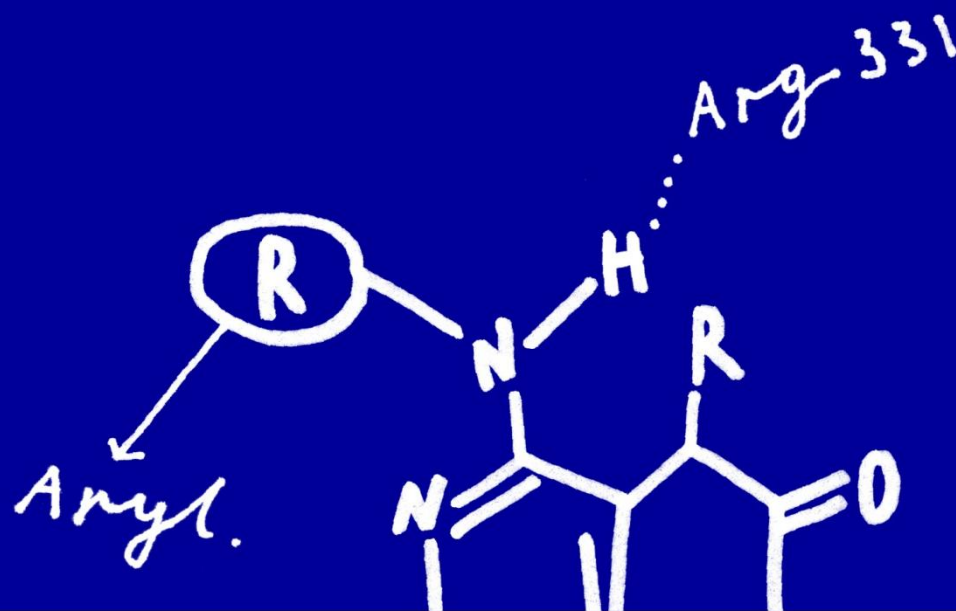

Instant JChem: Linking Chemistry & Biology Data



Agenda

- **Overview**

- Project data flow

- Schemas

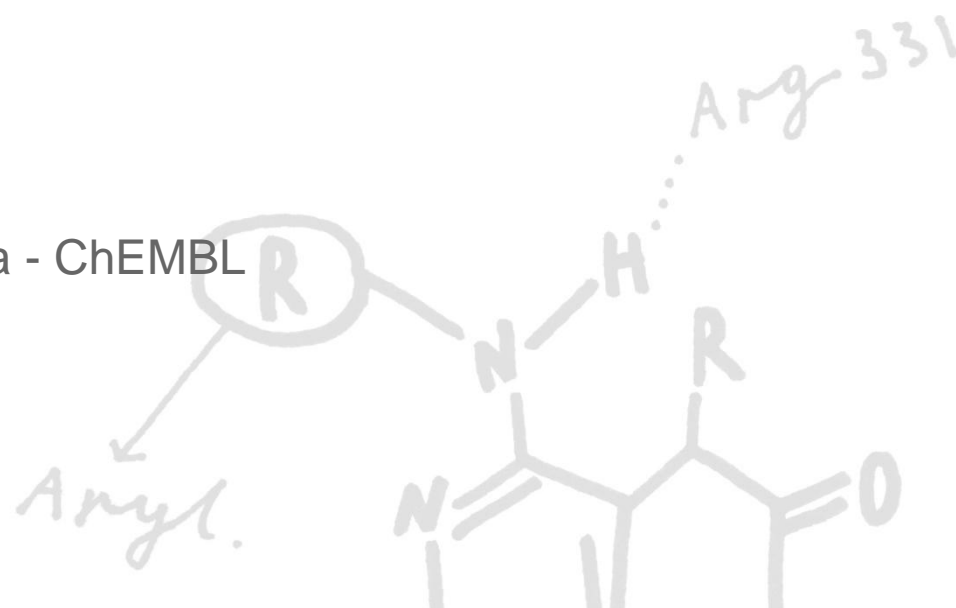
- Data model

- User security

- Deployment

- Accessing public data - ChEMBL

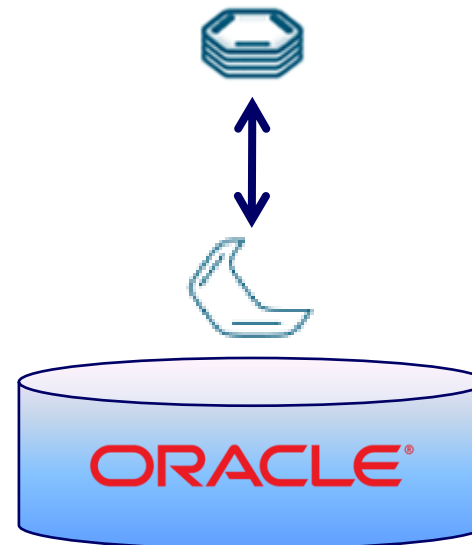
- Conclusion



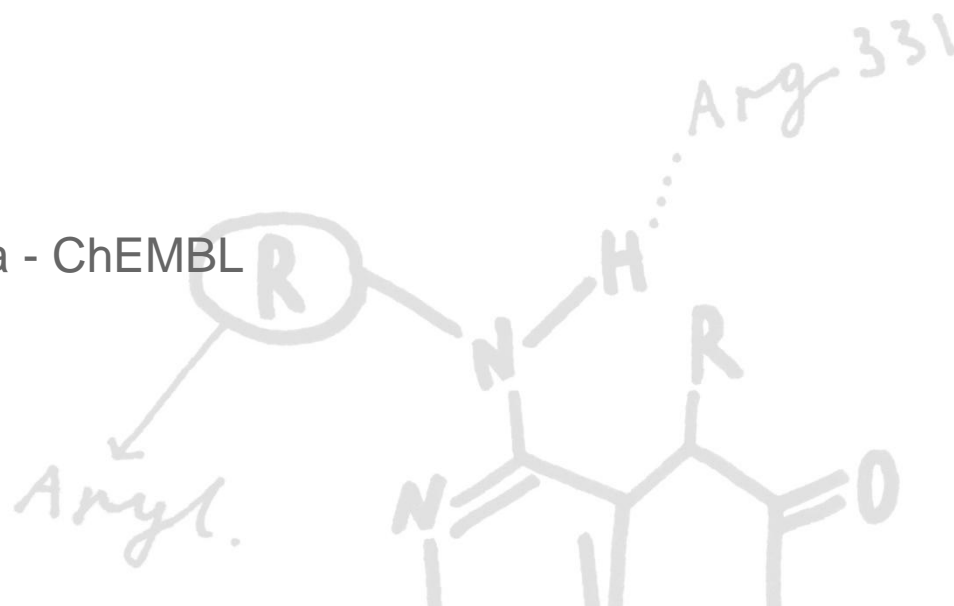
Overview

Shared IJC project database setup at Evotec

- Enterprise licenses for server use shared across Abingdon and Hamburg sites
- Dedicated IJC server at Abingdon with Oracle 10g and JChem cartridge
- One Oracle schema per drug discovery project containing project data (currently used 7 live IJC shared projects)
- A project IJC schema is connected to each Oracle schema
- Built-in IJC security & roles used for user management
- Deployment to project users via web server URLs

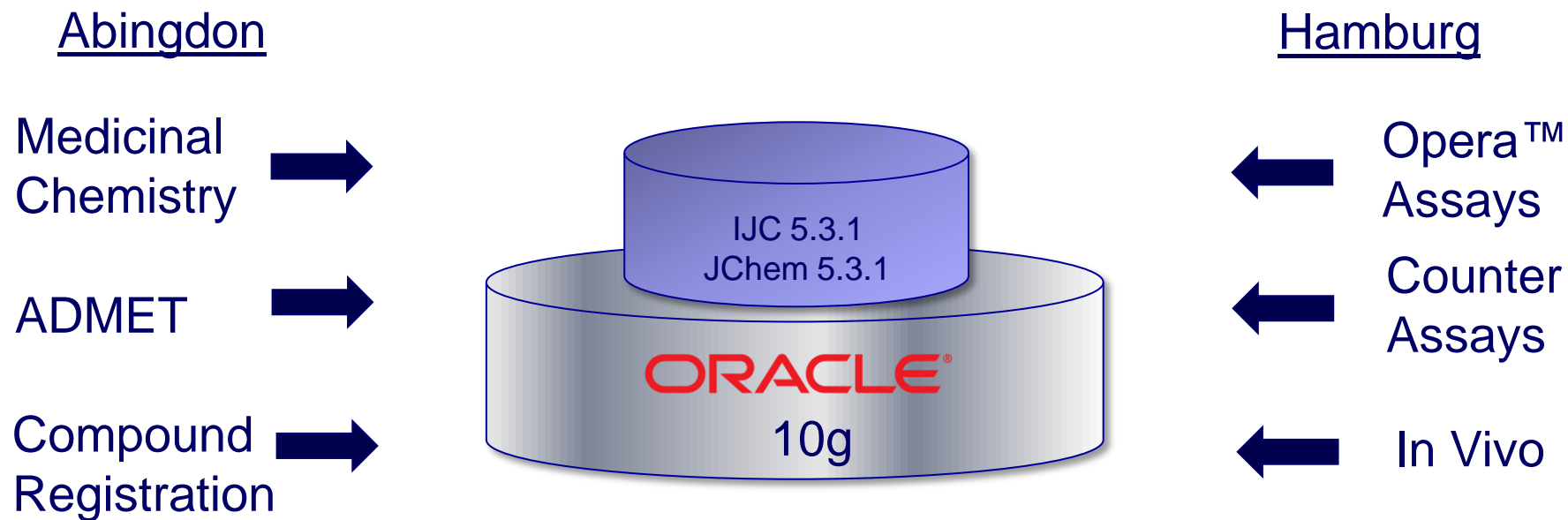


- Overview
- **Project data flow**
- Schemas
- Data model
- User security
- Deployment
- Accessing public data - ChEMBL
- Conclusion

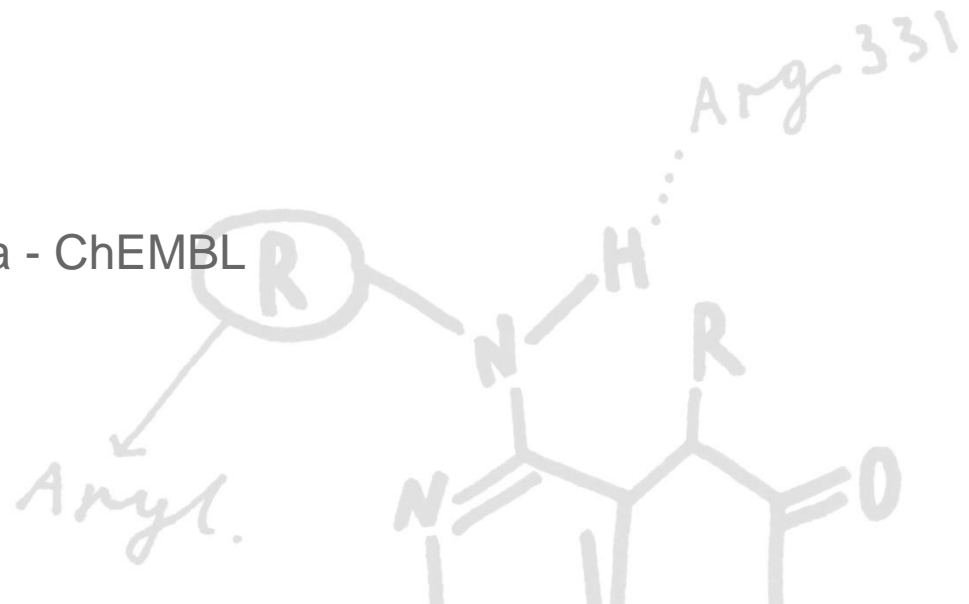


Project Data Flow

Drug Discovery IJC Project at Evotec



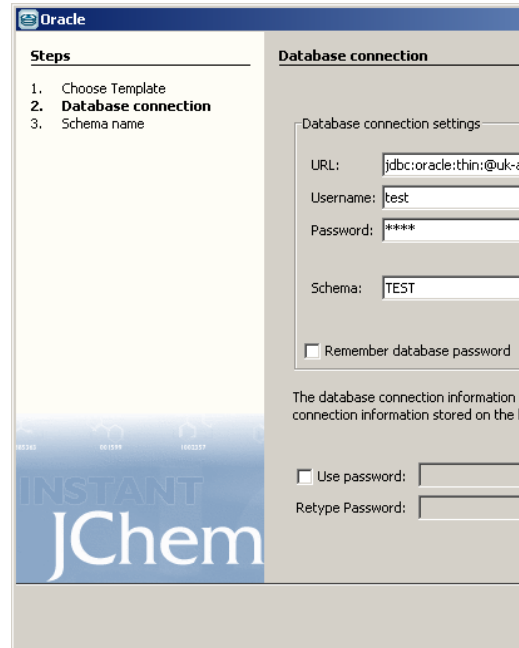
- Overview
- Project data flow
- **Schemas**
- Data model
- User security
- Deployment
- Accessing public data - ChEMBL
- Conclusion



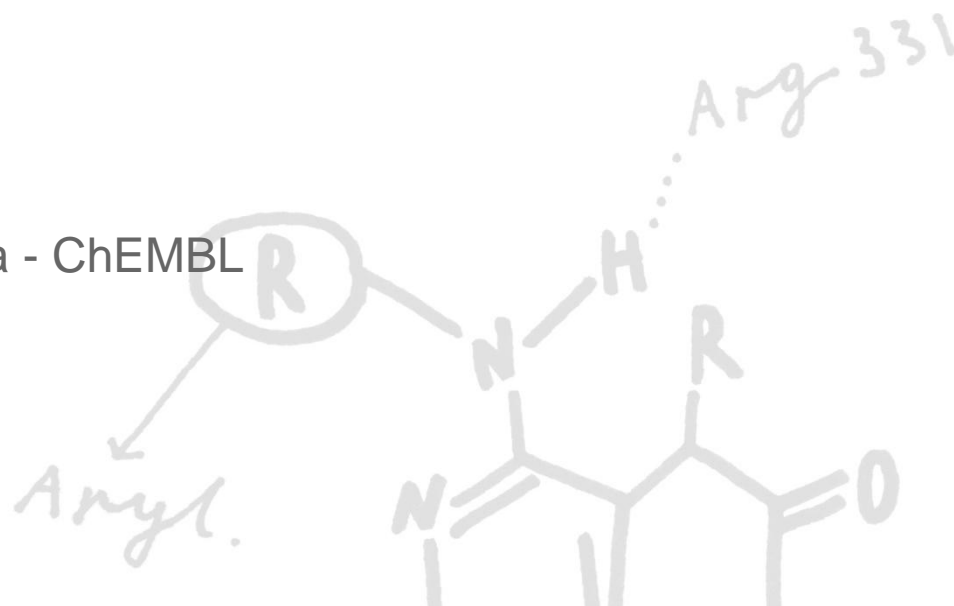
Schemas

Create Oracle & IJC Schemas

- Create project Oracle user
- Grant JCC_BASIC_ROLE to Oracle user
- Create a new empty IJC project
- Create a new schema in IJC & connect to the Oracle project schema



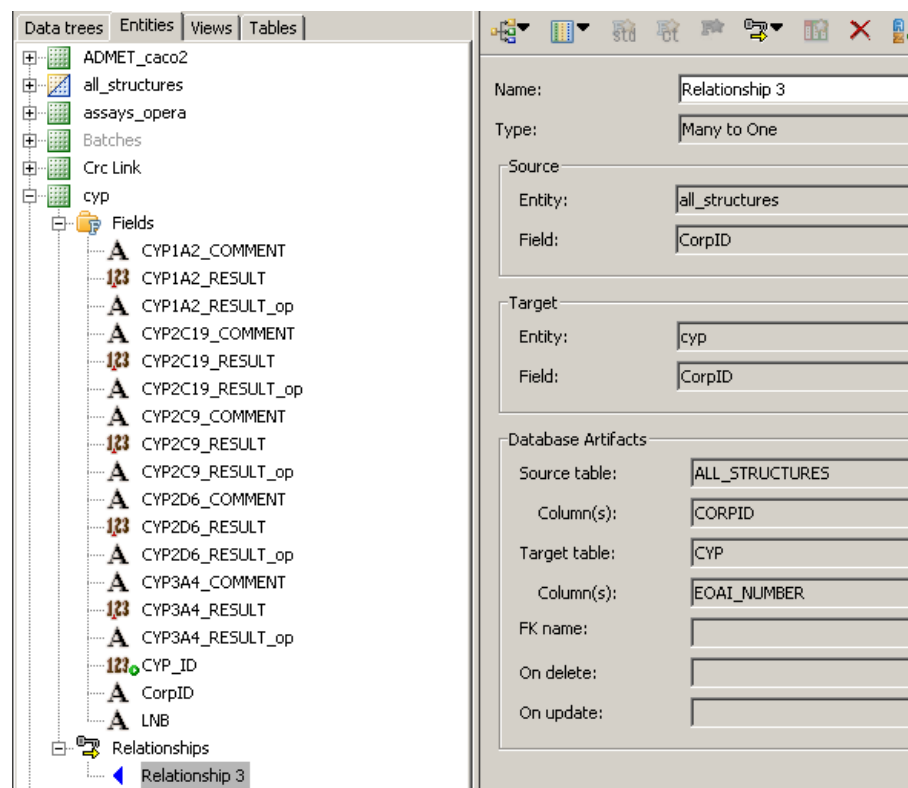
- Overview
- Project data flow
- Schemas
- **Data model**
- User security
- Deployment
- Accessing public data - ChEMBL
- Conclusion



Data Model

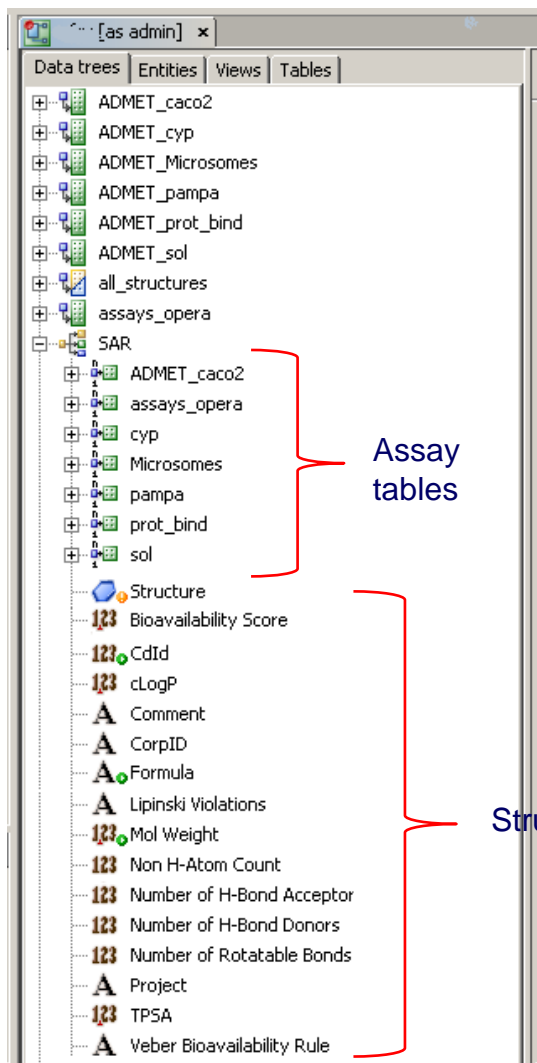
Build project databases in IJC

- Create tables
 - SDFFile import of existing data is a convenient method
 - Choose cartridge-indexed type for structure tables
- Create relationships, indexes, edges & datatrees
- Link tables on CorpID field
- Set constraints on tables & fields
- Create grid & form views
- Create project users and assign roles



Data Model

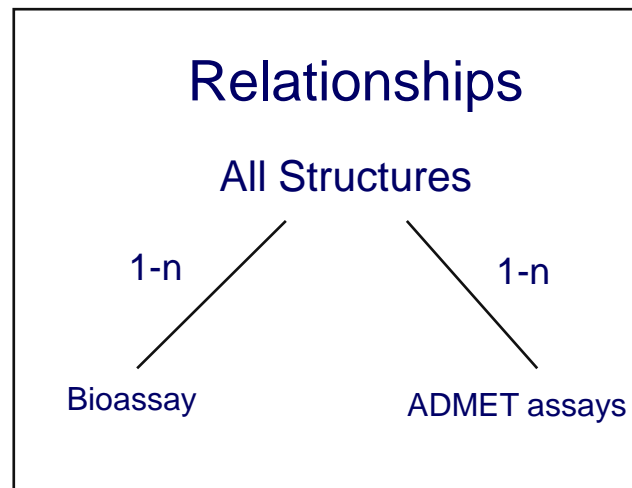
Schema Editor view: Data Trees



Assay tables

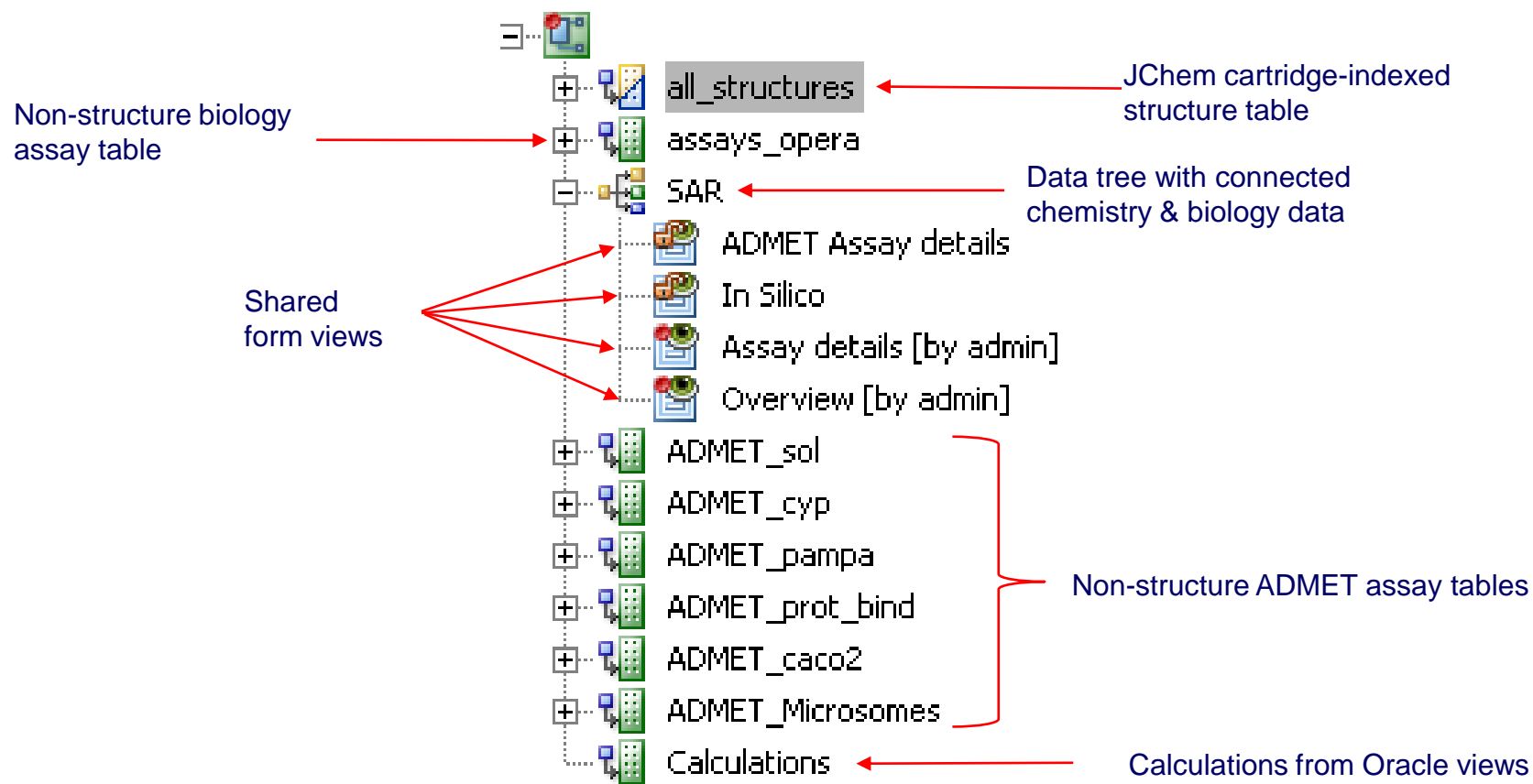
Structure table fields

- SAR data tree contains joined chemistry & biology tables, allowing many assay results for each assay type per compound structure



Data Model

User view



Data Model

User view

Overview with CRC x ADMET Assay details x In Silico x Assay details x

Design Query Browse Entity: all_struct...

Structure

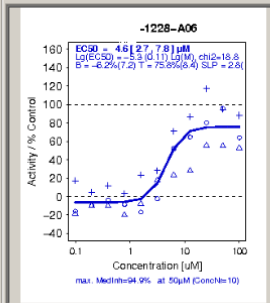
assays_opera

...	Plate	Well_ID	LNB	FPN date	Test_Date_reported
1	OP-29288B	A06	EV0865-010-006		

Ferroportin assay

...	op	IC50 [µM]	comment	op	retest 1	comment	op	retest 2	comment	op	retest 3	comment
1		4.59	efficacy 82%									

Ferroportin assay IC50 test 1



Ferroportin assay IC50 test 1 link to image

...	Fpn R0 Mfluo Pngchart	Link public 2 CRC MFLUO R0
1	Fpn.MembFluo-Batch20-100311_VIT-1228-A06.png	file:///vgc-ham-ora14\JC_SHARED\yifor_images\Fp

Ferroportin ligand efficiency -RT ln(IC50)/NHA

...	LE	retest 1 LE	retest 2 LE
1	0.38		

MDCK ATP

...	op	IC50	op	retest
1		123.90		

HEK ATP

...	op	IC50	op	retest
1	>	100.00		

Transferrin counter assay

...	op	IC50 [µM]	op	retest
1				

Ferritin counter assay

...	op	IC50
1		

Cytotox HEK cells

...	op	IC50 [µM]	comment
1			

CorpID: EV0865-010-006
 NotebookNo: OP-29288B
 PlateID: A06
 Well: A06
 Date: 04-Mar-2010
 Mol Weight:
 Formula:
 FreeMW:
 Salt_Weight:
 Salt_Formula:
 saltname:
 saltcount:
 NMR:
 LCMS: Y
 Purity: 100
 Supplier: ChemBridge
 Cat_no:
 Comment:
 Series:
 Cluster 9

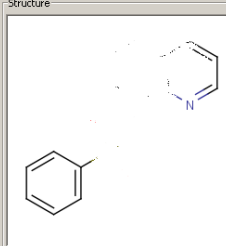
Overview with CRC x ADMET Assay details x In Silico x Assay details x

Design Query Browse Entity: all_struct...

In Silico Physical property data

Evotec no:
 Mol Weight:
 Formula:
 Lipinski Violations:
 cLogP: 2.94
 Bioavailability Score: 0.55
 Non H-Atom Count: 20
 Number of H-Bond Acceptors: 4
 Number of H-Bond Donors: 1
 Number of Rotatable Bonds: 3
 Weber Bioavailability Rule: Passes

Structure



Physicochemical property data

LNB	Pampr_pH_5-2_up	pH=6.2-cm=1	Pampr_pH_7-4_up	Pampr_pH_7-4	Pampr_pH_5_up	Pampr_pH_5	Comment
1	EV0865-010-006	96.59	131.30				

Case2

LNB	Case2_A-B	Case2_B-A	Case2_comment
1	EV0865-010-006		

Protein Binding

LNB	Prot_bind_near_unbound	Prot_bind_near_unbound_op	Prot_bind_Comment
1	EV0865-010-006	3	High Binding, h=2

Agonist/antagonist

LNB	Ag_Sol_mg_op	Ag_Sol_mg_of	Ag_Sol_Comments
1	EV0865-010-006	c	0.01 Low

Cp

LNB	CP1A2_op	CP1A2_RESULT	CP1A2_COMMENT	CP2C9_op	CP2C9_RESULT	CP2C9_COMMENT
1	EV0865-010-006	>	50.00 [C50 > 10 µM probably no drug interaction]	2.93	[C50 of 1 - 10 µM drug interactions may be an issue]	

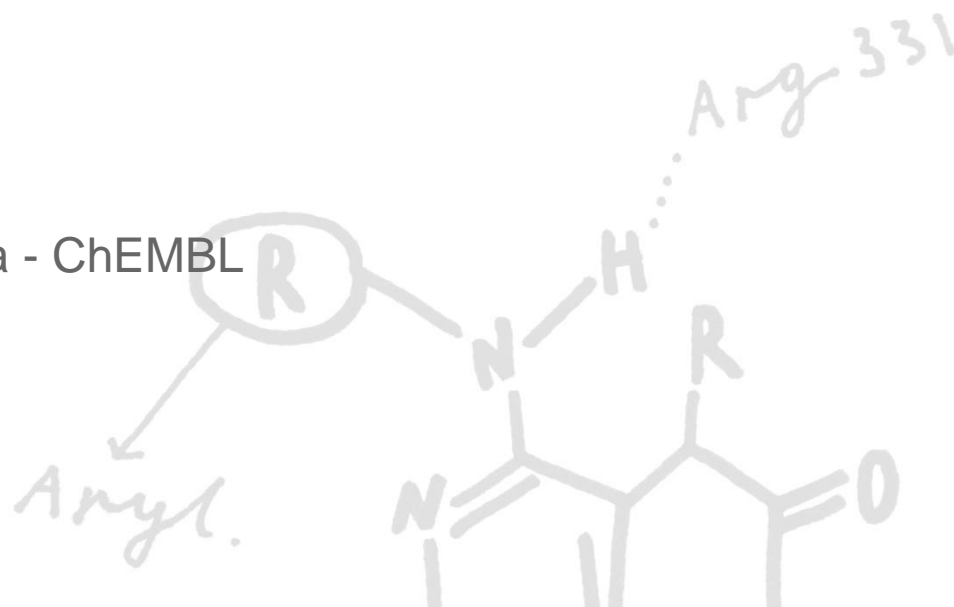
CP

LNB	CP3A4_op	CP3A4_RESULT	CP3A4_COMMENT	CP2C19_op	CP2C19_RESULT	CP2C19_COMMENT	CP2C6_op	CP2C6_RESULT	CP2C6_COMMENT
1	EV0865-010-006	>	90.00 [C50 > 10 µM probably]	5.27	[C50 of 1 - 10 µM]	>	50.00	[C50 > 10 µM probably]	

Microsomes

LNB	MFL_half-life_human_op	MFL_half-life_human	MFL_half-life_mouse_op	MFL_half-life_mouse	MFL_Comment
1	EV0865-010-006	1.31	2.69		

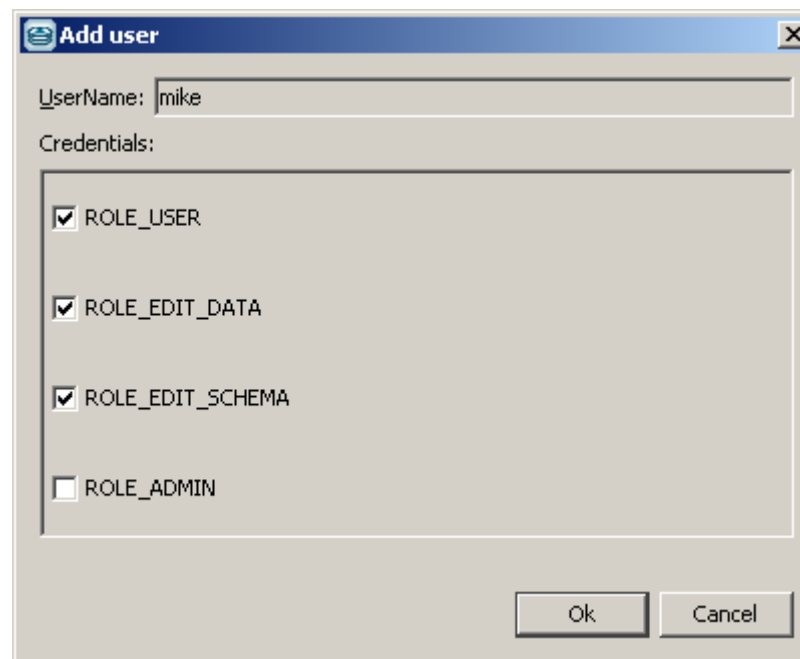
- Overview
- Project data flow
- Schemas
- Data model
- **User security**
- Deployment
- Accessing public data - ChEMBL
- Conclusion



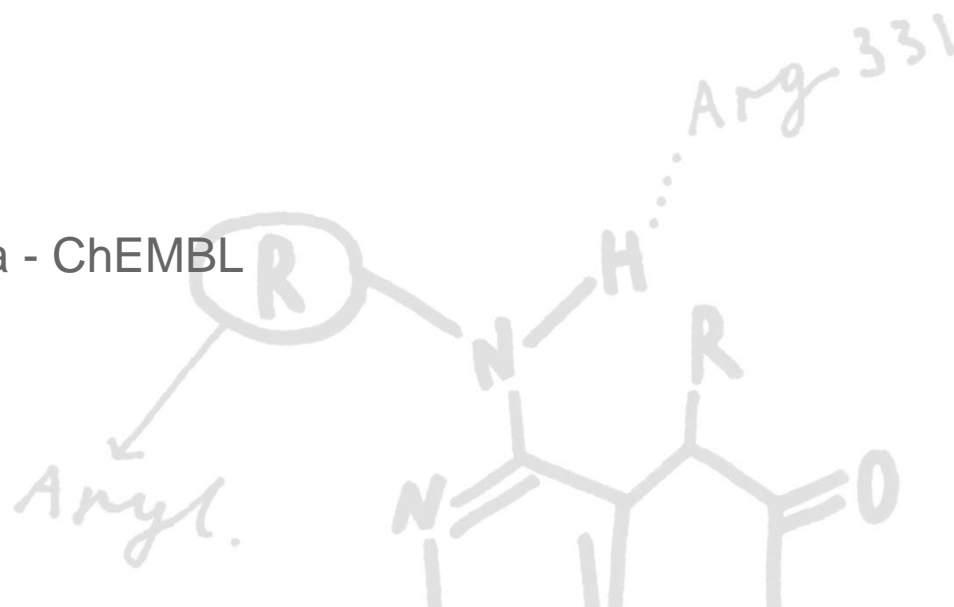
User Security

Typical IJC setup for a small drug discovery project

- 4 Scientists with Edit Data Role (2 in Abingdon & 2 in Hamburg)
- Updates made via SDfile (Edit Data Role) with auto mapping (SDfiles for loading formatted in the Evotec CCD)
- 14 Scientists with Read Only Role (ROLE USER) across Abingdon and Hamburg
- 2 Administrators (1 in Abingdon & 1 in Hamburg) with Edit Schema Role
- Built in Admin user used for rebuilding JChem tables after software updates & managing users and roles

A screenshot of a Windows-style dialog box titled "Add user". It has a blue title bar with a close button (X) in the top right corner. The dialog contains a text input field for "UserName:" with the value "mike". Below it is a section labeled "Credentials:" containing a list of roles with checkboxes: ROLE_USER, ROLE_EDIT_DATA, ROLE_EDIT_SCHEMA, and ROLE_ADMIN. At the bottom right, there are two buttons: "Ok" and "Cancel".

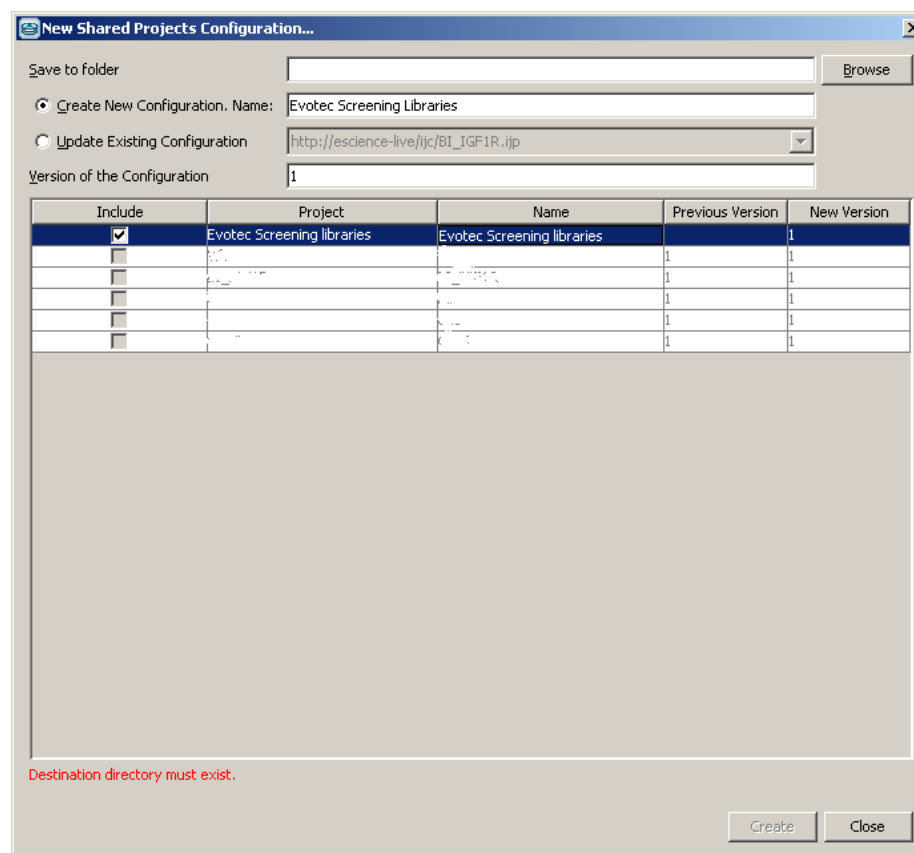
- Overview
- Project data flow
- Schemas
- Data model
- User security
- **Deployment**
- Accessing public data - ChEMBL
- Conclusion



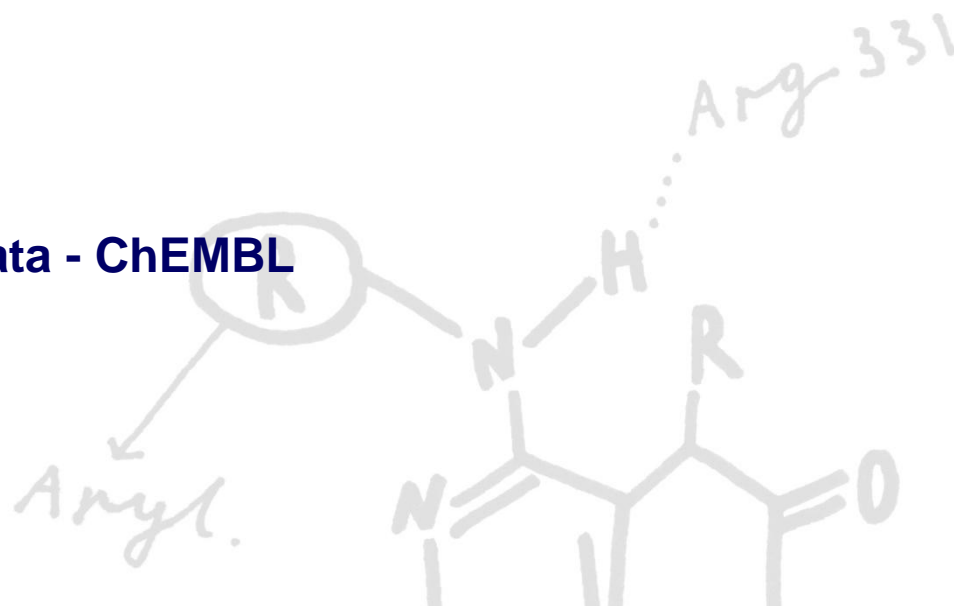
Deployment

Via shared URLs

- Create users and roles & make all views you want to deploy shared
- Use Create Shared Projects wizard to generate new configuration files (.ijp and .zip files)
- Place configuration files on a web server: <http://escience/ijc/>
- Provide user account with initial password and URL e.g. <http://escience/ijc/project.ijp> to all of the users

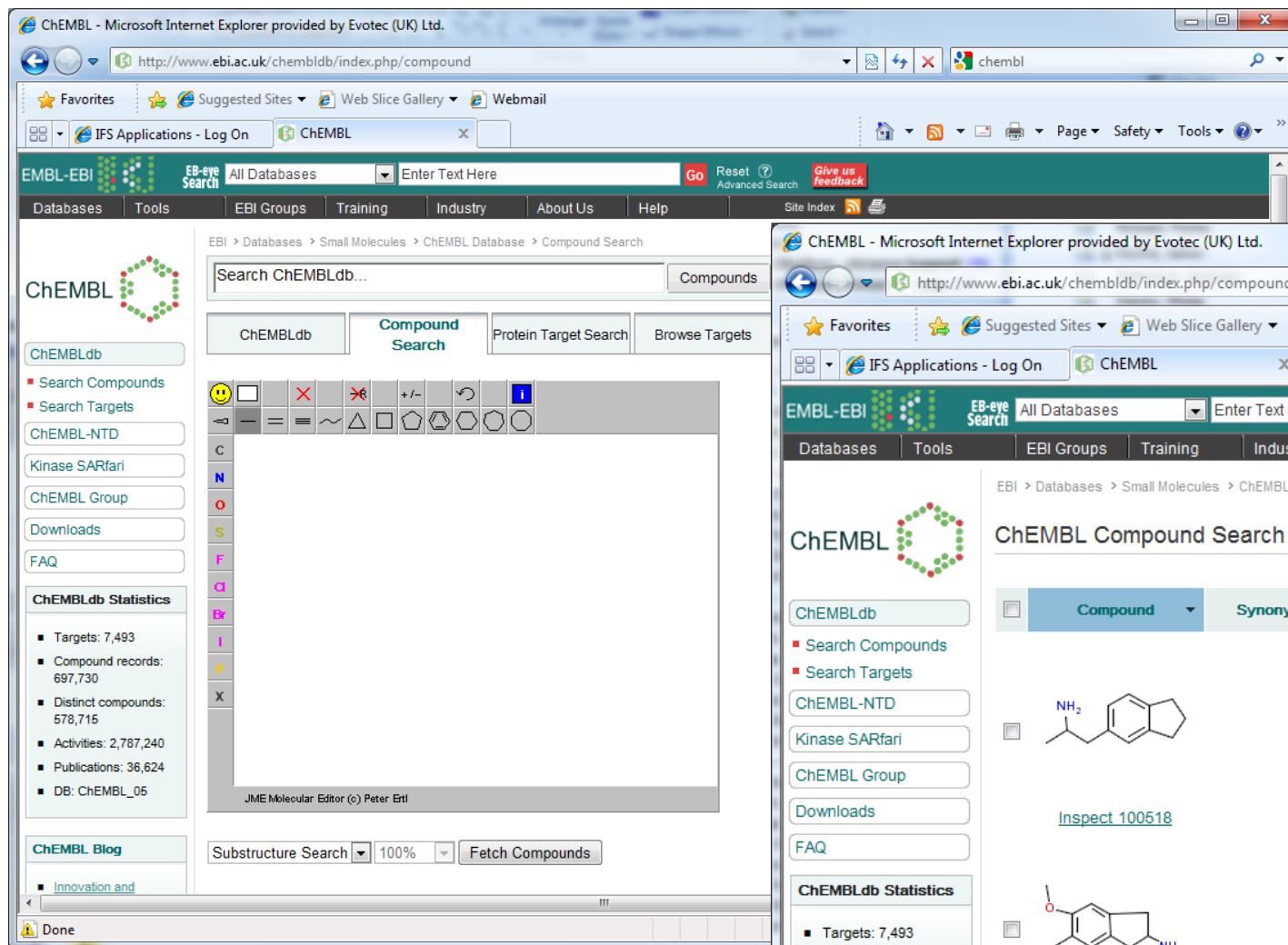


- Overview
- Project data flow
- Schemas
- Data model
- User security
- Deployment
- **Accessing public data - ChEMBL**
- Conclusion



Accessing public data

Viewing ChEMBL data



ChEMBL - Microsoft Internet Explorer provided by Evotec (UK) Ltd.

http://www.ebi.ac.uk/chembl/index.php/compound

EMBL-EBI EB-eye Search All Databases Enter Text Here Go Reset Advanced Search Give us feedback

Databases Tools EBI Groups Training Industry About Us Help Site Index

EBI > Databases > Small Molecules > ChEMBL Database > Compound Search

Search ChEMBLdb... Compounds

ChEMBLdb Compound Search Protein Target Search Browse Targets

ChEMBLdb

- Search Compounds
- Search Targets

ChEMBL-NTD

Kinase SARfari

ChEMBL Group

Downloads

FAQ

ChEMBLdb Statistics

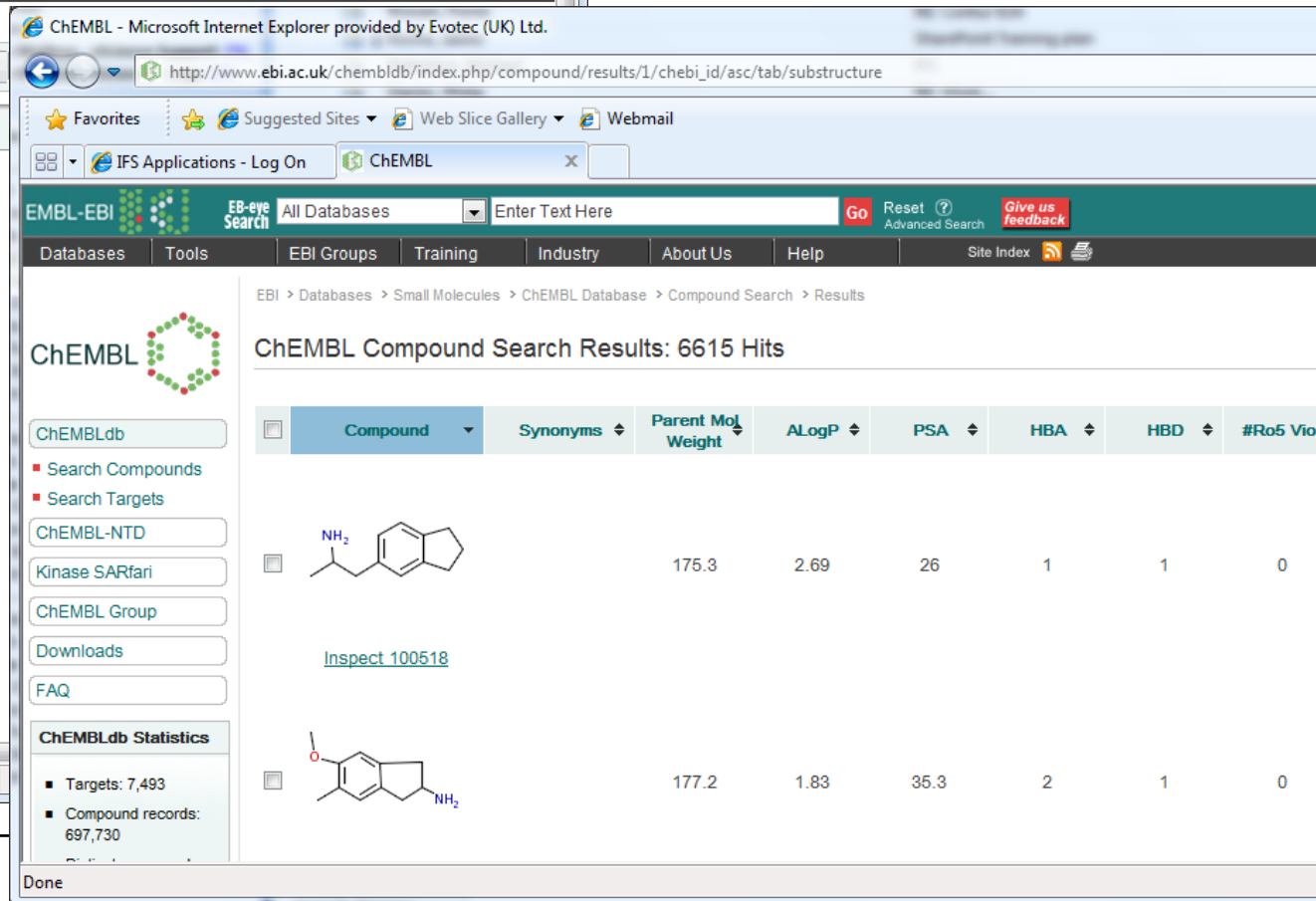
- Targets: 7,493
- Compound records: 697,730
- Distinct compounds: 578,715
- Activities: 2,787,240
- Publications: 36,624
- DB: ChEMBL_05

ChEMBL Blog

- Innovation and

JME Molecular Editor (c) Peter Ertl

Substructure Search 100% Fetch Compounds



ChEMBL - Microsoft Internet Explorer provided by Evotec (UK) Ltd.

http://www.ebi.ac.uk/chembl/index.php/compound/results/1/chebi_id/asc/tab/substructure

EMBL-EBI EB-eye Search All Databases Enter Text Here Go Reset Advanced Search Give us feedback

Databases Tools EBI Groups Training Industry About Us Help Site Index

EBI > Databases > Small Molecules > ChEMBL Database > Compound Search > Results

ChEMBL Compound Search Results: 6615 Hits

ChEMBLdb

- Search Compounds
- Search Targets

ChEMBL-NTD

Kinase SARfari

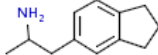
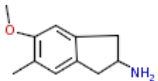
ChEMBL Group

Downloads

FAQ

ChEMBLdb Statistics

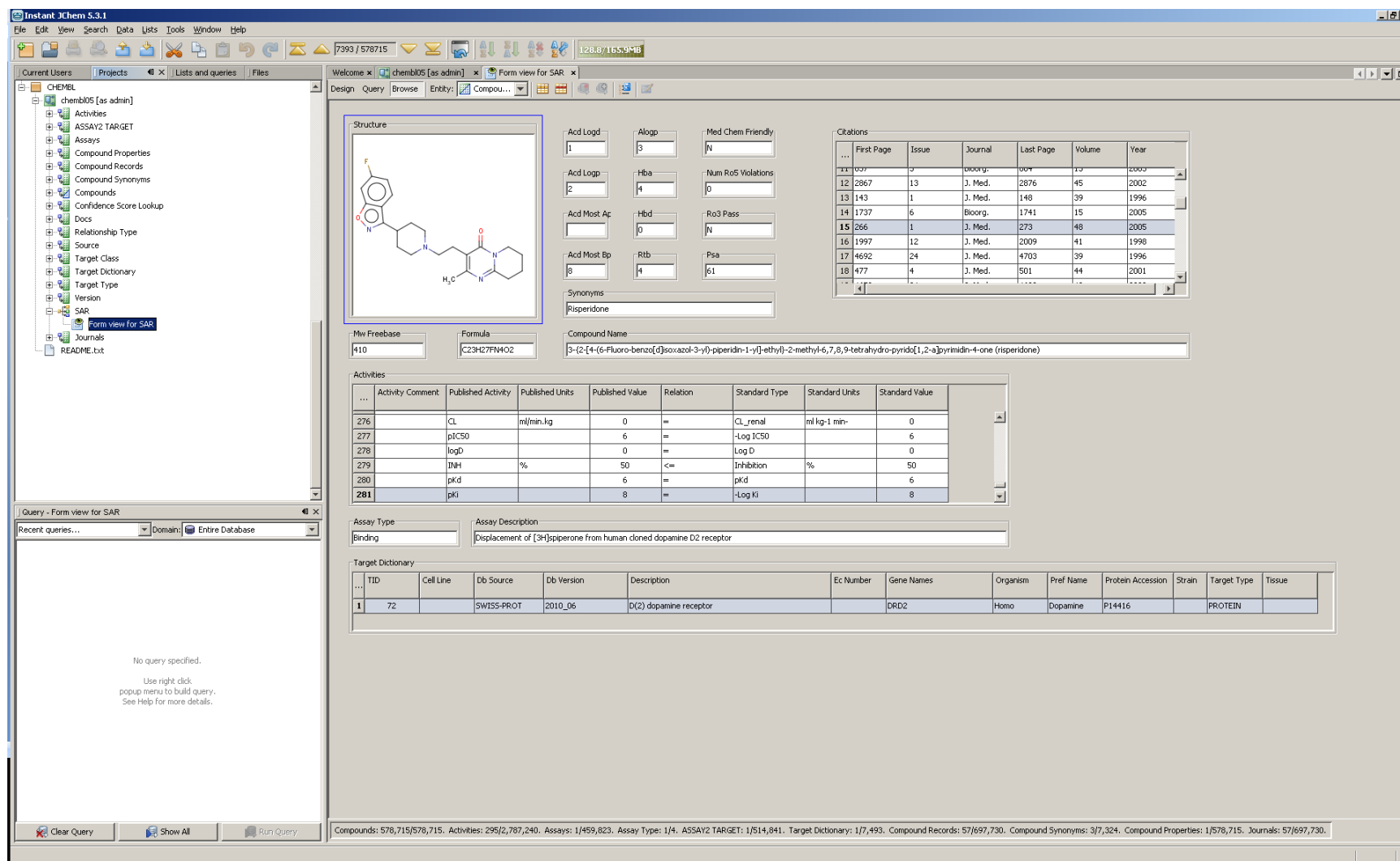
- Targets: 7,493
- Compound records: 697,730

Compound	Synonyms	Parent Mol Weight	ALogP	PSA	HBA	HBD	#Ro5 Viol
		175.3	2.69	26	1	1	0
Inspect 100518							
		177.2	1.83	35.3	2	1	0

Done

Accessing public data

Viewing ChEMBL data through IJC



The screenshot displays the Instant JChem 5.3.1 interface. The left sidebar shows a tree view of ChEMBL data categories. The main window is titled 'Form view for SAR' and contains the following sections:

- Structure:** A chemical structure diagram of Risperidone.
- Form Fields:**
 - Acid Logp: 1
 - Allogp: 3
 - Med Chem Friendly: N
 - Acid Logp: 2
 - Hiba: 4
 - Num Ro5 Violations: 0
 - Acid Most Ac: 0
 - Hbdb: 0
 - Ro3 Pass: N
 - Acid Most Bp: 8
 - Rtb: 4
 - Psa: 61
 - Synonyms: Risperidone
- Citations:**

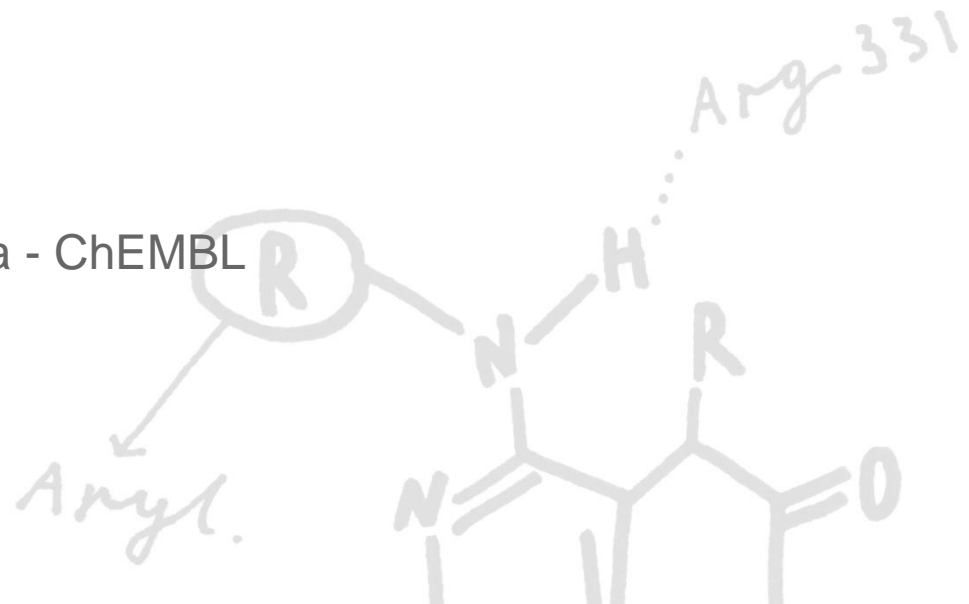
...	First Page	Issue	Journal	Last Page	Volume	Year
12	2867	13	J. Med.	2876	45	2002
13	143	1	J. Med.	148	39	1996
14	1737	6	Bioorg.	1741	15	2005
15	266	1	J. Med.	273	48	2005
16	1997	12	J. Med.	2009	41	1998
17	4692	24	J. Med.	4703	39	1996
18	477	4	J. Med.	501	44	2001
- Freebase/Formula/Compound Name:**
 - Freebase: 410
 - Formula: C23H27FM4O2
 - Compound Name: [3-[-2-[4-(6-Fluoro-benzo[d]isoxazol-3-yl)-piperidin-1-yl]-ethyl]-2-methyl-6,7,8,9-tetrahydro-pyrido[1,2-a]pyrimidin-4-one (risperidone)
- Activities:**

...	Activity Comment	Published Activity	Published Units	Published Value	Relation	Standard Type	Standard Units	Standard Value
276	CL		ml/min.kg	0	=	CL_renal	ml.kg-1.min-	0
277	pIC50			6	=	-Log IC50		6
278	logD			0	=	Log D		0
279	INH	%		50	<=	Inhibition	%	50
280	pKd			6	=	pKd		6
281	pKi			8	=	-Log Ki		8
- Assay Type/Description:**
 - Assay Type: Binding
 - Assay Description: Displacement of [3H]spiperone from human cloned dopamine D2 receptor
- Target Dictionary:**

...	TID	Cell Line	Db Source	Db Version	Description	Ec Number	Gene Names	Organism	Pref Name	Protein Accession	Strain	Target Type	Tissue
1	72		SWISS-PROT	2010_06	D(2) dopamine receptor		DRD2	Homo	Dopamine	P14416		PROTEIN	

At the bottom, a status bar provides summary statistics: Compounds: 578,715/578,715; Activities: 295/2,787,240; Assays: 1/459,823; Assay Type: 1/4; ASSAY2 TARGET: 1/514,841; Target Dictionary: 1/7,493; Compound Records: 57/697,730; Compound Synonyms: 3/7,324; Compound Properties: 1/578,715; Journals: 57/697,730.

- Overview
- Project data flow
- Schemas
- Data model
- User security
- Deployment
- Accessing public data - ChEMBL
- **Conclusion**



Conclusion

Overview and future

- IJC is providing Evotec a valuable resource for storing and sharing Chemistry and Biology data for our client projects
- We are using IJC as a replacement for ISIS/Base databases with positive feedback from the users
- We are still finding new uses for it!

Future Plans

- Integrate our own calculators & methods into Marvin / IJC
- Integrate more ways of displaying results (graphing, etc.) (coming!)
- Implement user dependent permissions on viewing & editing entities (coming!)
- Improved ways of sharing IJC data with our clients

Thank you

- Mike Potterton
- Bob Marmon
- Catherine Reisser
- Jane Couzens
- Anne Pintat
- Steve Courtney
- Mike Mazanetz
- Robert Godemann
- Christian Kirchhoff
- Alex Böcker
- Tim Dudgeon
- IJC Developers

Your contact:

Ian Berry
eScience Group Leader

+44.(0).1235.441451
ian.berry@evotec.com

