WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Nov 6, 2023 – 11:35 AM EST

PDB ID	:	6PZ6
Title	:	Co-assembly of VIQKI D452(beta-L-homoaspartic acid) with human parain-fluenza virus type 3 (HPIV3) fusion glycoprotein N-terminal heptad repeat domain
Authors	:	Outlaw, V.K.; Gellman, S.H.
Deposited on	:	2019-07-31
Resolution	:	1.70 Å(reported)

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

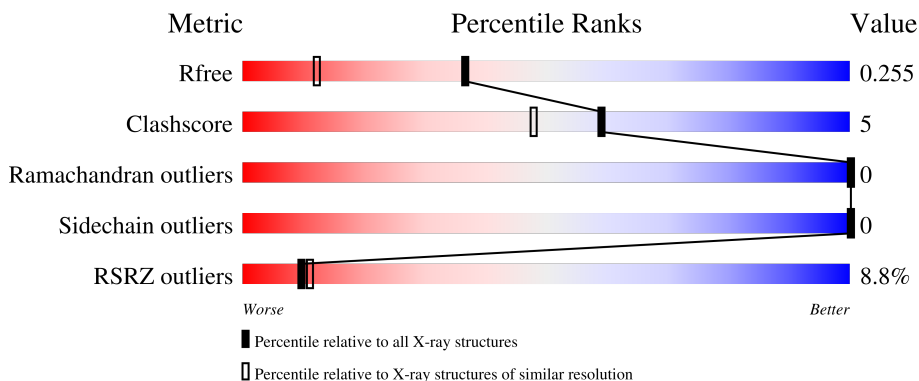|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry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

1   Overall quality at a glance ⓘ

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

The reported resolution of this entry is 1.70 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	Percentile Ranks	Value
Rfree		0.255
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		0
RSRZ outliers		8.8%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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	53	4% · 85% · 6% · 9%
1	C	53	9% · 81% · 8% · 11%
1	E	53	9% · 77% · 9% · 13%
2	B	38	8% · 71% · 18% · 8%
2	D	38	3% · 66% · 18% · 13%

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	38	13% · · · 68% · · · 16% · 13%

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	B3D	D	452	-	-	-	X
2	B3D	F	452	-	-	-	X

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 3936 atoms, of which 1991 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 Fusion glycoprotein F1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	48	Total	C	H	N	O	0	0	0
			735	225	376	62	72			
1	C	47	Total	C	H	N	O	0	0	0
			721	221	366	61	73			
1	E	46	Total	C	H	N	O	0	0	0
			714	218	364	60	72			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	138	ACE	-	acetylation	UNP P06828
A	190	NH2	-	amidation	UNP P06828
C	138	ACE	-	acetylation	UNP P06828
C	190	NH2	-	amidation	UNP P06828
E	138	ACE	-	acetylation	UNP P06828
E	190	NH2	-	amidation	UNP P06828

- Molecule 2 is a protein called VIQKI D4(beta-L-homoaspartic acid) synthetic peptid derived from Fusion glycoprotein F1e.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	35	Total	C	H	N	O	0	0	1
			585	179	303	49	54			
2	D	33	Total	C	H	N	O	0	0	0
			564	173	291	47	53			
2	F	33	Total	C	H	N	O	0	0	0
			564	173	291	47	53			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	448	ACE	-	acetylation	UNP P06828

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	459	VAL	GLU	engineered mutation	UNP P06828
B	463	ILE	ALA	engineered mutation	UNP P06828
B	466	GLN	ASP	engineered mutation	UNP P06828
B	479	LYS	GLN	engineered mutation	UNP P06828
B	480	ILE	LYS	engineered mutation	UNP P06828
B	485	NH2	-	amidation	UNP P06828
D	448	ACE	-	acetylation	UNP P06828
D	459	VAL	GLU	engineered mutation	UNP P06828
D	463	ILE	ALA	engineered mutation	UNP P06828
D	466	GLN	ASP	engineered mutation	UNP P06828
D	479	LYS	GLN	engineered mutation	UNP P06828
D	480	ILE	LYS	engineered mutation	UNP P06828
D	485	NH2	-	amidation	UNP P06828
F	448	ACE	-	acetylation	UNP P06828
F	459	VAL	GLU	engineered mutation	UNP P06828
F	463	ILE	ALA	engineered mutation	UNP P06828
F	466	GLN	ASP	engineered mutation	UNP P06828
F	479	LYS	GLN	engineered mutation	UNP P06828
F	480	ILE	LYS	engineered mutation	UNP P06828
F	485	NH2	-	amidation	UNP P06828

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	9	Total 9	O 9	0	0
3	B	6	Total 6	O 6	0	0
3	C	10	Total 10	O 10	0	0
3	D	16	Total 16	O 16	0	0
3	E	8	Total 8	O 8	0	0
3	F	4	Total 4	O 4	0	0

3 Residue-property plots ⓘ

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 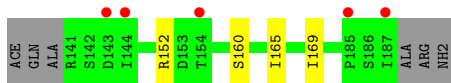residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: Fusion glycoprotein F1
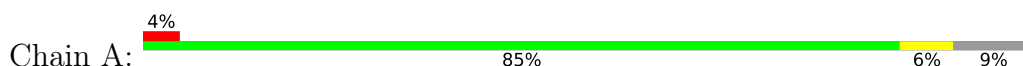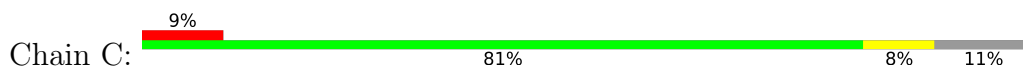
Chain A:

• Molecule 1: Fusion glycoprotein F1

Chain C:

• Molecule 1: Fusion glycoprotein F1

Chain E:

• Molecule 2: VIQKI D4(beta-L-homoaspartic acid) synthetic peptid derived from Fusion glycoprotein F1e

Chain B:

• Molecule 2: VIQKI D4(beta-L-homoaspartic acid) synthetic peptid derived from Fusion glycoprotein F1e

Chain D:

• Molecule 2: VIQKI D4(beta-L-homoaspartic acid) synthetic peptid derived from Fusion glycoprotein F1e

Chain F:
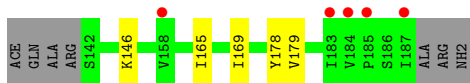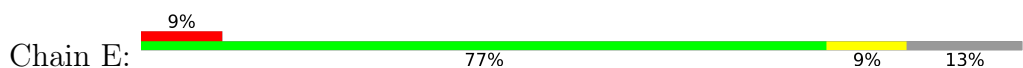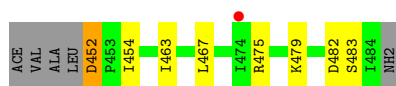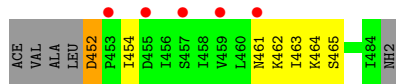
4 Data and refinement statistics (i)

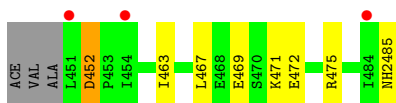| Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	39.64Å 52.45Å 55.62Å 90.00° 99.20° 90.00°	Depositor
Resolution (Å)	27.45 – 1.70 27.45 – 1.70	Depositor EDS
% Data completeness (in resolution range)	98.4 (27.45-1.70) 91.6 (27.45-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	0.94 (at 1.70Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.228 , 0.255 0.228 , 0.255	Depositor DCC
R_{free} test set	1999 reflections (8.16%)	wwPDB-VP
Wilson B-factor (Å^2)	33.9	Xtriage
Anisotropy	0.395	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 57.0	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3936	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.99% 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 ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: B3D, NH2

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	0/360	0.41	0/486
1	C	0.36	0/356	0.41	0/480
1	E	0.31	0/351	0.38	0/473
2	B	0.30	0/273	0.42	0/363
2	D	0.26	0/266	0.40	0/355
2	F	0.28	0/266	0.41	0/355
All	All	0.31	0/1872	0.40	0/2512

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
2	B	0	2
2	D	0	2
2	F	0	2
All	All	0	6

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	452	B3D	Peptide,Mainchain
2	D	452	B3D	Peptide,Mainchain
2	F	452	B3D	Peptide,Mainchain

5.2 Too-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 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	359	376	376	3	0
1	C	355	366	366	5	0
1	E	350	364	364	6	0
2	B	282	303	296	5	0
2	D	273	291	287	7	0
2	F	273	291	287	5	0
3	A	9	0	0	0	0
3	B	6	0	0	0	0
3	C	10	0	0	0	0
3	D	16	0	0	0	0
3	E	8	0	0	0	0
3	F	4	0	0	0	0
All	All	1945	1991	1976	18	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 5.

All (18) 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:152:ARG:NH2	2:D:482:ASP:OD1	2.17	0.77
1:E:169:ILE:HD11	2:F:463:ILE:HG22	1.84	0.60
2:F:462:LYS:O	2:F:465:SER:OG	2.15	0.59
1:E:179:VAL:HG11	2:F:454:ILE:HD11	1.84	0.59
2:D:483:SER:O	1:E:146:LYS:NZ	2.38	0.56
1:C:169:ILE:HD11	2:D:463:ILE:HG22	1.91	0.52
2:D:454:ILE:HD11	1:E:178:TYR:CD2	2.46	0.50
2:B:472:GLU:OE2	2:B:475:ARG:NH1	2.46	0.49
1:A:161:VAL:HG13	1:E:165:ILE:HD11	1.95	0.47
1:C:169:ILE:HD11	2:D:463:ILE:CG2	2.44	0.47
2:B:467:LEU:HG	2:B:471:LYS:HE3	1.97	0.46
1:A:169:ILE:HD11	2:B:463:ILE:HG22	1.98	0.44
1:C:165:ILE:HG22	2:D:467:LEU:HD11	2.00	0.43
1:A:148:LYS:HZ2	2:B:485:NH2:N	2.16	0.43
2:D:475:ARG:O	2:D:479:LYS:HD3	2.19	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:169:ILE:HD11	2:F:463:ILE:CG2	2.48	0.42
2:B:469:GLU:HG2	1:C:160:SER:OG	2.21	0.41
2:F:461:ASN:HA	2:F:464:LYS:HG2	2.02	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	46/53 (87%)	46 (100%)	0	0	100	100
1	C	45/53 (85%)	44 (98%)	1 (2%)	0	100	100
1	E	44/53 (83%)	44 (100%)	0	0	100	100
2	B	32/38 (84%)	30 (94%)	2 (6%)	0	100	100
2	D	31/38 (82%)	31 (100%)	0	0	100	100
2	F	31/38 (82%)	30 (97%)	1 (3%)	0	100	100
All	All	229/273 (84%)	225 (98%)	4 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	41/45 (91%)	41 (100%)	0	100	100
1	C	41/45 (91%)	41 (100%)	0	100	100
1	E	41/45 (91%)	41 (100%)	0	100	100
2	B	33/34 (97%)	33 (100%)	0	100	100
2	D	32/34 (94%)	32 (100%)	0	100	100
2	F	32/34 (94%)	32 (100%)	0	100	100
All	All	220/237 (93%)	220 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

3 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	B3D	D	452	2	8,8,9	1.45	2 (25%)	6,9,11	1.52	0
2	B3D	F	452	2	8,8,9	1.51	2 (25%)	6,9,11	1.41	0
2	B3D	B	452	2	8,8,9	1.52	2 (25%)	6,9,11	1.36	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.

Full wwPDB X-ray Structure Validation Report

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	B3D	D	452	2	-	4/7/7/8	-
2	B3D	F	452	2	-	3/7/7/8	-
2	B3D	B	452	2	-	1/7/7/8	-

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	452	B3D	CB-CA	2.90	1.57	1.53
2	B	452	B3D	CB-CA	2.89	1.57	1.53
2	D	452	B3D	CB-CA	2.77	1.56	1.53
2	B	452	B3D	CB-C	2.53	1.56	1.49
2	F	452	B3D	CB-C	2.46	1.56	1.49
2	D	452	B3D	CB-C	2.35	1.55	1.49

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	452	B3D	O-C-CB-CA
2	D	452	B3D	O-C-CB-CA
2	F	452	B3D	CB-CA-CG-CD
2	F	452	B3D	O-C-CB-CA
2	D	452	B3D	OE1-CD-CG-CA
2	D	452	B3D	OE2-CD-CG-CA
2	D	452	B3D	CG-CA-CB-C
2	F	452	B3D	N-CA-CG-CD

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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	48/53 (90%)	0.50	2 (4%) 36	40	32, 44, 63, 72	0
1	C	47/53 (88%)	0.75	5 (10%) 6	7	31, 41, 64, 74	0
1	E	46/53 (86%)	0.98	5 (10%) 5	6	31, 44, 73, 77	0
2	B	33/38 (86%)	0.76	3 (9%) 9	10	36, 53, 68, 125	0
2	D	32/38 (84%)	0.70	1 (3%) 49	53	33, 44, 64, 73	0
2	F	32/38 (84%)	1.15	5 (15%) 2	2	38, 59, 70, 77	0
All	All	238/273 (87%)	0.79	21 (8%) 10	11	31, 47, 72, 125	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	451	LEU	4.9
1	A	187	ILE	4.2
1	E	187	ILE	3.9
1	E	183	ILE	3.8
2	B	454	ILE	3.7
1	E	185	PRO	3.4
2	B	484	ILE	3.2
1	C	185	PRO	3.2
2	F	453	PRO	2.9
2	F	459	VAL	2.8
1	C	144	ILE	2.7
2	F	457	SER	2.6
1	C	187	ILE	2.3
1	C	143	ASP	2.2
1	A	186	SER	2.2
1	E	158	VAL	2.1
1	E	184	VAL	2.1
2	D	474	ILE	2.1
2	F	455	ASP	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	154	THR	2.0
2	F	461	ASN	2.0

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, 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	RSCC	RSR	B-factors($\mathring{A}^2$)	Q<0.9
2	B3D	F	452	9/10	0.32	0.46	87,93,178,178	0
2	B3D	D	452	9/10	0.51	0.42	107,123,215,249	0
2	B3D	B	452	9/10	0.57	0.24	75,83,182,191	0

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK