wwPDB X-ray Structure Validation Summary Report (i)

Nov 15, 2023 – 10:57 PM JST

PDB ID	:	6K3F
Title	:	Crystal Structure of beta-Arrestin 2 in Complex with CXCR7 Phosphopeptide
Authors	:	Min, K.J.; Yoon, H.J.; Lee, H.H.
Deposited on	:	2019-05-18
Resolution	:	2.30 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*
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 (1)) were used in the production of this report:

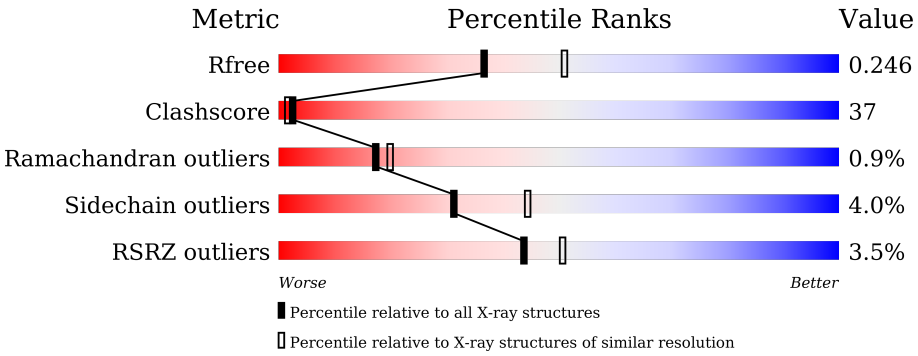
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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	377	43% 44% · 9%
1	B	377	3% 36% 51% · 10%
1	C	377	5% 37% 47% · · 10%
1	D	377	2% 43% 46% · 9%
1	E	377	4% 39% 49% · 10%
1	F	377	3% 41% 44% · 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	U	15	7% 27% 20% 40% 13%
2	V	15	13% 47% 20% 20%
2	W	15	13% 27% 13% 13% 7% 40%
2	X	15	7% 27% 33% 20% 20%
2	Y	15	7% 13% 47% 20% 20%
2	Z	15	7% 27% 13% 20% 40%

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
2	TPO	V	338	-	-	X	-
2	TPO	Z	341	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 17270 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 Beta-arrestin-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	343	Total	C	N	O	S	0	0	0
			2722	1733	481	497	11			
1	B	339	Total	C	N	O	S	0	0	0
			2680	1707	474	488	11			
1	C	338	Total	C	N	O	S	0	0	0
			2687	1711	476	489	11			
1	D	343	Total	C	N	O	S	0	0	0
			2722	1733	481	497	11			
1	E	339	Total	C	N	O	S	0	0	0
			2689	1711	474	493	11			
1	F	336	Total	C	N	O	S	0	0	0
			2676	1704	474	487	11			

There are 126 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	expression tag	UNP P29067
A	-19	GLY	-	expression tag	UNP P29067
A	-18	SER	-	expression tag	UNP P29067
A	-17	SER	-	expression tag	UNP P29067
A	-16	HIS	-	expression tag	UNP P29067
A	-15	HIS	-	expression tag	UNP P29067
A	-14	HIS	-	expression tag	UNP P29067
A	-13	HIS	-	expression tag	UNP P29067
A	-12	HIS	-	expression tag	UNP P29067
A	-11	HIS	-	expression tag	UNP P29067
A	-10	SER	-	expression tag	UNP P29067
A	-9	SER	-	expression tag	UNP P29067
A	-8	GLY	-	expression tag	UNP P29067
A	-7	LEU	-	expression tag	UNP P29067
A	-6	VAL	-	expression tag	UNP P29067
A	-5	PRO	-	expression tag	UNP P29067
A	-4	ARG	-	expression tag	UNP P29067

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP P29067
A	-2	SER	-	expression tag	UNP P29067
A	-1	HIS	-	expression tag	UNP P29067
A	0	MET	-	expression tag	UNP P29067
B	-20	MET	-	expression tag	UNP P29067
B	-19	GLY	-	expression tag	UNP P29067
B	-18	SER	-	expression tag	UNP P29067
B	-17	SER	-	expression tag	UNP P29067
B	-16	HIS	-	expression tag	UNP P29067
B	-15	HIS	-	expression tag	UNP P29067
B	-14	HIS	-	expression tag	UNP P29067
B	-13	HIS	-	expression tag	UNP P29067
B	-12	HIS	-	expression tag	UNP P29067
B	-11	HIS	-	expression tag	UNP P29067
B	-10	SER	-	expression tag	UNP P29067
B	-9	SER	-	expression tag	UNP P29067
B	-8	GLY	-	expression tag	UNP P29067
B	-7	LEU	-	expression tag	UNP P29067
B	-6	VAL	-	expression tag	UNP P29067
B	-5	PRO	-	expression tag	UNP P29067
B	-4	ARG	-	expression tag	UNP P29067
B	-3	GLY	-	expression tag	UNP P29067
B	-2	SER	-	expression tag	UNP P29067
B	-1	HIS	-	expression tag	UNP P29067
B	0	MET	-	expression tag	UNP P29067
C	-20	MET	-	expression tag	UNP P29067
C	-19	GLY	-	expression tag	UNP P29067
C	-18	SER	-	expression tag	UNP P29067
C	-17	SER	-	expression tag	UNP P29067
C	-16	HIS	-	expression tag	UNP P29067
C	-15	HIS	-	expression tag	UNP P29067
C	-14	HIS	-	expression tag	UNP P29067
C	-13	HIS	-	expression tag	UNP P29067
C	-12	HIS	-	expression tag	UNP P29067
C	-11	HIS	-	expression tag	UNP P29067
C	-10	SER	-	expression tag	UNP P29067
C	-9	SER	-	expression tag	UNP P29067
C	-8	GLY	-	expression tag	UNP P29067
C	-7	LEU	-	expression tag	UNP P29067
C	-6	VAL	-	expression tag	UNP P29067
C	-5	PRO	-	expression tag	UNP P29067
C	-4	ARG	-	expression tag	UNP P29067

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	expression tag	UNP P29067
C	-2	SER	-	expression tag	UNP P29067
C	-1	HIS	-	expression tag	UNP P29067
C	0	MET	-	expression tag	UNP P29067
D	-20	MET	-	expression tag	UNP P29067
D	-19	GLY	-	expression tag	UNP P29067
D	-18	SER	-	expression tag	UNP P29067
D	-17	SER	-	expression tag	UNP P29067
D	-16	HIS	-	expression tag	UNP P29067
D	-15	HIS	-	expression tag	UNP P29067
D	-14	HIS	-	expression tag	UNP P29067
D	-13	HIS	-	expression tag	UNP P29067
D	-12	HIS	-	expression tag	UNP P29067
D	-11	HIS	-	expression tag	UNP P29067
D	-10	SER	-	expression tag	UNP P29067
D	-9	SER	-	expression tag	UNP P29067
D	-8	GLY	-	expression tag	UNP P29067
D	-7	LEU	-	expression tag	UNP P29067
D	-6	VAL	-	expression tag	UNP P29067
D	-5	PRO	-	expression tag	UNP P29067
D	-4	ARG	-	expression tag	UNP P29067
D	-3	GLY	-	expression tag	UNP P29067
D	-2	SER	-	expression tag	UNP P29067
D	-1	HIS	-	expression tag	UNP P29067
D	0	MET	-	expression tag	UNP P29067
E	-20	MET	-	expression tag	UNP P29067
E	-19	GLY	-	expression tag	UNP P29067
E	-18	SER	-	expression tag	UNP P29067
E	-17	SER	-	expression tag	UNP P29067
E	-16	HIS	-	expression tag	UNP P29067
E	-15	HIS	-	expression tag	UNP P29067
E	-14	HIS	-	expression tag	UNP P29067
E	-13	HIS	-	expression tag	UNP P29067
E	-12	HIS	-	expression tag	UNP P29067
E	-11	HIS	-	expression tag	UNP P29067
E	-10	SER	-	expression tag	UNP P29067
E	-9	SER	-	expression tag	UNP P29067
E	-8	GLY	-	expression tag	UNP P29067
E	-7	LEU	-	expression tag	UNP P29067
E	-6	VAL	-	expression tag	UNP P29067
E	-5	PRO	-	expression tag	UNP P29067
E	-4	ARG	-	expression tag	UNP P29067

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	-3	GLY	-	expression tag	UNP P29067
E	-2	SER	-	expression tag	UNP P29067
E	-1	HIS	-	expression tag	UNP P29067
E	0	MET	-	expression tag	UNP P29067
F	-20	MET	-	expression tag	UNP P29067
F	-19	GLY	-	expression tag	UNP P29067
F	-18	SER	-	expression tag	UNP P29067
F	-17	SER	-	expression tag	UNP P29067
F	-16	HIS	-	expression tag	UNP P29067
F	-15	HIS	-	expression tag	UNP P29067
F	-14	HIS	-	expression tag	UNP P29067
F	-13	HIS	-	expression tag	UNP P29067
F	-12	HIS	-	expression tag	UNP P29067
F	-11	HIS	-	expression tag	UNP P29067
F	-10	SER	-	expression tag	UNP P29067
F	-9	SER	-	expression tag	UNP P29067
F	-8	GLY	-	expression tag	UNP P29067
F	-7	LEU	-	expression tag	UNP P29067
F	-6	VAL	-	expression tag	UNP P29067
F	-5	PRO	-	expression tag	UNP P29067
F	-4	ARG	-	expression tag	UNP P29067
F	-3	GLY	-	expression tag	UNP P29067
F	-2	SER	-	expression tag	UNP P29067
F	-1	HIS	-	expression tag	UNP P29067
F	0	MET	-	expression tag	UNP P29067

- Molecule 2 is a protein called Peptide from Atypical chemokine receptor 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	U	13	Total	C	N	O	P	0	0	0
			112	65	16	28	3			
2	V	12	Total	C	N	O	P	0	0	0
			104	61	15	25	3			
2	W	9	Total	C	N	O	P	0	0	0
			73	43	11	17	2			
2	X	12	Total	C	N	O	P	0	0	0
			104	61	15	25	3			
2	Y	12	Total	C	N	O	P	0	0	0
			104	61	15	25	3			
2	Z	9	Total	C	N	O	P	0	0	0
			73	43	11	17	2			

- Molecule 3 is water.

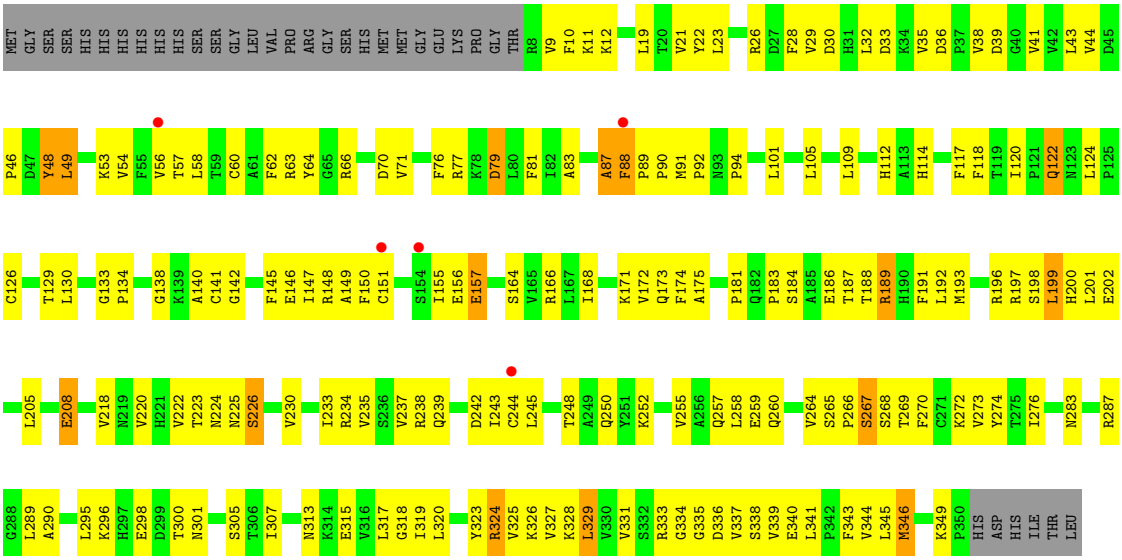
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	85	Total O 85 85	0	0
3	U	8	Total O 8 8	0	0
3	B	73	Total O 73 73	0	0
3	V	6	Total O 6 6	0	0
3	C	93	Total O 93 93	0	0
3	W	4	Total O 4 4	0	0
3	D	80	Total O 80 80	0	0
3	X	2	Total O 2 2	0	0
3	E	82	Total O 82 82	0	0
3	Y	4	Total O 4 4	0	0
3	F	87	Total O 87 87	0	0

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.

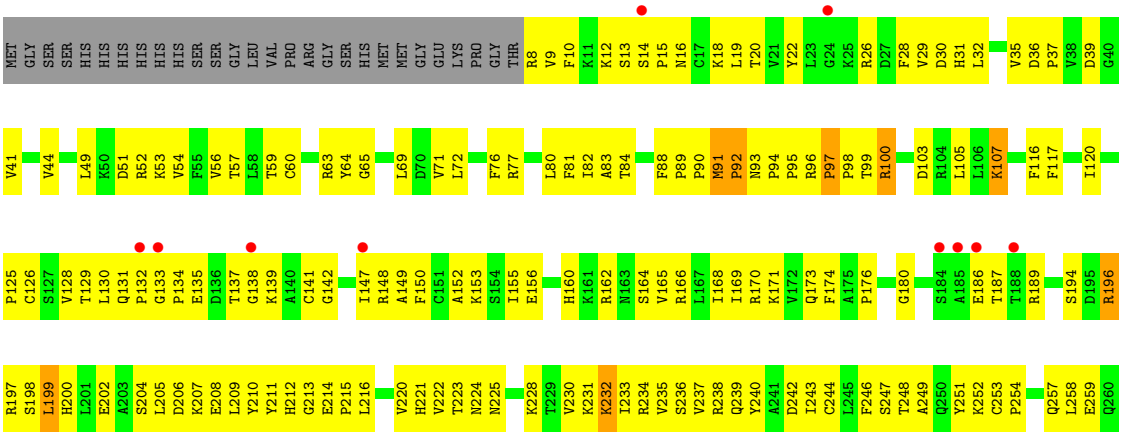
• Molecule 1: Beta-arrestin-2

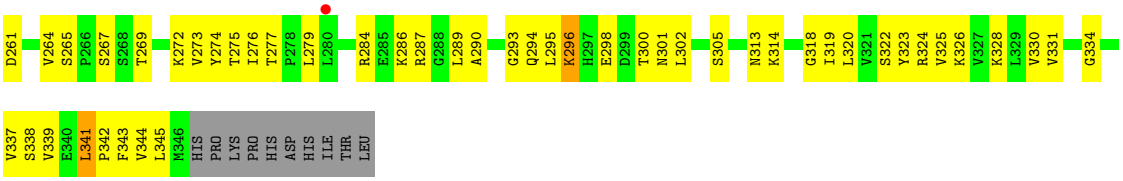
Chain A:



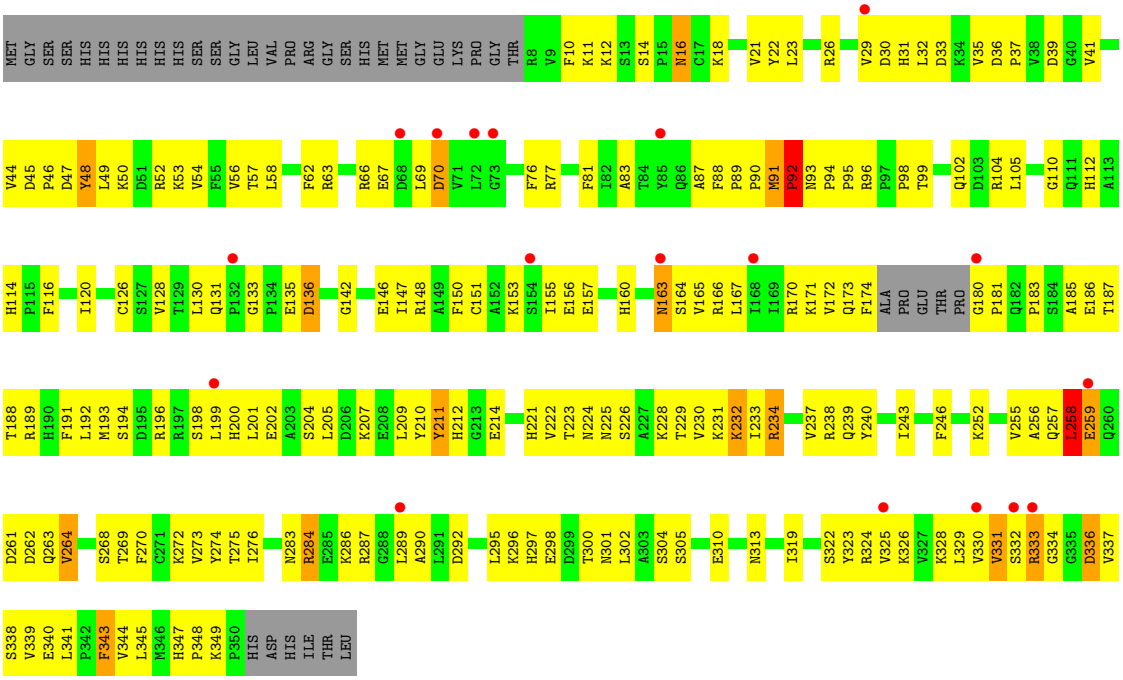
• Molecule 1: Beta-arrestin-2

Chain B:

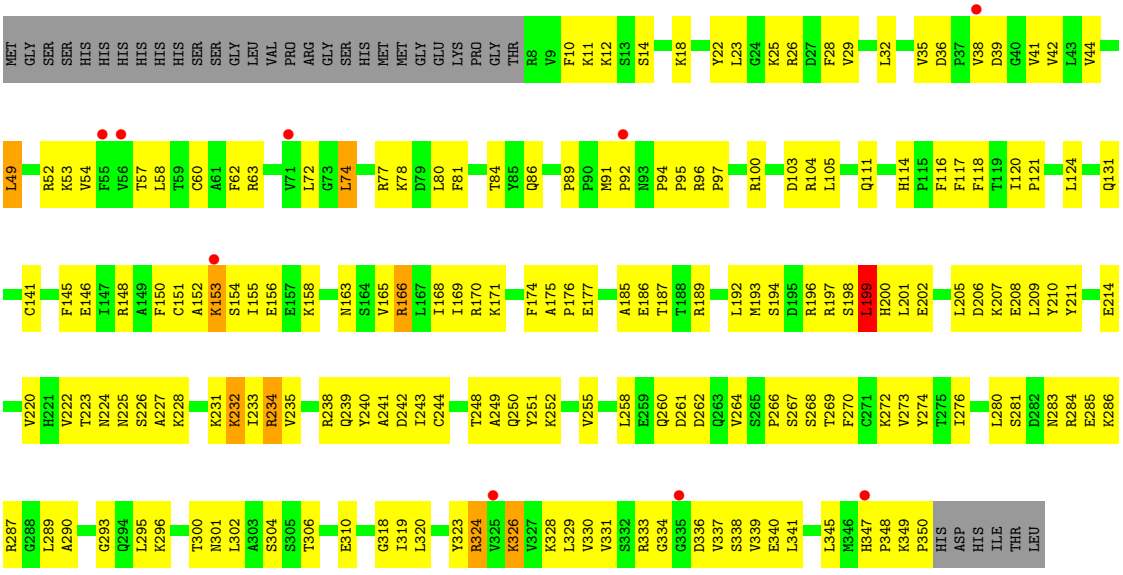




• Molecule 1: Beta-arrestin-2

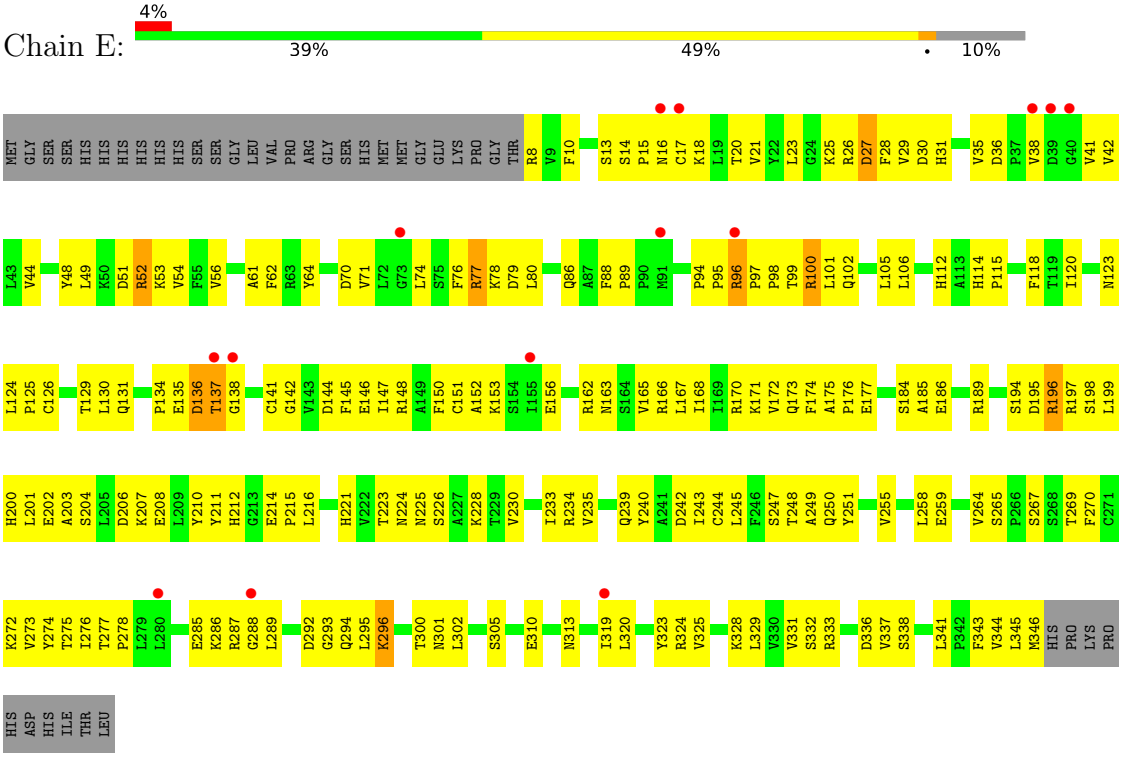


• Molecule 1: Beta-arrestin-2



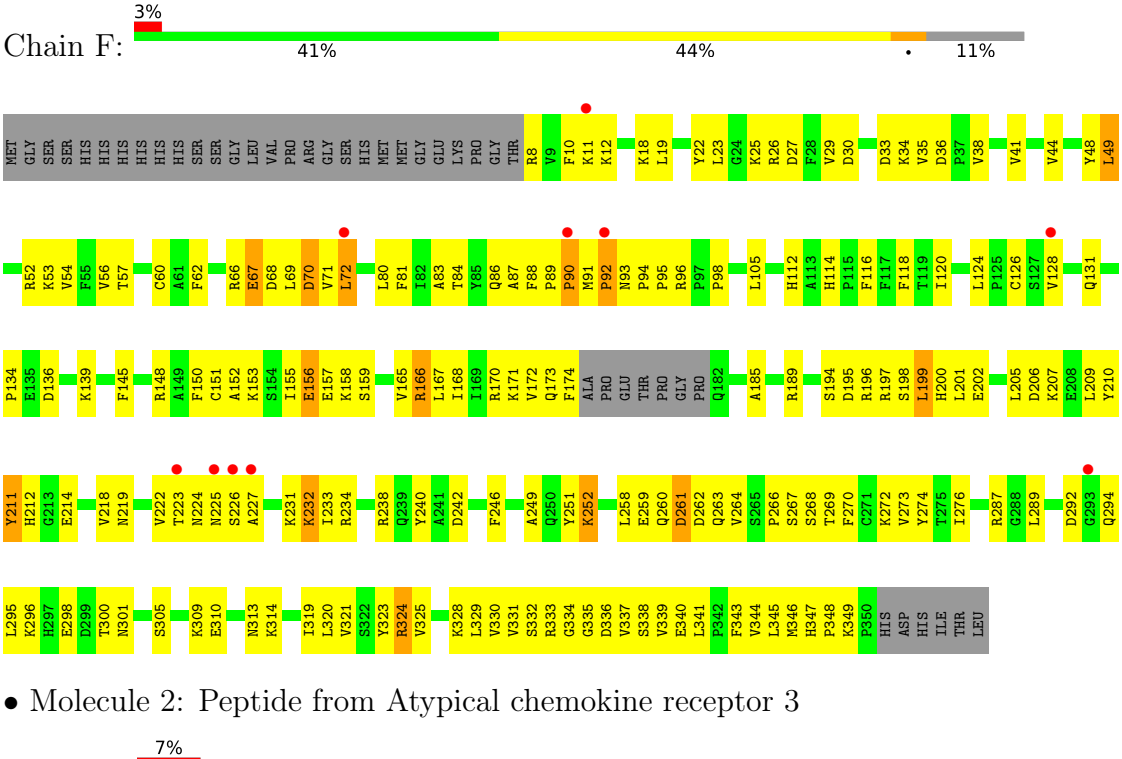
• Molecule 1: Beta-arrestin-2

Chain E:



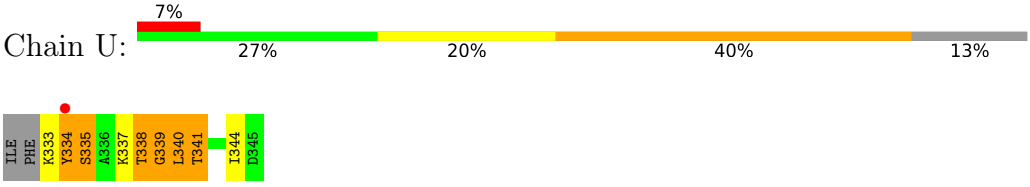
• Molecule 1: Beta-arrestin-2

Chain F:



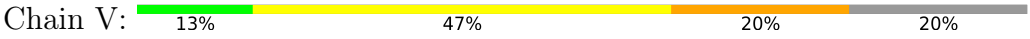
• Molecule 2: Peptide from Atypical chemokine receptor 3

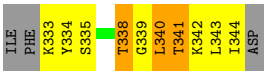
Chain U:



• Molecule 2: Peptide from Atypical chemokine receptor 3

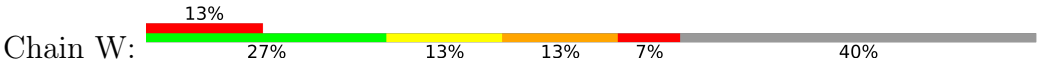
Chain V:

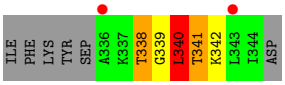




• Molecule 2: Peptide from Atypical chemokine receptor 3

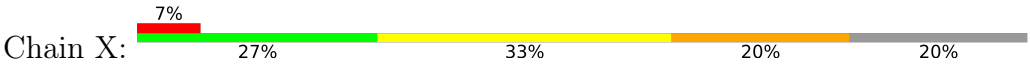
Chain W:

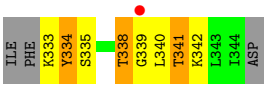




• Molecule 2: Peptide from Atypical chemokine receptor 3

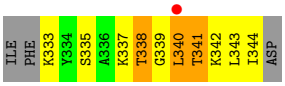
Chain X:





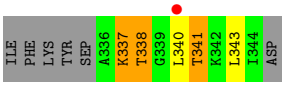
• Molecule 2: Peptide from Atypical chemokine receptor 3

Chain Y:



• Molecule 2: Peptide from Atypical chemokine receptor 3

Chain Z:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	91.17Å 127.91Å 206.04Å 90.00° 90.04° 90.00°	Depositor
Resolution (Å)	46.80 – 2.30 46.80 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.80-2.30) 100.0 (46.80-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	26.44 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.233 , 0.246 0.229 , 0.246	Depositor DCC
R_{free} test set	5309 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	28.9	Xtriage
Anisotropy	0.200	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 74.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	0.032 for -h,-k,l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	17270	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.27% 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)

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

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	0.29	0/2784	0.62	1/3773 (0.0%)
1	B	0.33	0/2739	0.63	1/3711 (0.0%)
1	C	0.51	3/2746 (0.1%)	0.68	3/3717 (0.1%)
1	D	0.30	0/2784	0.64	1/3773 (0.0%)
1	E	0.40	2/2748 (0.1%)	0.74	4/3723 (0.1%)
1	F	0.31	0/2734	0.62	1/3700 (0.0%)
2	U	0.39	0/77	0.83	0/95
2	V	0.34	0/69	0.95	0/84
2	W	0.23	0/48	1.15	1/58 (1.7%)
2	X	0.33	0/69	0.74	0/84
2	Y	0.30	0/69	0.74	0/84
2	Z	0.24	0/48	0.48	0/58
All	All	0.36	5/16915 (0.0%)	0.66	12/22860 (0.1%)

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	3
1	B	0	2
1	C	0	6
1	D	0	1
1	E	0	1
1	F	0	4
2	U	0	1
2	W	0	1
2	X	0	1
All	All	0	20

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	259	GLU	CG-CD	15.36	1.75	1.51
1	C	259	GLU	CD-OE2	9.68	1.36	1.25
1	E	96	ARG	CD-NE	-6.83	1.34	1.46
1	E	96	ARG	NE-CZ	-6.24	1.25	1.33
1	C	264	VAL	C-N	5.22	1.46	1.34

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	96	ARG	NE-CZ-NH1	-18.28	111.16	120.30
1	C	259	GLU	CA-CB-CG	10.36	136.20	113.40
1	E	96	ARG	NH1-CZ-NH2	10.09	130.50	119.40
1	E	96	ARG	CG-CD-NE	-7.96	95.09	111.80
2	W	340	LEU	CA-CB-CG	6.47	130.19	115.30

There are no chirality outliers.

5 of 20 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	199	LEU	Peptide
1	A	226	SER	Peptide
1	A	91	MET	Peptide
1	B	91	MET	Peptide
2	U	337	LYS	Peptide

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	2722	0	2773	235	0
1	B	2680	0	2724	259	1
1	C	2687	0	2740	227	0
1	D	2722	0	2773	222	1
1	E	2689	0	2739	239	3
1	F	2676	0	2730	207	3
2	U	112	0	105	14	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	V	104	0	101	23	0
2	W	73	0	76	10	0
2	X	104	0	101	7	0
2	Y	104	0	101	12	0
2	Z	73	0	76	13	0
3	A	85	0	0	3	0
3	B	73	0	0	4	0
3	C	93	0	0	8	0
3	D	80	0	0	10	0
3	E	82	0	0	7	0
3	F	87	0	0	11	1
3	U	8	0	0	0	0
3	V	6	0	0	0	0
3	W	4	0	0	1	0
3	X	2	0	0	0	0
3	Y	4	0	0	1	0
All	All	17270	0	17039	1255	5

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 37.

The worst 5 of 1255 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:C:259:GLU:CD	1:C:259:GLU:CG	1.75	1.54
1:A:337:VAL:HB	1:B:135:GLU:OE2	1.21	1.31
1:A:233:ILE:HD13	1:A:329:LEU:CG	1.76	1.14
1:F:91:MET:HG3	1:F:92:PRO:HD3	1.25	1.14
1:C:238:ARG:HE	1:C:326:LYS:HD3	1.08	1.12

All (5) 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	Interatomic distance (Å)	Clash overlap (Å)
1:E:197:ARG:NH2	1:F:158:LYS:NZ[1_455]	2.04	0.16
1:B:96:ARG:NH2	1:F:234:ARG:N[3_445]	2.09	0.11
1:E:197:ARG:CZ	1:F:158:LYS:NZ[1_455]	2.09	0.11
1:E:196:ARG:NH2	3:F:459:HOH:O[1_455]	2.15	0.05
1:D:349:LYS:N	1:D:349:LYS:O[2_655]	2.18	0.02

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	341/377 (90%)	282 (83%)	54 (16%)	5 (2%)	10 10
1	B	337/377 (89%)	275 (82%)	60 (18%)	2 (1%)	25 31
1	C	334/377 (89%)	270 (81%)	61 (18%)	3 (1%)	17 20
1	D	341/377 (90%)	283 (83%)	58 (17%)	0	100 100
1	E	337/377 (89%)	277 (82%)	60 (18%)	0	100 100
1	F	332/377 (88%)	280 (84%)	50 (15%)	2 (1%)	25 31
2	U	8/15 (53%)	3 (38%)	2 (25%)	3 (38%)	0 0
2	V	7/15 (47%)	3 (43%)	2 (29%)	2 (29%)	0 0
2	W	5/15 (33%)	3 (60%)	2 (40%)	0	100 100
2	X	7/15 (47%)	2 (29%)	4 (57%)	1 (14%)	0 0
2	Y	7/15 (47%)	3 (43%)	3 (43%)	1 (14%)	0 0
2	Z	5/15 (33%)	0	5 (100%)	0	100 100
All	All	2061/2352 (88%)	1681 (82%)	361 (18%)	19 (1%)	17 20

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	U	340	LEU
2	V	334	TYR
2	V	340	LEU
2	Y	340	LEU
1	A	87	ALA

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	310/339 (91%)	298 (96%)	12 (4%)	32	46
1	B	303/339 (89%)	297 (98%)	6 (2%)	55	72
1	C	306/339 (90%)	288 (94%)	18 (6%)	19	27
1	D	310/339 (91%)	299 (96%)	11 (4%)	36	50
1	E	306/339 (90%)	297 (97%)	9 (3%)	42	58
1	F	305/339 (90%)	288 (94%)	17 (6%)	21	29
2	U	8/10 (80%)	8 (100%)	0	100	100
2	V	7/10 (70%)	7 (100%)	0	100	100
2	W	5/10 (50%)	5 (100%)	0	100	100
2	X	7/10 (70%)	7 (100%)	0	100	100
2	Y	7/10 (70%)	6 (86%)	1 (14%)	3	3
2	Z	5/10 (50%)	4 (80%)	1 (20%)	1	1
All	All	1879/2094 (90%)	1804 (96%)	75 (4%)	31	44

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	8	ARG
1	F	252	LYS
1	F	49	LEU
1	F	195	ASP
1	C	153	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	221	HIS
1	C	294	GLN
1	F	212	HIS
1	E	102	GLN
1	F	160	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

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

16 non-standard protein/DNA/RNA residues 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	TPO	U	338	2	8,10,11	1.66	1 (12%)	10,14,16	2.36	2 (20%)
2	TPO	U	341	2	8,10,11	1.66	1 (12%)	10,14,16	1.83	1 (10%)
2	TPO	V	341	1,2	8,10,11	1.16	0	10,14,16	1.47	2 (20%)
2	TPO	W	338	2	8,10,11	1.69	1 (12%)	10,14,16	1.20	1 (10%)
2	TPO	X	338	2	8,10,11	1.73	1 (12%)	10,14,16	1.41	1 (10%)
2	TPO	Y	341	2	8,10,11	1.67	1 (12%)	10,14,16	1.13	0
2	TPO	Z	338	2	8,10,11	1.64	1 (12%)	10,14,16	1.59	1 (10%)
2	SEP	U	335	2	8,9,10	1.54	1 (12%)	8,12,14	1.63	2 (25%)
2	TPO	Y	338	2	8,10,11	1.14	0	10,14,16	1.60	1 (10%)
2	TPO	Z	341	2	8,10,11	1.78	3 (37%)	10,14,16	1.51	1 (10%)
2	TPO	V	338	2	8,10,11	1.76	1 (12%)	10,14,16	1.30	1 (10%)
2	SEP	V	335	2	8,9,10	1.54	1 (12%)	8,12,14	1.45	2 (25%)
2	SEP	Y	335	2	8,9,10	1.56	1 (12%)	8,12,14	1.41	2 (25%)
2	TPO	X	341	2	8,10,11	1.15	0	10,14,16	1.29	2 (20%)
2	SEP	X	335	2	8,9,10	1.57	1 (12%)	8,12,14	1.67	2 (25%)
2	TPO	W	341	2	8,10,11	1.60	1 (12%)	10,14,16	1.00	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
2	TPO	U	338	2	-	3/9/11/13	-
2	TPO	U	341	2	-	3/9/11/13	-
2	TPO	V	341	1,2	-	4/9/11/13	-
2	TPO	W	338	2	-	7/9/11/13	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	TPO	X	338	2	-	7/9/11/13	-
2	TPO	Y	341	2	-	5/9/11/13	-
2	TPO	Z	338	2	-	5/9/11/13	-
2	SEP	U	335	2	-	2/5/8/10	-
2	TPO	Y	338	2	-	5/9/11/13	-
2	TPO	Z	341	2	-	2/9/11/13	-
2	TPO	V	338	2	-	6/9/11/13	-
2	SEP	V	335	2	-	4/5/8/10	-
2	SEP	Y	335	2	-	2/5/8/10	-
2	TPO	X	341	2	-	7/9/11/13	-
2	SEP	X	335	2	-	2/5/8/10	-
2	TPO	W	341	2	-	3/9/11/13	-

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	X	338	TPO	P-O1P	3.52	1.61	1.50
2	Y	335	SEP	P-O1P	3.46	1.61	1.50
2	V	338	TPO	P-O1P	3.46	1.61	1.50
2	U	338	TPO	P-O1P	3.45	1.61	1.50
2	Z	338	TPO	P-O1P	3.42	1.61	1.50

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	U	338	TPO	P-OG1-CB	-6.02	105.01	123.21
2	U	341	TPO	P-OG1-CB	-5.22	107.43	123.21
2	Z	338	TPO	P-OG1-CB	-4.45	109.77	123.21
2	Y	338	TPO	P-OG1-CB	-3.98	111.19	123.21
2	Z	341	TPO	P-OG1-CB	-3.82	111.66	123.21

There are no chirality outliers.

5 of 67 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	U	335	SEP	CA-CB-OG-P
2	U	338	TPO	N-CA-CB-OG1
2	U	341	TPO	N-CA-CB-OG1
2	U	341	TPO	CB-OG1-P-O3P

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	V	335	SEP	N-CA-CB-OG

There are no ring outliers.

13 monomers are involved in 36 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	U	338	TPO	3	0
2	U	341	TPO	4	0
2	V	341	TPO	1	0
2	W	338	TPO	1	0
2	X	338	TPO	1	0
2	Y	341	TPO	1	0
2	Z	338	TPO	1	0
2	U	335	SEP	1	0
2	Y	338	TPO	1	0
2	Z	341	TPO	9	0
2	V	338	TPO	8	0
2	X	341	TPO	2	0
2	W	341	TPO	3	0

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

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, 95^{th} 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	343/377 (90%)	0.55	5 (1%) 73 79	14, 22, 36, 58	0
1	B	339/377 (89%)	0.49	11 (3%) 47 54	17, 23, 35, 54	0
1	C	338/377 (89%)	0.60	18 (5%) 26 33	17, 25, 41, 63	0
1	D	343/377 (90%)	0.54	9 (2%) 56 63	17, 24, 34, 46	0
1	E	339/377 (89%)	0.60	14 (4%) 37 44	17, 25, 41, 66	0
1	F	336/377 (89%)	0.50	10 (2%) 50 57	17, 24, 38, 64	0
2	U	10/15 (66%)	0.35	1 (10%) 7 10	15, 24, 30, 44	0
2	V	9/15 (60%)	0.71	0 100 100	20, 25, 33, 35	0
2	W	7/15 (46%)	0.95	2 (28%) 0 0	22, 35, 42, 63	0
2	X	9/15 (60%)	0.83	1 (11%) 5 7	22, 31, 42, 44	0
2	Y	9/15 (60%)	0.85	1 (11%) 5 7	26, 39, 50, 75	0
2	Z	7/15 (46%)	1.28	1 (14%) 2 3	25, 37, 55, 56	0
All	All	2089/2352 (88%)	0.55	73 (3%) 44 51	14, 24, 39, 75	0

The worst 5 of 73 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	73	GLY	13.0
1	A	88	PHE	6.8
1	F	92	PRO	6.8
1	C	259	GLU	5.7
1	D	92	PRO	5.1

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

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. The B-factors column lists the minimum,

median, 95^{th} 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	B-factors(Å^2)	Q<0.9
2	TPO	Z	341	11/12	0.68	0.24	26,34,46,50	0
2	TPO	Y	341	11/12	0.76	0.28	45,57,68,69	0
2	TPO	X	341	11/12	0.83	0.23	47,61,64,67	0
2	TPO	U	338	11/12	0.85	0.23	15,26,40,43	0
2	SEP	Y	335	10/11	0.86	0.18	29,40,46,47	0
2	TPO	W	338	11/12	0.87	0.16	44,64,75,79	0
2	TPO	Z	338	11/12	0.91	0.12	40,45,58,60	0
2	TPO	W	341	11/12	0.91	0.21	25,29,37,48	0
2	TPO	Y	338	11/12	0.92	0.16	32,40,49,49	0
2	TPO	V	341	11/12	0.92	0.23	16,25,41,44	0
2	TPO	V	338	11/12	0.94	0.17	23,31,34,36	0
2	TPO	U	341	11/12	0.94	0.18	15,18,28,29	0
2	TPO	X	338	11/12	0.94	0.14	30,34,40,41	0
2	SEP	X	335	10/11	0.95	0.16	24,31,36,41	0
2	SEP	V	335	10/11	0.95	0.16	16,25,27,28	0
2	SEP	U	335	10/11	0.97	0.13	15,20,26,30	0

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.