



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 11:24 PM GMT

PDB ID : 1HCY
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>

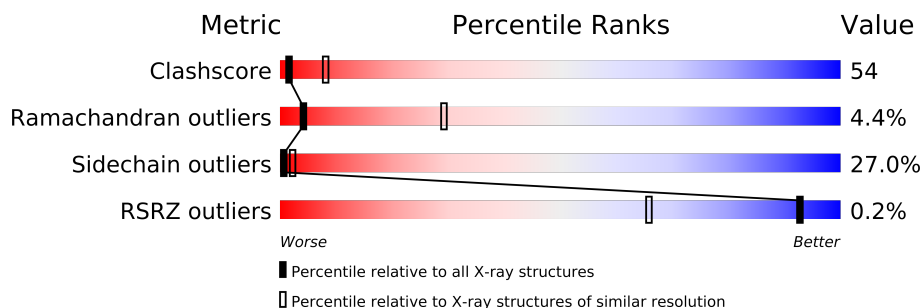
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
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)
RSRZ outliers	66119	1825 (3.30-3.10)

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.

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 4 unique types of molecules in this entry. The entry contains 31790 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	644	Total	C	N	O	S	0	0	0
			5239	3322	904	992	21			
1	B	644	Total	C	N	O	S	0	0	0
			5239	3322	904	992	21			
1	C	644	Total	C	N	O	S	0	0	0
			5239	3322	904	992	21			
1	D	644	Total	C	N	O	S	0	0	0
			5239	3322	904	992	21			
1	E	644	Total	C	N	O	S	0	0	0
			5239	3322	904	992	21			
1	F	644	Total	C	N	O	S	0	0	0
			5239	3322	904	992	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 a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	2	Total	C	N	O	0	0
			28	16	2	10		
2	B	2	Total	C	N	O	0	0
			28	16	2	10		
2	C	2	Total	C	N	O	0	0
			28	16	2	10		
2	D	2	Total	C	N	O	0	0
			28	16	2	10		
2	E	2	Total	C	N	O	0	0
			28	16	2	10		
2	F	2	Total	C	N	O	0	0
			28	16	2	10		

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

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	514	SER	LYS	CONFLICT	UNP P04254
E	32	ASP	GLU	CONFLICT	UNP P04254
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 3 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total Cu 2 2	0	0

- Molecule 4 is water.

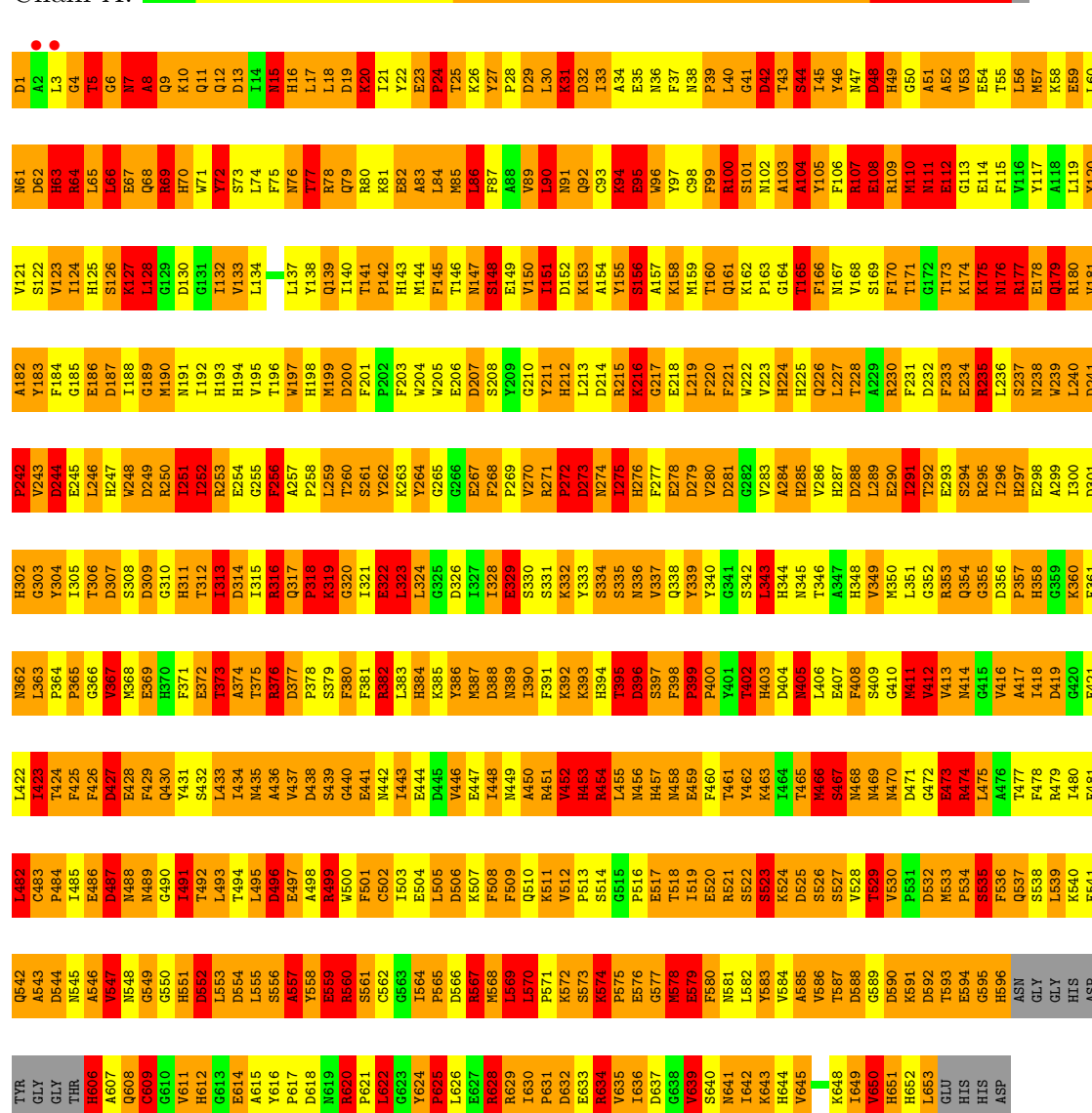
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	176	Total O 176 176	0	0
4	B	4	Total O 4 4	0	0
4	C	2	Total O 2 2	0	0
4	E	3	Total O 3 3	0	0
4	F	1	Total O 1 1	0	0

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.

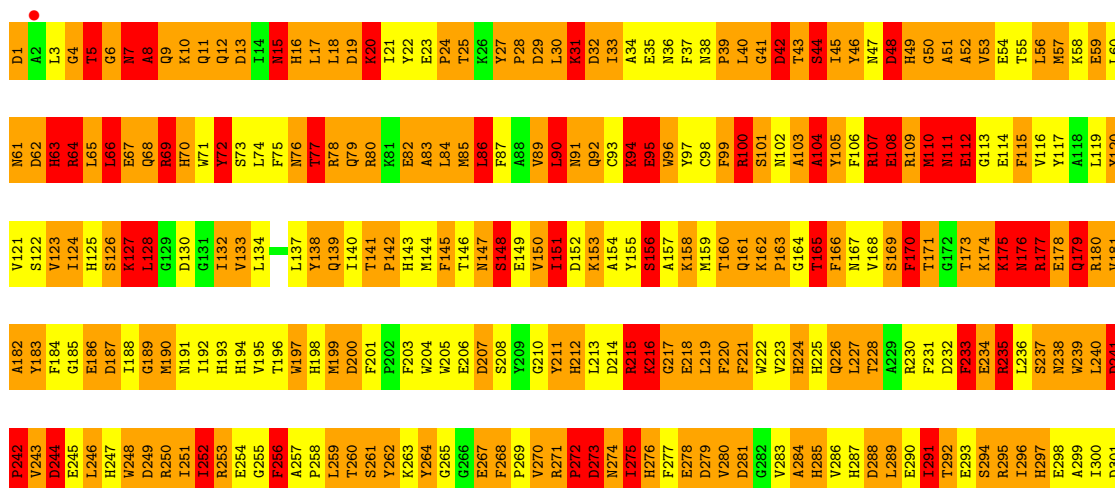
• Molecule 1: ARTHROPODAN HEMOCYANIN

Chain A:



• Molecule 1: ARTHROPODAN HEMOCYANIN

Chain B:



TYR	GLY	THR	H606	Q608	C609	G610	V611	H612	G613	E614	A615	V616	P617	D618	N619	R620	P621	L622	G623	Y624	P625	L626	E627	R628	R629	L630	P631	D632	E633	R634	V635	L636	D637	G638	R639	S640	N641	L642	K643	H644	V645
-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

• Molecule 1: ARTHROPODAN HEMOCYANIN

Chain E:

Q542	L482	L422	N362	H302	P242	A182	V121	N61	D1
A543	C483	L423	L363	G303	Y243	Y183	S122	D62	A2
D544	P484	T424	P364	Y304	D244	F184	S123	H63	L3
N545	L485	F425	P365	T305	E245	G185	I124	R64	G4
A546	E486	F426	G366	T306	L246	E186	H125	L65	T5
V547	D487	D427	V367	D307	H247	D187	S126	L66	G6
N548	N488	F428	N368	S308	W248	I188	K127	E67	H7
G549	A489	E429	E369	D309	D249	G189	L128	Q68	A8
G550	L490	G430	H370	G310	R250	M190	G129	R69	Q9
H551	L491	Y431	F371	H311	L251	N191	D130	H70	K10
L552	T492	S432	E372	T312	L252	I192	G131	W71	Q11
L553	L493	L433	T373	L313	R253	H193	I132	Y72	Q12
D554	T494	I434	A374	D314	E254	H194	V133	S73	D13
L495	L495	A435	T375	L315	G255	V195	L134	L74	I14
D496	A497	A436	R376	R316	F256	T196		F75	N15
E497	V437	V437	D377	Q317	A257	W197	L137	N76	H16
A498	L498	D438	P378	P318	P258	H198	Y138	T77	L17
E559	R499	S439	S379	K319	L259	M199	Q139	R78	L18
N500	W500	G440	F380	G320	T260	D200	I140	Q79	D19
F501	S561	E441	F381	I321	S261	F201	T141	R80	R20
C502	C502	N442	R382	E322	Y262	P202	P142	K81	I21
G563	L503	L443	L383	L323	K263	F203	H143	E82	Y22
E504	E504	E444	H384	L324	W264	W204	M144	A83	E23
P565	L505	V445	K385	G326	G265	W205	F145	L84	P24
D566	D506	V446	Y386	D326	G266	E206	T146	M85	T25
R567	K507	E447	K387	L327	E267	D207	N147	L86	K26
M568	F508	L448	D388	L328	F268	S208	S148	F87	Y27
L569	F509	N449	N389	E329	P269	Y209	E149	A88	P28
L570	Q510	A450	L390	S330	V270	G210	V150	V89	D29
P571	K511	R451	F391	S331	R271	Y211	I151	L80	L30
K572	V512	L462	K392	K332	P272	H212	D152	N91	K31
S573	P513	H453	K393	Y333	D273	L213	K153	Q92	D32
S514	G514	R454	H394	S334	N274	D214	A154	C93	L33
P575	G515	L455	T395	S335	L275	R215	Y155	K94	A34
E576	P516	N456	D396	N336	H276	K216	S156	E95	E35
E577	P517	H457	S397	V337	F277	G217	A157	W96	N36
T578	T518	N458	F398	Q338	E278	E218	K158	Y97	F37
E579	T519	E459	P399	Y339	D279	L219	M159	C98	N38
F580	E520	F460	P400	Y340	V280	F220	T160	F99	P39
N581	R521	T461	Y401	G341	D281	F221	K161	R100	L40
L582	S522	Y462	T402	S342	G282	W222	K162	S101	G41
F583	S523	K463	H403	L343	V283	V223	P163	N102	D42
V584	K524	L464	D404	H344	A284	H224	G164	A103	T43
E585	D525	T465	N405	N345	E285	H225	T165	A104	S44
V586	S526	H466	L406	T346	V286	Q226	F166	Y105	L45
S527	T527	D467	E407	A347	H287	L227	P167	F106	Y46
D588	V528	N468	F408	H348	D288	T228	V168	R107	N47
G589	T529	N469	S409	V349	L289	G229	S169	E108	D48
V530	V530	D470	G410	M350	E290	R230	F170	R109	H49
P531	F531	D471	N411	L351	T291	F231	T171	M110	G50
D592	D592	G472	V412	G352	T292	D32	G172	N111	A51
T593	T593	E473	V413	R353	E293	F233	K173	E112	A52
E594	P534	R474	N414	Q354	S294	E234	K174	G113	V53
G595	S535	L475	G415	G355	R295	R235	K175	E114	E54
H596	F536	A476	V416	D356	L296	L236	R176	F115	T55
ASN	Q537	T477	A417	P357	H297	S237	T177	Y116	L56
GLY	S538	F478	L418	H358	E298	E178	R177	V116	L57
GLY	L539	R479	D419	G359	A299	W239	Q179	A117	M57
H15	K540	L480	G420	K360	T300	L240	R180	L119	E59
ASP	E541	F481	E421	F361	R201	D240	V181	Y100	L60

TYR	GLY	THR	H606	Q608	C609	G610	V611	H612	G613	E614	A615	V616	P617	D618	N619	R620	P621	L622	G623	Y624	P625	L626	E627	R628	R629	L630	P631	D632	E633	R634	V635	L636	D637	G638	R639	S640	N641	L642	K643	H644	V645
-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

• Molecule 1: ARTHROPODAN HEMOCYANIN

Chain F:

D1	A2	L3	G4	T5	G6	H7	A8	Q9	K10	Q11	Q12	D13	T14	N15	H16	T17	L18	D19	K20	I21	Y22	E23	P24	M25	L26	Y27	R28	Q29	C30	G31	N32	Y33	L34	D35	K36	L37	T38	P39	R40	L41	D42	N43	T44	A45	S46	L47	H48	G49	T50	M51	E52	G53	E54	T55	L56	Y57	A58	R59	L60
N61	D62	R63	H64	L65	L66	E67	Q68	R69	H70	W71	Y72	S73	L74	F75	N76	T77	R78	Q79	R80	K81	E82	A83	P84	M85	L86	Y87	R88	Q89	C90	G91	N92	Y93	L94	D95	K96	L97	T98	P99	R100	L101	D102	N103	T104	A105	S106	L107	H108	G109	T110	M111	E112	G113	E114	T115	L116	Y117	A118	R119	L120
V121	S122	V123	H124	H125	S126	K127	L128	G129	D130	G131	I132	V133	L134	L135	L137	Y138	Q139	T140	T141	P142	H143	M144	F145	T146	N147	S148	E149	V150	I151	D152	K153	A154	V155	S156	A157	K158	M159	T160	Q161	K162	P163	G164	T165	F166	N167	V168	S169	F170	T171	G172	T173	K174	R175	N176	R177	E178	Q179	R180	V181

TYR	Q542	L482	L422	N362	H302	P242	A182
GLY	A543	C483	I423	L363	G303	V243	Y183
GLY	D544	P484	T424	P364	Y304	D244	F184
THR	N545	I485	F425	P365	I305	E245	G185
H606	A546	E486	F426	G366	T306	L246	E186
A607	V547	D487	D427	V367	D307	H247	L187
Q608	N548	N488	E428	M368	S308	V248	I188
C609	G549	N489	F429	E369	D309	D249	G189
G610	G550	G490	Q430	H370	G310	R250	M190
V611	H551	I491	Y431	F371	H311	I251	N191
H612	D552	T492	S432	E372	T312	L252	I192
G613	L553	L493	L433	T373	T313	R253	H193
D614	D554	T494	T434	A374	D314	E254	H194
A615	L555	L495	M435	T375	I315	G255	V195
V616	S556	D496	A436	R376	R316	P256	T196
P617	A557	E497	V437	D377	Q317	A257	W197
D618	Y558	A498	D438	P378	P318	P258	H198
H619	E559	R499	S439	S379	K319	L259	M199
R620	R560	W500	G440	F380	G320	T260	D200
P621	S561	F501	E441	F381	I321	S261	F201
L622	C562	C502	M442	R382	E322	Y262	F202
G623	G563	I503	T443	L383	L323	K263	F203
Y624	I564	E504	E444	H384	L324	Y264	W204
P625	P565	L505	D445	K385	G325	G265	W205
L626	D566	D506	V446	Y386	D326	G266	E206
E627	R567	K507	E447	M387	I327	E267	D207
R628	M568	F508	T448	D388	I328	P268	S208
R629	L569	F509	M449	N389	E329	F269	S209
I630	L570	Q510	A450	I390	S330	V270	G210
P631	P571	K511	R451	F391	S331	R271	Y211
D632	K572	V512	V452	K392	K332	P272	H212
E633	S573	P513	R453	K393	Y333	D273	L213
R634	K574	S514	R454	H394	S334	M274	D214
V635	P575	G515	L455	T395	S335	L275	R215
I636	E576	P516	M456	D396	K336	H276	K216
D637	G577	E517	H457	S397	V337	G277	G217
G638	N578	T518	N458	F398	Q338	E278	E218
R639	E579	I519	E459	P399	Y339	D279	L219
S640	F580	E520	F460	P400	T340	V280	F220
N641	N581	R521	T461	Y401	G341	D281	F221
I642	L582	S522	Y462	T402	S342	G282	W222
K643	Y583	S523	K463	H403	L343	V283	V223
H644	V584	K524	L464	D404	H344	A284	H224
V645	A585	D525	T465	M405	N345	H285	H225
	V586	S526	M466	L406	T346	V286	Q226
K648	T587	S527	S467	E407	A347	H287	L227
I649	D588	V528	M468	F408	H348	D288	T228
V650	G589	T529	N469	S409	V349	L289	L229
H651	D590	V530	N470	G410	M350	E290	A229
H652	K591	P531	D471	M411	L351	I291	F231
L653	D592	D532	G472	V412	G352	T292	D232
GLU	T593	M533	E473	V413	R353	E293	F233
HIS	E594	P534	R474	M414	Q354	S294	E234
HIS	G595	S535	L475	G415	G355	R295	R235
ASP	H596	F536	A476	V416	D356	I296	L236
	ASN	Q537	T477	A417	P357	H297	S237
	GLY	S538	F478	I418	H358	E298	N238
	GLY	L539	R479	D419	G359	A299	W239
	HIS	K540	I480	G420	I360	L300	L240
	ASP	E541	F481	E421	F361	D301	F241

4 Data and refinement statistics

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 8.00 – 3.16	Depositor EDS
% Data completeness (in resolution range)	(Not available) (8.00-3.20) 75.5 (8.00-3.16)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$	-	Xtriage
Refinement program	CORELS, PROLSQ	Depositor
R, R_{free}	0.221 , (Not available) 0.218 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	42.7	Xtriage
Anisotropy	0.361	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.23 , 0.6	EDS
Estimated twinning fraction	0.095 for -h-l,k,h 0.095 for l,k,-h-l 0.105 for h,-k,-h-l 0.097 for -h-l,-k,l 0.104 for l,-k,h	Xtriage
L-test for twinning	$\langle L \rangle = 0.36$, $\langle L^2 \rangle = 0.19$	Xtriage
Outliers	0 of 59193 reflections	Xtriage
F_o, F_c correlation	0.84	EDS
Total number of atoms	31790	wwPDB-VP
Average B, all atoms (Å ²)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.12% of the height of the origin peak. No significant pseudotranslation is detected.*

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NDG, NAG, 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	1.67	39/5385 (0.7%)	3.87	1158/7301 (15.9%)
1	B	1.67	40/5385 (0.7%)	3.87	1154/7301 (15.8%)
1	C	1.67	38/5385 (0.7%)	3.87	1154/7301 (15.8%)
1	D	1.67	38/5385 (0.7%)	3.87	1151/7301 (15.8%)
1	E	1.67	39/5385 (0.7%)	3.87	1156/7301 (15.8%)
1	F	1.67	39/5385 (0.7%)	3.87	1154/7301 (15.8%)
All	All	1.67	233/32310 (0.7%)	3.87	6927/43806 (15.8%)

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	7
1	B	0	7
1	C	0	7
1	D	0	7
1	E	0	7
1	F	0	7
All	All	0	42

The worst 5 of 233 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	217	GLY	N-CA	9.44	1.60	1.46
1	D	217	GLY	N-CA	9.44	1.60	1.46
1	C	217	GLY	N-CA	9.42	1.60	1.46
1	B	217	GLY	N-CA	9.42	1.60	1.46
1	A	217	GLY	N-CA	9.41	1.60	1.46

The worst 5 of 6927 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	628	ARG	NE-CZ-NH2	-45.56	97.52	120.30
1	E	628	ARG	NE-CZ-NH2	-45.53	97.54	120.30
1	C	628	ARG	NE-CZ-NH2	-45.49	97.56	120.30
1	D	628	ARG	NE-CZ-NH2	-45.48	97.56	120.30
1	A	628	ARG	NE-CZ-NH2	-45.47	97.56	120.30

There are no chirality outliers.

5 of 42 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	100	ARG	Sidechain
1	A	177	ARG	Sidechain
1	A	215	ARG	Sidechain
1	A	454	ARG	Sidechain
1	A	499	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	5239	0	4942	566	1
1	B	5239	0	4942	549	0
1	C	5239	0	4942	561	0
1	D	5239	0	4942	580	0
1	E	5239	0	4942	558	1
1	F	5239	0	4943	539	0
2	A	28	0	25	6	0
2	B	28	0	25	6	0
2	C	28	0	25	7	0
2	D	28	0	25	7	0
2	E	28	0	25	6	0
2	F	28	0	25	6	0
3	A	2	0	0	0	0
4	A	176	0	0	31	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	4	0	0	1	0
4	C	2	0	0	0	0
4	E	3	0	0	1	0
4	F	1	0	0	0	0
All	All	31790	0	29803	3314	1

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 54.

The worst 5 of 3314 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:140:ILE:HG22	1:B:141:THR:HG23	1.20	1.18
1:A:140:ILE:HG22	1:A:141:THR:HG23	1.20	1.17
1:F:140:ILE:HG22	1:F:141:THR:HG23	1.20	1.17
1:E:140:ILE:HG22	1:E:141:THR:HG23	1.20	1.15
1:D:140:ILE:HG22	1:D:141:THR:HG23	1.20	1.15

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	Distance(Å)	Clash(Å)
1:A:419:ASP:OD2	1:E:547:VAL:O[1.556]	1.86	0.34

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	640/657 (97%)	524 (82%)	88 (14%)	28 (4%)	4	29
1	B	640/657 (97%)	523 (82%)	89 (14%)	28 (4%)	4	29
1	C	640/657 (97%)	523 (82%)	89 (14%)	28 (4%)	4	29
1	D	640/657 (97%)	523 (82%)	89 (14%)	28 (4%)	4	29
1	E	640/657 (97%)	524 (82%)	88 (14%)	28 (4%)	4	29
1	F	640/657 (97%)	523 (82%)	89 (14%)	28 (4%)	4	29

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	3840/3942 (97%)	3140 (82%)	532 (14%)	168 (4%)	4	29

5 of 168 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	7	ASN
1	A	147	ASN
1	A	176	ASN
1	A	473	GLU
1	A	552	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	570/580 (98%)	416 (73%)	154 (27%)	1	2
1	B	570/580 (98%)	416 (73%)	154 (27%)	1	2
1	C	570/580 (98%)	416 (73%)	154 (27%)	1	2
1	D	570/580 (98%)	416 (73%)	154 (27%)	1	2
1	E	570/580 (98%)	416 (73%)	154 (27%)	1	2
1	F	570/580 (98%)	416 (73%)	154 (27%)	1	2
All	All	3420/3480 (98%)	2496 (73%)	924 (27%)	1	2

5 of 924 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	474	ARG
1	D	259	LEU
1	F	375	THR
1	C	529	THR
1	D	53	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 134 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	453	HIS
1	D	225	HIS
1	F	338	GLN
1	C	469	ASN
1	D	15	ASN

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 ⓘ

12 carbohydrates 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	NDG	A	658	1,2	12,14,15	19.26	2 (16%)	15,19,21	10.39	11 (73%)
2	NAG	A	659	2	12,14,15	22.96	3 (25%)	15,19,21	10.18	4 (26%)
2	NDG	B	658	1,2	12,14,15	19.26	2 (16%)	15,19,21	10.39	11 (73%)
2	NAG	B	659	2	12,14,15	22.96	3 (25%)	15,19,21	10.18	4 (26%)
2	NDG	C	658	1,2	12,14,15	19.26	2 (16%)	15,19,21	10.39	11 (73%)
2	NAG	C	659	2	12,14,15	22.97	3 (25%)	15,19,21	10.18	4 (26%)
2	NDG	D	658	1,2	12,14,15	19.25	2 (16%)	15,19,21	10.39	11 (73%)
2	NAG	D	659	2	12,14,15	22.96	3 (25%)	15,19,21	10.18	4 (26%)
2	NDG	E	658	1,2	12,14,15	19.26	2 (16%)	15,19,21	10.39	11 (73%)
2	NAG	E	659	2	12,14,15	22.96	3 (25%)	15,19,21	10.18	4 (26%)
2	NDG	F	658	1,2	12,14,15	19.27	2 (16%)	15,19,21	10.39	11 (73%)
2	NAG	F	659	2	12,14,15	22.95	3 (25%)	15,19,21	10.18	4 (26%)

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	NDG	A	658	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	659	2	-	0/6/23/26	0/1/1/1
2	NDG	B	658	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	659	2	-	0/6/23/26	0/1/1/1
2	NDG	C	658	1,2	-	0/6/23/26	0/1/1/1
2	NAG	C	659	2	-	0/6/23/26	0/1/1/1
2	NDG	D	658	1,2	-	0/6/23/26	0/1/1/1
2	NAG	D	659	2	-	0/6/23/26	0/1/1/1
2	NDG	E	658	1,2	-	0/6/23/26	0/1/1/1
2	NAG	E	659	2	-	0/6/23/26	0/1/1/1
2	NDG	F	658	1,2	-	0/6/23/26	0/1/1/1
2	NAG	F	659	2	-	0/6/23/26	0/1/1/1

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	659	NAG	C8-C7	77.81	3.14	1.50
2	B	659	NAG	C8-C7	77.78	3.14	1.50
2	D	659	NAG	C8-C7	77.78	3.14	1.50
2	E	659	NAG	C8-C7	77.77	3.14	1.50
2	A	659	NAG	C8-C7	77.76	3.14	1.50

The worst 5 of 90 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	659	NAG	C8-C7-N2	-33.15	51.30	116.11
2	D	659	NAG	C8-C7-N2	-33.14	51.33	116.11
2	A	659	NAG	C8-C7-N2	-33.14	51.33	116.11
2	F	659	NAG	C8-C7-N2	-33.13	51.34	116.11
2	E	659	NAG	C8-C7-N2	-33.13	51.34	116.11

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.6 Ligand geometry ⓘ

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.

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(Å ²)	Q<0.9
1	A	644/657 (98%)	-0.43	2 (0%) 91 58	2, 11, 53, 85	0
1	B	644/657 (98%)	-0.41	1 (0%) 93 66	2, 11, 53, 85	0
1	C	644/657 (98%)	-0.39	1 (0%) 93 66	2, 11, 53, 85	0
1	D	644/657 (98%)	-0.38	0 100 100	2, 11, 53, 85	0
1	E	644/657 (98%)	-0.38	0 100 100	2, 11, 53, 85	0
1	F	644/657 (98%)	-0.35	1 (0%) 93 66	2, 11, 53, 85	0
All	All	3864/3942 (98%)	-0.39	5 (0%) 93 74	2, 11, 54, 85	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2	ALA	2.8
1	F	2	ALA	2.3
1	A	3	LEU	2.2
1	B	2	ALA	2.0
1	C	2	ALA	2.0

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 ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	NDG	C	658	14/15	0.33	19.58	52,57,68,70	0
2	NDG	E	658	14/15	0.36	10.17	52,57,68,70	0
2	NDG	A	658	14/15	0.35	8.76	52,57,68,70	0
2	NDG	B	658	14/15	0.34	7.72	52,57,68,70	0
2	NDG	D	658	14/15	0.35	7.53	52,57,68,70	0
2	NDG	F	658	14/15	0.32	5.16	52,57,68,70	0
2	NAG	B	659	14/15	0.39	-	73,82,87,89	0
2	NAG	A	659	14/15	0.43	-	73,82,87,89	0
2	NAG	F	659	14/15	0.47	-	73,82,87,89	0
2	NAG	E	659	14/15	0.45	-	73,82,87,89	0
2	NAG	D	659	14/15	0.60	-	73,82,87,89	0
2	NAG	C	659	14/15	0.66	-	73,82,87,89	0

6.4 Ligands

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	CU	A	665	1/1	0.06	-2.18	9,9,9,9	0
3	CU	A	666	1/1	0.06	-2.28	10,10,10,10	0

6.5 Other polymers

There are no such residues in this entry.