

Bioinformatics, 31(17), 2015, 2860–2866

doi: 10.1093/bioinformatics/btv285

Advance Access Publication Date: 5 May 2015

Original Paper

OXFORD

Databases and ontologies

The SwissLipids knowledgebase for lipid biology

Lucila Aimo^{1,†}, Robin Liechti^{2,†}, Nevila Hyka-Nouspikel¹,
Anne Niknejad¹, Anne Gleizes¹, Lou Götz², Dmitry Kuznetsov²,
Fabrice P.A. David^{3,4}, F. Gisou van der Goot⁵, Howard Riezman^{6,7},
Lydie Bougueleret¹, Ioannis Xenarios^{1,2,6,8} and Alan Bridge^{1,*}

¹Swiss-Prot Group, SIB Swiss Institute of Bioinformatics, CMU, 1 rue Michel-Servet, CH-1211 Geneva 4, Switzerland, ²Vital-IT, SIB Swiss Institute of Bioinformatics, Quartier Sorge, Bâtiment Génopode, CH-1015 Lausanne, Switzerland, ³Bioinformatics and Biostatistics Core Facility, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, ⁴SIB Swiss Institute of Bioinformatics, CH-1015 Lausanne, Switzerland, ⁵Global Health Institute, École Polytechnique Fédérale de Lausanne, Station 19, CH-1015 Lausanne, Switzerland, ⁶Department of Biochemistry, University of Geneva, CH-1211 Geneva, Switzerland, ⁷Switzerland National Centre of Competence in Research “Chemical Biology”, University of Geneva, CH-1211 Geneva, Switzerland and ⁸Centre for Integrative Genomics, University of Lausanne, CH-1015 Lausanne, Switzerland

*To whom correspondence should be addressed.

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

Associate Editor: John Hancock

Received on March 3, 2015; revised on April 28, 2015; accepted on April 29, 2015

Abstract

Motivation: Lipids are a large and diverse group of biological molecules with roles in membrane formation, energy storage and signaling. Cellular lipidomes may contain tens of thousands of structures, a staggering degree of complexity whose significance is not yet fully understood. High-throughput mass spectrometry-based platforms provide a means to study this complexity, but the interpretation of lipidomic data and its integration with prior knowledge of lipid biology suffers from a lack of appropriate tools to manage the data and extract knowledge from it.

Results: To facilitate the description and exploration of lipidomic data and its integration with prior biological knowledge, we have developed a knowledge resource for lipids and their biology—SwissLipids. SwissLipids provides curated knowledge of lipid structures and metabolism which is used to generate an *in silico* library of feasible lipid structures. These are arranged in a hierarchical classification that links mass spectrometry analytical outputs to all possible lipid structures, metabolic reactions and enzymes. SwissLipids provides a reference namespace for lipidomic data publication, data exploration and hypothesis generation. The current version of SwissLipids includes over 244 000 known and theoretically possible lipid structures, over 800 proteins, and curated links to published knowledge from over 620 peer-reviewed publications. We are continually updating the SwissLipids hierarchy with new lipid categories and new expert curated knowledge.

Availability: SwissLipids is freely available at <http://www.swisslipids.org/>.

Contact: alan.bridge@isb-sib.ch

Supplementary information: [Supplementary data](#) are available at *Bioinformatics* online.

1 Introduction

Lipids are a large and diverse group of biological molecules that perform a wide variety of biological functions (Sud *et al.*, 2007). They comprise the bulk of membrane bilayers (Holthuis and Menon, 2014; van Meer *et al.*, 2008) and energy stores (Lass *et al.*, 2011), provide markers for the recognition and sorting of distinct organelles (Balla, 2013), and regulate processes such as cell proliferation and death (Morad and Cabot, 2013), immunity (Knowlden and Georas, 2014; Simmons and Ishii, 2014), angiogenesis, cardiogenesis and neurogenesis (Mendelson *et al.*, 2014) and inflammation (Ji *et al.*, 2011). The lipid complement or ‘lipidome’ of individual cells may include hundreds of thousands of lipid structures whose occurrence is tightly regulated in response to changes in the cellular environment and lipid availability (Gross and Han, 2011; Shevchenko and Simons, 2010). Dysregulation of the lipidome is a feature of many pathological conditions including cardiovascular disease (Siguener *et al.*, 2014), hypertension (Graessler *et al.*, 2009), diabetes (Meikle *et al.*, 2014) and Alzheimer (Bennett *et al.*, 2014; Mapstone *et al.*, 2014), for which lipids may comprise an important source of biomarkers. Pathogens also manipulate the host lipidome in order to facilitate their own persistence and replication (Loizides-Mangold *et al.*, 2014; Lyn *et al.*, 2014).

Elucidating the roles of lipids in biological systems requires the integration of quantitative measurements of lipidome composition with knowledge of lipid metabolic pathways, enzymes and interacting proteins. Lipidome composition may be analyzed using mass spectrometry (MS) and tandem mass spectrometry (MS/MS)-based approaches which are able to quantify hundreds of lipid species over a broad dynamic range (da Silveira Dos Santos *et al.*, 2014). These approaches provide information on the atomic composition of intact lipids and their component parts (such as acyl chains) which can be mapped to prior knowledge of lipid structures using a standardized hierarchical notation (Liebisch *et al.*, 2013). One of the foremost sources of such knowledge is the LIPID MAPS Structure Database (LMSD), which provides a comprehensive classification of over 37 000 lipid structures grouped in 8 major lipid categories—fatty acyls, glycerolipids, glycerophospholipids, sphingolipids, sterol lipids, prenol lipids, polyketides and saccharolipids (Fahy *et al.*, 2007). Other resources and tools for the analysis of lipidomic data include LipidBLAST, which provides a library of 120 000 lipid structures (some derived from LIPID MAPS) and associated theoretical tandem mass spectra (Kind *et al.*, 2013), LipidHome, a database of over 36 million hypothetical glycerolipid and glycerophospholipid structures specifically designed for MS data integration (Foster *et al.*, 2013) and Visualization and phospholipid identification (VaLID) which provides a search engine for a library of hypothetical glycerophospholipid structures somewhat similar to those of LipidHome (Blanchard *et al.*, 2013). Specialized software for annotating MS data include LipidXplorer (Herzog *et al.*, 2013), mzMine (Pluskal *et al.*, 2010) and Lipid Data Analyzer (Hartler *et al.*, 2011). A common feature of all these resources and tools is that they focus mainly on the classification and identification of lipids. Knowledge of lipid biology, including links to metabolic pathways and enzymes, may be found in public knowledge resources such as BRENDA (Chang *et al.*, 2015), HMDB (Wishart *et al.*, 2013), KEGG (Kanehisa *et al.*, 2014), MetaCyc (Caspi *et al.*, 2014) and Reactome (Croft *et al.*, 2014). Although KEGG has been used as basis for computational approaches for matching hypothetical structures to pathways (Yetukuri *et al.*, 2007), these knowledge resources do not aim to provide comprehensive coverage of the hundreds of thousands of lipids that may exist in nature.

Here, we describe the development of a new knowledge resource that is designed to help connect lipidomics and lipid biology—SwissLipids. Our aim in creating SwissLipids was to provide a comprehensive reference database that can be used to describe mass spectrometry-based lipid identifications and link them to curated knowledge of lipid structures, metabolic reactions, enzymes and interacting proteins. SwissLipids was created through an iterative process in which prior knowledge of lipid structures and metabolism curated from published literature is used to generate an *in silico* library of all feasible structures for common lipid categories. These structures form the basis for a hierarchical lipid classification that is consistent with currently accepted lipidomic data notation and the classification of LIPID MAPS. The result is a reference namespace for lipidomic data publication, data exploration and hypothesis generation, allowing users to ask questions such as ‘where has this lipid species been seen before?’, ‘what structures might this lipid species have?’, ‘what are the relevant metabolic reactions?’ and ‘what proteins are known to interact with these putative structures?’

2 Methods

2.1 Curation of lipid structures and metabolism

We performed systematic curation of experimental knowledge of lipid structures using the chemical ontology ChEBI (Hastings *et al.*, 2013) and of lipid metabolism using the annotated reaction database Rhea (Morgat *et al.*, 2015) for several of the most common lipid categories including glycerophospholipids, glycerolipids, fatty acyls and sphingolipids. We focused mainly on lipids and enzymes of *Homo sapiens*, *Mus musculus*, *Caenorhabditis elegans* and *Saccharomyces cerevisiae*, but information from other species was also included when required to complete metabolic pathways (as in the desaturation of polyunsaturated fatty acids or PUFAs). Enzymatic reactions were curated in the SwissLipids database by linking UniProtKB (UniProt Consortium, 2014) protein accession numbers to Rhea reaction identifiers. All annotations in SwissLipids are associated with supporting evidence which includes a mandatory evidence type, represented by a code from the Evidence Codes Ontology (ECO) (Chibucos *et al.*, 2014) and supporting source text and literature citation where applicable (articles from the scientific literature are represented as PubMed records). All annotations were checked for accuracy and consistency by a second curator.

2.2 Generation and annotation of lipid structures

Curated knowledge of experimentally characterized lipid structures and metabolism was used to design and create a library of theoretically feasible lipid structures *in silico*. Lipid structures were generated using the Java application SMILIB v2.0 (Schüller *et al.*, 2007) which combines SMILES (Weininger, 1988) representations of chemical substructures. Each structure was annotated with a standard name and abbreviations (created using a custom Perl script that follows accepted naming conventions of LIPID MAPS), as well as lipid class and fatty-acyl/alkyl components (including mappings to ChEBI) and their positions on the glycerol backbone. Formula, InChI and InChI key (Heller *et al.*, 2013) and masses were calculated using the MolConvert and cxcalc tools (<http://www.chemaxon.com>), and exact *m/z* values for the following selected adducts were also calculated: [M.]⁺, [M+H]⁺, [M+K]⁺, [M+Na]⁺, [M+Li]⁺, [M+NH₄]⁺, [M-H]⁻, [M+Cl]⁻, [M+OAc]⁻. Mappings to

Table 1. The hierarchical classification used in SwissLipids

Level	Example
Category	Glycerophospholipid
Class	Glycerophosphocholine
Class	Monoalkylmonoacylglycerophosphocholine
Species	PC(O-36:5)
Molecular subspecies	PC(O-16:1_20:4)
Structural subspecies	PC(P-16:0/20:4)
Isomeric subspecies	PC(P-16:0/20:4(5Z,8Z,11Z,14Z))

The hierarchy includes seven levels that are illustrated below with a single example. The hierarchy is compatible with that of LipidHome (Foster et al., 2013) but uses only known components in the generation of the base *Isomeric subspecies*. The hierarchy is fully mapped to ChEBI at all levels. The prefix ‘O–’ indicates an alkyl bond, the prefix ‘P–’ a 1Z-alkenyl bond and the absence of a prefix an ester bond. PC, phosphatidylcholine.

corresponding structures in ChEBI, HMDB and LMSD were established by matching InChI keys.

2.3 Generation of the lipid hierarchy

Glycerolipid and glycerophospholipid structures were used as the basis for the generation of a hierarchy of analytical outputs from mass spectrometry experiments (Liebisch et al., 2013). The lipid hierarchy was generated using custom software that takes as input a set of structures and a template file that specifies the required annotations for each of the corresponding higher levels and the relations that link them (Table 1). These annotations include lipid nomenclature and human readable descriptions, ChEBI identifiers for lipid parent classes and parent components, SMILES representations, formula, mass and *m/z* values for adducts, and InChI and InChI keys where applicable.

2.4 Curation of the lipid hierarchy

The lipid hierarchy was curated using experimental literature on the occurrence of lipids in biological systems. Qualitative annotations take the form of a link between a lipid identifier (generally one of *Species*, *Molecular subspecies* or *Structural subspecies*), a term from the Gene Ontology (GO) (Blake et al., 2013) or the cross-species tissue and cell ontology Uberon (Mungall et al., 2012), and a taxonomic identifier (NCBI Resource Coordinators., 2014). As with other annotation types, all annotations of lipid occurrence in SwissLipids are associated with supporting evidence represented by a code from the Evidence Codes Ontology (ECO) and supporting source text and literature citation where applicable.

2.5 Databases

SwissLipids data is organized in a set of MySQL tables. Data from external reference resources such as ChEBI (Hastings et al., 2013), Rhea (Morgat et al., 2015), UniProtKB (UniProt Consortium, 2014), GO (Blake et al., 2013), Cellosaurus and Uberon (Mungall et al., 2012) is sourced from local copies synchronized on a weekly basis.

2.6 Website development

The SwissLipids website has been developed using the AngularJS (<http://www.angularjs.org>) framework and server side PHP scripts (<http://www.php.net>). SVG images of molecules are generated from their *mol* representations using MolConvert (<http://www.chemaxon.com>).

3 Results and discussion

3.1 SwissLipids content

SwissLipids includes over 2000 curated enzymatic reactions which directly link more than 800 proteins (UniProtKB) and 1500 reactions (Rhea) and supporting evidence. These annotations cover a variety of lipid categories including glycerophospholipids, glycerolipids, sphingolipids, sterols, fatty acids, fatty alcohols and wax esters. As the number of experimentally characterized lipid structures constitutes only a small fraction of the number of possible structures that may exist in nature (Gross and Han, 2011; Shevchenko and Simons, 2010), the curated knowledge available in SwissLipids was used to design and generate an *in silico* library of feasible lipid structures for selected common lipid categories. The current version of this library features 244 155 structures—hereafter referred to as *Isomeric subspecies*—from 38 classes of glycerophospholipid and 13 classes of glycerolipid (Supplementary Table S1) including 80 fatty acids and 18 fatty alcohols (Supplementary Table S2). Forthcoming versions of SwissLipids will generalize this approach to other lipid categories for which curated data is available, including more classes of glycerophospholipid as well as sphingolipids. Structures are arranged in a hierarchical classification consistent with the latest lipid notation (Liebisch et al., 2013) that includes 91 124 *Structural subspecies* (which assume knowledge of lipid class and the composition and positions of individual fatty acyl groups), 45 571 *Molecular subspecies* (which assume knowledge of lipid class and the composition of individual fatty acyl groups) and 5508 *Species* (which assume knowledge of lipid class and the sum composition of all fatty acyl groups) (Table 1). The SwissLipids hierarchy includes curated lipid identifications from published literature for over 300 *Species*, *Molecular subspecies* and *Structural subspecies*.

The combinatorial generation of lipid structures described here and elsewhere (Blanchard et al., 2013; Foster et al., 2013) may produce lipid structures that do not exist in nature, and SwissLipids includes significantly more glycerophospholipid and glycerolipid structures than reference databases such as LMSD (which leverages the expert knowledge of the LIPID MAPS consortium). Although the existence of these extra structures is currently unproven, the use of known structural building blocks from the reference ontology ChEBI ensures that they are at least relevant to our current understanding of lipid structures and metabolism, and the procedure for brute force enumeration of structures is completely transparent. The SwissLipids library also provides a means to link structures from other databases such as HMDB and LMSD to published lipid identifications and to metabolic reactions (described using Rhea) and enzymes (UniProtKB), and is cross-referenced by the experimental lipidomics platform of the LipidX consortium (www.lipidomes.org).

3.2 SwissLipids website and access

The SwissLipids website at www.swisslipids.org provides a simple interface to search for lipids (by identifiers, names, synonyms, abbreviations, formulae, *m/z* values for common adducts and SMILES, InChI and InChI key representations). Users may also browse the SwissLipids hierarchy at <http://www.swisslipids.org/#/browse>, beginning at *Species* level with the selection of the desired lipid class and the number of carbon atoms and double bonds in the acyl chains (Fig. 1). This returns a list of all possible *Molecular subspecies*, *Structural subspecies* and *Isomeric subspecies*. Exploration of the result lists is facilitated by a simple icon system that indicates the available information for each lipid entry (Fig. 1). Individual lipids

SwissLipids Search Browse ID Mapping About BETA version

Browse SwissLipids
Select a class and specify the number of carbons and double bonds to list molecular subspecies, structural subspecies and isomers (described here)

Class: PC(O-): O-Phosphatidylcholine Nb of carbons: 36 Nb of double bonds: 5 Search

Phosphatidylcholine (O-36:5)

7 molecular subspecies show all

SLM:000065681	Phosphatidylcholine (O-14:0_22:5)	ch cl
SLM:000065745	Phosphatidylcholine (O-14:1_22:4)	ch cl
SLM:000065934	Phosphatidylcholine (O-16:0_20:5)	ch cl
SLM:000065998	Phosphatidylcholine (O-16:1_20:4)	ch cl
SLM:000066252	Phosphatidylcholine (O-18:1_18:4)	ch cl
SLM:000066316	Phosphatidylcholine (O-18:2_18:3)	ch cl
SLM:000066380	Phosphatidylcholine (O-18:3_18:2)	ch cl

9 structural subspecies

20 isomeric subspecies

SLM:000028457	1-O-tetradecyl-2-(4Z,7Z,10Z,13Z,16Z-docosapentaenoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000028466	1-O-tetradecyl-2-(7Z,10Z,13Z,16Z,19Z-docosapentaenoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000049213	1-O-(1Z-tetradecenyl)-2-(7Z,10Z,13Z,16Z-docosatetraenoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000028142	1-O-hexadecyl-2-(5Z,8Z,11Z,14Z,17Z-eicosapentaenoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000048885	1-O-(1Z-hexadecenyl)-2-(5Z,8Z,11Z,14Z-eicosatetraenoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000048896	1-O-(1Z-hexadecenyl)-2-(8Z,11Z,14Z,17Z-eicosatetraenoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000027346	1-O-(11Z-octadecenyl)-2-(6Z,9Z,12Z,15Z-octadecatetraenoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000027586	1-O-(13Z-octadecenyl)-2-(6Z,9Z,12Z,15Z-octadecatetraenoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000027826	1-O-(9Z-octadecenyl)-2-(6Z,9Z,12Z,15Z-octadecatetraenoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000049048	1-O-(1Z-octadecenyl)-2-(6Z,9Z,12Z,15Z-octadecatetraenoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000027747	1-O-(9Z,12Z-octadecadienyl)-2-(6Z,9Z,12Z-octadecatrienoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000027759	1-O-(9Z,12Z-octadecadienyl)-2-(9Z,12Z,15Z-octadecatrienoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000048089	1-O-(1Z,11Z-octadecadienyl)-2-(6Z,9Z,12Z-octadecatrienoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000048101	1-O-(1Z,11Z-octadecadienyl)-2-(9Z,12Z,15Z-octadecatrienoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000048329	1-O-(1Z,13Z-octadecadienyl)-2-(6Z,9Z,12Z-octadecatrienoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000048341	1-O-(1Z,13Z-octadecadienyl)-2-(9Z,12Z,15Z-octadecatrienoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000048569	1-O-(1Z,9Z-octadecadienyl)-2-(6Z,9Z,12Z-octadecatrienoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000048581	1-O-(1Z,9Z-octadecadienyl)-2-(9Z,12Z,15Z-octadecatrienoyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000048498	1-O-(1Z,9Z,12Z-octadecatrienyl)-2-(9Z,11E-octadecadienyl)-sn-glycero-3-phosphocholine	ch cl
SLM:000048503	1-O-(1Z,9Z,12Z-octadecatrienyl)-2-(9Z,12Z-octadecadienyl)-sn-glycero-3-phosphocholine	ch cl

SwissLipids is brought to you by members of the Swiss-Prot and Vital-IT teams of SIB Swiss Institute of Bioinformatics.
Contact the SwissLipids team at swisslipids@sib-sib.ch.
SwissLipids is supported by funding from the Swiss Initiative for Systems Biology SystemsX.ch and the Swiss State Secretariat for Education, Research and Innovation through SIB Swiss Institute of Bioinformatics.

Logit SystemsX.ch The Swiss Initiative in Systems Biology

Fig. 1. Browsing SwissLipids. Users can enter the lipid hierarchy at *Species* level by specifying the desired lipid class and number of carbon atoms and double bonds. The result lists consist of the corresponding *Molecular subspecies*, *Structural subspecies* and *Isomeric subspecies*. Color-coded icons with abbreviations provide an overview of the available information for lipids—their meaning can be revealed by moving the mouse over them. In this example the lipid *Species* Phosphatidylcholine (O-36:5) has 7 *Molecular subspecies*, 9 *Structural subspecies* (for which the list is not expanded) and 20 *Isomeric subspecies*. The *Molecular subspecies* Phosphatidylcholine (O-16:0_20:5) and Phosphatidylcholine (O-16:1_20:4) have experimental data on their location in specific tissues or taxa (indicated by the brown “lo” icon), while the *Isomeric subspecies* 1-O-(1Z-hexadecenyl)-2-(5Z,8Z,11Z,14Z-eicosatetraenoyl)-sn-glycero-3-phosphocholine has experimental data relating to metabolism (indicated by the purple “me” icon). All lipids have cheminformatic descriptors (green “ch” icon) and are classified (orange “cl” icon) in the SwissLipids hierarchy

as well as complete result lists may be selected for the identifier mapping service (<http://www.swisslipids.org/#/mapper>), which maps lipid identifiers from SwissLipids, ChEBI, HMDB and LMSD to each other as well as to the corresponding biochemical reactions from Rhea and curated enzymes from SwissLipids. Individual lipid entry pages provide key information including structure, cheminformatics descriptors, nomenclature, classification, reactions, enzymes, and subcellular and tissular location (Fig. 2). Enzyme names link to protein entry pages that list all curated information from SwissLipids, such as Rhea reactions and citations (not shown). SwissLipids may also be searched and accessed through the ExPASy

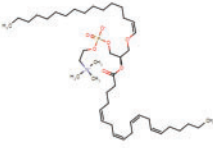
bioinformatics resource portal (<http://www.expasy.org>) (Artimo et al., 2012).

4 Conclusions

SwissLipids provides a core of curated information on lipid structures and metabolism which serves as the base for the generation of a library of all feasible lipid structures (according to our current knowledge of the relevant lipid structural classes) and their corresponding analytical outputs for MS and MS/MS. The curation of

SwissLipids Search Browse ID Mapping About BETA version

1-O-(1Z-hexadecenyl)-2-(5Z,8Z,11Z,14Z-eicosatetraenyl)-sn-glycero-3-phosphocholine



download mol file

ID SLM:000048885
Name 1-O-(1Z-hexadecenyl)-2-(5Z,8Z,11Z,14Z-eicosatetraenyl)-sn-glycero-3-phosphocholine
Abbreviation PC(P-16:0/20:4(5Z,8Z,11Z,14Z)) (SwissLipids)
Synonyms 1-(1Z)-hexadecenyl-2-(5Z,8Z,11Z,14Z)-eicosatetraenyl-sn-glycero-3-phosphocholine (ChEBI)
 1-(1Z)-hexadecenyl-2-(5Z,8Z,11Z,14Z)-eicosatetraenyl-sn-glycero-3-phosphocholine (ChEBI)
Cross-references PC(O-16:1(1Z)/20:4(5Z,8Z,11Z,14Z)) (ChEBI)
 LMGP01030010 (LipidMaps)
 77292 (ChEBI)
 HMDB11220 (HMDB)

ChemInfo Classification Metabolism Comments Gene ontology Interactions Cells and tissues Genetics Citations

ChemInfo

Characteristics of the major microspecies of this molecule at pH 7.3 (structure shown):

Formula C44H80NO7P
Charge 0
InChI 1S/C44H80NO7P/c1-6-8-10-12-14-16-18-20-22-23-24-25-27-29-31-33-35-37-44(46)52-43(42-51-53(47,48)50-40-38-45(3,4)5)41-49-39-36-34-32-30-28-26-21-19-17-15-13-11-9-7-2/h14,16,20,22,24-25,29,31,36,39,43H,6-13,15,17-19,21,23,26-28,30,32-35,37-38,40-42H2,1-5H3/b16-14-,22-20-,25-24-,31-29-,39-36-/43-m/1/s1
InChI key IOYKZPNDXIXLN-LOQSCQKMSA-N
SMILES CCCCCCCCCCCC/C=C/OC[C@H](COP(=O)(=O)OCC[N+](C)(C)C)OC(=O)CCC/C=C/C/C=C/C/C=C/C/C=C/C/C=C/C/CCCC

Mass/charge ratios for protonated forms and adducts:

m/z exact mass = 765.567261

ChemInfo Classification Metabolism Comments Gene ontology Interactions Cells and tissues Genetics Citations

Classification

Isomeric subspecies: This is an isomeric subspecies of lipid, with fully resolved structure

Parents:
 SLM:000047888 1-O-(1Z-alkenyl)-2-acyl-sn-glycero-3-phosphocholine
 SLM:000093162 Phosphatidylcholine (P-16:0/20:4)

Components:
 SLM:000000202 hexadecan-1-ol position = sn1
 SLM:000000296 (5Z,8Z,11Z,14Z)-eicosatetraenoate position = sn2

ChemInfo Classification **Metabolism** Comments Gene ontology Interactions Cells and tissues Genetics Citations

Metabolism

Rhea reactions for this lipid that are linked to an enzyme:

[RHEA: 40580] - [PNPLA8 (Homo sapiens) [see evidence](#)]

1 H₂O + 1 1-O-(1Z-hexadecenyl)-2-(5Z,8Z,11Z,14Z-eicosatetraenyl)-sn-glycero-3-phosphocholine
 =>
 1 H⁺ + 1 (5Z,8Z,11Z,14Z)-eicosatetraenoate + 1 1-O-(1Z-hexadecenyl)-sn-glycero-3-phosphocholine

Fig. 2. A SwissLipids entry. SwissLipids entry pages provide information on lipid structure and nomenclature (top panel), cheminformatics descriptors (second panel), lipid classification and components (third panel), reactions (Rhea) and enzymes (UniProtKB) (fourth panel), and subcellular (GO) and tissular (Uberon) location (not shown) which can be navigated through distinct tabs. Links to other databases such as ChEBI, HMDB and LIPID MAPS are also provided. The lipid classification indicates the structural class of lipid (here, 1-O-(1Z-alkenyl)-2-acyl-sn-glycero-3-phosphocholine) and its 'analytical parent' (the Structural subspecies Phosphatidylcholine (P-16:0/20:4)) as well as the individual components (hexadecanol-1-ol and 5Z,8Z,11Z,14Z-eicosatetraenoate) and the annotations available for each. The underlying evidence for curated assertions can be viewed by clicking on 'see evidence' (see fourth panel)

information using ChEBI, Rhea and UniProtKB provides a ready-made mapping of SwissLipids to these widely used resources and others that reference them such as the MetaboLights repository for metabolomic studies (Haug et al., 2013), the IntEnz resource for

enzyme nomenclature and classification (Fleischmann et al., 2004), the MetaNetX.org platform for the analysis of metabolic models (Bernard et al., 2014; Ganter et al., 2013) and the genome annotation platform MicroScope (Vallet et al., 2013). As well as

enriching SwissLipids with new lipid categories, structures and curated information, we plan to provide enhanced search options that include common isotopes and fragmentation patterns, and to improve interoperability by providing annotations and structures using the W3C standard RDF, a mainstay of large bioinformatics resources like UniProtKB.

Acknowledgements

The authors thank members of the Swiss-Prot and Vital-IT groups including Nicolas Guex, Mark Ibberson, Thierry Lombardot, Kristian Axelsen, Anne Morgat and Marco Pagni as well as members of the LipidX consortium, in particular Thomas Hannich, Ursula Loizides-Mangold and Aline da Silveira dos Santos, for fruitful discussions and comments. We would also like to thank Marcus Ennis, Janna Hastings, Gareth Owen and Steve Turner of the Cheminformatics and Metabolism Team at the European Bioinformatics Institute for expert curation of ChEBI entries, Gerhard Liebisch for discussions and advice regarding the lipid hierarchy, and Andreas Schüller for advice and recommendations on the use of SMILIB. We also gratefully acknowledge the software contributions of ChemAxon [<https://www.chemaxon.com/products/marvin/>].

Funding

The development of SwissLipids was supported by the Swiss Federal Government through the State Secretariat for Education, Research and Innovation (SERI) and by the special opportunities fund of the SystemsX.ch.

Conflict of Interest: none declared.

References

- Artimo, P. *et al.* (2012) ExpASY: SIB bioinformatics resource portal. *Nucleic Acids Res.*, **40**, W597–W603.
- Balla, T. (2013) Phosphoinositides: tiny lipids with giant impact on cell regulation. *Physiol. Rev.*, **93**, 1019–1137.
- Bennett, S.A. *et al.* (2014) Using neurolipidomics to identify phospholipid mediators of synaptic (dys)function in Alzheimer's Disease. *Front. Physiol.*, **4**, 168.
- Bernard, T. *et al.* (2014) Reconciliation of metabolites and biochemical reactions for metabolic networks. *Brief. Bioinform.*, **15**, 123–135.
- Blake, J.A. *et al.* (2013) Gene Ontology annotations and resources. *Nucleic Acids Res.*, **41**, D530–D535.
- Blanchard, A.P. *et al.* (2013) Visualization and Phospholipid Identification (VaLID): online integrated search engine capable of identifying and visualizing glycerophospholipids with given mass. *Bioinformatics*, **29**, 284–285.
- Caspi, R. *et al.* (2014) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res.*, **42**, D459–D471.
- Chang, A. *et al.* (2015) BRENDA in 2015: exciting developments in its 25th year of existence. *Nucleic Acids Res.*, **43**, D439–D446.
- Chibucos, M.C. *et al.* (2014) Standardized description of scientific evidence using the Evidence Ontology (ECO). *Database (Oxford)*, **2014** doi:10.1093/database/bau075.
- Croft, D. *et al.* (2014) The Reactome pathway knowledgebase. *Nucleic Acids Res.*, **42**, D472–D477.
- da Silveira Dos Santos, A.X. *et al.* (2014) Systematic lipidomic analysis of yeast protein kinase and phosphatase mutants reveals novel insights into regulation of lipid homeostasis. *Mol. Biol. Cell*, **25**, 3234–3246.
- Fahy, E. *et al.* (2007) LIPID MAPS online tools for lipid research. *Nucleic Acids Res.*, **35**, W606–W612.
- Fleischmann, A. *et al.* (2004) IntEnz, the integrated relational enzyme database. *Nucleic Acids Res.*, **32**, D434–D437.
- Foster, J.M. *et al.* (2013) LipidHome: a database of theoretical lipids optimized for high throughput mass spectrometry lipidomics. *PLoS One*, **8**, e61951.
- Ganter, M. *et al.* (2013) MetaNetX.org: a website and repository for accessing, analysing and manipulating metabolic networks. *Bioinformatics*, **29**, 815–816.
- Graessler, J. *et al.* (2009) Top-down lipidomics reveals ether lipid deficiency in blood plasma of hypertensive patients. *PLoS One*, **4**, e6261.
- Gross, R.W. and Han, X. (2011) Lipidomics at the interface of structure and function in systems biology. *Chem. Biol.*, **18**, 284–291.
- Hartler, J. *et al.* (2011) Lipid Data Analyzer: unattended identification and quantitation of lipids in LC-MS data. *Bioinformatics*, **27**, 572–577.
- Hastings, J. *et al.* (2013) The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. *Nucleic Acids Res.*, **41**, D456–D463.
- Haug, K. *et al.* (2013) MetaboLights—an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Res.*, **41**, D781–D786.
- Heller, S. *et al.* (2013) InChI - the worldwide chemical structure identifier standard. *J. Cheminform.*, **5**, 7.
- Herzog, R. *et al.* (2013) LipidXplorer: a software for consensual cross-platform lipidomics. *PLoS One*, **7**, e29851.
- Holthuis, J.C. and Menon, A.K. (2014) Lipid landscapes and pipelines in membrane homeostasis. *Nature*, **510**, 48–57.
- Ji, R.R. *et al.* (2011) Emerging roles of resolvins in the resolution of inflammation and pain. *Trends Neurosci.*, **34**, 599–609.
- Kanehisa, M. *et al.* (2014) Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.*, **42**, D199–D205.
- Kind, T. *et al.* (2013) LipidBlast in silico tandem mass spectrometry database for lipid identification. *Nat. Methods*, **10**, 755–758.
- Knowlden, S. and Georas, S.N. (2014) The autotaxin-LPA axis emerges as a novel regulator of lymphocyte homing and inflammation. *J. Immunol.*, **192**, 851–857.
- Lass, A. *et al.* (2011) Lipolysis - a highly regulated multi-enzyme complex mediates the catabolism of cellular fat stores. *Prog. Lipid Res.*, **50**, 14–27.
- Liebisch, G. *et al.* (2013) Shorthand notation for lipid structures derived from mass spectrometry. *J. Lipid Res.*, **54**, 1523–1530.
- Loizides-Mangold, U. *et al.* (2014) HCV 3a core protein increases lipid droplet cholesteryl ester content via a mechanism dependent on sphingolipid biosynthesis. *PLoS One*, **9**, e115309.
- Lyn, R.K. *et al.* (2014) Stearoyl-CoA desaturase inhibition blocks formation of hepatitis C virus-induced specialized membranes. *Sci. Rep.*, **4**, 4549.
- Mapstone, M. *et al.* (2014) Plasma phospholipids identify antecedent memory impairment in older adults. *Nat. Med.*, **20**, 415–418.
- Meikle, P.J. *et al.* (2014) Plasma lipid profiling shows similar associations with prediabetes and type 2 diabetes. *PLoS One*, **8**, e74341.
- Mendelson, K. *et al.* (2014) Sphingosine 1-phosphate signalling. *Development*, **141**, 5–9.
- Morad, S.A. and Cabot, M.C. (2013) Ceramide-orchestrated signalling in cancer cells. *Nat. Rev. Cancer*, **13**, 51–65.
- Morgat, A. *et al.* (2015) Updates in Rhea—a manually curated resource of biochemical reactions. *Nucleic Acids Res.*, **43**, D459–D464.
- Mungall, C.J. *et al.* (2012) Uberon, an integrative multi-species anatomy ontology. *Genome Biol.*, **13**, R5.
- NCBI Resource Coordinators. (2014) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **42**, D7–D17.
- Pluskal, T. *et al.* (2010) MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics*, **11**, 395.
- Schüller, A. *et al.* (2007) SMIlib v2.0: a java-based tool for rapid combinatorial library enumeration. *QSAR Comb. Sci.*, **26**, 407–410.
- Shevchenko, A. and Simons, K. (2010) Lipidomics: coming to grips with lipid diversity. *Nat. Rev. Mol. Cell Biol.*, **11**, 593–598.
- Sigruener, A. *et al.* (2014) Glycerophospholipid and sphingolipid species and mortality: the Ludwigshafen Risk and Cardiovascular Health (LURIC) study. *PLoS One*, **9**, e85724.

- Simmons, S. and Ishii, M. (2014) Sphingosine-1-phosphate: a master regulator of lymphocyte egress and immunity. *Arch. Immunol. Ther. Exp. (Warsz)*, **62**, 103–115.
- Sud, M. et al. (2007) LMSD: LIPID MAPS structure database. *Nucleic Acids Res.*, **35**, D527–D532.
- UniProt Consortium. (2014) Activities at the universal protein resource (UniProt). *Nucleic Acids Res.*, **42**, D191–D198.
- Vallenet, D. et al. (2013) MicroScope—an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. *Nucleic Acids Res.*, **41**, D636–D647.
- van Meer, G. et al. (2008) Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.*, **9**, 112–124.
- Weininger, D. (1988) SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inform. Comput. Sci.*, **28**, 31–36.
- Wishart, D.S. et al. (2013) HMDB 3.0—The Human Metabolome Database in 2013. *Nucleic Acids Res.*, **41**, D801–D807.
- Yetukuri, L. et al. (2007) Bioinformatics strategies for lipidomics analysis: characterization of obesity related hepatic steatosis. *BMC Syst. Biol.*, **1**, 12.