wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 07:33 PM GMT

PDB ID	:	1G50
Title	:	CRYSTAL STRUCTURE OF A WILD TYPE HER ALPHA LBD AT 2.9 ANGSTROM RESOLUTION
Authors	:	Eiler, S.; Gangloff, M.; Duclaud, S.; Moras, D.; Ruff, M.; Structural Proteomics in Europe (SPINE)
Deposited on	:	2000-10-30
Resolution	:	2.90 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

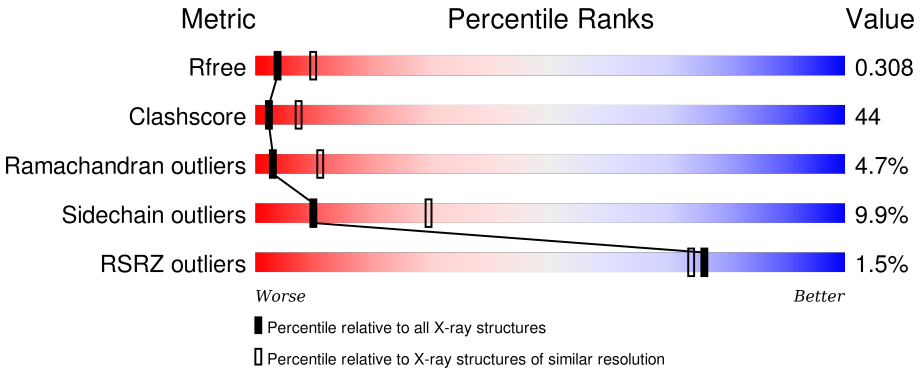
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.90 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}	91344	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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 >=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 <=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	247	29% 59% 11%
1	B	247	% 36% 57%
1	C	247	3% 37% 54% 6%

2 Entry composition

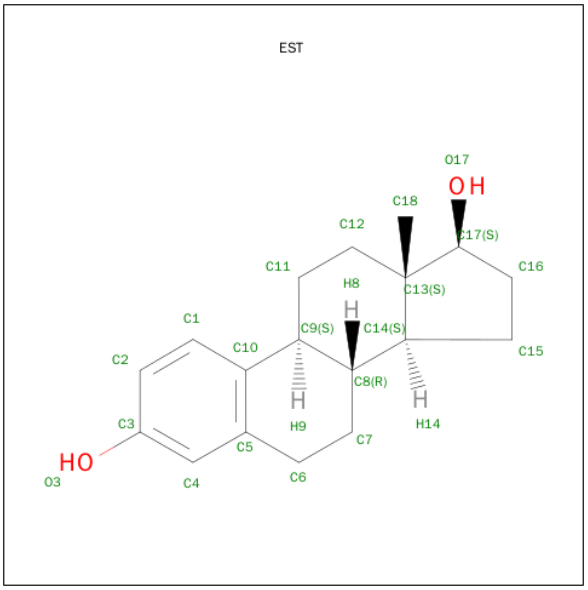
There are 3 unique types of molecules in this entry. The entry contains 6109 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 ESTROGEN RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	247	Total	C	N	O	S	0	0	0
			1975	1264	338	354	19			
1	B	246	Total	C	N	O	S	0	0	0
			1964	1258	335	352	19			
1	C	244	Total	C	N	O	S	0	0	0
			1945	1246	330	350	19			

- Molecule 2 is ESTRADIOL (three-letter code: EST) (formula: $C_{18}H_{24}O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			20	18	2		
2	B	1	Total	C	O	0	0
			20	18	2		
2	C	1	Total	C	O	0	0
			20	18	2		

- Molecule 3 is water.

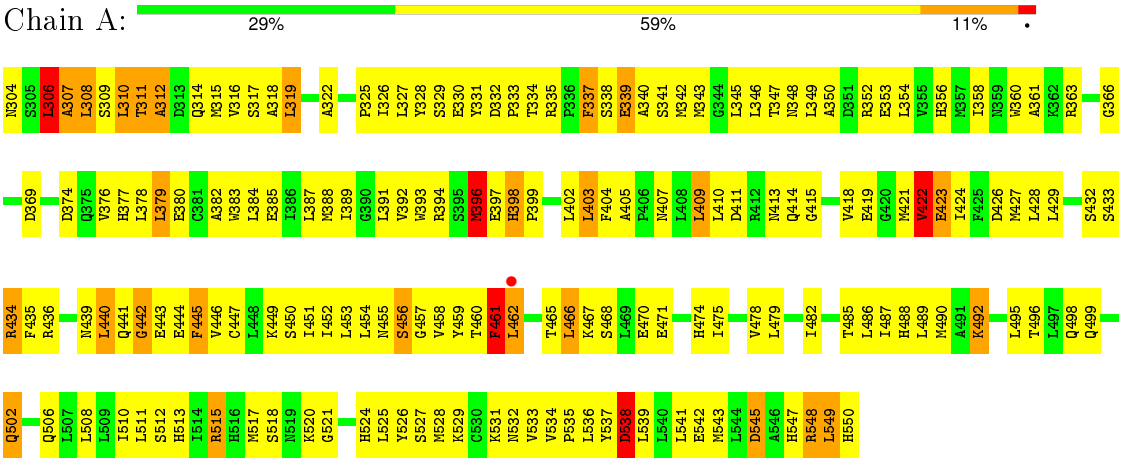
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	57	Total O 57 57	0	0
3	B	53	Total O 53 53	0	0
3	C	55	Total O 55 55	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

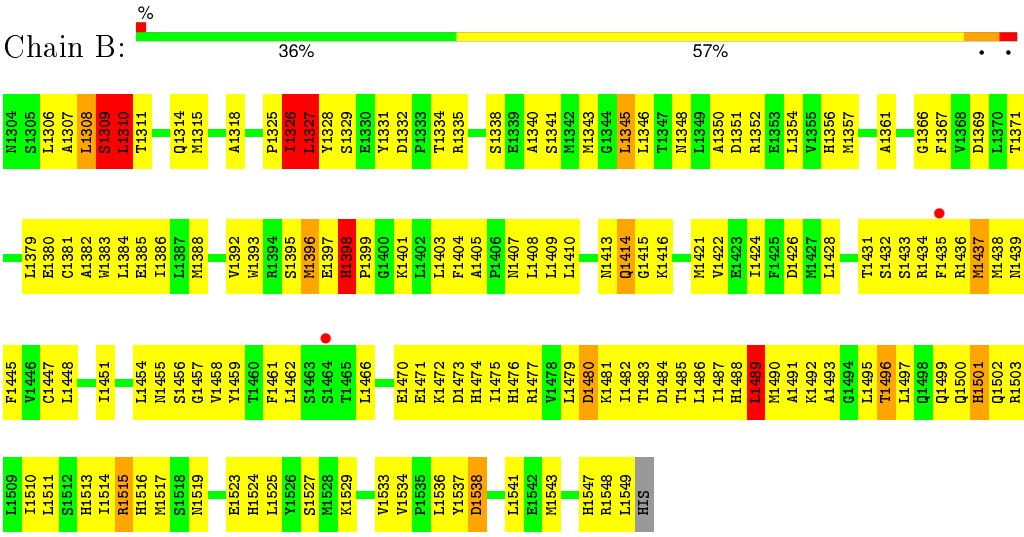
- Molecule 1: ESTROGEN RECEPTOR

Chain A:



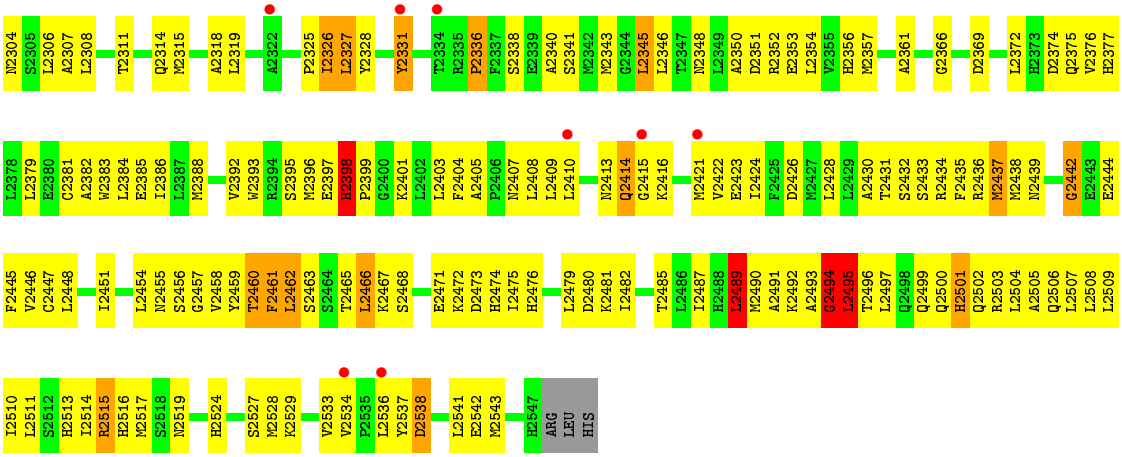
- Molecule 1: ESTROGEN RECEPTOR

Chain B:



- Molecule 1: ESTROGEN RECEPTOR

Chain C:


N2304 S2305 L2306 A2307 L2308 T2311 Q2314 M2315 A2318 L2319 A2322 P2325 I2326 L2327 Y2328 Y2331 T2334 R2335 P2336 F2337 S2338 E2339 A2340 S2341 M2342 M2343 G2344 L2345 L2346 T2347 N2348 L2349 A2350 D2351 R2352 E2353 L2354 V2355 H2356 M2357 A2361 G2366 D2369 L2372 H2373 D2374 Q2375 V2376 H2377
L2378 L2379 E2380 C2381 A2382 W2383 L2384 E2385 I2386 L2387 M2388 V2392 W2393 R2394 S2395 M2396 E2397 H2398 P2399 G2400 K2401 L2402 L2403 F2404 A2405 P2406 N2407 L2408 L2409 L2410 N2413 Q2414 G2415 K2416 M2421 V2422 E2423 I2424 F2425 D2426 M2427 L2428 L2429 A2430 T2431 S2432 S2433 R2434 F2435 R2436 M2437 M2438 N2439 G2442 E2443 E2444
F2445 V2446 C2447 L2448 I2451 L2454 N2455 S2456 G2457 V2458 Y2459 T2460 F2461 L2462 S2463 S2464 T2465 L2466 K2467 S2468 E2471 K2472 D2473 H2474 I2475 H2476 L2479 D2480 K2481 I2482 T2485 L2486 I2487 H2488 L2489 M2490 A2491 K2492 A2493 G2494 L2495 T2496 L2497 Q2498 Q2499 Q2500 H2501 Q2502 R2503 L2504 A2505 Q2506 L2507 L2508 L2509
I2510 L2511 S2512 H2513 I2514 R2515 H2516 M2517 S2518 N2519 H2524 S2527 M2528 K2529 V2533 V2534 P2535 L2536 Y2537 D2538 L2541 E2542 M2543 H2547 ARG LEU HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	105.50Å 105.50Å 136.08Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.18 – 2.90 49.18 – 2.90	Depositor EDS
% Data completeness (in resolution range)	91.0 (49.18-2.90) 91.0 (49.18-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	8.00	Depositor
$< I/\sigma(I) >$ [1]	2.61 (at 2.91Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.241 , 0.310 0.243 , 0.308	Depositor DCC
R_{free} test set	1778 reflections (9.80%)	DCC
Wilson B-factor (Å^2)	49.7	Xtriage
Anisotropy	0.700	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 43.4	EDS
Estimated twinning fraction	0.067 for -h,-k,l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Outliers	1 of 18142 reflections (0.006%)	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6109	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.88% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: EST

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	0.52	2/2013 (0.1%)	0.74	2/2720 (0.1%)
1	B	0.76	2/2001 (0.1%)	0.72	5/2705 (0.2%)
1	C	0.68	1/1982 (0.1%)	0.69	3/2680 (0.1%)
All	All	0.66	5/5996 (0.1%)	0.72	10/8105 (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	B	0	2
1	C	0	1
All	All	0	3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	2495	LEU	N-CA	26.00	1.98	1.46
1	B	1310	LEU	N-CA	23.00	1.92	1.46
1	B	1327	LEU	N-CA	18.79	1.83	1.46
1	A	548	ARG	N-CA	6.64	1.59	1.46
1	A	461	PHE	N-CA	5.12	1.56	1.46

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	2494	GLY	C-N-CA	-11.12	93.89	121.70
1	A	548	ARG	N-CA-CB	11.12	130.62	110.60
1	B	1309	SER	C-N-CA	-10.34	95.86	121.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	2495	LEU	N-CA-CB	-9.75	90.90	110.40
1	B	1310	LEU	N-CA-C	-8.91	86.94	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	1309	SER	Peptide
1	B	1326	ILE	Peptide
1	C	2494	GLY	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	1975	0	2025	206	0
1	B	1964	0	2018	175	0
1	C	1945	0	1994	170	0
2	A	20	0	24	2	0
2	B	20	0	24	2	0
2	C	20	0	24	1	0
3	A	57	0	0	10	0
3	B	53	0	0	2	0
3	C	55	0	0	5	0
All	All	6109	0	6109	525	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 44.

The worst 5 of 525 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:B:1327:LEU:CA	1:B:1327:LEU:N	1.83	1.38
1:B:1310:LEU:N	1:B:1310:LEU:CA	1.92	1.32
1:B:1327:LEU:N	1:B:1327:LEU:HD23	1.46	1.27

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:2495:LEU:CA	1:C:2495:LEU:N	1.98	1.27
1:B:1327:LEU:H	1:B:1327:LEU:CD2	1.62	1.12

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	245/247 (99%)	185 (76%)	43 (18%)	17 (7%)	1 4
1	B	244/247 (99%)	204 (84%)	31 (13%)	9 (4%)	4 17
1	C	242/247 (98%)	202 (84%)	32 (13%)	8 (3%)	5 20
All	All	731/741 (99%)	591 (81%)	106 (14%)	34 (5%)	3 11

5 of 34 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1308	LEU
1	A	306	LEU
1	A	308	LEU
1	A	339	GLU
1	A	403	LEU

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	223/223 (100%)	197 (88%)	26 (12%)	7 19
1	B	222/223 (100%)	203 (91%)	19 (9%)	13 36
1	C	220/223 (99%)	199 (90%)	21 (10%)	11 31
All	All	665/669 (99%)	599 (90%)	66 (10%)	10 29

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1345	LEU
1	B	1461	PHE
1	C	2489	LEU
1	B	1374	ASP
1	B	1414	GLN

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

Mol	Chain	Res	Type
1	B	1439	ASN
1	B	1455	ASN
1	C	2474	HIS
1	B	1359	ASN
1	B	1375	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)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

3 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	$\#\|Z\| > 2$
2	EST	A	600	-	23,23,23	2.39	9 (39%)	36,36,36	1.12	3 (8%)
2	EST	B	1600	-	23,23,23	2.47	11 (47%)	36,36,36	1.07	2 (5%)
2	EST	C	2600	-	23,23,23	2.53	9 (39%)	36,36,36	1.04	2 (5%)

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	EST	A	600	-	-	0/0/40/40	0/4/4/4
2	EST	B	1600	-	-	0/0/40/40	0/4/4/4
2	EST	C	2600	-	-	0/0/40/40	0/4/4/4

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1600	EST	C7-C6	2.04	1.56	1.52
2	B	1600	EST	C8-C14	2.08	1.57	1.53
2	B	1600	EST	C18-C13	2.08	1.58	1.54
2	B	1600	EST	C4-C3	2.09	1.42	1.39
2	B	1600	EST	C4-C5	2.21	1.43	1.39

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1600	EST	C11-C9-C8	-2.48	107.89	111.30
2	C	2600	EST	C11-C9-C8	-2.46	107.92	111.30
2	A	600	EST	C11-C9-C8	-2.40	108.01	111.30
2	C	2600	EST	C6-C7-C8	2.01	113.61	110.67
2	B	1600	EST	C16-C17-C13	2.32	106.39	104.58

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	600	EST	2	0
2	B	1600	EST	2	0
2	C	2600	EST	1	0

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, 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	247/247 (100%)	-0.30	1 (0%)	93	92	19, 45, 78, 98	0
1	B	246/247 (99%)	-0.01	2 (0%)	87	86	46, 70, 85, 95	0
1	C	244/247 (98%)	0.13	8 (3%)	50	42	47, 70, 92, 102	0
All	All	737/741 (99%)	-0.06	11 (1%)	76	74	19, 63, 89, 102	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	2334	THR	5.0
1	C	2415	GLY	2.9
1	C	2410	LEU	2.8
1	B	1464	SER	2.5
1	C	2331	TYR	2.5

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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^{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	LLDF	B-factors(Å^2)	Q<0.9
2	EST	A	600	20/20	0.97	0.22	1.48	30,33,35,39	0
2	EST	B	1600	20/20	0.96	0.21	-0.16	34,36,38,39	0
2	EST	C	2600	20/20	0.96	0.19	-0.61	31,36,38,38	0

6.5 Other polymers (i)

There are no such residues in this entry.