Description of the PDBSite fields.

Field name	Description.
ID[#]	Entry identifier.
PDBID[#]	PDB ID code.
SITE_TYPE[#]	Functional site type.
HEADER[#]	PDB classification for the entry. The field content corresponds to that in PDB.
TITLE[#]	The title for the experiment or analysis described in the entry. The field content corresponds to the one in PDB.
KEYWORD[#]	Keywords describing the macromolecule. The content corresponds to the one of KEYWDS of PDB.
MOLECULE[#]	Contains names of macromolecules from the COMPND of PDB, designed to search for entries by the names of macromolecules.
MOL_CHAINS	Identifier list of molecule chains from which the site was extracted.
NUM_SITE_CHAINS[#]	Number of different chains to which the residues of the site belong.
SITE_DESCR[#]	Description of the site. The content corresponds to the one of the SITE_DESCRIPTION sub-field of REMARK 800 field of PDB.
RESIDUE_NOTAA[#]	Contains the names of residues that are not amino acids but contained by the site.
NUMBER_OF_AA[#]	Residue number in the site and its environment.
EXPOSURE[#]	Average solvent accessibility of the site and its environment [DSSP (Kabsch & Sander, 1983)].
DISCONTINUITY[#]	Site discontinuity in the primary structure.
SITE_CHAINS	Contains the chain identifiers for each site residue.
POS	Contains the numbering of site residues and their environment. The numbering is the same as in the PDB database.
RESNAME	Contains residue names of the site and its environment in single-letter amino acid code.
2DSTRUCTURE_COMPOSITION	Describes the distribution of the secondary structure types in site and its environment. Contains the proportions for the site residue assigned to the coil, helix, sheet, bridge [DSSP (Kabsch & Sander, 1983)].
2DSTRUCTURE	Contains secondary structure type for each site residue and its environment [DSSP (Kabsch & Sander, 1983)].
ACCESSIBILITY	Contains the solvent accessibility value for each site residue and its environment [DSSP (Kabsch & Sander, 1983)].
CONTACT_NUMBER	Contains residue contact number in a protein for site residues. The contact number was calculated using the distances between the residue Cα atoms. The threshold for the distance was 10 Å.
POLAR_FRACTION	The accessible polar fraction score (Bowie

	et al, 1991).
ENVIR_CHAINS	Contains chain identifiers for each residue in the site environment.
AVERAGE	Contains the mean values for the physicochemical characteristics of the site and its environment. The physicochemical characteristics are ordered as follows: hydropathy (Kyte & Doolittle, 1982), hydrophilicity (Hopp & Woods, 1983), hydrophobicity (Eisenberg, 1984), charge (Bogardt et al, 1980), isoelectric point (Bogardt et al, 1980), polarity (Bogardt et al, 1980), volume (Bogardt et al, 1980), hydrophobicity (Bogardt et al, 1980).
SPATIAL_MOMENT	Contains the spatial moment values for the physicochemical characteristics of the site and its environment. The same physicochemical characteristics as in the AVERAGE field.
ATOM	Contains atomic coordinates of site residues. The values are extracted from the PDB database without changes.
END	End of the entry.

Remark: # the field can be used for search queries in SRS

References

Bogardt,R.A., Jones B.N., Dwulet F.E., Garner W.H., Lehman L.D. and Gurd F.R. (1980) Evolution of the amino acid substitution in the mammalian myoglobin gene. *J Mol Evol*, **15**, 197–218.

Bowie,J.U., Luthy,R. and Eisenberg,D. (1991) A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, **253**, 164-170.

Eisenberg,D., Schwarz,E., Komaromy,M. and Wall,R. (1984) Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Mol Biol*, **179**,125-42.

Hopp,T.P. and Woods,K.R. (1983) A computer program for predicting protein antigenic determinants. *Mol. Immunol*. **20**, 483–489.

Kabsch,W. and Sander,C. (1983) Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577-637.

Kyte,J. and Doolittle,R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol*, **157**, 105-132.