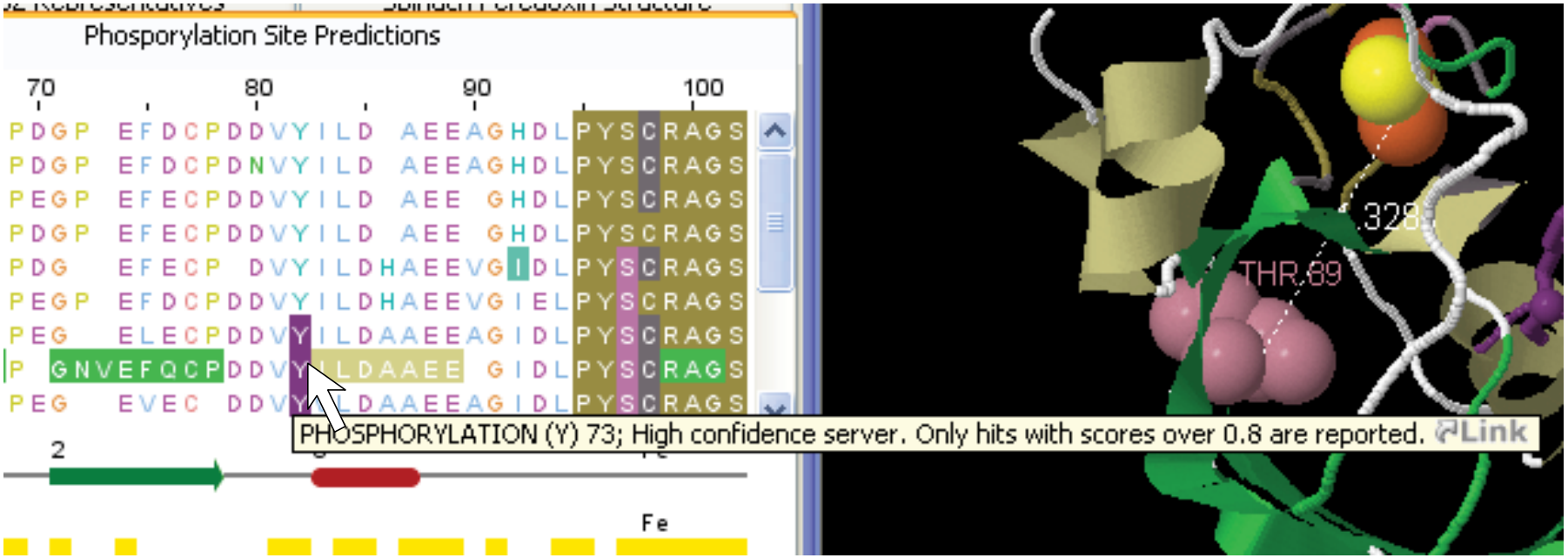
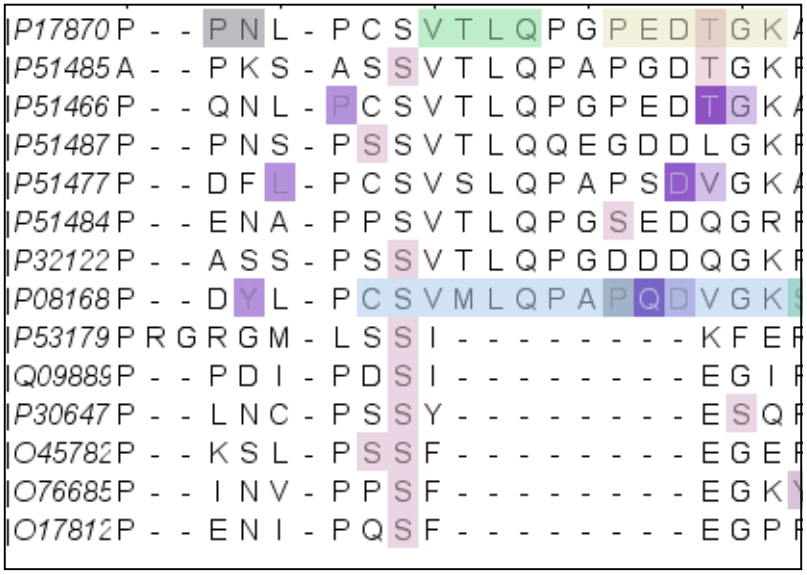
- End of Session 2
  - Alignment & analysis

- Session 3
  - Sequence DB refs and Sequence Features
  - Protein secondary structure prediction
  - Protein disorder prediction

Sequence Features

Manual section 2.8





Sequence Feature Settings

Feature Settings | DAS Settings

- [x] **uniprot**
- [ ] **Pfam Other Features**
- [x] **PDBsum_protprot**
- [x] **cbs_total**

Feature Type	Colour	Display
DISULFID		☑
Protein-protein contact		☑
MOD_RES		☑
ISOFORM		☑
PHOSPHORYLATION (S)		☑
PHOSPHORYLATION (Y)		☑
INIT_MET		☑
PHOSPHORYLATION (T)		☑
NES-SIGNAL		☑
CONFLICT		☑
ACETYLATION		☑
TURN		☑
HELIX		☑
STRAND		☑
SIGNAL		☑
COMPBIAS		☑
PROSITE		☑
PRINTS		☑
PROPEPTIDE		☐
OUTSIDE		☐
CHAIN		☐
Pfam		☐
ProDom		☐

Invert Selection

OK | Cancel | Load Colours | Save Colours

Creating Sequence Features Section 2.8.1-3 & Ex 28

- Tasks
  - Visualise, create, modify, import and export features.

Sequence Features
Section 2.8.1-3 & Ex 28

- Tasks
  - Visualise, create, modify, import and export features.
- Questions
  - What are the different types of file formats available for import and export ?

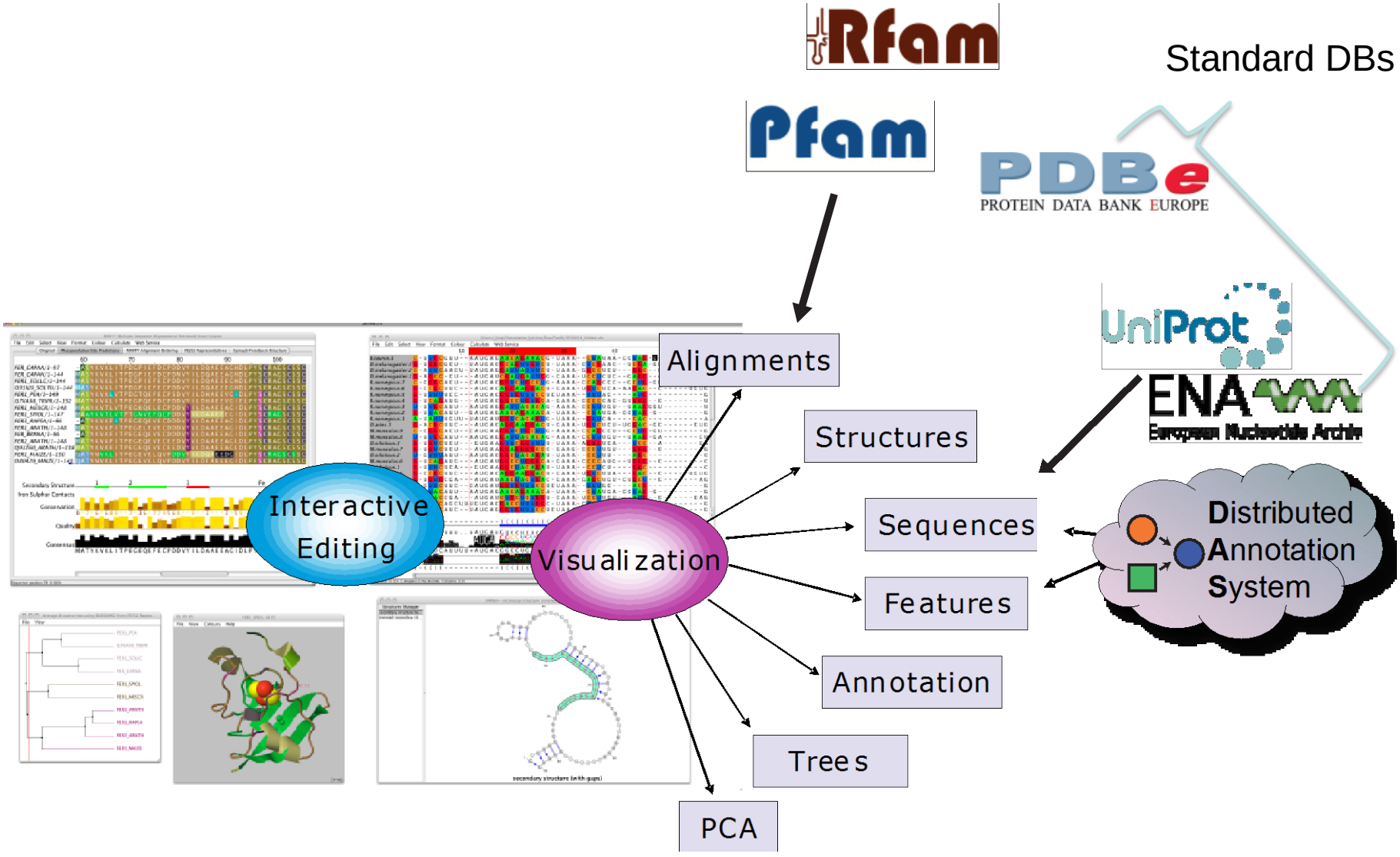
Files supporting sequence feature data

- Jalview sequence feature files
- GFF files
- Certain ‘rich’ alignment formats
  - Stockholm
  - AMSA

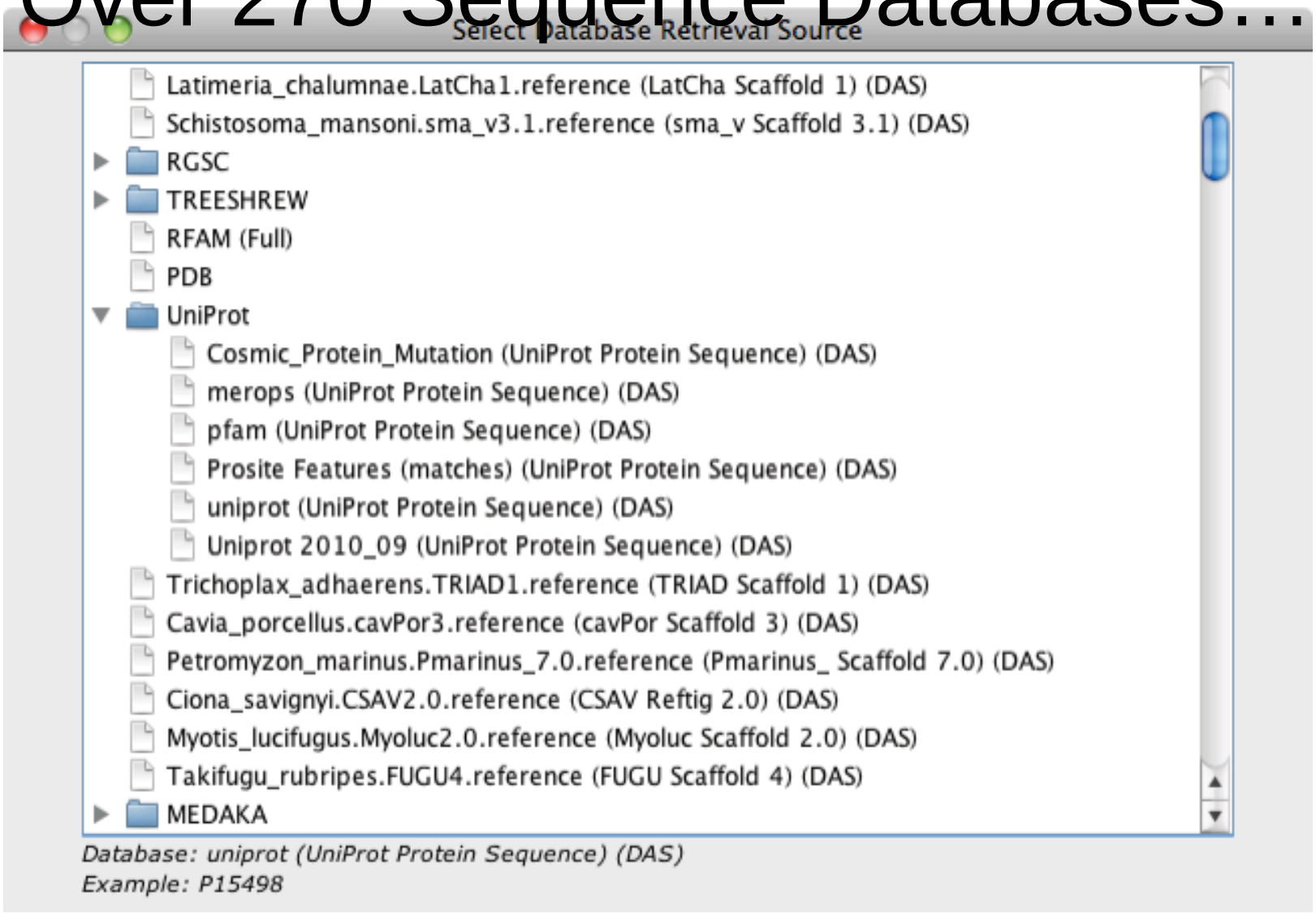
Sequence Features
Section 2.8.1-3 & Ex 28

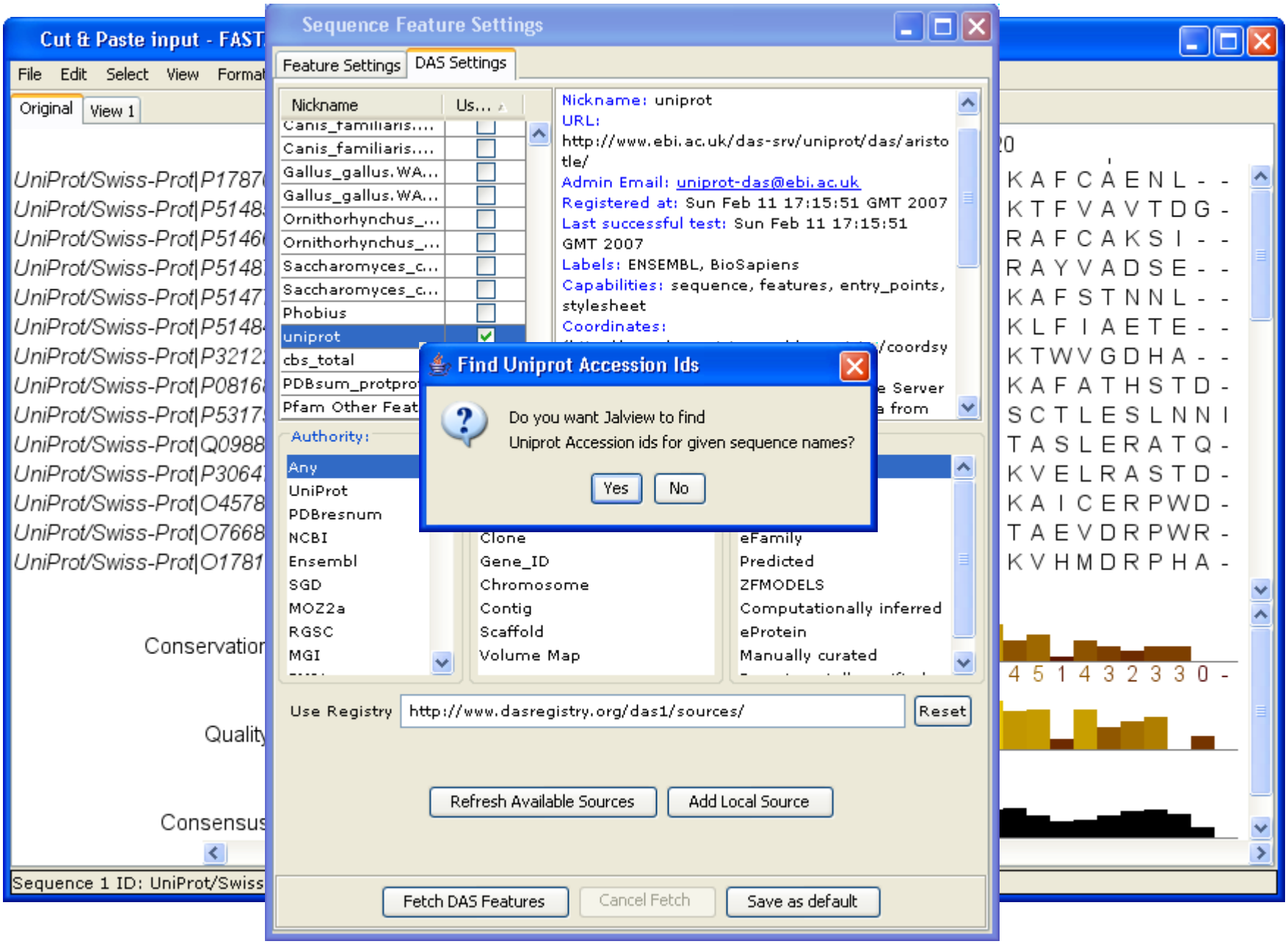
- Tasks
  - Visualise, create, modify, import and export features.
- Questions
  - What are the different types of file formats available for import and export ?
  - What services allow you to discover annotation and sequence features for sequence ?

Retrieval from External Databases

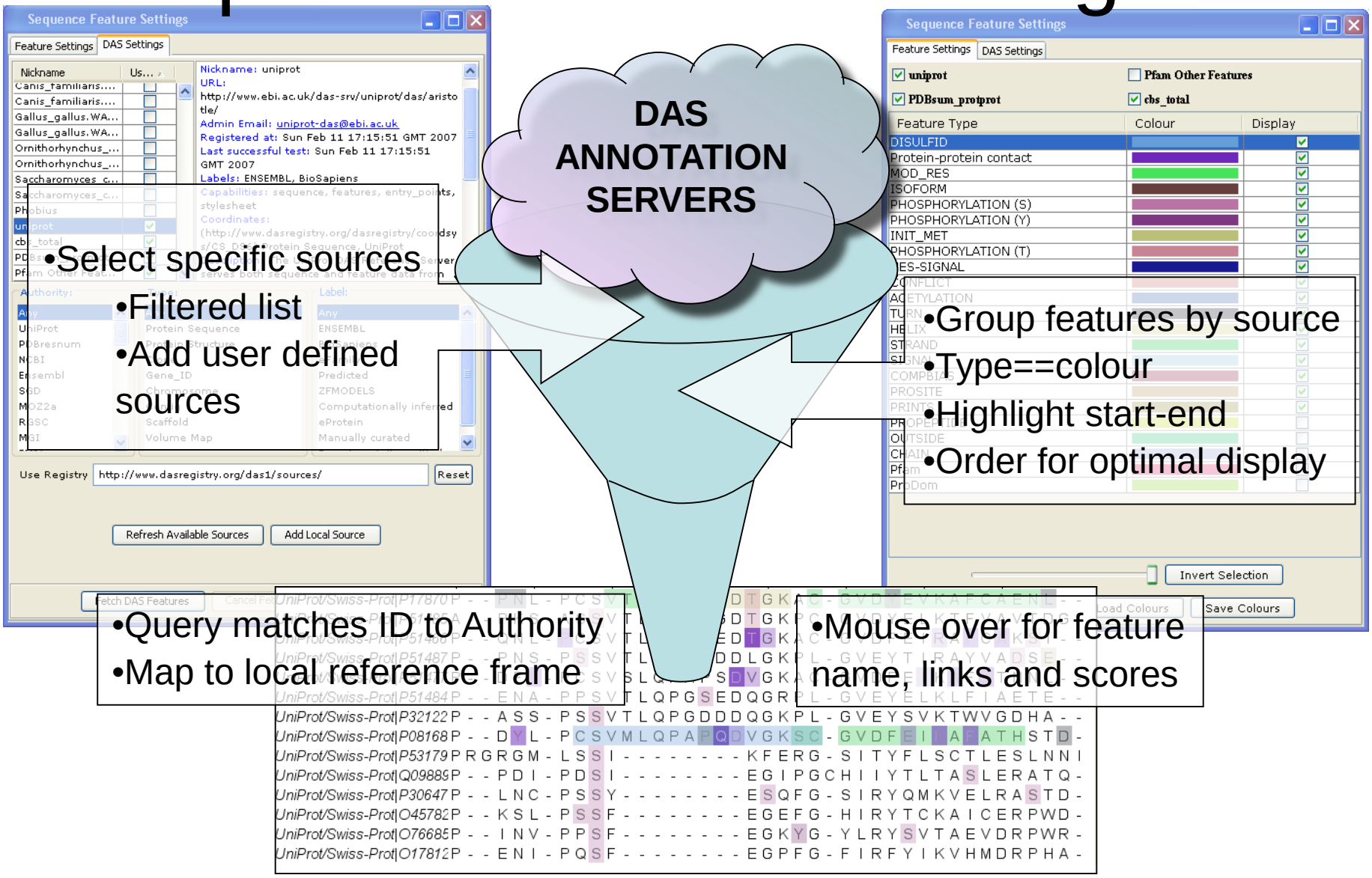


DAS allows Jalview access to Over 270 Sequence Databases…




Cut & Paste input - FAST
File Edit Select View Format
Original View 1
UniProt/Swiss-Prot|P1787
UniProt/Swiss-Prot|P5148
UniProt/Swiss-Prot|P5146
UniProt/Swiss-Prot|P5148
UniProt/Swiss-Prot|P5147
UniProt/Swiss-Prot|P5148
UniProt/Swiss-Prot|P3212
UniProt/Swiss-Prot|P0816
UniProt/Swiss-Prot|P5317
UniProt/Swiss-Prot|Q0988
UniProt/Swiss-Prot|P3064
UniProt/Swiss-Prot|O4578
UniProt/Swiss-Prot|O7668
UniProt/Swiss-Prot|O1781
Conservation
Quality
Consensus
Sequence 1 ID: UniProt/Swiss
Sequence Feature Settings
Feature Settings
DAS Settings
Nickname
Us...
Canis_familiaris....
Canis_familiaris....
Gallus_gallus.WA...
Gallus_gallus.WA...
Ornithorhynchus_...
Ornithorhynchus_...
Saccharomyces_c...
Saccharomyces_c...
Phobius
uniprot
cbs_total
PDBsum_protpro
Pfam Other Feat
Nickname: uniprot
URL:
http://www.ebi.ac.uk/das-srv/uniprot/das/aristotle/
Admin Email: uniprot-das@ebi.ac.uk
Registered at: Sun Feb 11 17:15:51 GMT 2007
Last successful test: Sun Feb 11 17:15:51 GMT 2007
Labels: ENSEMBL, BioSapiens
Capabilities: sequence, features, entry_points, stylesheet
Coordinates:
Find Uniprot Accession Ids
Do you want Jalview to find
Uniprot Accession ids for given sequence names?
Yes
No
Authority:
Any
UniProt
PDBresnum
NCBI
Ensembl
SGD
MOZ2a
RGSC
MGI
Clone
Gene_ID
Chromosome
Contig
Scaffold
Volume Map
eFamily
Predicted
ZFMODELS
Computationally inferred
eProtein
Manually curated
Use Registry
http://www.dasregistry.org/das1/sources/
Reset
Refresh Available Sources
Add Local Source
Fetch DAS Features
Cancel Fetch
Save as default
K A F C A E N L - -
K T F V A V T D G -
R A F C A K S I - -
R A Y V A D S E - -
K A F S T N N L - -
K L F I A E T E - -
K T W V G D H A - -
K A F A T H S T D -
S C T L E S L N N I
T A S L E R A T Q -
K V E L R A S T D -
K A I C E R P W D -
T A E V D R P W R -
K V H M D R P H A -
4 5 1 4 3 2 3 3 0 -

Sequence Features Dialog box



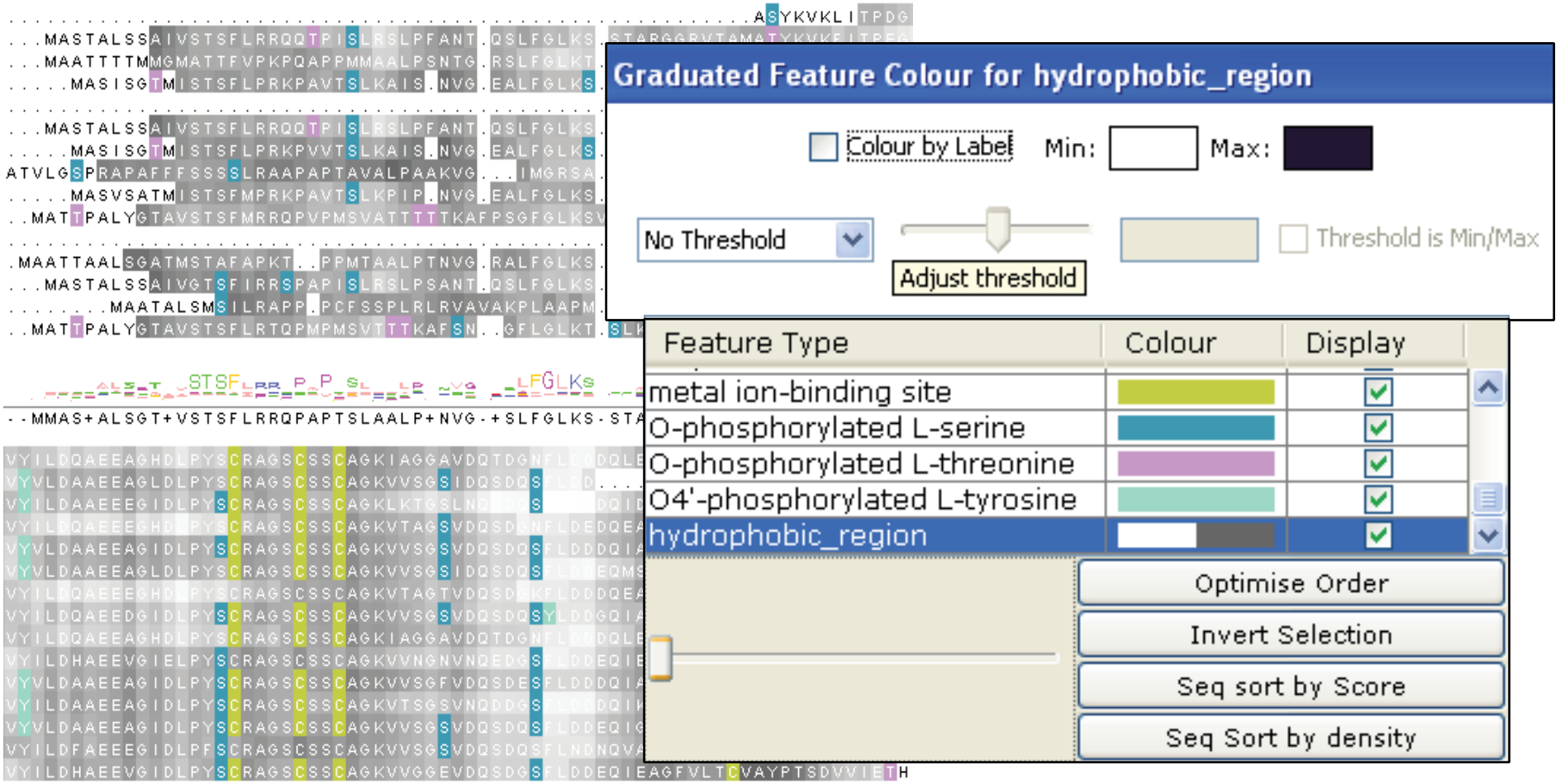
The Distributed Annotation System Section 2.9.2, Exercise 31

- Task
  - Browse available DAS sources for protein sequences
  - Retrieve annotation for the ferredoxin alignment.

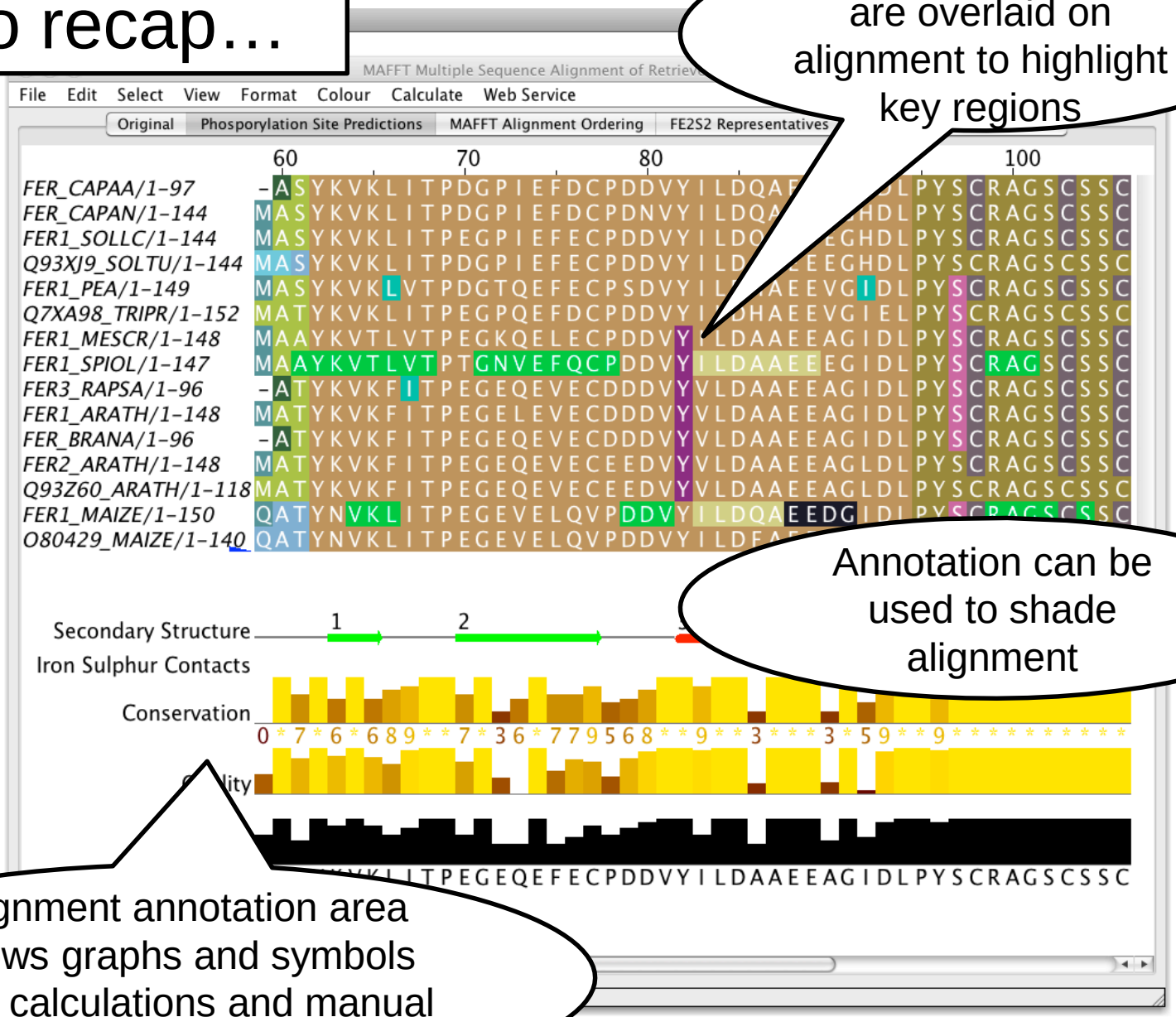
The Distributed Annotation System Section 2.9.2, Exercise 31

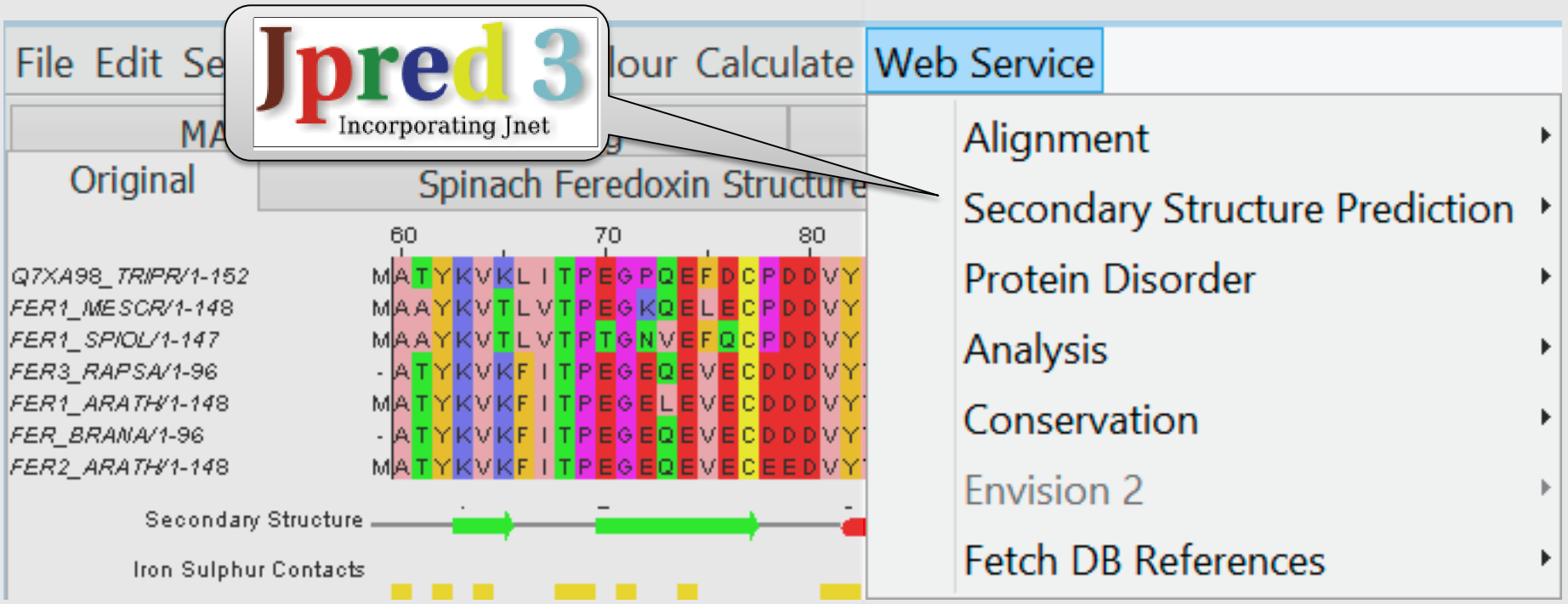
- Task
  - Browse available DAS sources for protein sequences
  - Retrieve annotation for the ferredoxin alignment.
- Question
  - What does the ‘optimise order’ button do ?

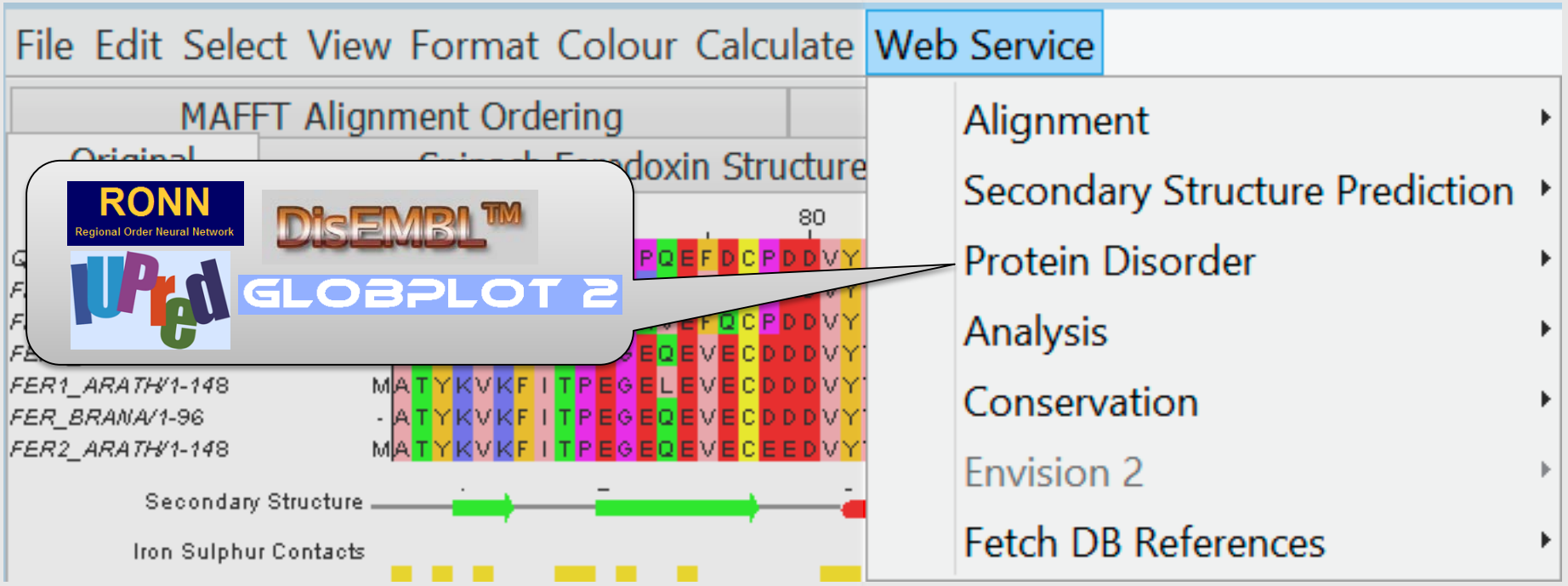
Shading, thresholding, colour by label.



Just to recap…




Jpred 3
Incorporating Jnet
File Edit Se
lour Calculate
Web Service
Alignment
Secondary Structure Prediction
Protein Disorder
Analysis
Conservation
Envision 2
Fetch DB References
Original
Spinach Feredoxin Structure
60
70
80
Q7XA98_TRIPR/1-152
FER1_MESCR/1-148
FER1_SPIOL/1-147
FER3_RAPSA/1-96
FER1_ARATH/1-148
FER_BRANA/1-96
FER2_ARATH/1-148
Secondary Structure
Iron Sulphur Contacts


File Edit Select View Format Colour Calculate
Web Service
MAFFT Alignment Ordering
RONN
Regional Order Neural Network
DisEMBL™
GLOBPLOT 2
80
FER1_ARATH/1-148
FER_BRANA/1-96
FER2_ARATH/1-148
Secondary Structure
Iron Sulphur Contacts
Alignment
Secondary Structure Prediction
Protein Disorder
Analysis
Conservation
Envision 2
Fetch DB References

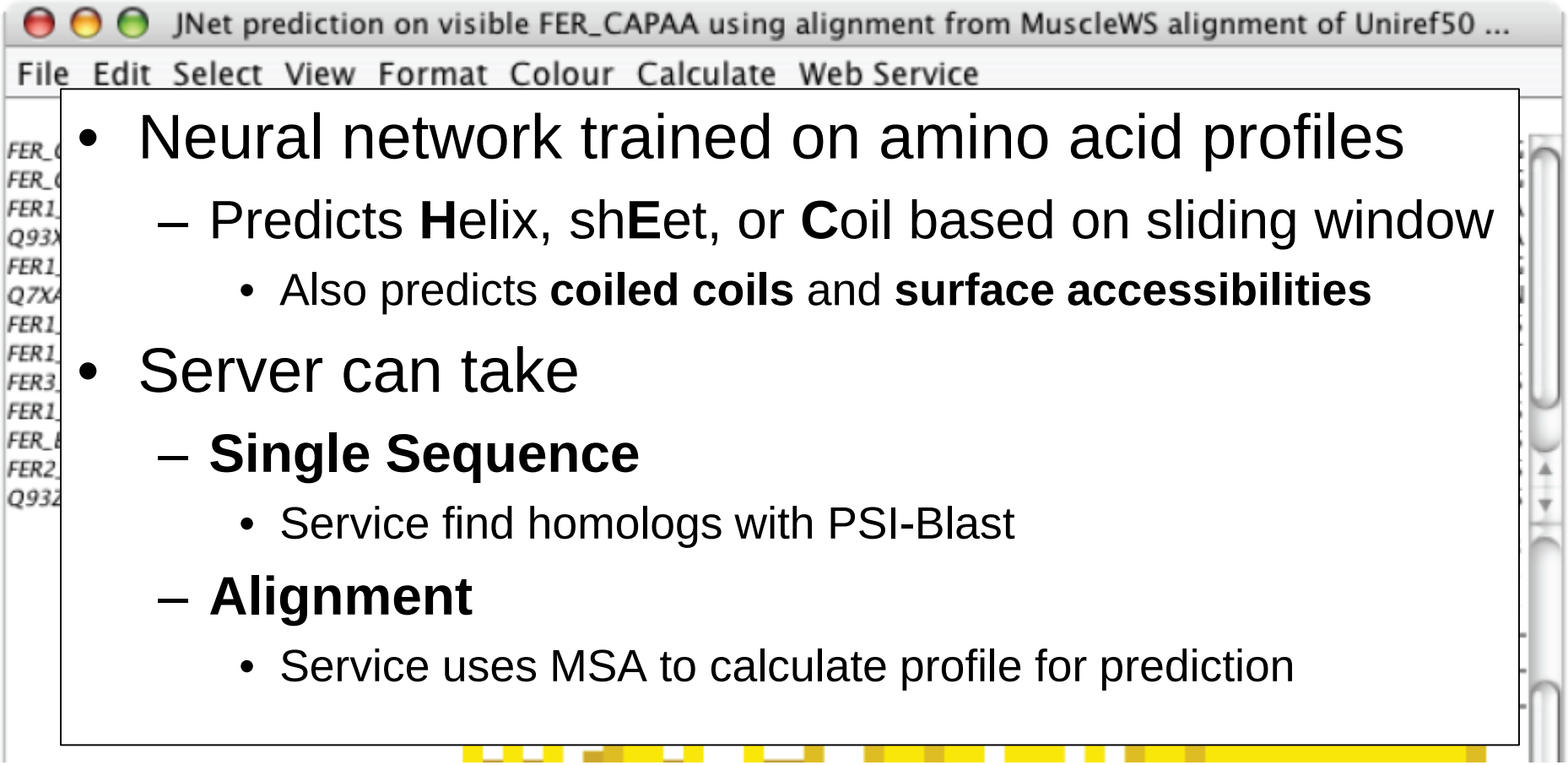
Protein secondary structure prediction

Section 2.6 onwards

in the Manual

Page 65

Protein Secondary Structure Prediction Sec. 2.6



- Neural network trained on amino acid profiles
  - Predicts **H**elix, sh**E**et, or **C**oil based on sliding window
    - Also predicts **coiled coils** and **surface accessibilities**
- Server can take
  - **Single Sequence**
    - Service find homologs with PSI-Blast
  - **Alignment**
    - Service uses MSA to calculate profile for prediction

Exercise 26

- Tasks
  - Perform a variety of Jnet predictions
    - Note the effect of hidden regions
    - Learn about sequence associated annotation
- Questions
  - …

Exercise 26

- Tasks
  - Perform a variety of Jnet predictions
    - Note the effect of hidden regions
    - Learn about sequence associated annotation

- Questions
  - What other data does Jnet provide ?
  - Which is better – a PSI blast prediction or an MSA based prediction ?
  - What happens when you have hidden regions ?