

G-SRS Software Development

TIM SHEILS

NATIONAL CENTER FOR ADVANCING TRANSLATIONAL SCIENCES
NATIONAL INSTITUTES OF HEALTH

8 SEPTEMBER 2015

NCATS



National Center
for Advancing
Translational Sciences



ginas ● February ginas meeting

Pros:

Data model complete

.9 was ‘feature complete’ in relation to a changing technical specification

Exchangeability of messages was good

- Users could send and retrieve data

Utility and validation features:

- import docs
- check structures
- validate names
- import proteins

API was used to retrieve information independent of software

ginas ● February ginas meeting

Cons:

How is it implemented with current systems?

- Useful tool as a standalone

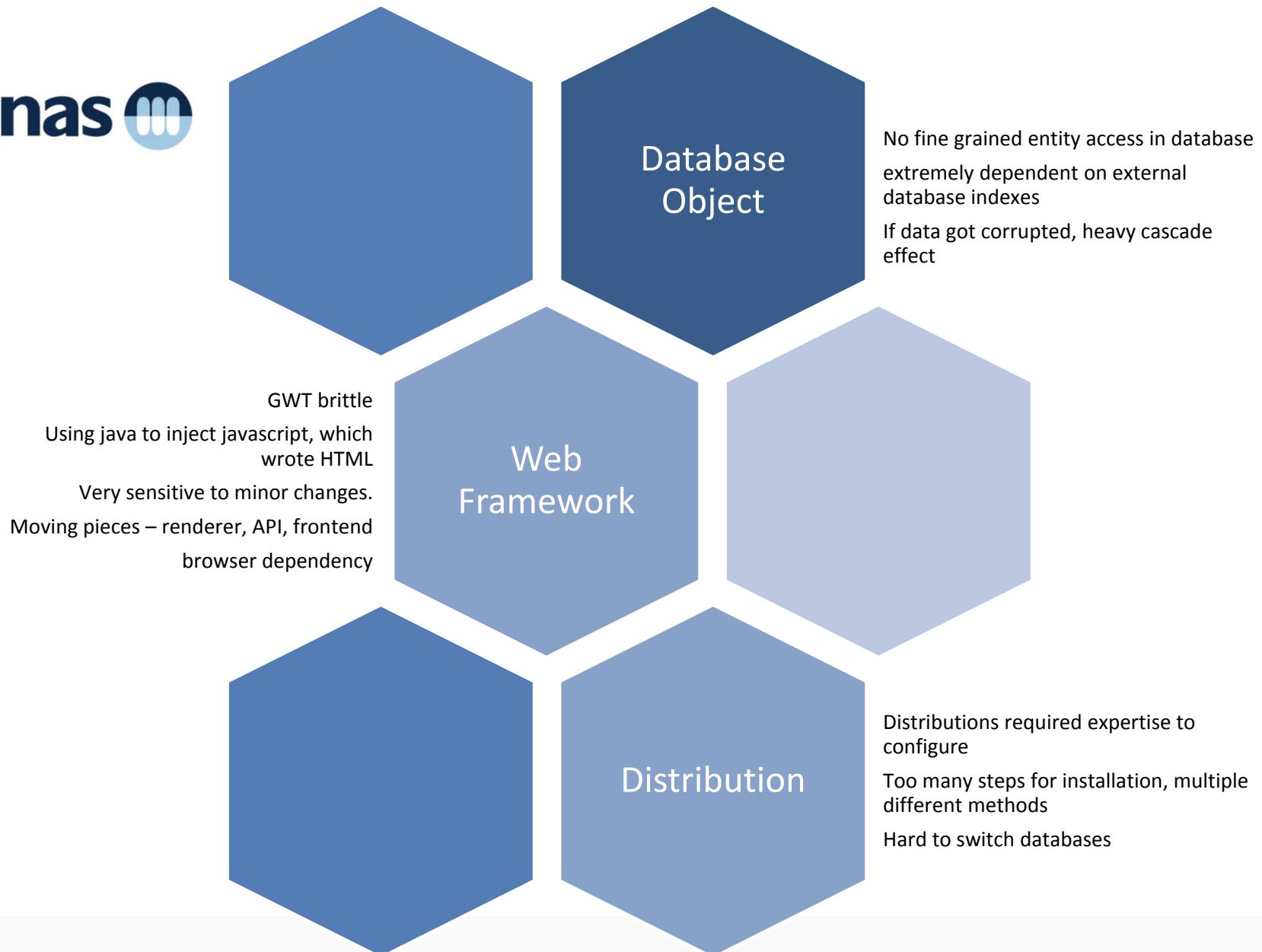
Meets technical requirements, but not usability

Forms complete, but hard to use, especially for non-experts

Portability of the data itself

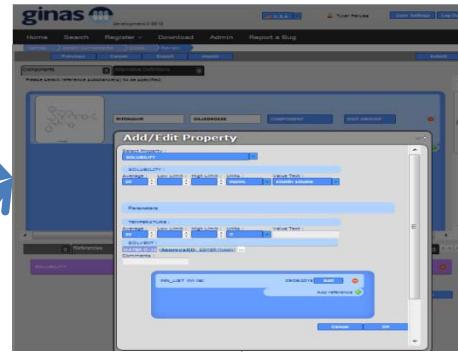
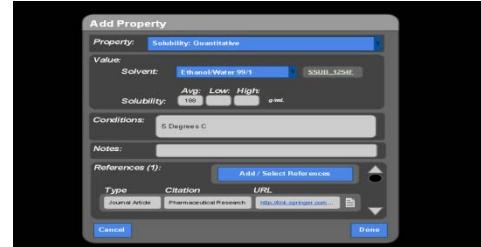
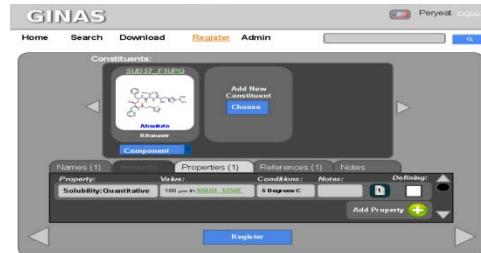
- ISO disc image used for distribution

Data not as discoverable – no browse functionality



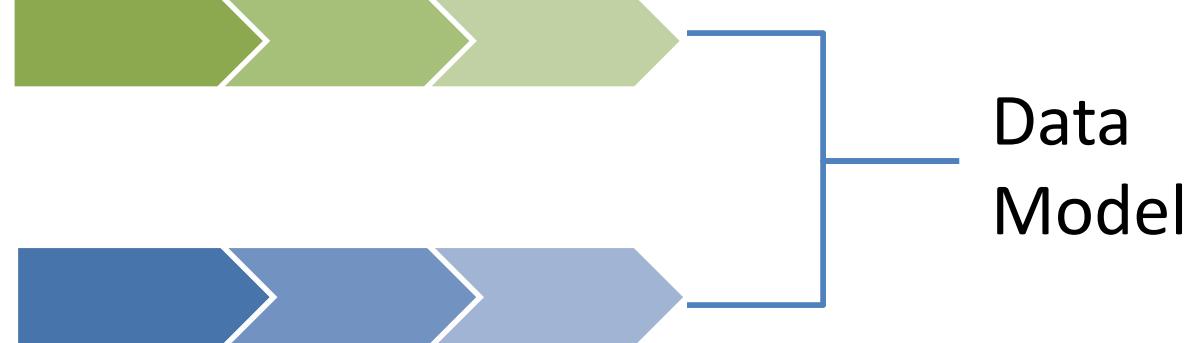
Main product to-date has been the data model

ginas



Data Model and API

Implementation Guide



G-SRS .9

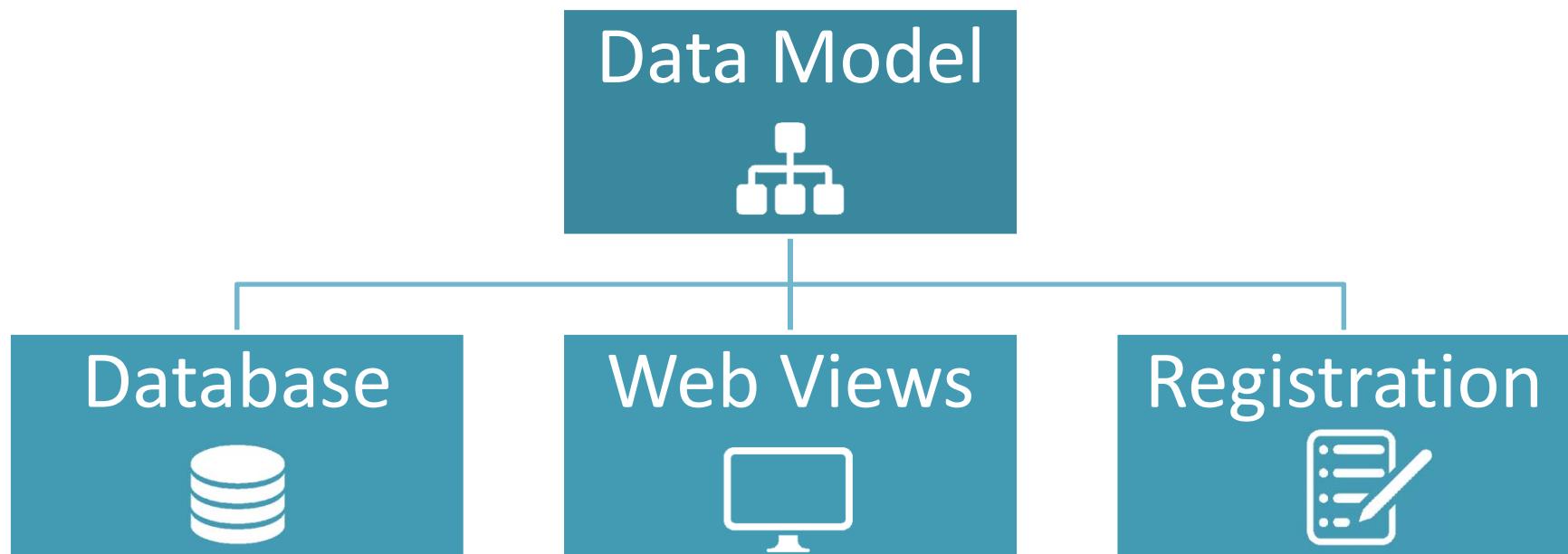
NCATS



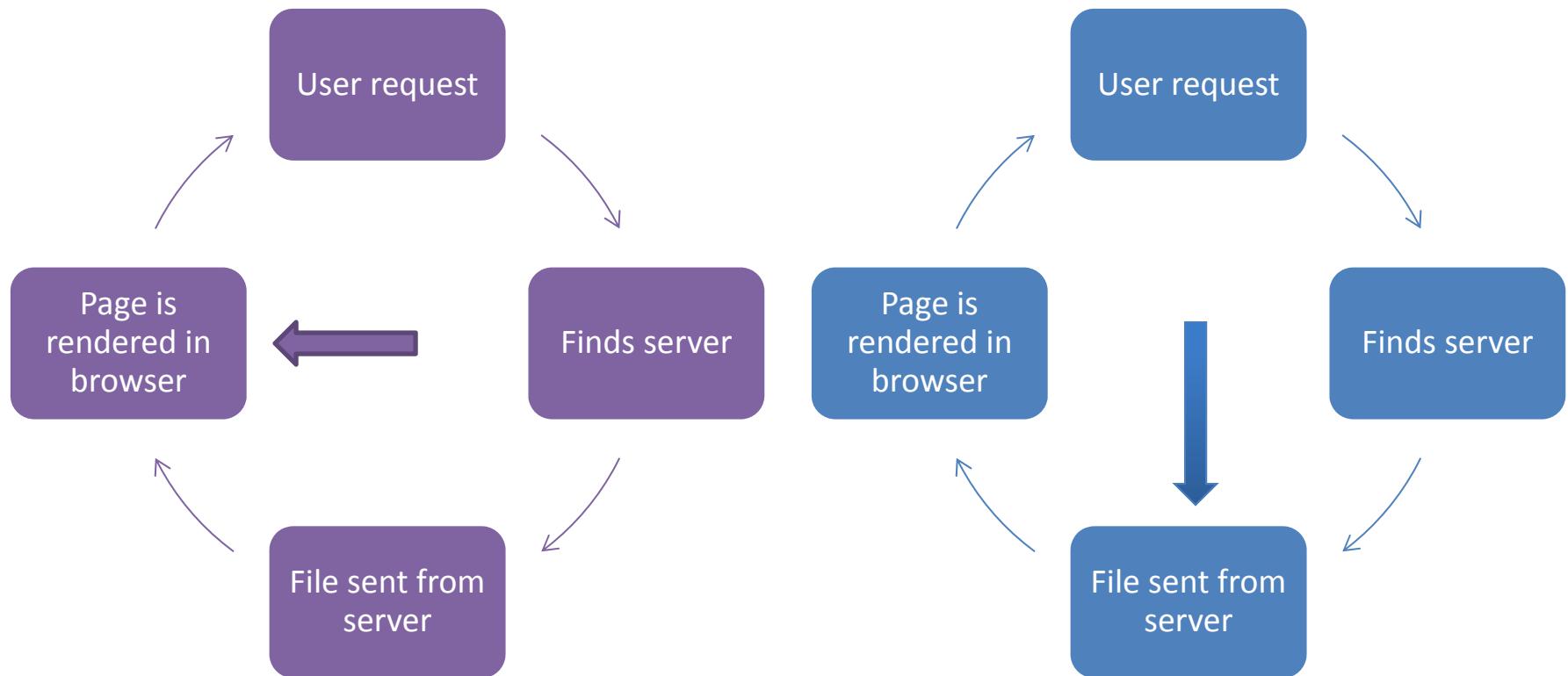
National Center
for Advancing
Translational Sciences



G-SRS 1.0 structure



ginas Client Side vs. Server Side



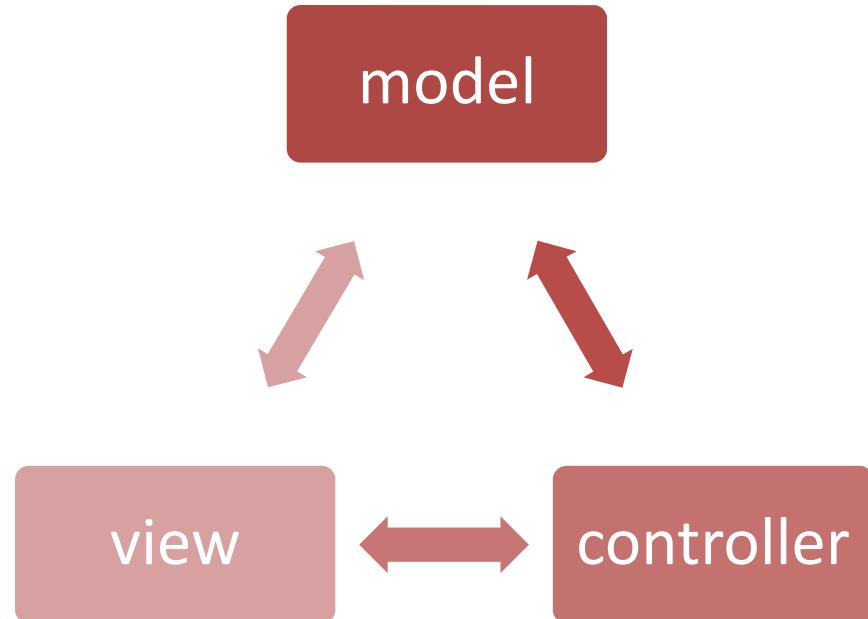


Play MVC Framework

Model: Java object
(database)

View: Scala, HTML,
Javascript

Controller: Java API
functions



ginas Additional technologies

- Codebase in Java at spotlite.nih.gov
- Scala HTML templating
- AngularJS for forms and UI
- Bootstrap for look and feel
- RDMS agnostic



ginas Development Environment

- Gitlab
 - version control system
 - Issue Tracking
 - Ongoing issues
 - Short term fixes
 - Assign issue to developers
 - Tags for bugs/ feature requests, etc.
 - Milestone for issues/deployment
 - Code is open to anyone
- Jenkins
 - continuous integration
 - Builds
 - Deploys
 - Tests



Jenkins



National Center
for Advancing
Translational Sciences

NCATS

ginas G-SRS 1.0 (Play) milestone:

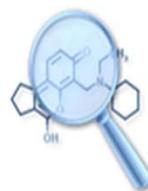
- Targeting November release:
 - MVC modules for:
 - Chemical
 - Protein
 - Nucleic Acid
 - Polymer
 - Structurally Diverse
 - Mixture
 - Group 1 Specified Substance



License:

- Open source / distributable
- All used libraries are open source (majority MIT)
- JSDraw licensed for distribution
- NCATS code public domain notice
 - Working on notice language
 - Written, awaiting approval



[Home](#)[Search](#)[Register ▾](#)[Download](#)[Admin](#)[Report a Bug](#)[Search](#)[Chemical Structure](#)[Nucleic Acid Sequence](#)[Amino Acid Sequence](#)



alpha version

Search ...



Home Register Substance Download Structure Search Sequence Search

Login



BROWSE
SUBSTANCES



REGISTER A
SUBSTANCE



National Center
for Advancing
Translational Sciences

 Chemical Protein Nucleic Acid Polymer

Entity Type :

 Structurally Diverse Mixture Specified Substance Group 1 Concept[Select](#)



alpha version

Home Register Substance Download Structure Search Sequence Search

Search ...



Login

Demo registration forms:



[Register a Chemical](#)

[Register a Protein](#)

[Register a Structurally Diverse Substance](#)

The current ginas registration pages can be accessed here:

tripod.nih.gov/dev/ginias/app/wizard?kind=chemical



National Center
for Advancing
Translational Sciences

G-SRS 1.0 Substance Views

Show All

1

Substance Class (1)

chemical

1

Stereochemistry (1)

ACHIRAL

1

Moieties (2)

CH4O3S

1

C29H31N7O

1

Moiety Count (1)

2

1

Last Editor (1)

tyler

1

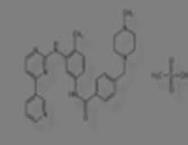
Query

query :

GLEEVEC

Import Set

ApprovalID: 8A1O1M485B



IMATINIB MESYLATE

Export

Export results:

```
[  
  {  
    "references": [  
      {  
        "tags": [  
          "NOMEN"  
        ],  
        "documentDate": "2014-06-03T16:00:00",  
        "citation": "ORANGE BOOK APPROVED 18-APR-2003",  
        "publicDomain": true,  
        "lastModified": 1421376233630,  
        "docType": "SRS",  
        "uuid": "4af070a-b602-4cae-8541-ca89e294d5f2"  
      },  
      {  
        "tags": [  
          "NOMEN"  
        ]  
      }  
    ]  
  }  
]
```

Close

Save File

Record Status

 Approved 58402 Non-approved 3043

Substance Class

 Chemical 49423 Structurally Diverse 5986 Concept 3043 Polymer 1130 Mixture 1111 Protein 752

Stereochemistry

 AQUEOUS 67560

74269

<< 1 2 3 4 5 6 7 ... 4641 4642 >>



RUBBER PLECO, COOKED

976DIQ3CZL



Other Names:

HYPOSTOMUS NIGRICANS MUSCLE, COOKED;
PARANCISTRUS NIGRICANS MUSCLE, COOKED;
RUBBER PLECO FLESH, COOKED;
PARANCISTRUS AURANTIACUS FLESH, COOKED;
HYPOSTOMUS VICINUS MUSCLE, COOKED;
PARANCISTRUS AURANTIACUS MUSCLE,
COOKED

Part:

MUSCLE

Date approved:

Mon Aug 31
09:56:55 EDT 2015

Last modified:

Mon Aug 31
09:56:55 EDT 2015

approved

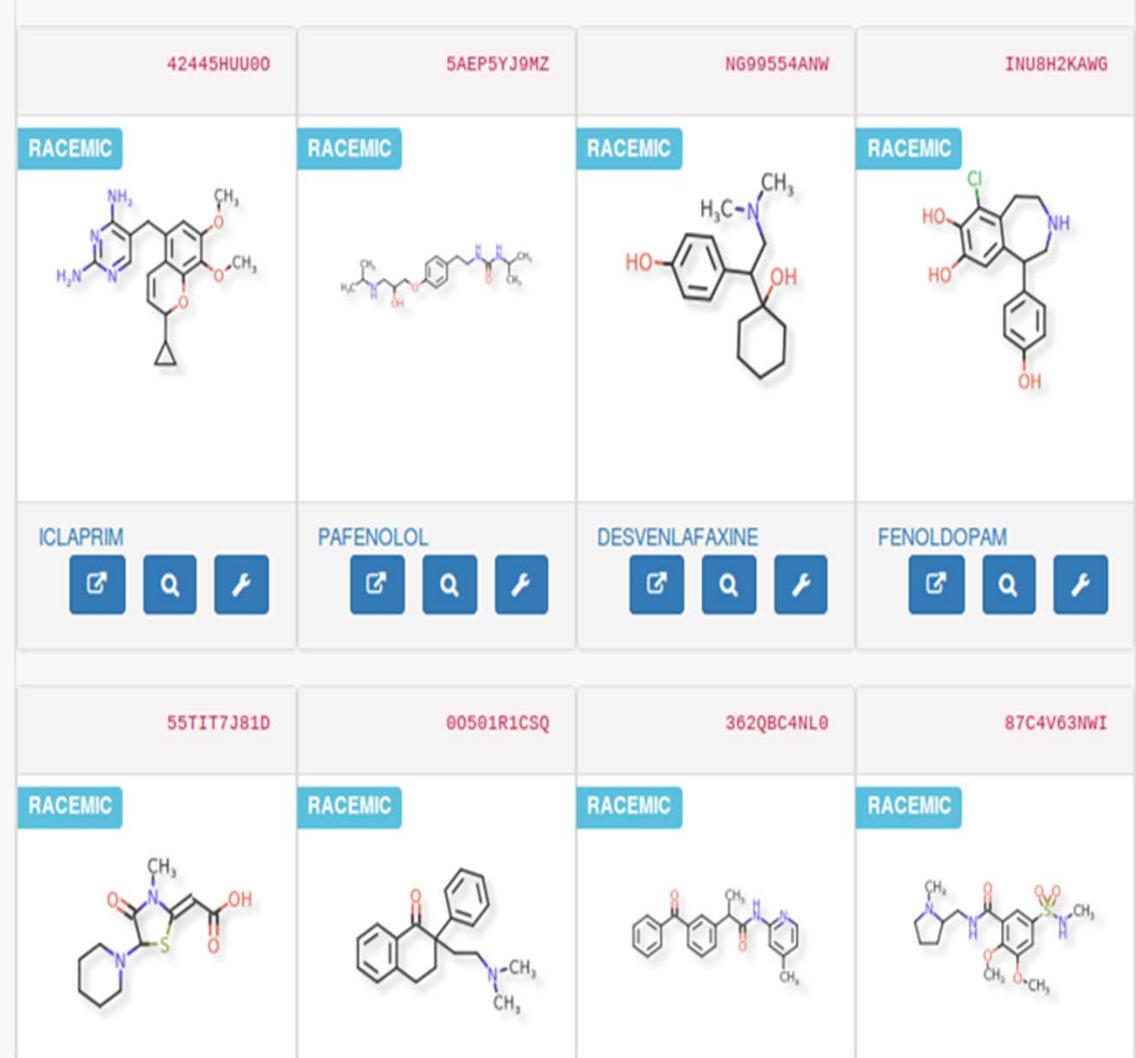
version: 1

BELLADONNA LEAF

6GZW20TIOI

<input checked="" type="checkbox"/> INN	1062
<input type="checkbox"/> WHO-DD	1055
<input type="checkbox"/> MI	537
<input type="checkbox"/> MART.	324
<input type="checkbox"/> USAN	232
<input type="checkbox"/> VANDF	193
<input type="checkbox"/> HSDB	93
<input type="checkbox"/> JAN	86
<input type="checkbox"/> ORANGE BOOK	84
<input type="checkbox"/> USP	69

Code System	
<input checked="" type="checkbox"/> EVMPD	1062
<input type="checkbox"/> CAS	1440
<input type="checkbox"/> INN	1440
<input type="checkbox"/> NCI_THESSAURUS	1256
<input type="checkbox"/> MESH	838



Substance Class

 Chemical 1440

Stereochemistry

 RACEMIC 1440
 ACHIRAL 3553
 ABSOLUTE 2477
 MIXED 177
 UNKNOWN 101
 EPIMERIC 59

Molecular Weight

 200:400 1085
 400:600 258
 0:200 184
 600:800 26
 800:1000 2

CROTETAMIDE

642I97LB5B

RACEMIC

Other Names: CROTETAMIDE; 2-BUTENAMIDE, N-((DIMETHYLAMINO)CARBONYL)PROPYL-N-ETHYL-, (E)-; 克罗乙胺; кротетамид; crotétamide; CROTEHAMIDE

Codes: SUB06829MIG; 6168-76-9; C81383; 604-64-8; 308; C015090

C12H22N2O2

Relationships: 1

Moieties: 1

Date approved:

Mon Aug 31
09:50:00 EDT 2015

Last modified:

Mon Aug 31
09:50:00 EDT 2015



G-SRS 1.0 Substance Details

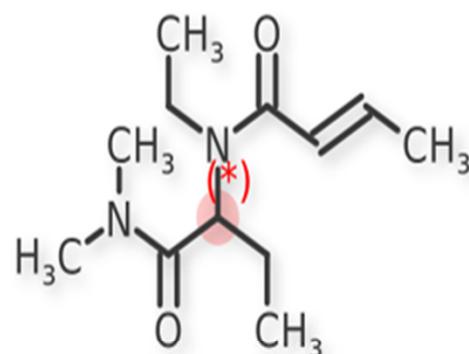
CROTETAMIDE

642197/LB5B

RECORD STATUS: APPROVED

VERSION: 1

Structure



Molecular Formula	Molecular Weight	Defined Stereocenters	EZ Centers	Charge	Reference
C12H22N2O2	226.3153	0 / 1	No	0	



[Subunits 1](#) [Disulfide Links 3](#) [Names 11](#) [Codes 8](#) [Relationships 1](#)[Properties 1](#) [References 24](#)

LEPIRUDIN

Y43GF64K34

RECORD STATUS: APPROVED

VERSION: 1

Subunits 1

Subunit 1

1-14 C:Cysteine
10 L T Y T D C T E S G 20 Q N L C L C E G S N 30 V C G Q G N K C I L 40 G S D G E K N Q C V 50 T G E G T P K P Q S 60 H N D G D F E E I P
65 E E Y L Q



Disulfide Links 3

Residue Index

1_6

1_14

1_16

1_28

1_22

1_39

Official Names 7

National Center
for Advancing
Translational Sciences

CROTETAMIDE

CROTETAMIDE	INN	MART.	WHO-DD	Drug	INN	English		7
crotetamida				Drug	INN	SPAIN		1
crotetamidum				Drug	INN	Latin		1
<hr/>								
Citation	Doc Type		Date Accessed					
INN proposed list:36	INN							
crotétamide				Drug	INN	FRANCE		1
кrottетамид				Drug	INN	RUSSIAN FEDERATION		1
كروتيناميد				Drug	INN	Arabic		1
克罗乙胺				Drug	INN	Chinese		1

www.who.int/entity/medicines/publications/druginformation/innlists/PL36.pdf

CROTETAMIDE

Official Names 7

Name	Domain	Jurisdiction	Naming Org	Language	Preferred	References
CROTETAMIDE <small>INN MART. WHO-DD</small>	Drug		INN	English		7
crotetamida	Drug		INN	SPAIN		1
crotetamidum	Drug		INN	Latin		1
crotétamide	Drug		INN	FRANCE		1
кротетамид	Drug		INN	RUSSIAN FEDERATION		1
كروتاتاميد	Drug		INN	Arabic		1
克罗乙胺	Drug		INN	Chinese		1

Names 4



National Center
for Advancing
Translational Sciences

CROTETAMIDE

References 15

Codes 7

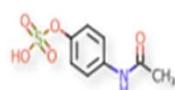
Code	Code System	Type	Description	References
SUB06829MIG	EVMPD	PRIMARY		1
6168-76-9	CAS	ALTERNATIVE		1
C81383	NCI_THESSAURUS	PRIMARY		1
604-64-8	CAS	PRIMARY		1
308	INN	PRIMARY		1
C015090	MESH	PRIMARY		1
228-208-9	EC (EINECS)	PRIMARY		1

Relationships 1

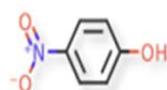
National Center
for Advancing
Translational Sciences

ACETAMINOPHEN**Relationships 13**

Name	Type	Comments	Interaction Type	Qualification	Amount	References
ACETAMINOPHEN SULFATE	METABOLITE INACTIVE (PARENT)		MAJOR	URINE	32.0	1

**Citation**<http://www.tylenolprofessional.com/pharmacology.html>**Doc Type**

DRUG PRODUCT LABEL

Date Accessed**P-NITROPHENOL****IMPURITY (PARENT)**www.tylenolprofessional.com/pharmacology.htmlNational Center
for Advancing
Translational Sciences

CROTETAMIDE

References 15**References 15**

Citation	Doc Type	Date Accessed
CROTEHAMIDE [MI]	SRS_LOCATOR	
CROTETAMIDE [INN]	SRS_LOCATOR	
SRS import [642I97LB5B]	SRS	Mon Aug 31 09:50:00 EDT 2015
MERCK INDEX	SRS	
USP DICTIONARY 2008	SRS	
CROTETAMIDE [MART.]	SRS_LOCATOR	
MERCK	SRS	
MARTINDALE 2011	SRS	
jsonDump2015-09-02.part2.txt.gz	BATCH IMPORT	Wed Sep 02 23:41:23 EDT 2015
INN proposed list:36	INN	
FDA SRS	SRS	

G-SRS 1.0 Substance Search

G-SRS 1.0 Substance Registration

Add / Edit Name

Name*

IMATINIB

Name Type*

Official Name

Language*

English - Enter values

Domain*

Drug - Enter values

United States Adopted Name



Add Organization +

Official Jurisdiction :

UNITED STATES -

Enter values

INN_LIST INN list 84

09/03/2015

Edit



Access Control :

protected

Cancel

OK



Names > Sequence > Codes > Review

Previous

Cancel

Export

Import

Next

Amino Acid Sequence

Total number of
Subunits :

Sequence Type*

- COMPLETE
- INCOMPLETE
- UNKNOWN

Subunit 1

(Length: 17)

SDFSDTSDGS DHDGDGS

[Edit Sequence](#)

(N-Terminus)

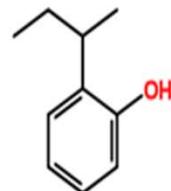
(duplicate)

(C-Terminus)

Structure



H No JSDraw License Found

C
N
O
S
P
F
Cl
Br
...
•

Molecular Formula

Molecular Weight

Defined Stereocenters

EZ
Centers

Charge

Reference ID

C10H14O

150.2176

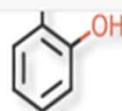
0/1

0



Moieties 1

National Center
for Advancing
Translational Sciences



Official Names

Name	Domain	Jurisdiction	Naming Org.	Language	Preferred	Reference ID
------	--------	--------------	-------------	----------	-----------	--------------



Substance name is required



Names

Name	Type	Language	Preferred	Reference ID
------	------	----------	-----------	--------------



Official Names

Name	Domain	Jurisdiction	Naming Org.	Language	Preferred	Reference ID
<input type="text"/>	<input type="text"/> Domain	<input type="text"/> Jurisdiction	<input type="text"/> Org	<input type="text"/> language	<input type="checkbox"/>	

Names

Name	Type	Language	Preferred	Reference ID
<input type="text"/>	<input type="text"/> Type	<input type="text"/> language	<input type="checkbox"/>	

Codes

Code	Code System	Type	Description	Reference ID

Select substance

Relationships

aspirin

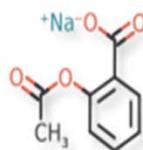


35



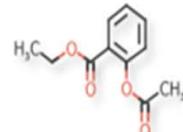
Notes

Note



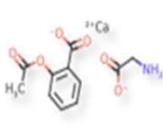
ASPIRIN SODIUM

Select



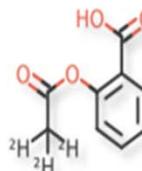
ETHYL ACETYLSALICYLATE

Select



ASPIRIN GLYCINE CALCIUM

Select



ASPIRIN CD3

Select

Selected:

OK

Cancel



Reference ID



Reference ID



Properties

Property Type	Property Name	Property Value	Parameters	Defining	Reference ID
<input type="button" value=""/>	<input type="text"/>			<input type="checkbox"/>	

[Show JSON](#)[Validate](#)[Submit](#)

ERROR

Substances must have names



WARNING

Substances should have at least one (1) preferred name



G-SRS 1.0 API data

JSON_Structure_Substances

[Page History](#)[Edit](#)

Entities are defined, exchanged, parsed, and indexed via their JSON structure. There are some JSON elements that are common to all kinds of entities, and some that are specific to certain entity classes.

There are 11 types of entities within substance-like space. The following table explains these entity types:

Entity	Category	ISO 11238	"substanceClass"
Chemical	Substance	Yes	chemical
Protein	Substance	Yes	protein
Nucleic Acid	Substance	Yes	nucleicAcid
Polymer	Substance	Yes	polyer
Structurally Diverse	Substance	Yes	structurallyDiverse
Mixture	Substance	Yes	mixture
Specified Substance Group 1	Specified Substance	Yes	specifiedSubstance
Specified Substance Group 2	Specified Substance	Yes	specifiedSubstanceG2
Specified Substance Group 3	Specified Substance	Yes	specifiedSubstanceG3
Specified Substance Group 4	Specified Substance	Yes	specifiedSubstanceG4
Virtual	Substance "Concept"	No	virtual

Names



Structur

Previous

Next

Enter or import

Export

Export :

```
{  
  "names": [  
    {  
      "type": "CN",  
      "publicDomain": true,  
      "domains": [],  
      "Drug"  
    ],  
    "languages": [  
      "en"  
    ],  
    "name": "ASPIRIN",  
    "nameOrg": [  
    ],  
    "references": [  
      "6aae1017-826c-4e92-a769-040f4abbdeb"  
    ]  
  ]  
}
```

Close

Save File



National Center
for Advancing
Translational Sciences

```
},
- {
  name: "principal",
  kind: "ix.core.models.Principal",
  href: "http://tripod.nih.gov/dev/ginias/app/api/v1/principal",
  description: "Users, groups and organizations"
},
- {
  name: "publications",
  kind: "ix.core.models.Publication",
  href: "http://tripod.nih.gov/dev/ginias/app/api/v1/publications",
  description: "Resource for handling of publications"
},
- {
  name: "references",
  kind: "ix.ginas.models.v1.Reference",
  href: "http://tripod.nih.gov/dev/ginias/app/api/v1/references",
  description: "Resource for handling of GInAS references"
},
- {
  name: "structures",
  kind: "ix.core.models.Structure",
  href: "http://tripod.nih.gov/dev/ginias/app/api/v1/structures",
  description: "Resource for handling chemical structures"
},
- {
  name: "substances",
  kind: "ix.ginas.models.v1.Substance",
  href: "http://tripod.nih.gov/dev/ginias/app/api/v1/substances",
  description: "Resource for handling of GInAS substances"
},
- {
  name: "values",
  kind: "ix.core.models.Value",
  href: "http://tripod.nih.gov/dev/ginias/app/api/v1/values",
  description: "Resource for handling Value's"
},
tripod.nih.gov/dev/ginias/app/api/v1/substances
```

```
{  
  id: 267,  
  version: 1,  
  created: 1441287623751,  
  modified: 1441287623765,  
  deprecated: false,  
  etag: "981bfa7f8711b4e6",  
  uri: "http://tripod.nih.gov/dev/ginias/app/api/v1/substances",  
  path: "/dev/ginias/app/api/v1/substances",  
  method: "GET",  
  sha1: "b3f4a3833b047bc393ea10c748f768ec7da32349",  
  total: 74269,  
  count: 10,  
  skip: 0,  
  top: 10,  
  status: null,  
  query: "",  
  filter: null,  
  _namespace: null,  
- content: [  
  - {  
    uuid: "98314be9-0231-4f2b-b22a-ad9ce354c790",  
    created: 1441263656139,  
    lastEdited: 1441263685502,  
    deprecated: false,  
    substanceClass: "structurallyDiverse",  
    status: "approved",  
    version: "1",  
    approved: 1441029415091,  
    - modifications: {  
      uid: "8a3d1a84-fe4c-46c3-a8ab-9f63a80a3d6a",  
      created: 1441287623749,  
      lastEdited: 1441263685482,  
      deprecated: false,  
      agentModifications: [ ],  
      - physicalModifications: [  
        ]  
      ]  
    }  
  }  
]
```

sha1

```
{  
  id: 266,  
  version: 1,  
  created: 1441287538233,  
  modified: 1441287538233,  
  deprecated: false,  
  etag: "26540b8550a07c6f",  
  uri: "http://tripod.nih.gov/dev/ginias/app/api/v1/substances/search?q=%22aspirin%22",  
  path: "/dev/ginias/app/api/v1/substances/search",  
  method: "GET",  
  sha1: "01a4f1a0f1038276113c306d8e756ed0ede4111a",  
  total: 35,  
  count: 10,  
  skip: 0,  
  top: 10,  
  status: null,  
  query: "\"aspirin\"",  
  filter: null,  
  _namespace: null,  
  sideway: [ ],  
  - facets: [  
    - {  
      name: "Approval ID",  
      - values: [  
        - {  
          label: "6QT214X4XU",  
          count: 1  
        },  
        - {  
          label: "5DR11472UI",  
          count: 1  
        },  
        - {  
          label: "2JJ274J145",  
          count: 1  
        },  
        - {  
          label: "5DR11472UI",  
          count: 1  
        }  
      ]  
    }  
  ]  
}
```

count



National Center
for Advancing
Translational Sciences



Record Status

Approved 4

Substance Class

Chemical 4

Stereochemistry

ACHIRAL 4

Molecular Weight

200:400 4

0:200 2

GInAS Tag

tripod.nih.gov/dev/ginias/app# MI 4

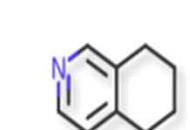
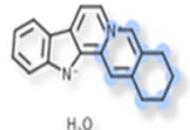
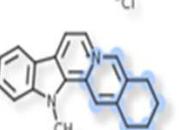
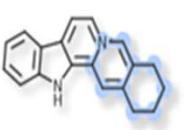
Substructure Query: C1CCC2=CN=CC=C2C1

4

F809288136

318DQW2HSM UX0WA3X59B

ACHIRAL ACHIRAL ACHIRAL ACHIRAL

SEMPERVIRINE SEMPERVIRINE MONOH... SEMPERVIRINE METHO... SEMPERVIRINE CATION

Sequence Search

Search Identity:

0.5

Query Sequence:

```
APDVQDCPECTLQEDPFFSOPGAPILOCMGCCFSRAYPTPLRSKKTMLVQKNVTSESTCCVAKSYNRVTVMGGFKVENHT  
ACHCSTCYYHKS|
```

VARFOLLITROPIN ALFA

3SD08DNC80



Other Names: VARFOLLITROPIN ALFA

Codes: 9013

Subunits: 2

Date approved:

Mon Aug 31
09:56:41 EDT 2015

Last modified:

Mon Aug 31
09:56:41 EDT 2015

approved

version: 1

Subunit 2 17a2a738-a15b-4039-b1cd-f31d2b6ca56a

identity = 0.978

APDVQDCPECTLQEDPFFSQPGAPILQCMGCCFSRAYPTPLRSKKMLVQKNVTSESTCCVAKSYNRVTVMGGFKVENHTACHCSTCYYHKS	0 - 92
---	--------

--	--

APDVQDCPECTLQENPFFSQPGAPILQCMGCCFSRAYPTPLRSKKMLVQKNVTSESTCCVAKSYNRVTVMGGFKVENHTACNCSTCYYHKS	0 - 92
---	--------



National Center
for Advancing
Translational Sciences



Properties

Property Type	Property Name	Property Value	Parameters	Defining	Reference ID
<input type="button" value=""/>	<input type="text"/>			<input type="checkbox"/>	

[Show JSON](#)[Validate](#)[Submit](#)

```
{  
  "names": [  
    {  
      "type": "cn",  
      "publicDomain": true,  
      "domains": [  
        "Drug"  
      ],  
      "languages": [  
        "en"  
      ]  
    }  
  ]  
}
```

[paste](#)

Acknowledgements

U.S. Food and Drug Administration

Yulia Borodina	Archana Newatia
Larry Callahan	Vada Perkins
Ta-Jen Chen	Mary-Ann Slack
Ramez Ghazzaoui	Frank Switzer
Elaine Johansen	Alex Welsch

U.S. Pharmacopeial Convention

Fouad Atouf	Andrej Wilk
Tina Morris	

Federal Institute for Drugs and Medical Devices (Germany)

Thomas Balzer

SwissMedic

Philipp Weyermann

European Directorate Quality of Medicines

Christopher Jarvis

NIH/NCATS

Chris Austin	Tyler Peryea
Ajit Jadhav	Tim Sheils
Dac-Trung Nguyen	Tongan Zhao

Medicines Evaluation Board (Netherlands)

Herman Diederik	Burt Kroes
Marcel Hoefnagel	Ciska Matai
Joris Kampmeijer	

Health Canada

Vikesh Srivastava

Royal Botanic Gardens, Kew (UK)

Bob Allkin	Elizabeth Dauncey
------------	-------------------

Dow Corning

Katherine Ulman

European Medicines Agency

Paolo Alcini	Telonis Pangiotis
Sabine Brosch	Ilaria Del Seppia

Uppsala Montintoring Centre / WHO

Malin Jakobsson