A Semantic Assistant (SA) for Lipidomics Researchers

Alexandre Kouznetsov¹,
Jonas B. Laurila¹,
Rene Witte²,

Christopher J. O. Baker¹





Motivation

- Semantic data integration is necessary for lipid research yet this is poorly achievable due to an absence of a single *unified*, *consistent*, and *universally accepted* lipid classification system.
- Lipid nomenclature is highly heterogeneous.
 - Not semantically explicit
 - Many conflicting nomenclatures and multiple synonyms
 - Graphically dependent definitions
 - Lack of universal systematic nomenclature

LIPID Nomenclature

IUPAC

- Open to erroneous interpretation by scientists
- Too bulky for adoption
- No informatics implementation
- Re-interpretations and implementations not scientifically robust.

LIPIDMAPS

- Class names inconsistent with instances
- Classes with no instances
- Lack semantic & textual definitions
- Use of "dumping ground" classes to hold lipid instances that do not "fit in"
- Not all instances classified by structure; some by functions & biosynthetic origin (lack of consistency in classification)

Upto (2009)

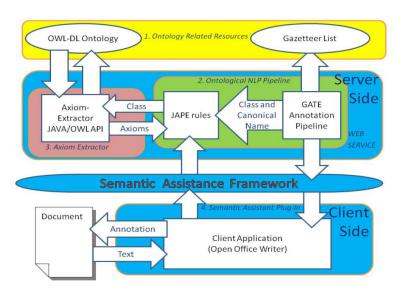
Objectives

- Formalize and represent lipid nomenclature & classification hierarchy in the web ontology language(OWL-DL)
- Provide access to lipid definitions that are:
 - Independent of graphical descriptions
 - Semantically explicit
 - Amenable to inference / classification based reasoning
- To make available:
 - a systematic and formalized OWL-DL definitions of lipids for testing appropriateness of existing nomenclature to lipid structures.
 - serve as a reusable standard for lipid researchers and the lipid bioinformatics community
 - Formal annotations to users doing Knowledge discovery tasks

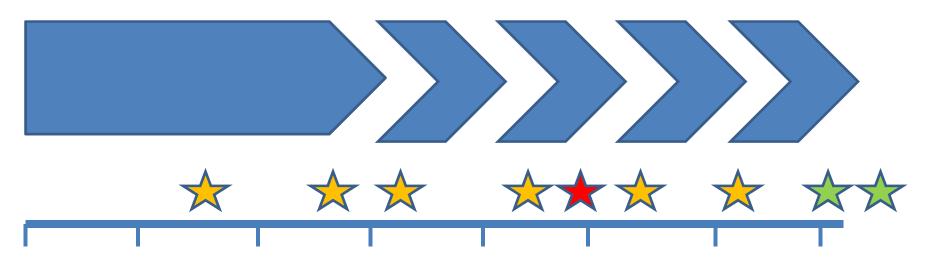
National Institutes of Health

SA: Core System Components

- OWL-DL Lipid Ontology
- Natural Language Processing (GATE/JAPE)
- Semantic Assistant Framework
- Ontology Axiom-Extractor

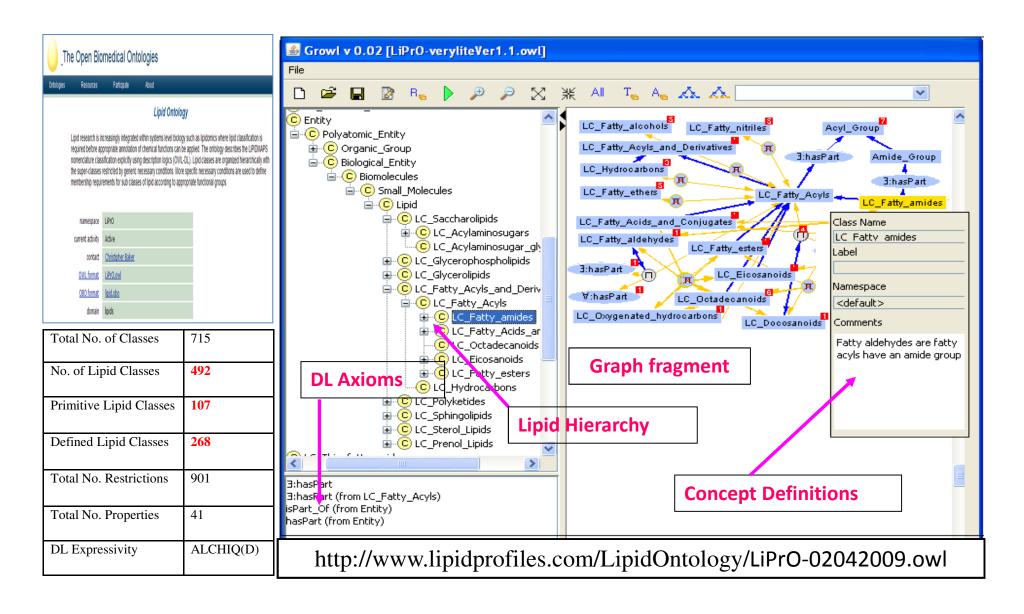


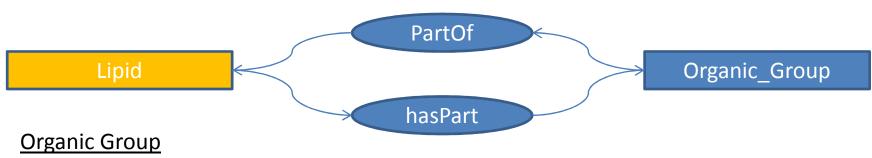
Lipid Ontology: a history



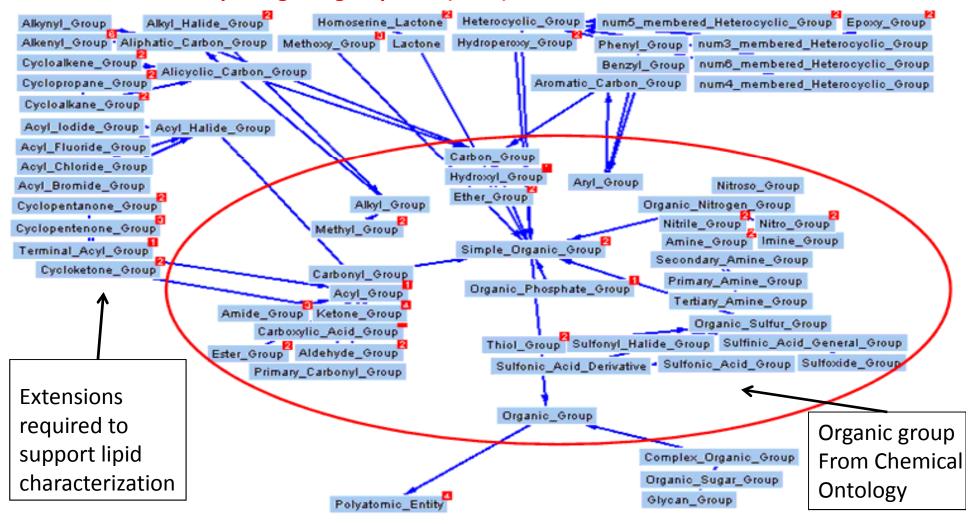
2007	2008	2009	2010
FIRST-ISWC 2007	- Early Text Mining	/ Simple Ontology /	modeling lipid nomenclature
BMC Bioinf. 2008	- Baseline Ontology	/ Text Mining / Visi	ual Query
ISMB 2008	- Text Mining / Apo	ptosis / Data Mining	for Transitive Relations
OBO 2009	 Online Listing 		
AMIA 2009	- Text Mining / Ova	rian Cancer	
HS Low MSc 2009	- Formal Lipid Onto	logy	
ICBO 2009	- OWL-DL Ontology	for classification of I	-ipids
IntOnt 2010	 Lipid Ontologies 		
ACS 2010	- Semantic Chemist	ry - Corpus Annotation	on with Lipid Ontology Axioms

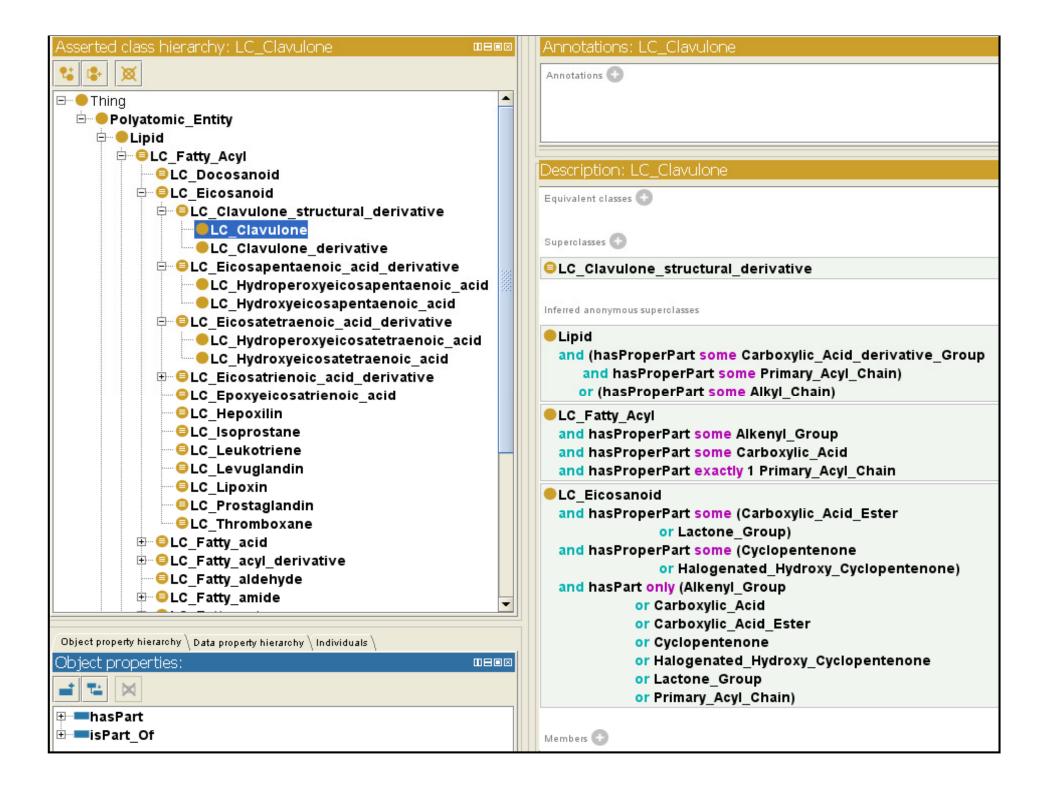
Lipid Ontology





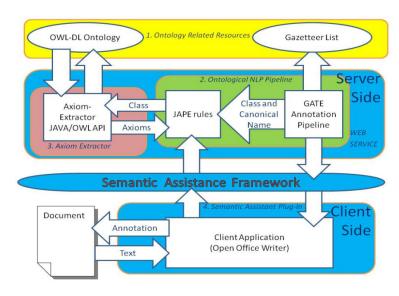
Total no. of simple organic group = 95 (2009)



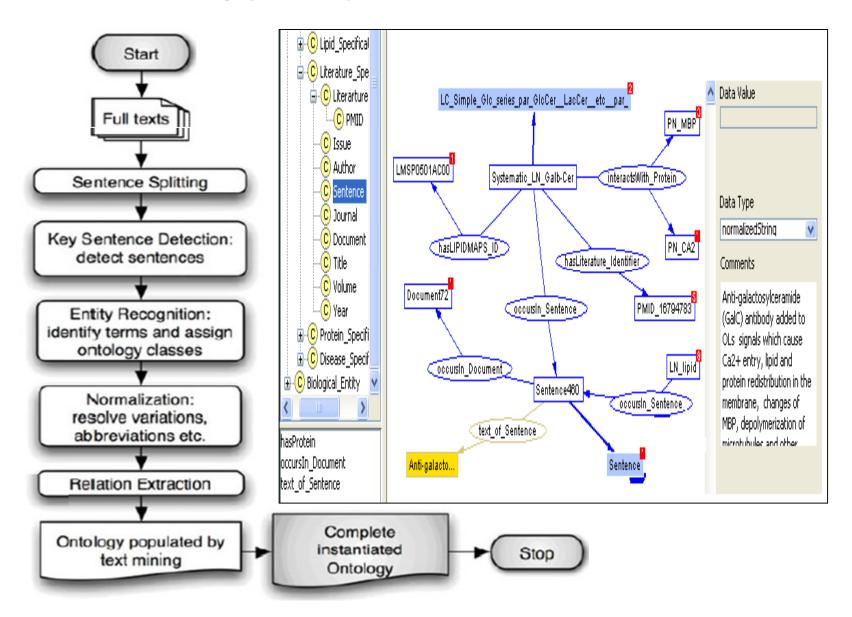


SA: Core System Components

- OWL-DL Lipid Ontology
- Natural Language Processing (GATE/JAPE)
- Semantic Assistant Framework
- Ontology Axiom-Extractor



Ontology Population Workflow

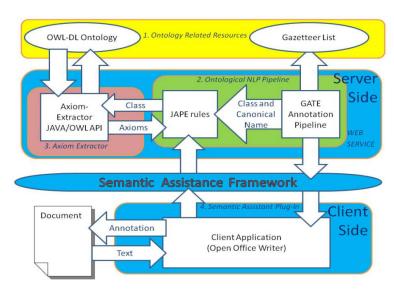


Mapping Named Entities to Axioms

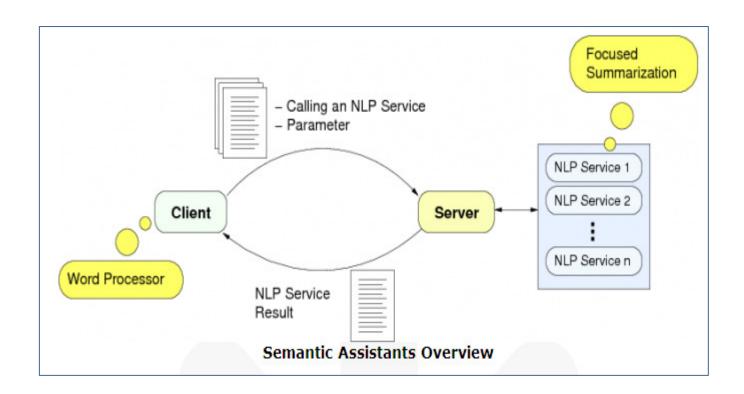
	Α	R		n n	Е	G	H
1	LMID	Common name	Systematic name	lupac name	Pubchem ID	LipidMaps Class	Lipid Ontology Class
1421	LMFA01060161	5-oxopentanoic acid	5-0x0-pentanoic acid	5-oxopentanoic acid	7982783	Fatty Acids and Conjugates < <fa01></fa01>	LC_Fatty_Acyl
1422	LMFA01060160	9-oxononanoic acid	9-oxo-nonanoic acid	9-oxononanoic acid	7982782	Fatty Acids and Conjugates < <fa01>></fa01>	LC Fatty A
1423	LMFA01060159	8-oxononanoic acid	8-oxo-nonanoic acid	8-oxononanoic acid	7982781	Fatty Acids and Conjugates < <fa01>></fa01>	Lo
1424	LMFA01060158	3-oxononanoic acid	3-oxo-nonanoic acid	3-oxononanoic acid	7982780	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1425	LMFA01060157	7-methyl-6-oxo-octanoic acid	7-methyl-6-oxo-octanoic acid	7-methyl-6-oxo-octanoic acid	7982779	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1426	LMFA01060156	4,7-dioxooctanoic acid	4,7-dioxo-octanoic acid	4,7-dioxooctanoic acid	7982778	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1427	LMFA01060150	(Z)-19-oxooctacos-22-enoic acid	19-oxo-22Z-octacosenoic acid	(Z)-19-oxooctacos-22-enoic acid	7982776	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1428	LMFA01060149	(Z)-17-oxohexacos-20-enoic acid	17-oxo-20Z-hexacosenoic acid	(Z)-17-oxohexacos-20-enoic acid	7982775	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1429	LMFA01060148	(Z)-15-oxotetracos-18-enoic acid	15-oxo-18Z-tetracosenoic acid	(Z)-15-oxotetracos-18-enoic acid	7982774	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1430	LMFA01060147	3-oxotetracosanoic acid	3-oxo-tetracosanoic acid	3-oxotetracosanoic acid	7982773	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1431	LMFA01060146	3-oxotricosanoic acid	3-oxo-tricosanoic acid	3-oxotricosanoic acid	7982772	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1432	LMFA01060155	7-methyl-4-oxo-octanoic acid	7-methyl-4-oxo-octanoic acid	7-methyl-4-oxo-octanoic acid	7982777	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1433	LMFA01060136	20-oxohenicosanoic acid	20-oxo-heneicosanoic acid	20-oxohenicosanoic acid	7982762	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1434	LMFA01060135	2-oxohenicosanoic acid	2-oxo-heneicosanoic acid	2-oxohenicosanoic acid	7982761	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1435	LMFA01060134	3-oxoicosanoic acid	3-oxo-eicosanoic acid	3-oxoicosanoic acid	7982760	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1436	LMFA01060133	2-oxoicosanoic acid	2-oxo-eicosanoic acid	2-oxoicosanoic acid	7982759	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1437	LMFA01060132	19-oxoicosanoic acid	19-oxo-eicosanoic acid	19-oxoicosanoic acid	7982758	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1438	LMFA01060131	3-oxononadecanoic acid	3-oxo-nonadecanoic acid	3-oxononadecanoic acid	7982757	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1439	LMFA01060130	2-oxononadecanoic acid	2-oxo-nonadecanoic acid	2-oxononadecanoic acid	7982756	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1440	LMFA01060120	8-oxooctadecanoic acid	8-oxo-octadecanoic acid	8-oxooctadecanoic acid	7982746	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1441	LMFA01060119	2-oxooctadecanoic acid	2-oxo-octadecanoic acid	2-oxooctadecanoic acid	7982745	Fatty Acids and Conjugates < <fa01>></fa01>	LC Fatty Acyl
1442	LMFA01060118	17-oxooctadecanoic acid	17-oxo-octadecanoic acid	17-oxooctadecanoic acid	7982744	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1443	LMFA01060117	16-oxooctadecanoic acid	16-oxo-octadecanoic acid	16-oxooctadecanoic acid	7982743	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1444	LMFA01060116	15-oxooctadecanoic acid	15-oxo-octadecanoic acid	15-oxooctadecanoic acid	7982742	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1445	LMFA01060115	14-oxooctadecanoic acid	14-oxo-octadecanoic acid	14-oxooctadecanoic acid	7982741	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1446	LMFA01060114	13-oxooctadecanoic acid	13-oxo-octadecanoic acid	13-oxooctadecanoic acid	7982740	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1447	LMFA01060113	12-oxooctadecanoic acid	12-oxo-octadecanoic acid	12-oxooctadecanoic acid	7982739	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1448	LMFA01060084	(E)-5-oxodec-7-enoic acid	5-oxo-7E-decenoic acid	(E)-5-oxodec-7-enoic acid	7982710	Fatty Acids and Conjugates < <fa01>></fa01>	LC Fatty Acyl
1449	LMFA01060082	(E)-9-oxodec-2-enoic acid	9-oxo-2E-decenoic acid	(E)-9-oxodec-2-enoic acid	7982708	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1450	LMFA01060080	6,9-dioxodecanoic acid	6,9-dioxo-decanoic acid	6,9-dioxodecanoic acid	7982706	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1451	LMFA01060079	3,6-dioxodecanoic acid	3,6-dioxo-decanoic acid	3,6-dioxodecanoic acid	7982705	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1452	LMFA01060110	2-methyl-4-oxo-heptadecanoic acid	2-methyl-4-oxo-heptadecanoic acid	2-methyl-4-oxo-heptadecanoic acid	7982736	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1453	LMFA01060167	-	12-oxo-9(Z)-dodecenoic acid	-	-	Fatty Acids and Conjugates < <fa01>></fa01>	LC Fatty Acyl
1454	LMFA01060176)xodecanoic acid##(2S)-2-amino-8-(oxiran-2	2-amino-8-oxo-9,10-epoxy-decanoic acid	-	14184	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1455	LMFA01060175	cid##4-Oxo-cis-9,trans-11,trans-13-octadec	4-oxo-9Z,11E,13E-octadecatrienoic acid	-	10517	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1456	LMFA01060174	ic acid##2-Oxohex-trans-4-enoate##(E)-2-ox	2-Oxo-4E-hexenoic acid	-	8981	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1457	LMFA01060171	acid##2-Amino-4-oxopentanoate##(2R)-2-an	2-amino-4-oxo-pentanoic acid	-	6184	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1458	LMFA01060177	ODE##9-KODE##(10E,12Z)-9-Oxooctadeca	9-oxo-10E,12Z-octadecadienoic acid	-	-	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1459	LMFA01060170	cid##4-Methylthio-2-oxobutanoate##4-methy	2-oxo-4-methylthio-butanoic acid	-	4407	Fatty Acids and Conjugates < <fa01>></fa01>	LC Fatty Acyl
1460	LMFA01060169	oate##2-0xo-5-aminopentanoate##alpha-Ke	2-oxo-5-amino-pentanoic acid	-	4342	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1461	LMFA01060168	ate##8-Amino-7-oxononanoic acid##8-amin	7-oxo-8-amino-nonanoic acid	-	4327	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1462	LMFA01060173	acid##(S)-5-Amino-3-oxohexanoate##(5S)-5	3-oxo-5S-amino-hexanoic acid	-	6434	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl
1463	LMFA01060172	ate##L-2-Amino-acetoacetate##(S)-2-Amino	2S-amino-3-oxo-butanoic acid	-	6318	Fatty Acids and Conjugates < <fa01>></fa01>	LC_Fatty_Acyl

SA: Core System Components

- OWL-DL Lipid Ontology
- Natural Language Processing (GATE/JAPE)
- Semantic Assistant Framework
- Ontology Axiom-Extractor



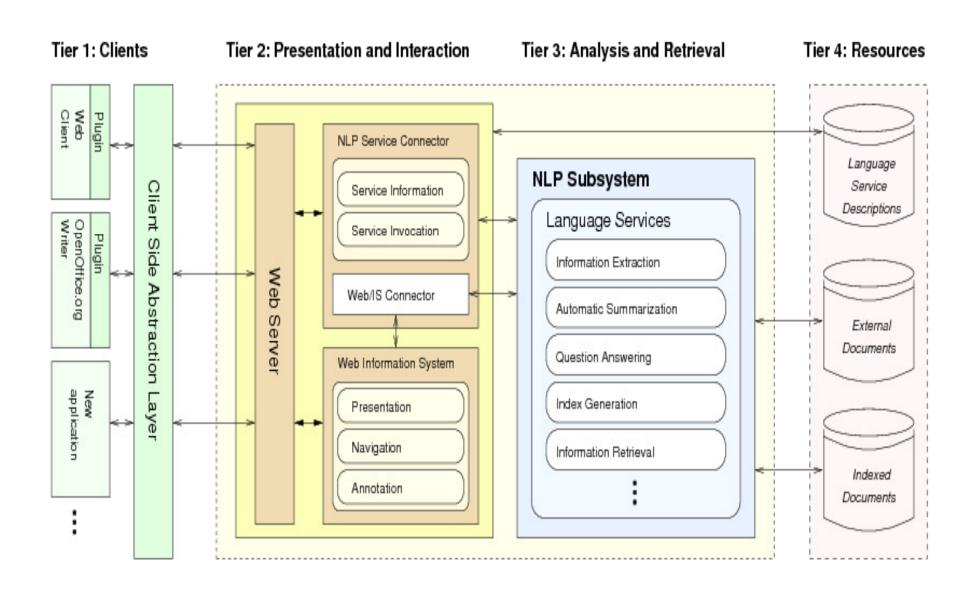
Semantic Desktop Assistant



René Witte and Thomas Gitzinger.

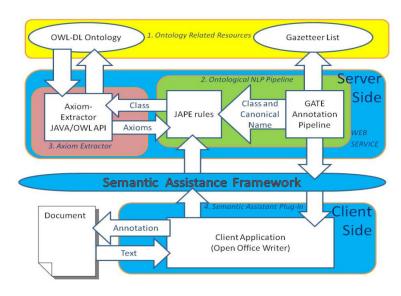
<u>Semantic Assistants – User-Centric Natural Language Processing Services for Desktop Clients.</u> 3rd Asian Semantic Web Conference (ASWC 2008), February 2–5, 2009, Bangkog, Thailand. Springer LNCS 5367, pp. 360–374. (Acceptance rate: 31%)

Semantic Assistant Framework

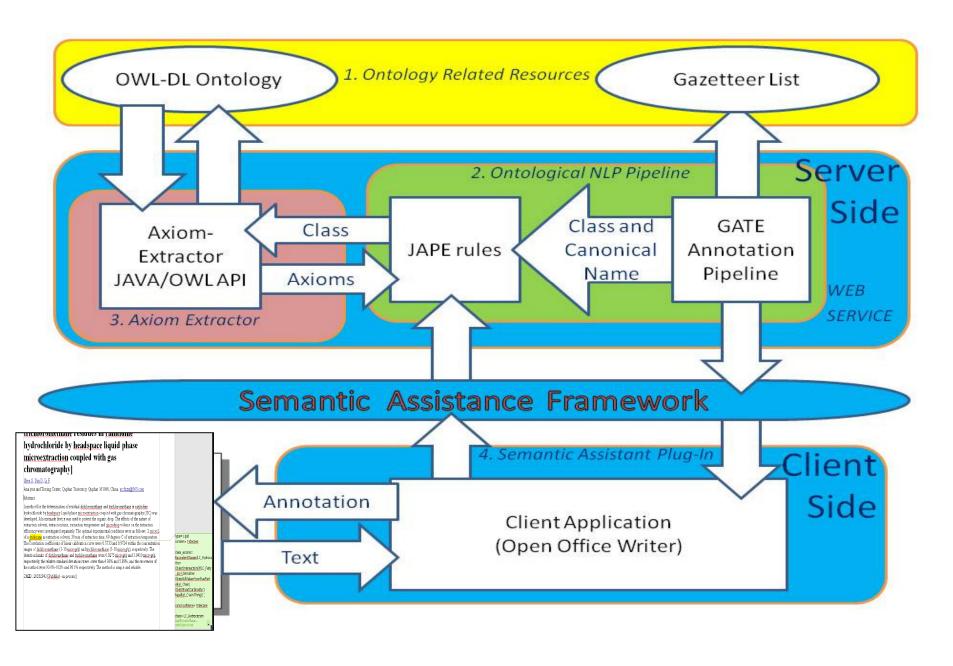


Core System Components

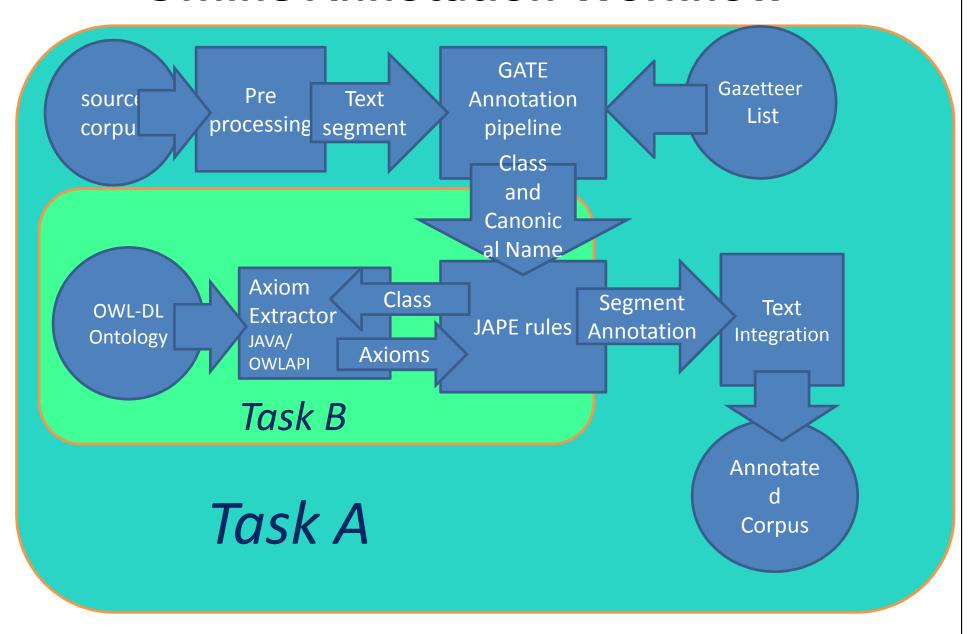
- OWL-DL Lipid Ontology
- Natural Language Processing (GATE/JAPE)
- Semantic Assistant Framework
- Ontology Axiom-Extractor



Online Annotation workflow

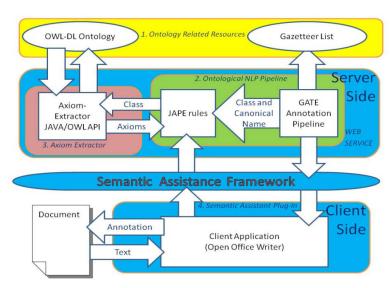


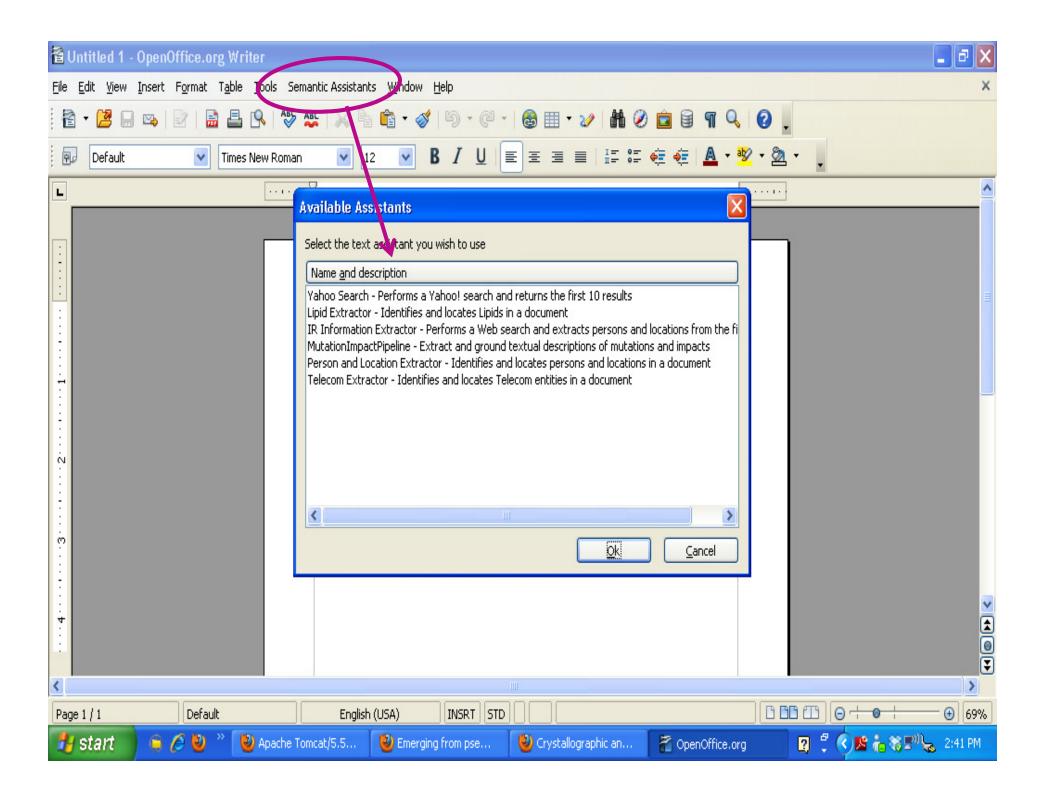
Offline Annotation Workflow



Core System Components

- OWL-DL Lipid Ontology
- Natural Language Processing (GATE/JAPE)
- Semantic Assistant Framework
- Ontology Axiom-Extractor
- Clients:
 - Open Office
 - Firefox







Dual inhibition of mycobacterial fatty acid biosynthesis and degradation by 2-alkynoic acids.

2-Hexadecynoic acid and 2-octadecynoic acid have cidal activity against Mycobacterium smegmatis and Mycobacterium bovis BCG. At subinhibitory concentrations, M. smegmatis rapidly transformed [1-(14)C]-2-hexadecynoic acid into endogenous fatty acids and elongated them into mycolic acids. Toxic concentrations of 2-hexadecynoic acid resulted in accumulation of 3-ketohexadecanoic acid, which blocked fatty acid biosynthesis, and 3-hexadecynoic acid, an inhibitor of fatty acid degradation. The combination of these metabolites is necessary to achieve the inhibition of M. smegmatis. Our conclusion is that 2- and 3-hexa/octadecynoic acids inhibit mycolic acid biosynthesis

Annotates Lipids

- Cannonical Name
- Lipid Ontology Class
- Functional groups allowed for specific lipid class i.e.

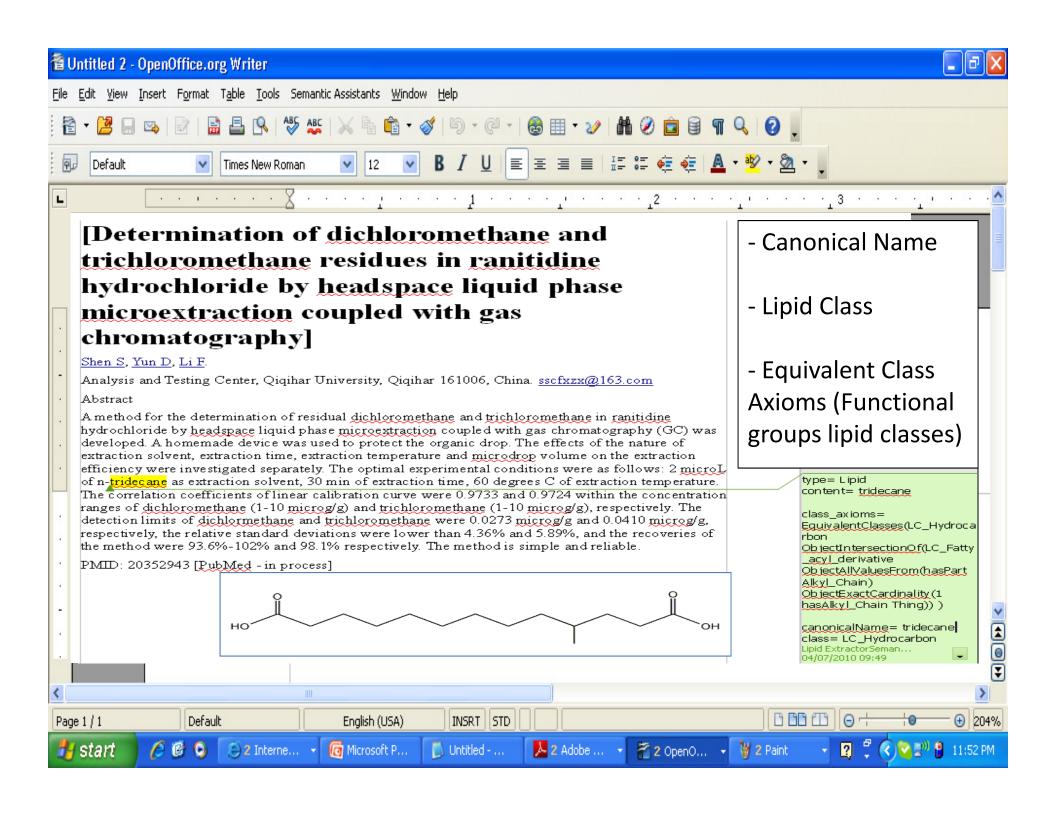
Kouznetsov, Witte, Baker (2010)

Here is some text (+/-)-20-hydroxy-4Z,7Z,10Z,13Z,16Z,18E-docosahexaenoic acid after that some other text (+/-)-20-HDoHE

canonica|Name= 14octadecynoic acidclass= LC_Fatty_Acids_and_Conj Lipid ExtractorSeman... Today, 22:10

type= Lipid
content= (+/-)-20-hydro
4Z,7Z,10Z,13Z,16Z,18Edocosahexaenoic acid
canonicalName= (+/-)-10
dihydroxy-4Z,7Z,10Z,13Z
docosapentaenoic acidcla
LC_Docosanoids
Lipid ExtractorSeman...
Today, 21:05

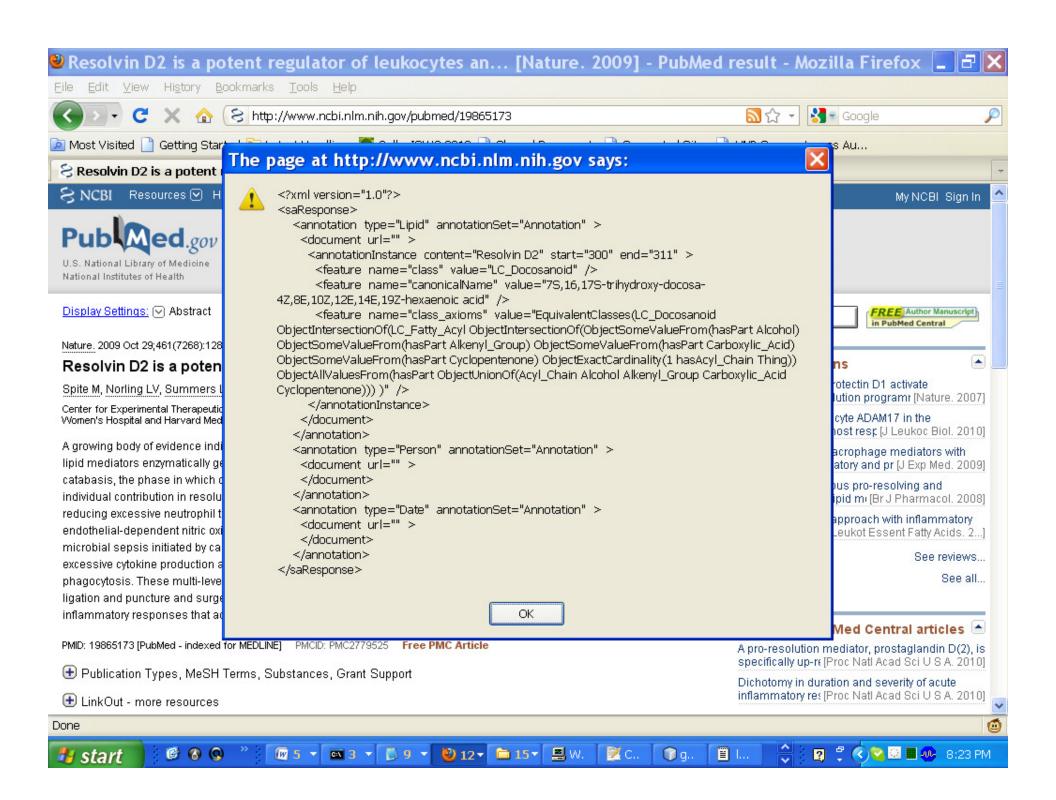
type= Lipid content= (+/-)-20-HDoH canonicalName= (+/-)-10 dihydroxy-4Z,7Z,10Z,13Z docosapentaenoic acidcla LC_Docosanoids Lipid ExtractorSeman... Today, 21:05

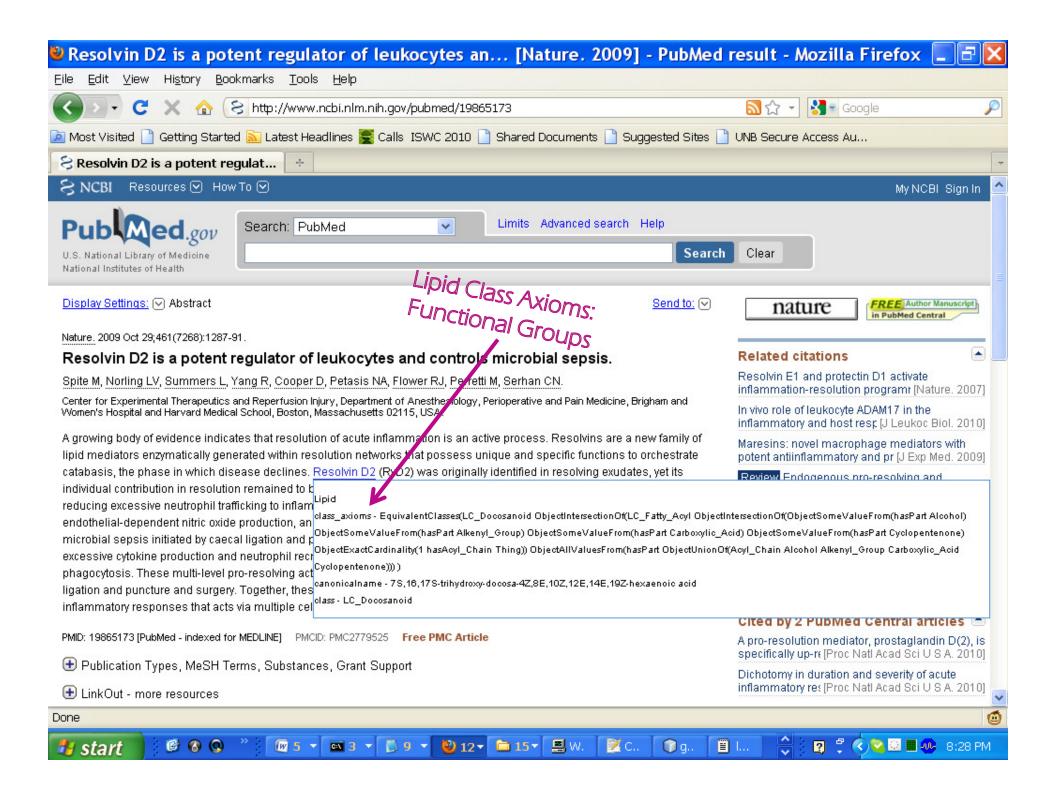


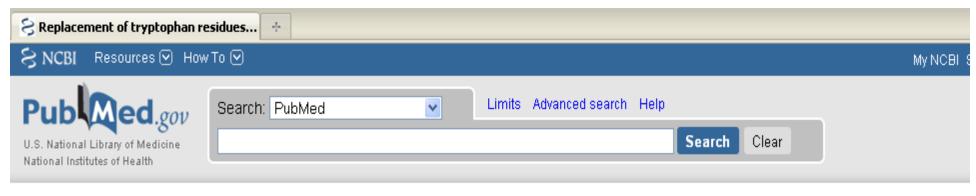
Migration to Firefox

- 1) Firefox
- 2) GreaseMonkey Plugin
- 3) Install Lipid SA service (Java Script / Tomcat)









haswildtyperesidue - W

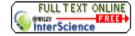
hasmutantresidue - F

Display Settings:

✓ Abstract

Point Mutations

Send to: ▽



Eur J Biochem. 1995 Mar 1;228(2):403-7.

Replacement of tryptophan residues in haloalkane dehalogenase reduces halide binding and catalytic activity.

Kennes C, Pries F, Krooshof GH, Bokma E, Kingma J, Janssen DB.

Department of Biochemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, The Netherlands.

Haloalkane dehalogenase catalyzes the hydrolytic cleavage of carbon-halogen bond in short-chain haloalkanes. Two tryptophan residues of the enzyme (Trp125 and Trp175) form a halide-binding site in the active-site cavity, and were proposed to play a role in catalysis. The function of these residues was studied by replacing Trp125 with phenylalanine, glutamine or arginine and Trp175 by glutamine using site-directed mutagenesis. All mutants 10-fold reduced kcat and much higher Km values with 1,2-dichloroethane and 1 PointMutation 10-fold reduced kcat and much higher Km values with 1,2-dichloroethane and 1 PointMutation 10-fold reduced with the wild-type enzyme in deuterium oxide was lost in hasmentionedposition - 125 pt the Trp125-->Phe enzyme. The results indicate that both tryptophans are involved in stabilizing the hascorrectposition - 125 is prounded to - QBQ3HO

PMID: 7705355 [PubMed - indexed for MEDLINE] Free Article

- ⊕ Publication Types, MeSH Terms, Substances, Secondary Source ID
- LinkOut more resources

Related citations

Kinetic analysis and X-ray structure of haloalkane dehalogenase (Biochemistry

Repositioning the catalytic triad aspartic a haloalkane dehalogenase: [Biochemistry

Kinetic characterization and X-ray structur mutant of haloalkane dehal [Biochemistry

Improved catalytic properties of halohydrir dehalogenase by modificati [Biochemistr,

Review Evolving haloalkane dehalogena (Curr Opin Chem Bio

See ret

Cited by 5 PubMed Central articl

Functionally relevant motions of haloalkar dehalogenases occur in the s [Protein Sc

Persistently conserved positions in struct similar, sequence dissimilar | [Protein Sc

The importance of reactant positioning in



Display Settings:

✓ Abstract

hasmutantresidue - A

Point Mutations

Send to: <

Biochemistry, 1999 Sep 14;38(37):12052-61.

Crystallographic and kinetic evidence of a collision complex formed during halide import in haloalkane dehalogenase.

Pikkemaat MG, Ridder IS, Rozeboom HJ, Kark KH, Dijkstra BW, Janssen DB.

Laboratory of Biochemistry, BIOSON Research Institute, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, The Netherlands.

Haloalkane dehalogenase (DhIA) converts haloalkanes to their corresponding alcohols and halide ions. The rate-limiting step in the reaction of DhIA is the release of the halide ion. The kinetics of halide release have been analyzed by measuring halide. binding with stopped-flow Morescence experiments. At high halide concentrations, halide import occurs predominantly via the rapid formation of a weak initial collision complex, followed by transport of the ion to the active site. To obtain more insight in this collision complex, we determined the X-ray structure of DhIA in the presence of bromide and investigated the kinetics of mutants that were construited on the basis of this structure. The X-ray structure revealed one bromide ion firmly bound in the active site and two bromize ions weakly bound on the surface of the enzyme. One of the weakly bound ions is close to Thr197 and Phe294, near the elitance of the earlier proposed tunnel for substrate import. Kinetic analysis of bromide import by the Thr197Ala and Phe294Ala mutants of DhIA at high halide concentration showed that the rate constants for halide binding no longer displayed a hcrease with increasing bromide concentrations. This is in agreement with an elimination or a Point Mutation hasmentionedposition - 294; surface-located halide-binding site. Likewise, chloride binding kinetics of the mutants indicated ith wild-type enzyme. The results indicate that Thr197 and Phe294 are involved in the formation of an hascorrectposition - 294 or halide import in DhIA and provide experimental evidence for the role of the tunnel in substrate and isgroundedto - Q6Q3H0 haswildtyperesidue - F

Funding Sources

- NBIF, New Brunswick, Canada.
- Quebec -New Brunswick University Co-operation in Advanced Education - Research Program, Government of New Brunswick, Canada
- NSERC Discovery Grant, Canada 2009