



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 01:24 am GMT

PDB ID : 2A73  
Title : Human Complement Component C3  
Authors : Janssen, B.J.C.; Huizinga, E.G.; Raaijmakers, H.C.A.; Roos, A.; Daha, M.R.;  
Nilsson-Ekdahl, K.; Nilsson, B.; Gros, P.  
Deposited on : 2005-07-04  
Resolution : 3.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

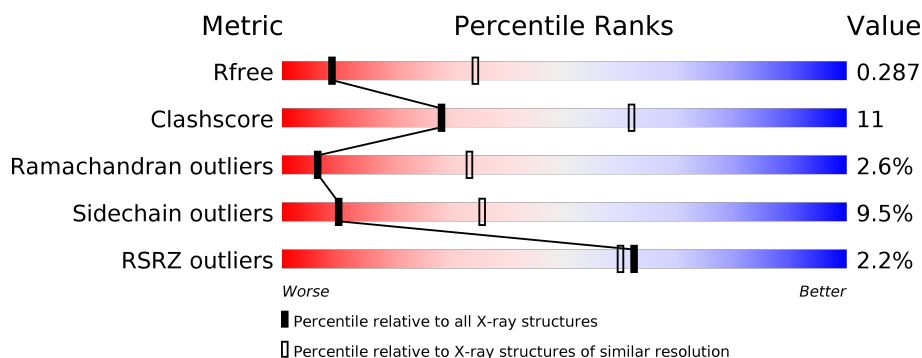
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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))
$R_{free}$	100719	1034 (3.36-3.24)
Clashscore	112137	1100 (3.36-3.24)
Ramachandran outliers	110173	1081 (3.36-3.24)
Sidechain outliers	110143	1080 (3.36-3.24)
RSRZ outliers	101464	1039 (3.36-3.24)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$ . 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	643	<div> <div>4%</div> <div> <div></div> <div>73%</div> <div>22%</div> <div>••</div> </div> </div>
2	B	991	<div> <div>%</div> <div> <div></div> <div>62%</div> <div>29%</div> <div>7%</div> <div>••</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	MAN	B	3	X	-	-	-
3	MAN	B	5	X	-	-	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 12860 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	635	Total	C	N	O	S	0	0	0
			4950	3153	837	945	15			

- Molecule 2 is a protein called Complement C3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	976	Total	C	N	O	S	0	0	0
			7821	4943	1331	1501	46			

- Molecule 3 is a polymer of unknown type called SUGAR (5-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	2	Total	C	N	O	0	0
			28	16	2	10		

These plots are drawn for all protein, RNA and DNA 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 ( $\text{RSRZ} > 2$ ). 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.

Chain A:

73%

22%

Chain B:

62%

29%

7%

••

Legend: SER, ASN, L729, D730, E731, D732, I733, I734, E735, E736, I739, R742, F745, P746, L750, N752, W753, E754, D755, P760, K685, Q686, R687, R688, T689, E696, K700, V701, D704, C705, C706, I709, T710, E711, L712, R713, R714, Q715, H716, A717, R718, A719, SER, H1S, L1U, G1Y, L1U, A1A, P1C

L930	D1007	A1120	Q1255	E1350	A1438	L1532
Q931	E1008	L1121	A1256	T1351	F1439	E1544
R932	K1014	V1125	K1262	R1354	K1440	D1552
E933	L1126	L1126	D1263	P1355	Y1444	Q1555
G934	E1018	Q1130	D1266	Q1356	E1448	Q1558
K937	L1024	A1131	H1267	D1357	L1449	I1566
E938	K1029	A1132	Q1268	T1361	I1450	C1568
D939	K1033	D1133	E1269	R1371	Q1451	K1577
I940	Q1033	D1134	L1270	R1372	V1457	W1583
D944	P1040	E1137	D1273	D1373	Y1460	D1588
D947	S1041	L1143	L1276	Q1374	Y1461	G1591
Q948	S1042	P1144	L1277	D1375	W1462	E1592
V949	A1043	I1147	L1278	A1376	L1463	K1593
P950	F1044	D1152	K1284	T1377	E1464	P1594
D951	P1052	N1157	I1285	M1378	R1469	M1595
T952	T1057	I1169	W1291	S1379	P1473	L1596
E953	S1065	L1180	E1292	I1380	E1474	I1599
S954	D1074	R1197	W1291	L1381	K1475	G1602
E955	Q1076	D1200	S1293	D1382	E1476	D1603
T956	C1079	Q1204	A1294	M1385	D1477	T1604
Q961	K1083	L1205	S1295	T1387	G1478	W1605
G962	K1091	Y1206	R1298	A1390	K1479	D1613
P963	P1092	N1207	T1302	P1391	L1480	D1617
V965	D1093	V1208	T1302	D1392	N1481	E1618
A966	V1095	L1215	V1310	T1393	L1483	E1619
E970	F1096	L1221	G1314	D1394	C1484	D1626
D971	A1100	D1225	K1315	D1395	R1485	E1632
D974	I1103	F1226	G1315	D1404	D1486	S1633
H980	H1104	P1227	L1320	S1408	C1489	M1634
L981	Q1105	P1228	S1321	K1409	C1496	N1641
I982	E1106	Q1237	V1322	Y1410	F1497	
V983	ILE	R1238	V1323	E1411	I1498	
T984	GLY	Y1239	T1324	L1412	Q1499	
P985	LEU	F1229	M1325	D1413	K1500	
C988	ARG	P1229	Y1326	K1414	S1501	
G989	N1113	P1228	H1327	A1415	D1502	
E990	N1114	Q1248	A1330	F1416	D1503	
Q991	N1115	A1249	D1332	S1417	E1509	
A1002	E1116	T1250	K1331	D1418	D1512	
Y1005	K1117	F1251	D1332	R1419	E1516	
L1006	M1119	M1252	C1336	N1420	P1517	
		V1253	M1337	I1423	G1518	
		F1254	K1338	I1424	V1519	
			F1339	D1427	D1520	
			L1341	K1428	K1524	
			I1345	E1433	L1527	
			P1347	D1434		
				D1435		

## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	116.98Å 156.26Å 271.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 3.30 39.12 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (40.00-3.30) 99.8 (39.12-3.30)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.58 (at 3.32Å)	Xtriage
Refinement program	REFMAC 5.2.0003	Depositor
R, $R_{free}$	0.230 , 0.289 0.228 , 0.287	Depositor DCC
$R_{free}$ test set	1882 reflections (5.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	83.5	Xtriage
Anisotropy	0.201	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 73.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	12860	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	94.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.26% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, NAG, MAN

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	0.21	0/5048	0.56	27/6859 (0.4%)
2	B	0.21	0/7973	0.59	54/10779 (0.5%)
All	All	0.21	0/13021	0.58	81/17638 (0.5%)

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
3	B	2	0

There are no bond length outliers.

The worst 5 of 81 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	602	ASP	CB-CG-OD2	5.59	123.33	118.30
2	B	674	ASP	CB-CG-OD2	5.58	123.32	118.30
2	B	730	ASP	CB-CG-OD2	5.51	123.26	118.30
2	B	1512	ASP	CB-CG-OD2	5.51	123.26	118.30
2	B	1413	ASP	CB-CG-OD2	5.49	123.25	118.30

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	B	3	MAN	C1
3	B	5	MAN	C1

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	4950	0	5012	83	0
2	B	7821	0	7743	211	0
3	B	61	0	52	4	0
4	A	28	0	25	0	0
All	All	12860	0	12832	286	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 11.

The worst 5 of 286 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:584:ASN:HB3	1:A:585:LYS:CA	1.85	1.07
1:A:584:ASN:CB	1:A:585:LYS:HA	1.84	1.06
2:B:908:LYS:HE3	2:B:1417:SER:HA	1.38	1.05
2:B:1484:CYS:HA	2:B:1489:CYS:HB2	1.37	1.04
1:A:584:ASN:HB3	1:A:585:LYS:HA	0.99	0.98

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	629/643 (98%)	564 (90%)	54 (9%)	11 (2%)	11	42
2	B	970/991 (98%)	853 (88%)	87 (9%)	30 (3%)	5	30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1599/1634 (98%)	1417 (89%)	141 (9%)	41 (3%)	6	33

5 of 41 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	203	LYS
1	A	204	GLU
1	A	424	TYR
2	B	923	ARG
2	B	930	LEU

### 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	560/567 (99%)	525 (94%)	35 (6%)	21	56
2	B	867/878 (99%)	766 (88%)	101 (12%)	6	27
All	All	1427/1445 (99%)	1291 (90%)	136 (10%)	10	35

5 of 136 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	908	LYS
2	B	1024	LEU
2	B	1503	ASP
2	B	922	VAL
2	B	980	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	679	ASN
2	B	686	GLN
2	B	1472	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	557	GLN
1	A	584	ASN

### 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 ⓘ

7 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	$\# Z  > 2$
4	NAG	A	644	1,4	14,14,15	0.54	0	15,19,21	0.60	0
4	NAG	A	645	4	14,14,15	0.51	0	15,19,21	0.60	0
3	NAG	B	1	3,2	14,14,15	0.62	0	15,19,21	1.25	2 (13%)
3	NAG	B	2	3	14,14,15	0.67	0	15,19,21	1.25	1 (6%)
3	MAN	B	3	3	11,11,12	0.83	0	13,15,17	0.70	0
3	BMA	B	4	3	11,11,12	0.63	0	13,15,17	0.71	0
3	MAN	B	5	3	11,11,12	0.63	0	13,15,17	0.60	0

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
4	NAG	A	644	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	645	4	-	2/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	B	1	3,2	-	0/6/23/26	0/1/1/1
3	NAG	B	2	3	-	0/6/23/26	0/1/1/1
3	MAN	B	3	3	1/1/4/5	0/2/19/22	0/1/1/1
3	BMA	B	4	3	-	0/2/19/22	0/1/1/1
3	MAN	B	5	3	1/1/4/5	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1	NAG	C3-C4-C5	2.41	114.47	110.22
3	B	1	NAG	C4-C3-C2	2.96	115.36	111.02
3	B	2	NAG	C4-C3-C2	4.04	116.94	111.02

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	B	3	MAN	C1
3	B	5	MAN	C1

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	645	NAG	C8-C7-N2-C2
4	A	645	NAG	O7-C7-N2-C2

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1	NAG	3	0
3	B	3	MAN	1	0
3	B	5	MAN	1	0

## 5.6 Ligand geometry

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	635/643 (98%)	0.33	27 (4%) 36 34	79, 97, 116, 123	0
2	B	976/991 (98%)	0.05	8 (0%) 86 85	71, 91, 108, 129	0
All	All	1611/1634 (98%)	0.16	35 (2%) 62 60	71, 93, 113, 129	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	739	ILE	3.7
1	A	101	VAL	3.7
1	A	20	MET	3.7
1	A	427	VAL	3.4
1	A	61	MET	3.1

### 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](#)

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, 95<sup>th</sup> 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	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	A	645	14/15	0.62	0.32	-	106,106,106,107	0
3	NAG	B	2	14/15	0.89	0.19	-	126,129,132,138	0
3	BMA	B	4	11/12	0.60	0.31	-	150,151,152,152	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MAN	B	3	11/12	0.84	0.19	-	141,144,147,149	0
3	NAG	B	1	14/15	0.90	0.22	-	109,113,116,121	0
4	NAG	A	644	14/15	0.71	0.30	-	107,108,108,109	0
3	MAN	B	5	11/12	0.68	0.33	-	147,148,148,149	0

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.