

wwPDB X-ray Structure Validation Summary Report (i)

Dec 2, 2024 – 02:08 PM JST

PDB ID : 8ZPU

Title : Crystal structure of the anti-phosphorylated peptide C7 scFv antibody with

peptide bound

Authors : Caaveiro, J.M.M.; Kawade, R.; Tsumoto, K.

Deposited on : 2024-05-31

Resolution : 2.80 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:

 $Mol Probity \quad : \quad 4.02b\text{--}467$

Mogul: 1.8.5 (274361), CSD as541be (2020)

Xtriage (Phenix) : 1.21 EDS : 3.0

Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)

CCP4 : 9.0.004 (Gargrove)

Density-Fitness : 1.0.11

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

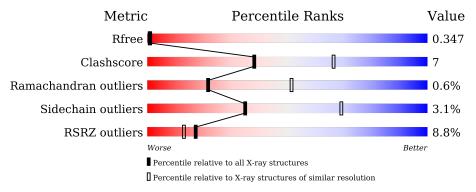
Validation Pipeline (wwPDB-VP) : 2.40

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 2.80 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	Similar resolution
Metric	$(\# \mathrm{Entries})$	$(\# ext{Entries}, ext{ resolution range}(ext{Å}))$
R_{free}	164625	3657 (2.80-2.80)
Clashscore	180529	4123 (2.80-2.80)
Ramachandran outliers	177936	4071 (2.80-2.80)
Sidechain outliers	177891	4073 (2.80-2.80)
RSRZ outliers	164620	3659 (2.80-2.80)

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	257	9%	7	5%	14% •	10%	
1	В	257	5%	72	%	17% •	10%	
2	С	12	17%	50%	8%	42%		
2	D	12	25%	50%	17%	33%		

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 crite-



ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	SO4	A	301	-	-	X	-



2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3648 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.

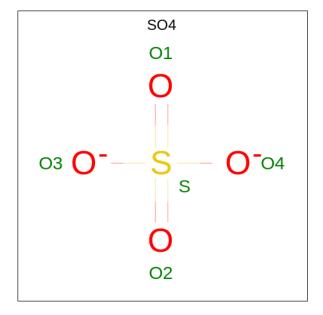
• Molecule 1 is a protein called scFv.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	A	232	Total 1730	C 1079	N 294	O 350	S 7	0	0	0
1	В	232	Total 1730		N 294	O 350	S 7	0	0	0

• Molecule 2 is a protein called PHE-PRO-GLN-PHE-SEP-TYR-SER.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
9	C	7	Total	С	N	О	Р	0	0	0
2		1	66	43	8	14	1	0		
9	D	0	Total	С	N	О	Р	0	0	0
2	ש	0	76	49	11	15	1	0	U	

• Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	В	1	Total O S 5 4 1	0	0
3	В	1	Total O S 5 4 1	0	0
3	В	1	Total O S 5 4 1	0	0

• Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Cl 1 1	0	0
4	В	1	Total Cl 1 1	0	0

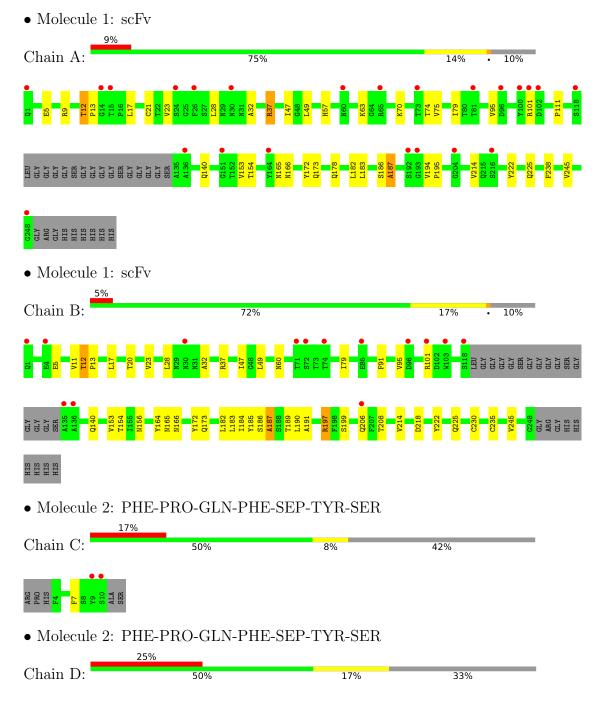
• Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	11	Total O 11 11	0	0
5	В	3	Total O 3 3	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.









4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants	126.61Å 126.61Å 82.36Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 120.00°	Depositor
Resolution (Å)	21.10 - 2.80	Depositor
rtesolution (A)	21.10 - 2.80	EDS
% Data completeness	99.7 (21.10-2.80)	Depositor
(in resolution range)	99.7 (21.10-2.80)	EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	2.66 (at 2.79Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.304 , 0.345	Depositor
it, it free	0.307 , 0.347	DCC
R_{free} test set	1971 reflections (10.36%)	wwPDB-VP
Wilson B-factor (Å ²)	46.9	Xtriage
Anisotropy	0.022	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.34, 31.7	EDS
L-test for twinning ²	$< L >=0.49, < L^2>=0.32$	Xtriage
Estimated twinning fraction	0.032 for -h,-k,l	Xtriage
F_o, F_c correlation	0.83	EDS
Total number of atoms	3648	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The analyses of the Patterson function reveals a significant off-origin peak that is 28.16 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.9345e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.

²Theoretical 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.



¹Intensities estimated from amplitudes.

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: SO4, CL, SEP

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		
IVIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	A	0.37	0/1766	0.77	0/2407	
1	В	0.39	0/1766	0.78	0/2407	
2	С	0.46	0/58	0.65	0/76	
2	D	0.48	0/69	0.84	0/91	
All	All	0.38	0/3659	0.77	0/4981	

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 maintenain group or atoms of a sidechain that are expected to be planar.

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

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	70	LYS	Peptide
1	A	9	ARG	Sidechain
1	В	197	ARG	Sidechain



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	1730	0	1664	23	0
1	В	1730	0	1664	26	0
2	С	66	0	50	1	0
2	D	76	0	57	2	0
3	A	15	0	0	3	0
3	В	15	0	0	1	0
4	A	1	0	0	0	0
4	В	1	0	0	0	0
5	A	11	0	0	0	0
5	В	3	0	0	0	0
All	All	3648	0	3435	48	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 7.

The worst 5 of 48 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	$\begin{array}{c} \text{Clash} \\ \text{overlap } (\text{\AA}) \end{array}$
1:A:23:VAL:HG21	1:A:28:LEU:HD21	1.68	0.75
1:B:23:VAL:HG21	1:B:28:LEU:HD21	1.69	0.72
1:B:230:CYS:SG	2:D:6:GLN:OE1	2.48	0.72
1:A:17:LEU:HD23	1:A:79:ILE:HD12	1.73	0.69
1:A:165:ASN:O	1:A:166:ASN:HB2	1.95	0.67

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	Perce	entiles	
1	A	$228/257 \ (89\%)$	215 (94%)	12 (5%)	1 (0%)	30	61
1	В	228/257~(89%)	212 (93%)	14 (6%)	2 (1%)	14	42
2	С	4/12 (33%)	4 (100%)	0	0	100	100
2	D	5/12 (42%)	5 (100%)	0	0	100	100
All	All	465/538~(86%)	436 (94%)	26 (6%)	3 (1%)	22	51

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	В	191	ALA
1	A	187	ALA
1	В	187	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	190/201 (94%)	185 (97%)	5 (3%)	41	75	
1	В	190/201~(94%)	183 (96%)	7 (4%)	29	63	
2	С	6/10 (60%)	6 (100%)	0	100	100	
2	D	7/10 (70%)	7 (100%)	0	100	100	
All	All	393/422~(93%)	381 (97%)	12 (3%)	35	69	

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	В	101	ARG
1	В	154	THR
1	В	245	VAL
1	В	199	SER
1	A	101	ARG



Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	215	GLN
1	В	60	ASN
1	В	178	GLN
1	В	206	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)

2 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	Trimo	Chain	Chain	Chain	Chain	Chain	Chain	Chain	Dag	Dec Link	Bond lengths			Bond angles		
MIOI	Type	Chain	Res	Link	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2						
2	SEP	D	8	2	8,9,10	0.67	0	8,12,14	0.89	0						
2	SEP	С	8	2	8,9,10	0.56	0	8,12,14	0.87	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	SEP	D	8	2	-	4/5/8/10	-
2	SEP	С	8	2	-	3/5/8/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 torsion outliers are listed below:



Mol	Chain	Res	Type	Atoms
2	С	8	SEP	CA-CB-OG-P
2	С	8	SEP	CB-OG-P-O2P
2	С	8	SEP	CB-OG-P-O3P
2	D	8	SEP	CA-CB-OG-P
2	D	8	SEP	CB-OG-P-O2P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 2 are monoatomic - leaving 6 for Mogul analysis.

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	Trino	Chain	Res	Link	Bond lengths			Bond angles		
Mol Type	Chain	nes	Lilik	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z > 2	
3	SO4	A	302	-	4,4,4	0.34	0	6,6,6	0.13	0
3	SO4	A	301	-	4,4,4	0.23	0	6,6,6	0.20	0
3	SO4	A	303	-	4,4,4	0.33	0	6,6,6	0.08	0
3	SO4	В	301	-	4,4,4	0.34	0	6,6,6	0.06	0
3	SO4	В	303	_	4,4,4	0.35	0	6,6,6	0.19	0
3	SO4	В	302	-	4,4,4	0.35	0	6,6,6	0.15	0

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.

3 monomers are involved in 4 short contacts:



Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	302	SO4	1	0
3	A	301	SO4	2	0
3	В	302	SO4	1	0

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^{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	$\langle { m RSRZ} \rangle$	$\# \mathrm{RSRZ}{>}2$	$OWAB(A^2)$	Q<0.9
1	A	232/257 (90%)	0.58	23 (9%) 14 11	25, 47, 68, 103	0
1	В	232/257 (90%)	0.40	14 (6%) 29 22	24, 41, 65, 81	0
2	С	6/12 (50%)	1.74	2 (33%) 1 1	55, 65, 79, 82	0
2	D	7/12 (58%)	1.74	3 (42%) 1 1	43, 51, 80, 90	0
All	All	477/538 (88%)	0.52	42 (8%) 17 13	24, 44, 68, 103	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	100	TYR	6.2
1	В	72	SER	4.2
1	В	118	SER	3.8
2	D	3	HIS	3.7
2	С	10	SER	3.6

6.2 Non-standard residues in protein, DNA, RNA chains (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. 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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
2	SEP	D	8	10/11	0.90	0.14	48,56,60,63	0
2	SEP	С	8	10/11	0.92	0.10	43,49,62,63	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.



6.4 Ligands (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. 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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
3	SO4	A	303	5/5	0.79	0.13	40,41,44,46	5
3	SO4	В	302	5/5	0.85	0.14	48,59,61,68	0
3	SO4	В	301	5/5	0.89	0.12	73,74,79,79	0
3	SO4	A	302	5/5	0.90	0.10	49,49,54,55	0
3	SO4	A	301	5/5	0.90	0.21	43,51,58,58	0
3	SO4	В	303	5/5	0.90	0.14	49,51,59,60	0
4	CL	A	304	1/1	0.93	0.11	54,54,54,54	0
4	CL	В	304	1/1	0.99	0.17	21,21,21,21	1

6.5 Other polymers (i)

There are no such residues in this entry.

