Full wwPDB X-ray Structure Validation Report (i)

Aug 2, 2023 – 11:57 PM EDT

PDB ID	:	1HZF
Title	:	C4ADG FRAGMENT OF HUMAN COMPLEMENT FACTOR C4A
Authors	:	van den Elsen, J.M.H.; Martin, A.; Wong, V.; Clemenza, L.; Rose, D.R.; Isenman, D.E.
Deposited on	:	2001-01-24
Resolution	:	2.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i)) were used in the production of this report:

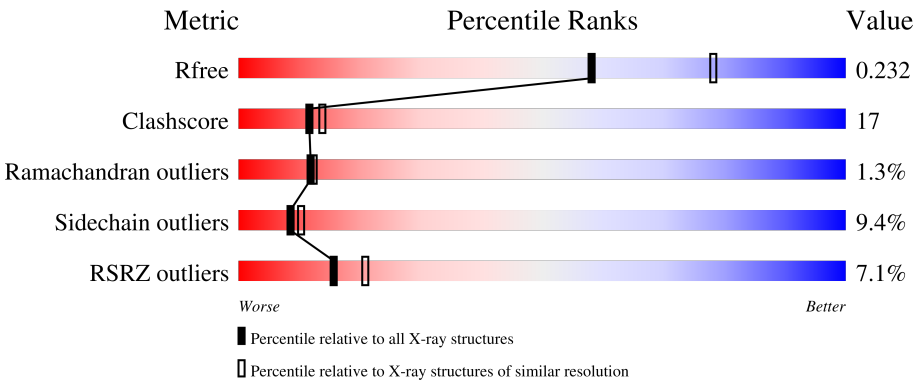
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.34
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

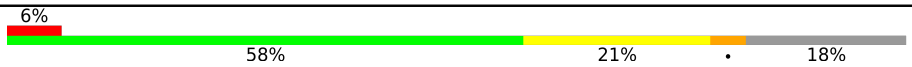
The reported resolution of this entry is 2.30 Å.

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	367	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2363 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 FACTOR C4A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	302	2286	1444	394	440	3	5	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	949	MSE	MET	modified residue	UNP P01028
A	996	MSE	MET	modified residue	UNP P01028
A	1109	MSE	MET	modified residue	UNP P01028
A	1157	SER	ASN	SEE REMRAK 999	UNP P01028
A	1182	THR	SER	SEE REMARK 999	UNP P01028
A	1188	ALA	VAL	SEE REMARK 999	UNP P01028
A	1191	ARG	LEU	SEE REMARK 999	UNP P01028
A	1199	MSE	MET	modified residue	UNP P01028
A	1201	MSE	MET	modified residue	UNP P01028
A	1235	MSE	MET	modified residue	UNP P01028
A	1262	MSE	MET	modified residue	UNP P01028

- Molecule 2 is water.

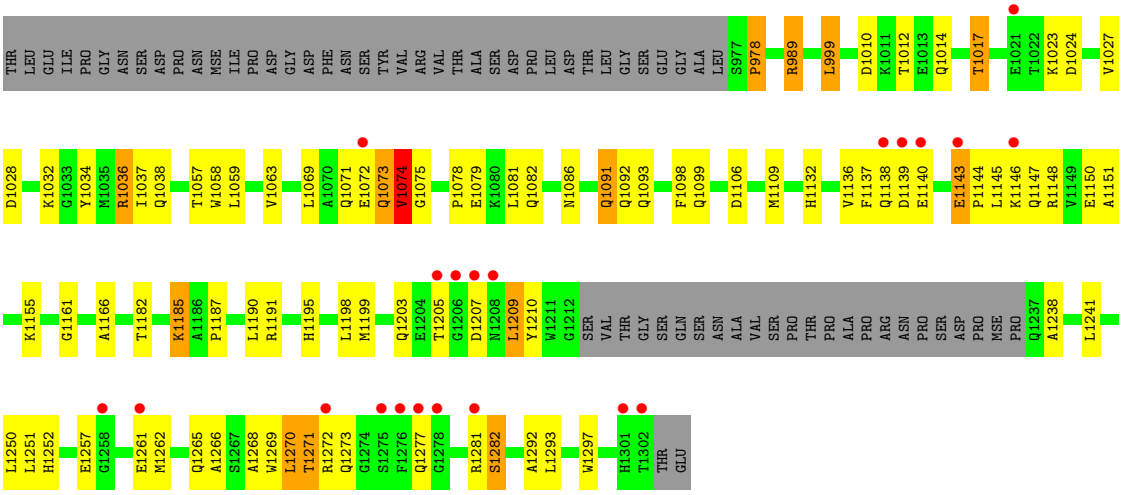
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	77	77	77	0	0

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 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 FACTOR C4A

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	56.51Å 71.43Å 85.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.84 – 2.30 26.84 – 2.30	Depositor EDS
% Data completeness (in resolution range)	93.0 (26.84-2.30) 93.0 (26.84-2.30)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.20 (at 2.31Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.215 , 0.233 0.214 , 0.232	Depositor DCC
R_{free} test set	724 reflections (4.62%)	wwPDB-VP
Wilson B-factor (Å^2)	34.9	Xtriage
Anisotropy	0.046	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 44.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	2363	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.17% 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	0.40	0/2330	0.53	0/3162

There are no bond length outliers.

There are no bond angle outliers.

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	2286	0	2235	79	0
2	A	77	0	0	10	0
All	All	2363	0	2235	79	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 17.

All (79) 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:1082:GLN:HE22	1:A:1138:GLN:HG2	1.31	0.96
1:A:1270:LEU:HD21	1:A:1293:LEU:HD21	1.52	0.91
1:A:1058:TRP:HZ3	1:A:1109:MSE:HE2	1.42	0.84
1:A:989:ARG:HD2	2:A:5:HOH:O	1.77	0.83

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1281:ARG:O	1:A:1282:SER:HB2	1.80	0.79
1:A:1140:GLU:HA	1:A:1143:GLU:OE2	1.81	0.79
1:A:1078:PRO:O	1:A:1082:GLN:HG3	1.84	0.78
1:A:1270:LEU:HD23	1:A:1271:THR:N	2.00	0.75
1:A:1199:MSE:HE2	1:A:1262:MSE:SE	2.40	0.72
1:A:1082:GLN:NE2	1:A:1138:GLN:HG2	2.06	0.70
1:A:1136:VAL:O	1:A:1136:VAL:HG12	1.94	0.68
1:A:1143:GLU:N	1:A:1144:PRO:HD2	2.07	0.68
1:A:1014:GLN:O	1:A:1017:THR:HG23	1.94	0.68
1:A:1058:TRP:CZ3	1:A:1109:MSE:HE2	2.28	0.64
1:A:1185:LYS:HA	1:A:1185:LYS:NZ	2.11	0.64
1:A:1082:GLN:HB2	2:A:23:HOH:O	2.00	0.60
1:A:1268:ALA:O	1:A:1272:ARG:HG3	2.01	0.60
1:A:1099:GLN:HB3	2:A:20:HOH:O	2.01	0.60
1:A:989:ARG:HA	1:A:1036:ARG:HE	1.67	0.60
1:A:1082:GLN:HE21	1:A:1137:PHE:HA	1.68	0.59
1:A:1073:GLN:O	1:A:1074:VAL:HG23	2.03	0.58
1:A:1261:GLU:O	1:A:1265:GLN:HG3	2.03	0.58
1:A:1057:THR:OG1	1:A:1091:GLN:HG2	2.04	0.57
1:A:1261:GLU:C	1:A:1265:GLN:HE21	2.08	0.56
1:A:1082:GLN:NE2	1:A:1138:GLN:H	2.02	0.56
1:A:1270:LEU:CD2	1:A:1293:LEU:HD21	2.31	0.56
1:A:1106:ASP:O	1:A:1109:MSE:HG2	2.05	0.56
1:A:1207:ASP:O	1:A:1207:ASP:OD1	2.24	0.55
1:A:1012:THR:OG1	1:A:1014:GLN:HG3	2.08	0.54
1:A:1195:HIS:CD2	1:A:1252:HIS:HE1	2.26	0.54
1:A:1270:LEU:HD23	1:A:1271:THR:H	1.70	0.54
1:A:1269:TRP:O	1:A:1273:GLN:HG2	2.09	0.53
1:A:1139:ASP:O	1:A:1140:GLU:HB3	2.09	0.53
1:A:1195:HIS:HD2	1:A:1252:HIS:CE1	2.26	0.53
1:A:1185:LYS:HA	1:A:1185:LYS:HZ3	1.74	0.52
1:A:1034:TYR:O	1:A:1037:ILE:HG22	2.09	0.52
1:A:1034:TYR:O	1:A:1038:GLN:HG2	2.09	0.52
1:A:1059:LEU:O	1:A:1063:VAL:HG23	2.10	0.51
1:A:1081:LEU:HD12	1:A:1136:VAL:HG11	1.94	0.50
1:A:1191:ARG:NH2	2:A:44:HOH:O	2.45	0.50
1:A:1195:HIS:HD2	1:A:1252:HIS:HE1	1.58	0.50
1:A:1146:LYS:O	1:A:1150:GLU:HG3	2.12	0.50
1:A:1151:ALA:O	1:A:1155:LYS:HG3	2.12	0.49
1:A:1199:MSE:CE	1:A:1262:MSE:SE	3.09	0.49
1:A:1209:LEU:HD11	1:A:1262:MSE:HG2	1.92	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1028:ASP:OD1	1:A:1032:LYS:HE3	2.12	0.49
1:A:1093:GLN:CG	2:A:20:HOH:O	2.60	0.49
1:A:1143:GLU:O	1:A:1147:GLN:HG3	2.13	0.49
1:A:1023:LYS:O	1:A:1027:VAL:HG23	2.14	0.48
1:A:1182:THR:OG1	1:A:1252:HIS:HD2	1.96	0.48
1:A:1069:LEU:HD21	1:A:1132:HIS:HD2	1.79	0.48
1:A:1182:THR:HG23	2:A:29:HOH:O	2.14	0.47
1:A:1036:ARG:HG2	1:A:1036:ARG:HH11	1.78	0.47
1:A:1209:LEU:CD1	1:A:1262:MSE:HG2	2.44	0.47
1:A:999:LEU:HD23	1:A:999:LEU:O	2.14	0.47
1:A:1143:GLU:N	1:A:1144:PRO:CD	2.76	0.46
1:A:1086:ASN:HD21	1:A:1148:ARG:HH12	1.63	0.46
1:A:1036:ARG:HH11	1:A:1036:ARG:CG	2.28	0.46
1:A:1251:LEU:HD21	1:A:1292:ALA:HA	1.98	0.46
1:A:1250:LEU:HD23	1:A:1292:ALA:HB1	1.98	0.45
1:A:1261:GLU:HB3	1:A:1265:GLN:NE2	2.31	0.45
1:A:1069:LEU:HD21	1:A:1132:HIS:CD2	2.52	0.44
1:A:1266:ALA:O	1:A:1270:LEU:HD22	2.17	0.44
1:A:1161:GLY:HA3	1:A:1190:LEU:HD22	2.00	0.44
1:A:1281:ARG:O	1:A:1282:SER:CB	2.56	0.44
1:A:1072:GLU:C	1:A:1073:GLN:HG3	2.38	0.44
1:A:1187:PRO:HD2	1:A:1190:LEU:HD12	2.00	0.43
1:A:1182:THR:OG1	1:A:1252:HIS:CD2	2.71	0.43
1:A:1024:ASP:CG	2:A:25:HOH:O	2.57	0.43
1:A:1038:GLN:HG3	2:A:1:HOH:O	2.19	0.43
1:A:1205:THR:OG1	1:A:1210:TYR:HE2	2.02	0.42
1:A:1270:LEU:HD23	1:A:1270:LEU:C	2.40	0.42
1:A:1238:ALA:HB3	1:A:1273:GLN:OE1	2.18	0.42
1:A:1092:GLN:HB2	1:A:1098:PHE:CE1	2.54	0.42
1:A:1136:VAL:O	1:A:1136:VAL:CG1	2.65	0.42
1:A:999:LEU:C	1:A:999:LEU:CD2	2.88	0.41
1:A:1257:GLU:HA	2:A:24:HOH:O	2.20	0.41
1:A:1166:ALA:HB3	2:A:42:HOH:O	2.20	0.40
1:A:1036:ARG:CG	1:A:1036:ARG:NH1	2.85	0.40

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	298/367 (81%)	277 (93%)	17 (6%)	4 (1%)	12 12

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	978	PRO
1	A	1074	VAL
1	A	1075	GLY
1	A	1282	SER

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	233/284 (82%)	211 (91%)	22 (9%)	8 10

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	978	PRO
1	A	989	ARG
1	A	999	LEU
1	A	1010	ASP
1	A	1017	THR
1	A	1036	ARG
1	A	1071	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1073	GLN
1	A	1074	VAL
1	A	1079	GLU
1	A	1091	GLN
1	A	1143	GLU
1	A	1145	LEU
1	A	1185	LYS
1	A	1198	LEU
1	A	1203	GLN
1	A	1209	LEU
1	A	1241	LEU
1	A	1270	LEU
1	A	1271	THR
1	A	1277	GLN
1	A	1297	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1031	GLN
1	A	1039	GLN
1	A	1071	GLN
1	A	1082	GLN
1	A	1086	ASN
1	A	1091	GLN
1	A	1116	ASN
1	A	1131	HIS
1	A	1195	HIS
1	A	1252	HIS
1	A	1256	HIS
1	A	1265	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)

There are no ligands in this entry.

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)

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	297/367 (80%)	0.20	21 (7%) 16 21	21, 39, 63, 71	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1302	THR	6.6
1	A	1276	PHE	4.4
1	A	1277	GLN	4.3
1	A	1301	HIS	3.7
1	A	1140	GLU	3.6
1	A	1208	ASN	3.4
1	A	1138	GLN	3.1
1	A	1205	THR	3.0
1	A	1272	ARG	2.8
1	A	1261	GLU	2.8
1	A	1139	ASP	2.6
1	A	1206	GLY	2.6
1	A	1072	GLU	2.5
1	A	1143	GLU	2.4
1	A	1281	ARG	2.4
1	A	1258	GLY	2.4
1	A	1146	LYS	2.3
1	A	1207	ASP	2.3
1	A	1278	GLY	2.3
1	A	1275	SER	2.1
1	A	1021	GLU	2.1

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.