



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2021 – 07:05 PM EDT

PDB ID : 1NBU
Title : 7,8-Dihydroneopterin Aldolase Complexed with Product From Mycobacterium Tuberculosis
Authors : Goulding, C.W.; Apostol, M.I.; Sawaya, M.R.; Phillips, M.; Parseghian, A.; Eisenberg, D.; TB Structural Genomics Consortium (TBSGC)
Deposited on : 2002-12-03
Resolution : 1.60 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 ⓘ symbol.

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 : 2.23.2
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) : 2.23.2

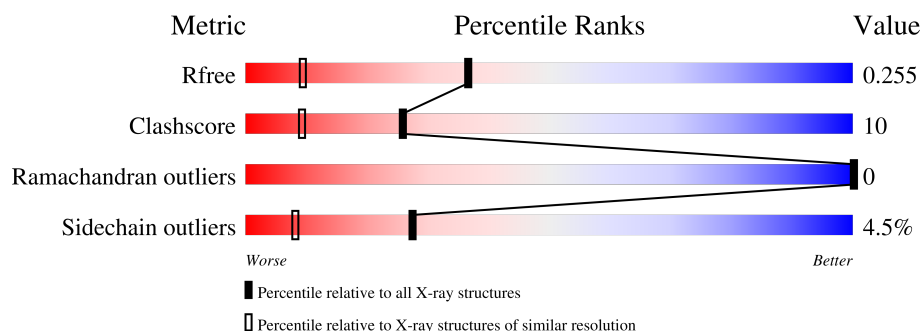
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.60 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	119	76% 22% ..
1	E	119	78% 18% ...
1	F	119	75% 18% 6% ..
1	G	119	80% 17% ..
1	H	119	80% 14% . ..
2	B	119	71% 24% . ..
2	D	119	73% 23% . ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	119	 76%21%..

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7988 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 Probable dihydroneopterin aldolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	118	Total	C	N	O	S	0	1	0
			929	577	177	174	1			
1	E	118	Total	C	N	O	S	0	2	0
			934	579	180	174	1			
1	F	118	Total	C	N	O	S	0	1	0
			929	577	177	174	1			
1	G	118	Total	C	N	O	S	0	0	0
			924	575	174	174	1			
1	H	118	Total	C	N	O	S	0	0	0
			924	575	174	174	1			

- Molecule 2 is a protein called Probable dihydroneopterin aldolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	118	Total	C	N	O	S	0	1	0
			926	576	177	172	1			
2	D	118	Total	C	N	O	S	0	0	0
			921	574	174	172	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	20	ALA	ASP	engineered mutation	UNP P0A580
D	20	ALA	ASP	engineered mutation	UNP P0A580

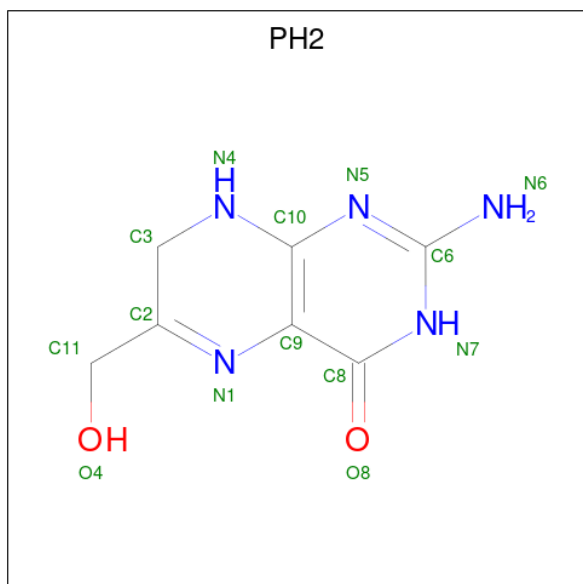
- Molecule 3 is a protein called Probable dihydroneopterin aldolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	118	Total	C	N	O	S	0	1	0
			919	570	177	171	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	19	ALA	TYR	engineered mutation	UNP P0A580
C	20	ALA	ASP	engineered mutation	UNP P0A580

- Molecule 4 is 2-AMINO-6-HYDROXYMETHYL-7,8-DIHYDRO-3H-PTERIDIN-4-ONE (three-letter code: PH2) (formula: C₇H₉N₅O₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	7	5	2		
4	B	1	Total	C	N	O	0	0
			14	7	5	2		
4	C	1	Total	C	N	O	0	0
			14	7	5	2		
4	D	1	Total	C	N	O	0	0
			14	7	5	2		
4	E	1	Total	C	N	O	0	0
			14	7	5	2		
4	F	1	Total	C	N	O	0	0
			14	7	5	2		
4	G	1	Total	C	N	O	0	0
			14	7	5	2		
4	H	1	Total	C	N	O	0	0
			14	7	5	2		

- Molecule 5 is water.

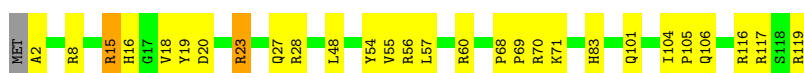
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	43	Total 43	O 43	0	0
5	B	69	Total 69	O 69	0	0
5	C	42	Total 42	O 42	0	0
5	D	52	Total 52	O 52	0	0
5	E	62	Total 62	O 62	0	0
5	F	61	Total 61	O 61	0	0
5	G	59	Total 59	O 59	0	0
5	H	82	Total 82	O 82	0	0

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. 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. 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: Probable dihydroneopterin aldolase

Chain A:  76% 22% ..



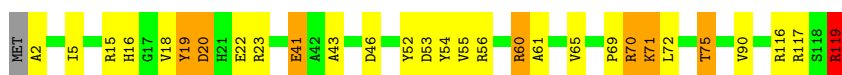
- Molecule 1: Probable dihydroneopterin aldolase

Chain E:  78% 18% ...




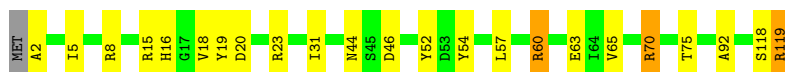
- Molecule 1: Probable dihydroneopterin aldolase

Chain F:  75% 18% 6% ..




- Molecule 1: Probable dihydroneopterin aldolase

Chain G:  80% 17% ..



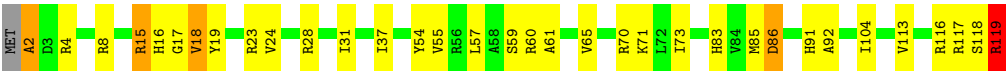
- Molecule 1: Probable dihydroneopterin aldolase

Chain H:  80% 14% . ..

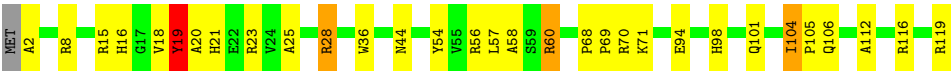


- Molecule 2: Probable dihydroneopterin aldolase

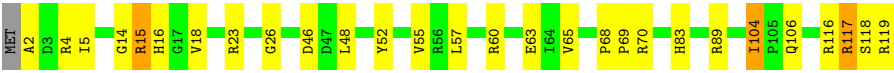
Chain B:  71% 24% . ..



● Molecule 2: Probable dihydroneopterin aldolase



● Molecule 3: Probable dihydroneopterin aldolase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	78.56Å 85.15Å 74.79Å 90.00° 112.69° 90.00°	Depositor
Resolution (Å)	10.00 – 1.60 53.61 – 1.60	Depositor EDS
% Data completeness (in resolution range)	93.8 (10.00-1.60) 96.4 (53.61-1.60)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.88 (at 1.60Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, R_{free}	0.166 , 0.257 0.214 , 0.255	Depositor DCC
R_{free} test set	5907 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	20.6	Xtriage
Anisotropy	0.192	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 57.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7988	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.42% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: PH2

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	0.39	0/952	1.20	5/1296 (0.4%)
1	E	0.44	0/963	1.28	9/1310 (0.7%)
1	F	0.42	0/952	1.24	7/1296 (0.5%)
1	G	1.15	1/941 (0.1%)	1.45	13/1282 (1.0%)
1	H	0.41	0/941	1.23	6/1282 (0.5%)
2	B	0.41	0/949	1.29	12/1292 (0.9%)
2	D	0.41	0/938	1.32	15/1278 (1.2%)
3	C	0.40	0/941	1.28	7/1281 (0.5%)
All	All	0.56	1/7577 (0.0%)	1.29	74/10317 (0.7%)

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	E	0	1
1	H	0	1
2	B	0	1
2	D	0	1
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	60	ARG	CB-CG	32.59	2.40	1.52

The worst 5 of 74 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	15	ARG	NE-CZ-NH1	-14.82	112.89	120.30
1	G	15	ARG	NE-CZ-NH2	12.28	126.44	120.30
1	G	119	ARG	NE-CZ-NH1	10.99	125.80	120.30
1	F	19	TYR	C-N-CA	10.67	148.38	121.70
2	D	8	ARG	NE-CZ-NH1	10.35	125.47	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	2	ALA	Peptide
2	D	19	TYR	Peptide
1	E	19	TYR	Peptide
1	H	19	TYR	Peptide

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	929	0	914	22	0
1	E	934	0	919	17	1
1	F	929	0	913	28	0
1	G	924	0	909	14	1
1	H	924	0	909	23	0
2	B	926	0	915	27	1
2	D	921	0	910	21	1
3	C	919	0	911	25	0
4	A	14	0	9	2	0
4	B	14	0	8	3	0
4	C	14	0	9	2	0
4	D	14	0	9	2	0
4	E	14	0	8	2	0
4	F	14	0	9	1	0
4	G	14	0	9	1	0
4	H	14	0	9	1	0
5	A	43	0	0	0	0
5	B	69	0	0	0	0
5	C	42	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	D	52	0	0	1	0
5	E	62	0	0	1	0
5	F	61	0	0	1	0
5	G	59	0	0	1	0
5	H	82	0	0	1	0
All	All	7988	0	7370	156	2

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 10.

The worst 5 of 156 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:G:60:ARG:CB	1:G:60:ARG:CG	2.40	1.00
1:E:16:HIS:HD2	1:E:70:ARG:H	1.23	0.85
1:G:16:HIS:HD2	1:G:70:ARG:H	1.23	0.85
2:D:18:VAL:HG23	2:D:19:TYR:HD1	1.42	0.84
3:C:16:HIS:HD2	3:C:70:ARG:H	1.26	0.83

All (2) 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)
2:B:19:TYR:OH	1:E:23[B]:ARG:NH1[2_656]	1.39	0.81
2:D:19:TYR:OH	1:G:23:ARG:NH2[2_555]	2.06	0.14

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	117/119 (98%)	115 (98%)	2 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	118/119 (99%)	117 (99%)	1 (1%)	0	100	100
1	F	117/119 (98%)	116 (99%)	1 (1%)	0	100	100
1	G	116/119 (98%)	116 (100%)	0	0	100	100
1	H	116/119 (98%)	116 (100%)	0	0	100	100
2	B	117/119 (98%)	116 (99%)	1 (1%)	0	100	100
2	D	116/119 (98%)	114 (98%)	2 (2%)	0	100	100
3	C	117/119 (98%)	116 (99%)	1 (1%)	0	100	100
All	All	934/952 (98%)	926 (99%)	8 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	97/97 (100%)	91 (94%)	6 (6%)	18	4
1	E	98/97 (101%)	94 (96%)	4 (4%)	30	9
1	F	97/97 (100%)	90 (93%)	7 (7%)	14	2
1	G	96/97 (99%)	93 (97%)	3 (3%)	40	15
1	H	96/97 (99%)	92 (96%)	4 (4%)	30	9
2	B	96/96 (100%)	91 (95%)	5 (5%)	23	6
2	D	95/96 (99%)	93 (98%)	2 (2%)	53	29
3	C	95/95 (100%)	91 (96%)	4 (4%)	30	9
All	All	770/772 (100%)	735 (96%)	35 (4%)	27	8

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	31	ILE
1	G	57	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	H	41	GLU
3	C	57	LEU
3	C	23	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	16	HIS
1	F	16	HIS
1	H	27	GLN
1	E	88	GLN
1	F	44	ASN

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](#)

8 ligands 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	PH2	F	720	-	10,15,15	3.92	6 (60%)	10,21,21	3.75	6 (60%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	PH2	G	520	-	10,15,15	3.97	5 (50%)	10,21,21	4.70	6 (60%)
4	PH2	A	420	-	10,15,15	3.77	6 (60%)	10,21,21	4.85	7 (70%)
4	PH2	E	820	-	10,15,15	3.72	5 (50%)	10,21,21	4.48	7 (70%)
4	PH2	C	120	-	10,15,15	3.73	5 (50%)	10,21,21	4.61	7 (70%)
4	PH2	B	320	-	10,15,15	4.05	5 (50%)	10,21,21	4.09	7 (70%)
4	PH2	D	220	-	10,15,15	3.92	6 (60%)	10,21,21	4.45	6 (60%)
4	PH2	H	620	-	10,15,15	4.04	6 (60%)	10,21,21	4.93	6 (60%)

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	PH2	F	720	-	-	0/0/11/11	0/2/2/2
4	PH2	G	520	-	-	0/0/11/11	0/2/2/2
4	PH2	A	420	-	-	0/0/11/11	0/2/2/2
4	PH2	E	820	-	-	0/0/11/11	0/2/2/2
4	PH2	C	120	-	-	0/0/11/11	0/2/2/2
4	PH2	B	320	-	-	0/0/11/11	0/2/2/2
4	PH2	D	220	-	-	0/0/11/11	0/2/2/2
4	PH2	H	620	-	-	0/0/11/11	0/2/2/2

The worst 5 of 44 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	320	PH2	C3-N4	-8.57	1.30	1.45
4	H	620	PH2	C3-N4	-8.44	1.31	1.45
4	G	520	PH2	C3-N4	-7.81	1.32	1.45
4	E	820	PH2	C3-N4	-7.62	1.32	1.45
4	D	220	PH2	C3-N4	-7.61	1.32	1.45

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	620	PH2	N6-C6-N5	9.22	131.60	117.25
4	A	420	PH2	N6-C6-N5	9.19	131.55	117.25
4	G	520	PH2	N6-C6-N5	8.61	130.64	117.25
4	H	620	PH2	N7-C6-N5	-8.37	112.29	125.42
4	C	120	PH2	N6-C6-N5	8.29	130.15	117.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	720	PH2	1	0
4	G	520	PH2	1	0
4	A	420	PH2	2	0
4	E	820	PH2	2	0
4	C	120	PH2	2	0
4	B	320	PH2	3	0
4	D	220	PH2	2	0
4	H	620	PH2	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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.