PubChem generates SD formatted data files using specific SD fields. The SD file format is described elsewhere and will not be described here. To learn more about the SD file format, go to: http://download.accelrys.com/freeware/ctfile-formats/ctfile-formats.zip

Provided below is documentation on the SD fields generated by PubChem. Please note that the SD fields may evolve over time as we modify, change, or revise the PubChem Data System. Please also note that PubChem SD fields may have data with text line lengths greater than 200 characters.

PubChem Substance Associated SD Fields

PUBCHEM_SUBSTANCE_ID
: PubChem Substance ID (SID) is the non-zero unsigned integer PubChem accession ID for a deposited substance. For example, "101".

PUBCHEM_SUBSTANCE_VERSION
: PubChem Substance ID (SID) version is the non-zero unsigned integer counter used to distinguish between and track changes in a deposited substance. For example, "1".

PUBCHEM_EXT_DATASOURCE_NAME
: External Source/Database Name assigned by PubChem to a Depositor. For example, "KEGG" stands for the Kyoto Encyclopedia of Genes and Genomes. For a list of data source name descriptions, go to "http://pubchem.ncbi.nlm.nih.gov/sources/sources.cgi".

PUBCHEM_EXT_DATASOURCE_REGID
: External Registry ID for a Substance that is unique for all substances provided by a Depositor. For example, "Plant Extract 451".

PUBCHEM_EXT_DATASOURCE_URL
: The Universal Resource Locator (URL) home page associated with the Depositor's External DataSource Name for a provided Substance. For example, "http://chem.sis.nlm.nih.gov/chemidplus/".

PUBCHEM_EXT_SUBSTANCE_URL
: The Universal Resource Locator (URL) associated with the Depositor's External Registry ID for a provided Substance. This link provides additional information about this substance at the depositor's website. For example, "http://chem.sis.nlm.nih.gov/chemidplus/direct.jsp?result=advanced&regno=003663829".

PUBCHEM_SUBSTANCE_SYNONYM
: Synonyms associated with a Substance provided by the Depositor, one per line. For example, "3-Methoxy-4-hydroxybenzalacetone".

PubChem Substance Associated SD Fields (cont.)

PUBCHEM_SUBSTANCE_COMMENT
Comments are textual annotations about a Substance provided by the Depositor. For example, "Is a reactant or product of enzyme EC 3.5.1.33".

PUBCHEM_DEPOSITOR_RECORD_DATE
Depositor Substance Record Date is the creation or modification date provided by the Depositor. The format of this date is modeled after the international standard date notation (ISO 8601) as either a complete date (YYY-MM-DD, e.g., "1997-07-16") or a complete date with hours and minutes (YYYY-MM-DDThh:mmTZD, where TZD is the time zone designator, Z or +hh:mm or -hh:mm, e.g., "1997-07-16T19:20+01:00").

PUBCHEM_XREF_EXT_ID
External Registry Name/IDs associated with a Substance provided by the Depositor, one per line. For example, "NSC 24520".

PUBCHEM_XREF_SUBSTANCE_ID
PubChem Substance IDs (SID) associated with a Substance provided by the Depositor, one per line. For example, "2345".

PUBCHEM_XREF_COMPOUND_ID
PubChem Compound IDs (CID) associated with a Substance provided by the Depositor, one per line. For example, "1234".

PUBCHEM_XREF_ASSAY_ID
PubChem BioAssay IDs (AID) associated with a Substance provided by the Depositor, one per line. For example, "169".

PUBCHEM_PUBMED_ID
NCBI/NLM/NIH PubMed IDs for an article or abstract associated with a Substance, one per line. For example, "12546981".

PUBCHEM_GENBANK_NUCLEOTIDE_ID
NCBI/NLM/NIH GenBank Nucleotide sequence IDs associated with a Substance, one per line. For example, "14567".

PUBCHEM_GENBANK_PROTEIN_ID
NCBI/NLM/NIH GenBank Protein sequence IDs associated with a Substance, one per line. For example, "2845".

PUBCHEM_NCBI_TAXONOMY_ID
NCBI/NLM/NIH Taxonomy IDs associated with a Substance by the Depositor, one per line. For example, "75846".

PubChem Substance Associated SD Fields (cont.)

PUBCHEM_NCBI_OMIM_ID
NCBI/NLM/NIH Online Mendelian Inheritance in Man (OMIM) IDs associated with a Substance by the Depositor, one per line. For example, "123".

PUBCHEM_NCBI_MMDB_ID
NCBI/NLM/NIH MMDB Structure IDs associated with a Substance by the Depositor, one per line. For example, "11993".

PUBCHEM_PUBMED_MESH_TERM
NLM/NIH Medical Subject Heading Terms associated with a Substance, one per line. For example, "Aspirin".

PUBCHEM_NCBI_GENE_ID
NCBI/NLM/NIH Entrez Gene IDs associated with a Substance by the Depositor, one per line. For example, "113".

PUBCHEM_NCBI_PROBE_ID
NCBI/NLM/NIH Entrez Probe IDs associated with a Substance by the Depositor, one per line. For example, "7".

PUBCHEM_NCBI_BIOSYSTEM_ID
NCBI/NLM/NIH Entrez BioSystem IDs associated with a Substance by the Depositor, one per line. For example, "22".

PUBCHEM_NCBI_GEO_GSE_ID
NCBI/NLM/NIH Entrez Gene Expression Omnibus Series Accession (GEO GSE) IDs associated with a Substance by the Depositor, one per line. For example, "22".

PUBCHEM_NCBI_GEO_GSM_ID
NCBI/NLM/NIH Entrez Gene Expression Omnibus Sample Accession (GEO GSM) IDs associated with a Substance by the Depositor, one per line. For example, "22".

PUBCHEM_MMDB_MOLECULE_NAME
The MMDB name(s) in an MMDB structure associated with the corresponding Substance, one per line.

PUBCHEM_MMDB_RESIDUE_ID
The MMDB residue ID in an MMDB structure associated with the corresponding Substance, one per line.

PUBCHEM_MMDB_RESIDUE_NAME
The MMDB residue name(s) in an MMDB structure associated with the corresponding Substance, one per line.

PubChem Substance Associated SD Fields (cont.)

PUBCHEM_MMDB_ATOM_ID
: The MMDB atom ID in an MMDB structure associated with the corresponding Substance, one per line.

PUBCHEM_MMDB_ATOM_NAME
: The MMD atom name(s) in an MMDB structure associated with the corresponding Substance, one per line.

PUBCHEM_HOLD_UNTIL_DATE
: Indicates a substance was not immediately made public. The date provided by this tag indicates when the substance was designated to be publically accessible.

PUBCHEM_SUBS_AUTO_STRUCTURE
: Indicates a substance was designated to have the deposited chemical record auto-generated by PubChem using provided synonyms. Text associated with this tag indicates the outcome of such processing.

PUBCHEM_REVOKE_SUBSTANCE
: Revoke Substance Record provided by the depositor with a comment as to why this substance was removed or suppressed from their data collection. For example, "Substance is no longer in stock".

PUBCHEM_CID_ASSOCIATIONS
: Associations between a substance SID and compound CIDs. A given substance, if standardized, is linked to one or more compounds listed in this SD field. The format for each line is an integer CID followed by an integer CompoundIdType, which has the same meaning as in the PUBCHEM_COMPOUND_ID_TYPE field described above. For example, "9876 3" indicates that the substance has CID 9876 as its neutralized form. Please note that, when this field is present, the Substance was standardized and there will be at least a single CID associated with the "Standardized Form". Additionally, there may be either a single "Neutralized Form" or one or more "Component of the Standardized Form".

PUBCHEM_GENERIC_REGISTRY_NAME
: Generic Registry Name/IDs associated with a Substance provided by the Depositor, one per line. For example, "1080-12-2".

PubChem Compound Associated SD Fields

PUBCHEM_COMPOUND_CID
PubChem Compound ID (CID) is the non-zero unsigned integer PubChem accession ID for a unique chemical structure. For example, "4231".

PUBCHEM_COMPOUND_ID_TYPE
PubChem Compound ID type qualifies the relationship between a compound CID and a substance SID. For example, a compound may be a component of a substance.

PubChem Compound ID type is in the format:
CompoundIdType

where CompoundIdType is an unsigned integer number.

CompoundIdType	Meaning
0	Deposited Compound
1	Standardized Form of the Deposited Compound
2	Component of the Standardized Form
3	Neutralized Form of the Standardized Form
4	Deposited Mixture Component
5	Alternate Tautomer Form of the Standardized Form
6	Ionized pKa Form of the Standardized Form
255	Unspecified or Unknown Compound Type

PUBCHEM_COORDINATE_TYPE
Coordinate Type provides additional information about the coordinates for a Compound. Only a single coordinate type qualifier is provided per line. For example, for 2D coordinates computed by PubChem with unspecified coordinate units:
1
5
255

CoordinateType	Meaning
1	2D Coordinates
2	3D Coordinates
3	Depositor provided coordinates
4	Experimentally determined coordinates
5	Coordinates were computed by PubChem
10	3D Coordinate units in Angstroms
11	3D Coordinate units in nanometers
12	2D Coordinate units in pixels
13	2D Coordinate units in points
14	2D Coordinate units in standard bond lengths (1.0)
255	Coordinate units are unknown or unspecified

PubChem Compound Associated SD Fields (cont.)

PUBCHEM_MOLECULAR_FORMULA
Calculated molecular formula in Hill-format for a Compound. For example, "C6H10N2O3S".

PUBCHEM_MOLECULAR_WEIGHT
Calculated molecular weight for a Compound, using atomic masses averaged according to naturally occurring abundances. For example, "204.243519".

PUBCHEM_MONOISOTOPIC_WEIGHT
Calculated monoisotopic molecular weight for a Compound, using only the masses of the most abundant naturally occurring isotopes or, if man-made, most stable isotopes. For example, "204.123583".

PUBCHEM_EXACT_MASS
Calculated exact mass corresponding to the most intense peak observed in high resolution mass spectrometers for a Compound. For example, "204.123583".

PUBCHEM_NONSTANDARDBOND
Non-Standard Bonds specify additional bonding information for a chemical structure that may not be readily encoded in the SD file format. The atoms involved in a non-standard bond do not have to have a bond record in the SD file format. If the atoms are already bonded in the SD file format, the non-standard bonds provided using this SD tag will supersede the bonding information encoded in the SD file format. Only a single Non-Standard Bond may be provided per line. Multiple Non-Standard Bonds may be provided for a chemical structure.

The format for a Non-Standard Bond is three non-zero unsigned numbers, separated by white-space, representing the AtomIDs of the two atoms, followed by the BondTypeID, respectively:
AtomID AtomID BondTypeID

For example, "8 4 5" is a dative bond between atoms 8 and 4.

BondTypeID	Meaning
1	Single Bond
2	Double Bond
3	Triple Bond
4	Quadruple Bond
5	Dative Bond
6	Complex Bond
7	Ionic Bond
255	Unspecified or Unknown Connectivity

PubChem Compound Associated SD Fields (cont.)

PUBCHEM_BONDANNOTATIONS
Bond Annotations allow additional structural annotations in a Compound to be provided between atoms. The atoms do not have to be explicitly bonded in the SD file format to have a bond annotation. Only a single Bond Annotation is provided per line. Multiple Bond Annotations may be provided for a chemical structure.

The format for a Bond Annotation is three non-zero unsigned integer numbers separated by white-space, representing the AtomIDs of the two atoms, followed by the AnnotationID, respectively:
AtomID AtomID AnnotationID

For example, "23 45 4" is a dotted line bond annotation between atoms 23 and 45.

AnnotationID	Meaning
1	Crossed Bond, a non-specific stereo double bond
2	Dashed Bond, a 3-D hydrogen bond
3	Wavy Bond, a non-specific stereo single bond
4	Dotted Bond, a complex or fractional bond
5	Wedge-up Bond, a solid wedge stereo bond
6	Wedge-down Bond, a dashed wedge stereo bond
7	Arrow Bond, a dative bond
8	Aromatic Bond, an aromatic bond
9	Resonance Bond, a resonating bond
10	Bold Bond, a thick bond
11	Fischer Bond, use Fischer stereo conventions
12	Close Contact, a 3-D atom-atom close contact
255	Unspecified or Unknown Atom-Atom Annotation

PUBCHEM_COMPOUND_CANONICALIZED
Boolean flag, indicated by a one or a zero, denoting whether the compound was subjected to valence-bond canonicalization procedure. Not all valid compounds can be subjected to this procedure. For example, "1".

PUBCHEM_TOTAL_CHARGE
Calculated total formal charge of a Compound. For example, "-1".

PUBCHEM_COMPONENT_COUNT
Count of covalently bonded units in a Compound. For example, "4".

PUBCHEM_HEAVY_ATOM_COUNT
Count of heavy atoms, elements with atomic number greater than one in a Compound. For example, "23".

PubChem Compound Associated SD Fields (cont.)

PUBCHEM_ISOTOPIC_ATOM_COUNT
Count of atoms with isotopic enrichment, that which is different from that found in nature, in a Compound. For example, "0".

PUBCHEM_ATOM_DEF_STEREO_COUNT
Count of the number of defined atom stereo centers in a Compound. For example, "0".

PUBCHEM_ATOM_UDEF_STEREO_COUNT
Count of the number of undefined atom stereo centers in a Compound. For example, "1".

PUBCHEM_BOND_DEF_STEREO_COUNT
Count of the number of defined stereo bonds in a Compound. For example, "2".

PUBCHEM_BOND_UDEF_STEREO_COUNT
Count of the number of undefined stereo bonds in a Compound. For example, "3".

PUBCHEM_CACTVS_TAUTO_COUNT
Count, to a maximum value of 1,000, of different tautomeric forms of a Compound, using the Xemistry GmbH's Cactvs implementation. For example, "2".

PUBCHEM_CACTVS_HBOND_ACCEPTOR
Calculated count of the number of Hydrogen Bond Acceptors in a Compound, using the Xemistry GmbH's Cactvs implementation. For example, "1".

PUBCHEM_CACTVS_HBOND_DONOR
Calculated count of the number of Hydrogen Bond Donors in a Compound, using the Xemistry GmbH's Cactvs implementation. For example, "0".

PUBCHEM_CACTVS_ROTATABLE_BOND
Calculated count of the number of rotatable bonds in a Compound, using the Xemistry GmbH's Cactvs implementation. For example, "3".

PUBCHEM_CACTVS_COMPLEXITY
Complexity measure, a floating point number, calculated for a Compound where less complex and more symmetrical molecules have smaller values and more complex and less symmetrical molecules have larger values, using the Xemistry GmbH's Cactvs implementation. For example, "921.222".

PubChem Compound Associated SD Fields (cont.)

PUBCHEM_CACTVS_SUBSKEYS
: Substructure Fingerprint calculated for a Compound using the Xemistry GmbH's Cactvs implementation. This is base64 encoded binary data where the first four binary bytes give the length of the fingerprint bit list. The description of the individual bits is provided elsewhere. To learn more about the PubChem substructure fingerprint, go to "ftp://ftp.ncbi.nlm.nih.gov/pubchem/specifications/pubchem_fingerprint.txt".

PUBCHEM_CACTVS_TPSA
: Topological Polar Surface Area (TPSA) measure, a floating point number, calculated for a Compound where the procedure is based on the summation of tabulated surface contributions of polar fragments (http://www.ncbi.nlm.nih.gov/pubmed/11020286), using the Xemistry GmbH's Cactvs implementation. For example, "52.93".

PUBCHEM_XLOGP3
: Calculated Log P for the relative solubility of a Compound in water and octanol, using the native implementation (http://www.sioc-ccbg.ac.cn/software/xlogp3/) of the XLogP v3.0 method (http://www.ncbi.nlm.nih.gov/pubmed/17985865). For example, "0.9".

PUBCHEM_XLOGP3_AA
: Calculated Log P for the relative solubility of a Compound in water and octanol, using the native implementation (http://www.sioc-ccbg.ac.cn/software/xlogp3/) of the XLogP v3.0 method (http://www.ncbi.nlm.nih.gov/pubmed/17985865). For example, "3.4".

PUBCHEM_IUPAC_INCHI
: Calculated InChI string for a Compound using the IUPAC standard InChI implementation (http://www.iupac.org/inchi/). For example, "InChI=1S/C9H8O4/c1-6(10)13-8-5-3-2-4-7(8)9(11)12/h2-5H,1H3,(H,11,12)".

PUBCHEM_IUPAC_INCHIKEY
: Calculated InChIKey string for a Compound using the IUPAC standard InChIKey implementation (http://www.iupac.org/inchi/). For example, "BSYNRYMUTXBXSQ-UHFFFAOYSA-N".

PUBCHEM_OPENEYE_CAN_SMILES
: Calculated Canonical SMILES for a Compound representation using OpenEye, Inc.'s OEChem implementation. For example, "C1CCN(C1)CNC(=O)C2=NC=CN=C2".

PubChem Compound Associated SD Fields (cont.)

PUBCHEM_OPENEYE_ISO_SMILES
: Calculated Isomeric SMILES for a Compound representation using OpenEye, Inc.'s OEChem implementation. For example, "CCOC(=O)[C@H](CSCC)N".

PUBCHEM_IUPAC_NAME
: Calculated IUPAC Preferred Name-variant, based on the latest IUPAC standard, for a Compound using OpenEye, Inc.'s LexiChem implementation and the IUPAC-style setting. For example, "ethyl 2-amino-3-ethylsulfanyl-propanoate".

PUBCHEM_IUPAC_OPENEYE_NAME
: Calculated IUPAC Acceptable Name-variant, using a previous IUPAC standards, for a Compound using OpenEye, Inc.'s LexiChem implementation and the OpenEye-style setting. For example, "1-acetyl-2-sulfanyl-5-(2-thioxobutyl)-1,5-dihydroimidazol-4-one".

PUBCHEM_IUPAC_CAS_NAME
: Calculated IUPAC CAS-style Name-variant, typically calculated by other IUPAC naming applications, for a Compound using OpenEye, Inc.'s LexiChem implementation and the CAS-style setting. For example, "ethyl 2-amino-3-ethylthio-propanoate".

PUBCHEM_IUPAC_TRADITIONAL_NAME
: Calculated IUPAC Name-variant, using a more traditional name, for a Compound using OpenEye, Inc.'s LexiChem implementation and the Traditional-style setting. For example, "2-amino-3-ethylthio-propanoic acid ethyl ester".

PUBCHEM_IUPAC_SYSTEMATIC_NAME
: Calculated IUPAC Name-variant, using a more systematic naming approach that attempts to be predictive of where future IUPAC naming conventions are headed, for a Compound using OpenEye, Inc.'s LexiChem implementation and the Systematic-style setting. For example, "3-ethylsulfanyl-2-methanoylamino-propanoic acid".

PUBCHEM_OPENEYE_TAUTO_COUNT
: Count, to a maximum value of 1,000, of different tautomeric forms of a Compound, using the OpenEye, Inc.'s QuACPAC implementation. For example, "5".

PubChem3D Compound Associated SD Fields

PUBCHEM_CONFORMER_ID_LIST
List of local conformer IDs (LIDs) for a Compound. One Conformer ID is provided per line. See also documentation for PUBCHEM_CONFORMER_ID.

PUBCHEM_CONFORMER_DIVERSEORDER
List of local conformer IDs (LIDs) for a Compound in diverse order, where the diversity metric used is the sum of shape and feature similarity and where the first LID is the lowest energy conformer used by default by PubChem. One local conformer ID is provided per line. See also documentation for PUBCHEM_CONFORMER_ID.

PUBCHEM_PHARMACOPHORE_FEATURES
Calculated pharmacophore features computed using the OpenEye, Inc.'s OEShape implementation and post processed to merge common proximate features (with 1 Angstrom). The first line is the count of features. Each subsequent line contains a feature description, which includes a count of atoms that comprise the feature, atom IDs, and a feature type, each space separated. The feature type can be either "rings", "anion", "cation", "acceptor", "donor", or "hydrophobe". For example:
4
1 1 acceptor
3 2 3 12 anion
3 15 17 20 hydrophobe
5 6 7 8 10 11 rings

PUBCHEM_EFFECTIVE_ROTOR_COUNT
Calculated count of the effective number of rotatable bonds of a Compound, which takes into account the flexibility of rings. For example, "5.6".

PUBCHEM_MMFF94_PARTIAL_CHARGES
Calculated partial charges computed with the MMFF94 (Merck Molecular Force Field) using the OpenEye, Inc.'s OEChem implementation. The first line is the count of non-zero partial charges. Each subsequent line contains an atom ID and partial charge, space separated. For example:
1
1 -0.28

PUBCHEM_CONFORMER_RMSD
Calculated minimum RMSD (root-mean-squared-deviation) spacing between Conformer coordinates in a conformer model for a Compound. For example, "0.8".

PubChem3D Compound Conformer Associated SD Fields

PUBCHEM_CONFORMER_ID
PubChem Conformer ID is a global non-zero 64-bit unsigned integer PubChem accession ID for a conformer (substance or compound). Three numbers are stored in these 64-bits: the (compound or substance) identifier, (substance) version, and local conformer ID (LID). The bit masks for the identifier, version, and LID are: 0xFFFFFFFF00000000, 0x00000000FFFF0000, and 0x000000000000FFFF. If the version is zero, the identifier refers to a compound, else it refers to a substance. This conformer ID is stored as a 16-character hex-encoded number. For example, "0103E23500000026".

PUBCHEM_SHAPE_VOLUME
Calculated steric volume of a Conformer of a Compound using the OpenEye, Inc.'s OEShape implementation. For example, "100.3".

PUBCHEM_MMFF94_ENERGY
Calculated MMFF94s (Merck Molecular Force Field) energy computed without electrostatic (coulombic) terms for a Conformer of a Compound using the OpenEye, Inc.'s CASE implementation. For example, "-190.1234".

PUBCHEM_SHAPE_MULTIPOLES
Calculated zero order, first order and second order moments (monopole, quadrupoles and octopoles, respectively) of the steric shape of a Conformer of a Compound using the OpenEye, Inc.'s OEShape implementation. Each value is a floating point number, one per line, with the first line being the monopole, the next three lines being the quadrupoles, in the order X, Y, Z, and the next ten being the octopoles, in the order XXX, YYY, ZZZ, XXY, XXZ, YYX, YYZ, ZZX, ZZY, and XYZ.

PUBCHEM_SHAPE_FINGERPRINT
Computed shape fingerprint of a conformer for rapid pre-alignment of 3D shapes (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1994057). A description of the specific reference shapes used is found elsewhere (ftp://ftp.ncbi.nlm.nih.gov/pubchem/Compound_3D/FP/). Each line is comprised of space-separated reference shape CID, LID, and a 64-bit unsigned number, representing the packed rotational/translational matrix to align the reference shape to the conformer.

PUBCHEM_SHAPE_SELFOVERLAP
Computed self-overlap of a conformer (Grant and Pickup, J. Phys. Chem. 1995, 99, 3503) ignoring hydrogen atoms and using Bondi Van der Waals radii.

PubChem3D Compound Conformer Associated SD Fields (cont.)

PUBCHEM_FEATURE_SELFOVERLAP
Computed self-overlap of conformer fictitious feature atoms (radius 1.08265), whose 3D location are the steric center of the atoms they comprise. See also documentation for PUBCHEM_SHAPE_SELFOVERLAP and PUBCHEM_PHARMACOPHORE_FEATURES.

Document Version History

V2.0.1 – 2011Oct17 - Updated URL to SD file format document.
V2.0.0 - 2010Aug12 - Added new fields PUBCHEM_NCBI_GEO_GSE_ID, PUBCHEM_NCBI_GEO_GSM_ID, PUBCHEM_HOLD_UNTIL_DATE, PUBCHEM_SUBS_AUTO_STRUCTURE, and PUBCHEM_SHAPE_FINGERPRINT; modified description of PUBCHEM_PHARMACOPHORE_FEATURES; and renamed fields PUBCHEM_IUPAC_INCHI and PUBCHEM_IUPAC_INCHIKEY (replaced NIST moniker for IUPAC). Organized SD tags into separate sections.
V1.8.0 - 2010May13 - Added new fields PUBCHEM_CONFORMER_DIVERSEORDER, PUBCHEM_SHAPE_SELFOVERLAP, and PUBCHEM_FEATURE_SELFOVERLAP and modified description of PUBCHEM_CONFORMER_ID, PUBCHEM_CONFORMER_ID_LIST, and PUBCHEM_SHAPE_MULTIPOLES.
V1.7.0 - 2009Jul23 - Added new fields PUBCHEM_NCBI_BIOSYSTEM_ID, PUBCHEM_COORDINATE_TYPE and PUBCHEM_PHARMACOPHORE_FEATURES and modified description of fields PUBCHEM_MMFF94_PARTIAL_CHARGES and PUBCHEM_EFFECTIVE_ROTOR_COUNT.
V1.6.0 - 2009Feb04 - Replaced field PUBCHEM_CACTVS_XLOGP with PUBCHEM_XLOGP3_AA, added new fields PUBCHEM_XLOGP3 and PUBCHEM_NIST_INCHIKEY, and modified the description of PUBCHEM_CACTVS_TPSA, PUBCHEM_NIST_INCHI, and PUBCHEM_CONFORMER_ID.
V1.5.0 - 2007Nov29 - Added new fields PUBCHEM_CONFORMER_RMSD, PUBCHEM_MMFF94_ENERGY, PUBCHEM_CONFORMER_ID, PUBCHEM_SHAPE_VOLUME, PUBCHEM_CONFORMER_ID_LIST, PUBCHEM_SHAPE_MULTIPOLES, PUBCHEM_EFFECTIVE_ROTOR_COUNT, and PUBCHEM_MMFF94_PARTIAL_CHARGES. Minor changes including wording of PUBCHEM_COMPOUND_CANONICALIZED.
V1.4.0 - 2007Sep04 - Changed the name and updated the description of SD fields for molecular formula, molecular weight, and monoisotopic weight. Updated the exact mass and xLogP SD field descriptions.
V1.3.2 - 2007Jul09 - Added new fields PUBCHEM_NCBI_PROBE_ID and PUBCHEM_COMPOUND_CANONICALIZED.
V1.3.1 - 2006Mar31 - Modified wording of "PUBCHEM_NIST_INCHI".
V1.3.0 - 2006Mar23 - Added new fields PUBCHEM_OPENEYE_MONOISOTOPICWT and PUBCHEM_CACTVS_EXACT_MASS. Numerous edits of existing descriptions.
V1.2.0 - 2005Nov30 - Added new fields PUBCHEM_NCBI_OMIM_ID, PUBCHEM_REVOKE_SUBSTANCE, PUBCHEM_DEPOSITOR_RECORD_DATE, PUBCHEM_CACTVS_TPSA, PUBCHEM_CACTVS_SUBSKEYS. Additional clarification added for some SD fields. Minor corrections made to reflect 3rd-party changes in property implementation(s). Minor reorganization of the SD fields.

Document Version History (cont.)

V1.1.0 - 2005Jun10 - Added new fields PUBCHEM_CID_ASSOCIATIONS, PUBCHEM_MMDB_MOLECULE_NAME, PUBCHEM_MMDB_RESIDUE_ID, PUBCHEM_MMDB_RESIDUE_NAME, PUBCHEM_MMDB_ATOM_ID, PUBCHEM_MMDB_ATOM_NAME.
V1.0.0 - 2005Apr20 - Initial release.