	501 (80%)	97 (16%)	28 (4%)	4	29
1	B	626/657 (95%)	506 (81%)	89 (14%)	31 (5%)	3	26
1	C	626/657 (95%)	517 (83%)	95 (15%)	14 (2%)	10	53
1	D	626/657 (95%)	516 (82%)	91 (14%)	19 (3%)	7	42
1	E	626/657 (95%)	512 (82%)	95 (15%)	19 (3%)	7	42
1	F	626/657 (95%)	503 (80%)	102 (16%)	21 (3%)	6	38
All	All	3756/3942 (95%)	3055 (81%)	569 (15%)	132 (4%)	6	37

5 of 132 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	42	ASP
1	A	147	ASN
1	A	176	ASN
1	A	441	GLU
1	A	471	ASP

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	564/580 (97%)	438 (78%)	126 (22%)	1	6
1	B	564/580 (97%)	450 (80%)	114 (20%)	2	9
1	C	564/580 (97%)	494 (88%)	70 (12%)	7	30
1	D	564/580 (97%)	489 (87%)	75 (13%)	6	27
1	E	564/580 (97%)	498 (88%)	66 (12%)	8	34
1	F	564/580 (97%)	498 (88%)	66 (12%)	8	34
All	All	3384/3480 (97%)	2867 (85%)	517 (15%)	4	18

5 of 517 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	596	HIS
1	C	461	THR
1	F	313	ILE
1	C	23	GLU
1	C	279	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 158 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	405	ASN
1	D	147	ASN
1	F	338	GLN
1	C	435	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	11	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 12 ligands modelled in this entry, 12 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.

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 ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.