

# Marvin and JChem + Proteax

- Benefits of integration

*Jan Holst Jensen*

*CEO, Biochemfusion*

***biochemfusion***  
- *Enabling biochemformatics*

# Bridging the gap

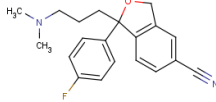
Cheminformatics

Bioinformatics

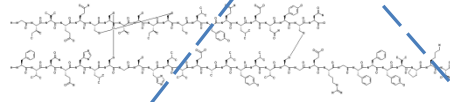
neither-nor

Molecule graphs

Sequences



```
1  GIVEQCCTSICSLYQLENYC  
21 NFVNQHLCGSHLVEALYLVC  
41  GERGFFYTPKT
```



```
1  MVSQALRLLCLLLCLGQCCLAAGGVAKASGCETRDMPWRKPC  
41  PHRVFTVQEEAHCVLHRRRRANAFLEELRPGSLERECKRE  
81  QCSFEEAREIFKDAERTKLFMISYSDGDCASSPCQNGGS  
121 CRDQLQSYICFCLPAFEGRNCETHRDDQLICVNEGCGCQ  
161 YCSDHTGTRRSCRCHEGYSLLADGVSCTPVEYPCGKIPI  
201 LEKRNASKPQGRIVGCKVCFKGCECPVQVLLLVNCAQLCGG  
241 TLINTIVVVSAAHCFDKIKNWRNLIAVLGEHDLSEHDGDE  
281 QSRRVAQVIIPSTYVPGTINHDIALLRLHQPVVLTDHVVE  
321 LCLPERTFSERTLAFVFSVLSGWCQLLDRCATALELMVE  
361 NVPRLHTQDCLQQSRKWCDSPNITETHFCACYSDGSKRDSC  
401 KGDSGCGPHATHYRGTWYLTGIVSMGCCCATVGHFGVYTRV  
441 SQYIEWLQKLMRSEPRFCVLLRAPFF
```

100

10k

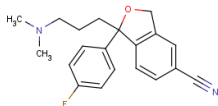
1M

MW  
Da

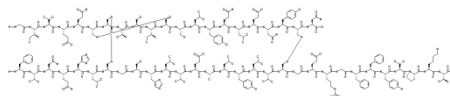
JChem for Excel  
 JChem Extensions for KNIME  
 Structures4MorphIt  
 JChem Cartridge  
 ...

Proteax for Spreadsheets  
 Proteax for KNIME  
 Proteax4MorphIt  
 Proteax Cartridge  
 ...

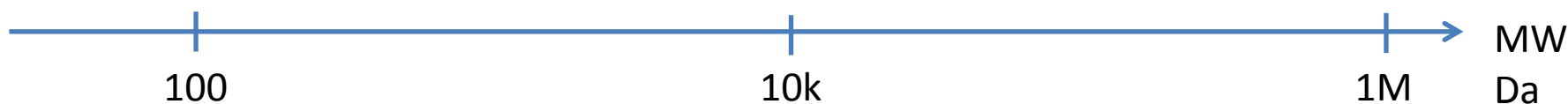
*product integration*



1 GIVEQCCTSLYQLENYC  
 21 NFVNQHLCGSHLVEALYLVC  
 41 GERGFFYTPKT



1 MVSQALRLLCLLLGLGCLAAAGVAKASGCETRDMPWRPG  
 41 PRRVFTQEEAHCVLHRRRANAFLLELRPGSLERICKRE  
 81 QCSFSEARRIFKDAERTKLFWISYSDGDCASSPCQNGCS  
 121 CRDQLQSYICFCLPAFEGRCNCETHRDDQLICVNEGGCEQ  
 161 YCSDHITGTRRSRCRCEGYSLLADCVSCTPTVEYPCCKIPF  
 201 LEKRNASKPQGRIVGCKVCFKGECPPQVLLLVNGAQLCGG  
 241 LLIMTIWVVSAAHCFDKIRNWRMLIAVLGEHDLSEHDGDE  
 281 QSRVAQVLIIPSTYVPGTIMHDIALLRLHQPVVLTDRHVV  
 321 LCLPERTFSERTLAFVVFSLVSGWQLLDRCATALELHVE  
 361 MVRRLHTQDCLQQSRKWCDSPNITZYHFCACYSDGSKDSC  
 401 KDGSGGPHATHYRGTWYLTGIVSWGQCCATVGHFGVYTRV  
 441 SQYIEWLQKLMRSEPRFCVLLRAPFF

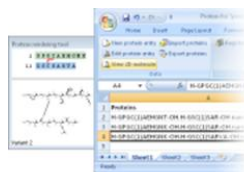


# Proteax<sup>®</sup> – easily integrated

## Proteax for Spreadsheets

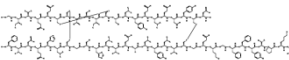
Spreadsheet functions,  
macro programming

MS Excel<sup>®</sup>  
OpenOffice.org<sup>®</sup>



Sequence	Molecular weights	Da
1 NEVQDQVETGK	Average	5887.51922
21 NEVQDQVETGK	Standard Unifrac	5795.64874
41 NEVQDQVETGK		

3D molecule



Java<sup>®</sup> applet, Flash<sup>®</sup>-based,  
JavaScript<sup>™</sup>, .NET, and Pascal  
visualizers

Structure and  
sequence visualization

## Proteax for KNIME

KNIME nodes for graphical  
processing of protein data

KNIME<sup>™</sup>  
graphical workflows



## Proteax Desktop

.NET, Java<sup>®</sup>, Python,  
C/C++, Pascal

Scripting,  
programming

Proteax toolkit

Proteax Cartridge  
Database integration  
registration and querying

Oracle<sup>®</sup>

```
select * from compounds  
where proteax.mw_avg(protein_text) > 5000;
```

PostgreSQL<sup>®</sup>

# Vendor example – Sysment, Hungary

- Using ChemAxon's JChem and Biochemfusion's Proteax toolkits together
- Using Proteax's .NET API and JavaScript renderers

Synthesis of  $\epsilon$ -NB29-dehydrochoyl insulin

Available Section Types

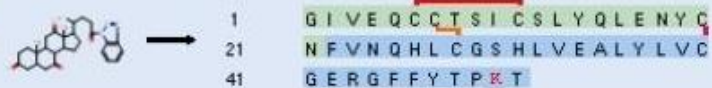
- [New Html Section](#)
- [New Text Section](#)
- [New Reaction Section](#)

Existing Sections

- [Synthesis of  \$\epsilon\$ -NB29-dehydrochoyl insulin](#)

Include in Extract

1 G I V E Q C C T S I C S L Y Q L E N Y C  
21 N F V N Q H L C G S H L V E A L Y L V C +  
41 G E R G F F Y T P K T

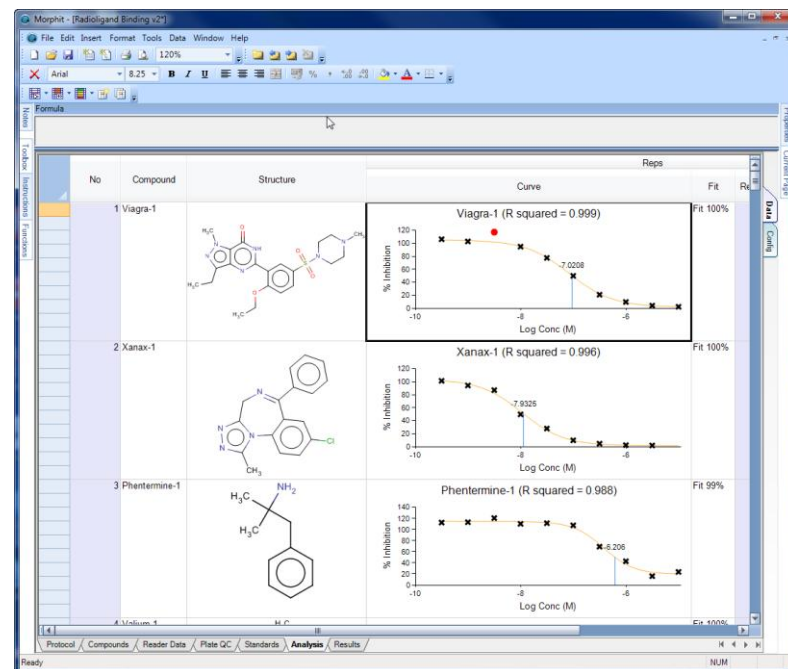


1 G I V E Q C C T S I C S L Y Q L E N Y C  
21 N F V N Q H L C G S H L V E A L Y L V C  
41 G E R G F F Y T P K T

- Resulting in the web-based *Sysment Notebook* ELN with full chemistry *and* biologics support

# Vendor example – The Edge, UK

- The Edge's *MorphIt* is a relational spreadsheet – specifically designed to handle research data
- *Structures4MorphIt*
  - a MorphIt plugin that uses ChemAxon renderers for chemical structure visualization





# Links...

- Proteax for Spreadsheets
  - <http://www.biochemfusion.com/downloads/#ProteaxForSpreadsheets>
  - <http://www.chemaxon.com/download/jchem-for-excel/>
  - Direct data link from Proteax for Spreadsheets to JChem for Excel
- Sysment ELN
  - <http://www.sysment.hu/>
- The Edge
  - [http://www.edge-ka.com/products/morphit/morphit\\_plugins](http://www.edge-ka.com/products/morphit/morphit_plugins)
  - *Structures4Morphit* - structure visualizer plug-in using ChemAxon renderers
  - *Proteax4Morphit* - protein plug-in using Biochemfusion's Proteax toolkit and renderers - available later this year