# Integrated design of antibodies for systems biology using AbDesigner

Trairak Pisitkun, MD

Chulalongkorn University Systems Biology (CUSB) Center
Epithelial Systems Biology Laboratory (ESBL), National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health, Maryland, USA
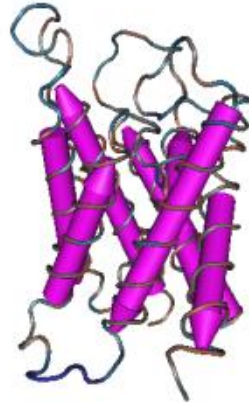InterPrET Research Center, Department of Biomedicine, Aarhus University, Denmark
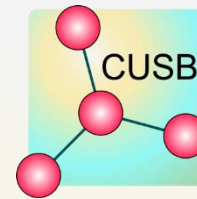
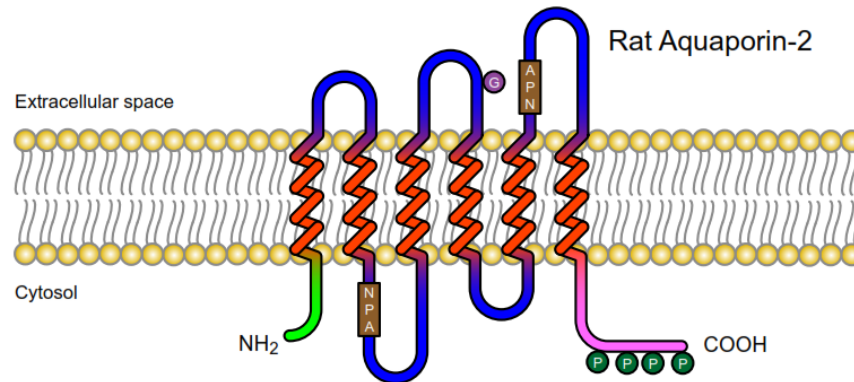# What problem does the *AbDesigner* solve?

- Lack of open-source, interactive, and artificial intelligence-based software for design of peptide-directed antibodies against whole proteome of organisms which facilitates studies at systems biology level

Custom antibody production

Protein of interest

AQP2
water channel

Antigen design

AbDesigner
http://helixweb.nih.gov/AbDesigner/

Peptide antigen synthesis and conjugation with carrier

 CWTSGQVEEYDLDADD

Immunization

Antibody purification and characterization

Custom antibody production

Protein of interest

AQP2
water channel

Antigen design

AbDesigner
http://helixweb.nih.gov/AbDesigner/

**Peptide antigen synthesis
and
conjugation with carrier**

 CWTSGQVEEYDLDADD

Immunization

**Antibody purification
and
characterization**
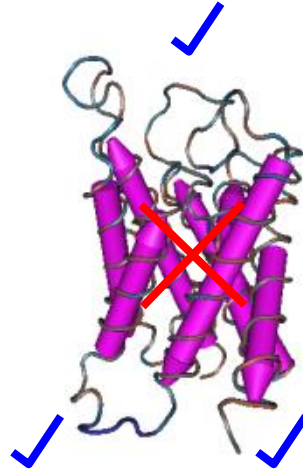
Antigen design

3D structure

Native, non-denatured epitopes
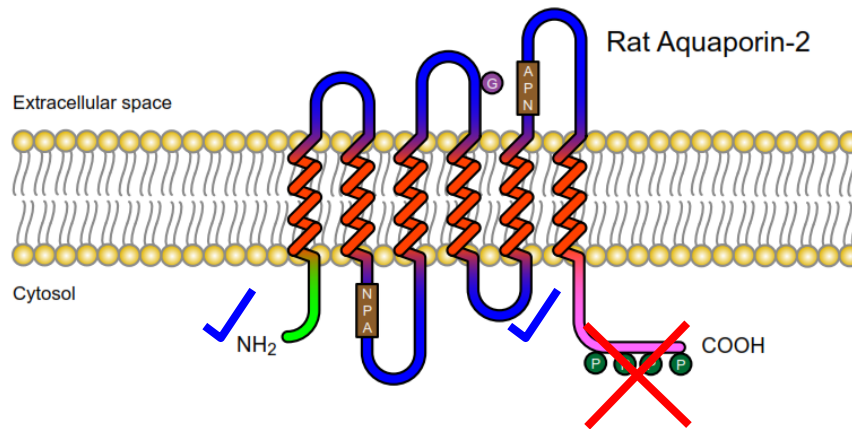Continuous or Discontinuous

Predicted structure from primary sequence

Antigen design

3D structure

Native, non-denatured epitopes
Continuous or Discontinuous

Predicted structure from primary sequence

Antigen design

- **AbDesigner**
  - http://helixweb.nih.gov/AbDesigner/

National **Heart Lung and Blood** Institute
AbDesigner

Input: AQP2_HUMAN   Swiss-Prot: P41181   Peptide Length: 15   Epitope Length: 7

Color   Export

```
        141      151      161      171      181      191      201      211      221      231      241      251      261      271
```

LTLQLVLCIFASTDERRGENPGTPALSIGFSVALGHLLGIHYTGCSMNPARSLAPAVVTGKFDDHWVFWIGPLVGAILGSLLYNYVLFPPAKSLSERLAVLKGLEPDTDWEEREVRRRQSVELHSPQSLPRGTKA

Chou-Fasman
K-D Hydropathy
Ig-score
Uniqueness
Conservation
Mature chain
Topological domain
MIP
Transmembrane re...
Variant
Conflict
Short sequence mo...
other
glycosylation
mutagenized
modified

Ig-score rank:

```
 1: PDTDWEEREVRRRQS [242-256]
 2: EPDTDWEEREVRRRQ [241-255]
 3: GLEPDTDWEEREVRR [239-253]
 4: LEPDTDWEEREVRRR [240-254]
 5: KGLEPDTDWEEREVR [238-252]
 6: DTDWEEREVRRRQSV [243-257]
 7: RRQSVELHSPQSLPR [253-267]
 7: RRRQSVELHSPQSLP [252-266]
 9: TDWEEREVRRRQSVE [244-258]
10: RQSVELHSPQSLPRG [254-268]
11: REVRRRQSVELHSPQ [249-263]
```

"Ig-score rank" shows peptides from most immunogenic to least.

Uniqueness-optimized rank:

```
 1: KGLEPDTDWEEREVR [238-252]
 2: RRQSVELHSPQSLPR [253-267]
 2: RRRQSVELHSPQSLP [252-266]
 4: RQSVELHSPQSLPRG [254-268]
 5: REVRRRQSVELHSPQ [249-263]
 6: EVRRRQSVELHSPQS [250-264]
 7: SVELHSPQSLPRGTK [256-270]
 8: EREVRRRQSVELHSP [248-262]
 9: QSVELHSPQSLPRGT [255-269]
10: LKGLEPDTDWEEREV [237-251]
10: VLKGLEPDTDWEERE [236-250]
```

"Uniqueness-optimized rank" shows Ig-score ranks for peptides predicted to be specific for target protein.

Conservation-optimized rank:

```
 1: PDTDWEEREVRRRQS [242-256]
 2: EPDTDWEEREVRRRQ [241-255]
 3: GLEPDTDWEEREVRR [239-253]
 4: LEPDTDWEEREVRRR [240-254]
 5: KGLEPDTDWEEREVR [238-252]
 6: DTDWEEREVRRRQSV [243-257]
 7: RRQSVELHSPQSLPR [253-267]
 7: RRRQSVELHSPQSLP [252-266]
 9: TDWEEREVRRRQSVE [244-258]
10: RQSVELHSPQSLPRG [254-268]
11: REVRRRQSVELHSPQ [249-263]
```

"Conservation-optimized rank" shows Ig-score ranks for peptides predicted to be recognized in target protein in multiple species.

A.

Rat Aquaporin-2

Extracellular space

Cytosol

NH₂

COOH

Secondary structure
Hydropathy
Ig-score
Uniqueness
Conservation

Protein features
Topological domain
Conflict
Short sequence motif
Glycosylation
Phosphorylation

COOH tail

B.

```
        231       241       251       261       271
Protein sequence  PSAKSLQERLAVLKGLEPDTDWEEREVRRRQSVELHSPQSLPRGSKA
```

Secondary structure
Hydropathy
Ig-score
Uniqueness
Conservation

Protein features
Topological domain
Conflict
Short sequence motif
Glycosylation
Phosphorylation

COOH tail

EPDTDWEEREVRRRQ 241-255
value: 10.61
rank: 2

EPDTDWEEREVRRRQ 241-255 AQP2

EPDTDWEEREVRRRQ 241-255 RAT
EPDTDWEEREVRRRQ 241-255 HUMAN
EPDTDWEEREVRRRQ 241-255 MOUSE

What is *AbDesigner* ?

AbDesigner is a new software tool that predicts optimal sites along the primary sequence of individual proteins for production of peptide-directed antibodies.

It calculates and extracts information needed to make a prediction and displays standard predictors in an easy-to-view graphical output.

National **Heart**
Lung and Blood Institute

Information calculated and extracted by *AbDesigner*

Immunogenicity score (Ig-score)

- Ig-score predicts immunogenicity of a peptide (ability of peptide to provoke an immune response).

National **Heart**
Lung and Blood Institute

A. Rat Aquaporin-2

B.

EPDTDWEEREVRRRQ 241-255
value: 10.61
rank: 2

EPDTDWEEREVRRRQ 241-255 AQP2

EPDTDWEEREVRRRQ 241-255 RAT
EPDTDWEEREVRRRQ 241-255 HUMAN
EPDTDWEEREVRRRQ 241-255 MOUSE

Information calculated and extracted by *AbDesigner*

Immunogenicity score (Ig-score)
- Ig-score predicts immunogenicity of a peptide (ability of peptide to provoke an immune response).

Uniqueness score
- Uniqueness score predicts specificity (to a targeted protein) of an antibody produced by a particular peptide.

National **Heart**
Lung and Blood Institute

A. Rat Aquaporin-2

Extracellular space

Cytosol

NH₂ COOH

Secondary structure
Hydropathy
Ig-score
Uniqueness
Conservation

Protein features
- Topological domain
- Conflict
- Short sequence motif
- Glycosylation
- Phosphorylation

COOH tail

B.

```
            231        241        251        261        271
Protein sequence  PSAKSLQERLAVLKGLEPDTDWEEREVRRRQSVELHSPQSLPRGSKA
```

Secondary structure
Hydropathy
Ig-score
Uniqueness
Conservation

Protein features
- Topological domain
- Conflict
- Short sequence motif
- Glycosylation
- Phosphorylation

```
EPDTDWEEREVRRRQ 241-255
value: 10.61
rank: 2
```

```
EPDTDWEEREVRRRQ 241-255 AQP2
```

```
EPDTDWEEREVRRRQ 241-255 RAT
EPDTDWEEREVRRRQ 241-255 HUMAN
EPDTDWEEREVRRRQ 241-255 MOUSE
```
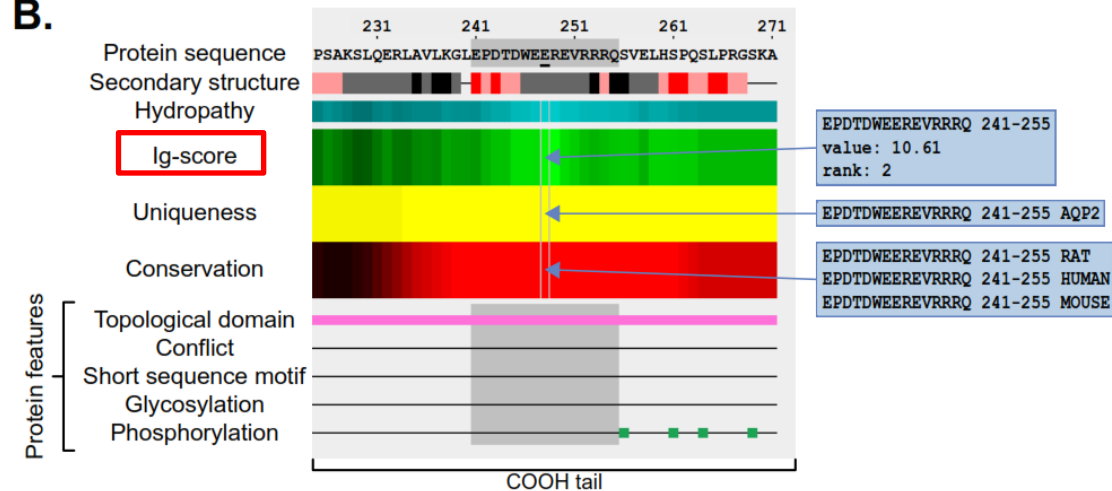
COOH tail

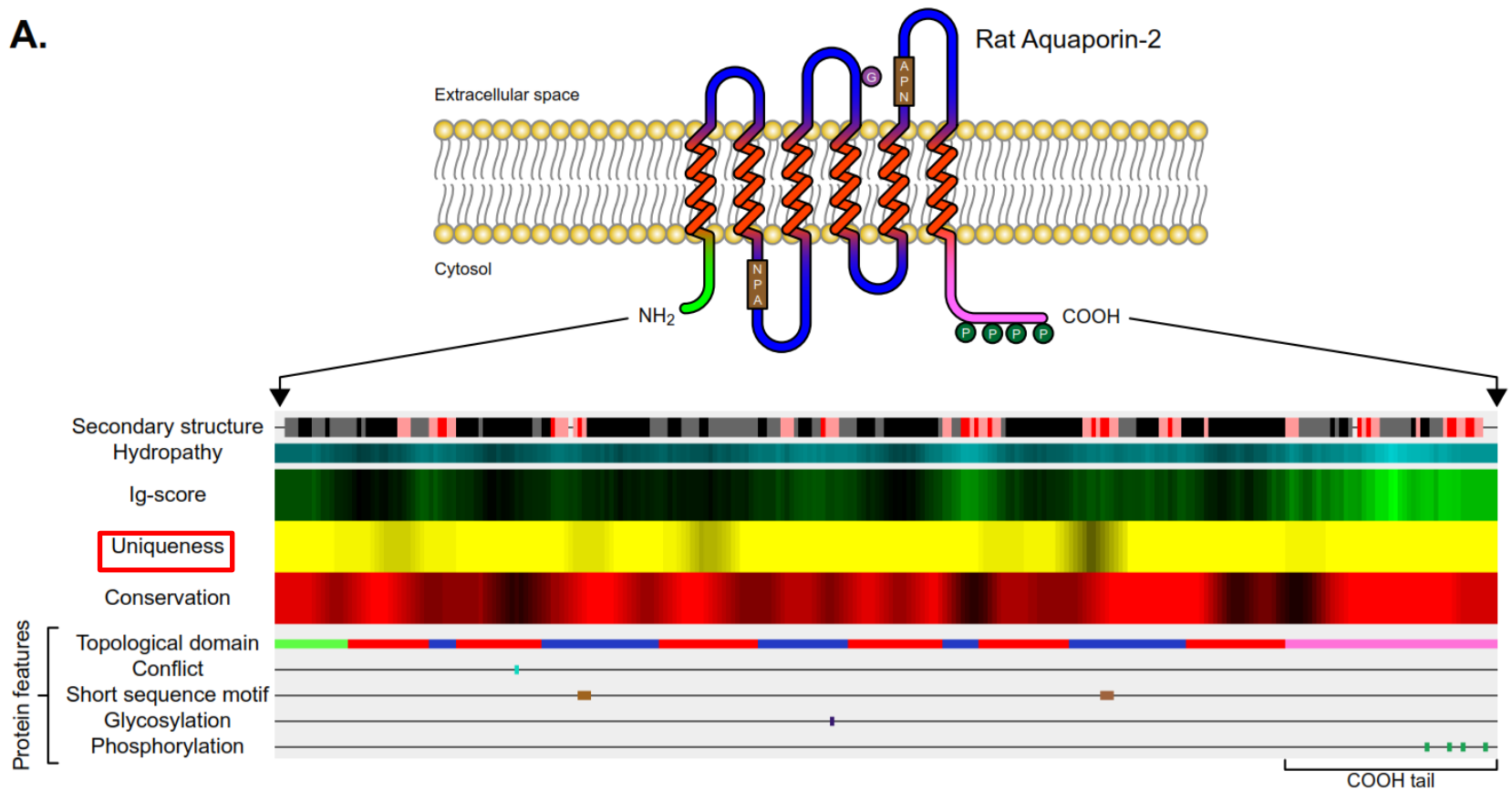# Information calculated and extracted by *AbDesigner*

Immunogenicity score (Ig-score)
- Ig-score predicts immunogenicity of a peptide (ability of peptide to provoke an immune response).
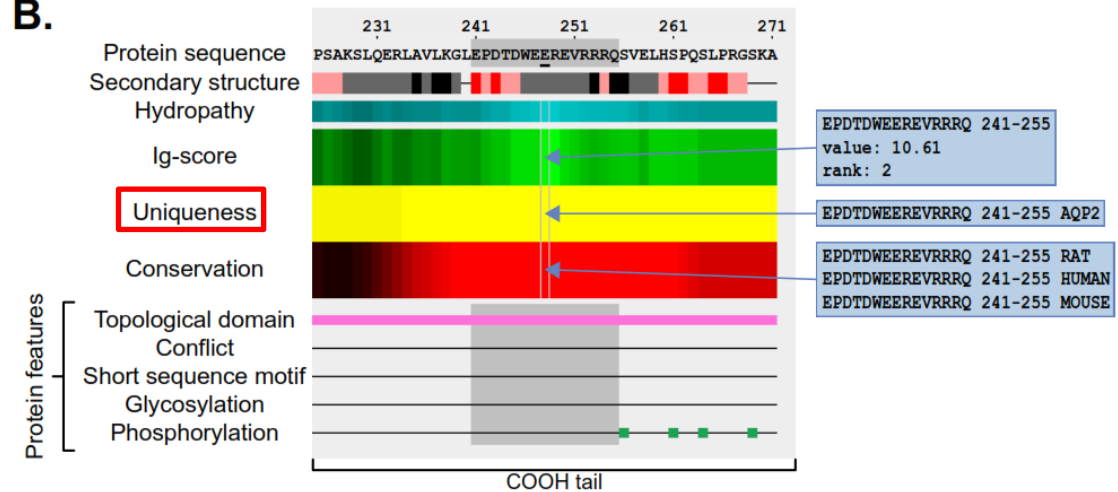
Uniqueness score
- Uniqueness score predicts specificity (to a targeted protein) of an antibody produced by a particular peptide.

Conservation score
- Conservation score predicts likelihood of multi-species recognition of an antibody produced by a particular peptide.

National **Heart**
Lung and Blood Institute

A.

Rat Aquaporin-2

Extracellular space

Cytosol

NH₂

COOH

Secondary structure
Hydropathy
Ig-score
Uniqueness
Conservation

Protein features
- Topological domain
- Conflict
- Short sequence motif
- Glycosylation
- Phosphorylation

COOH tail

B.

231   241   251   261   271

Protein sequence
PSAKSLQERLAVLKGLEPDTDWEEREVRRRQSVELHSPQSLPRGSKA

Secondary structure
Hydropathy
Ig-score
Uniqueness
Conservation

Protein features
- Topological domain
- Conflict
- Short sequence motif
- Glycosylation
- Phosphorylation

EPDTDWEEREVRRRQ 241-255
value: 10.61
rank: 2

EPDTDWEEREVRRRQ 241-255 AQP2

EPDTDWEEREVRRRQ 241-255 RAT
EPDTDWEEREVRRRQ 241-255 HUMAN
EPDTDWEEREVRRRQ 241-255 MOUSE

COOH tail

Information calculated and extracted by *AbDesigner*
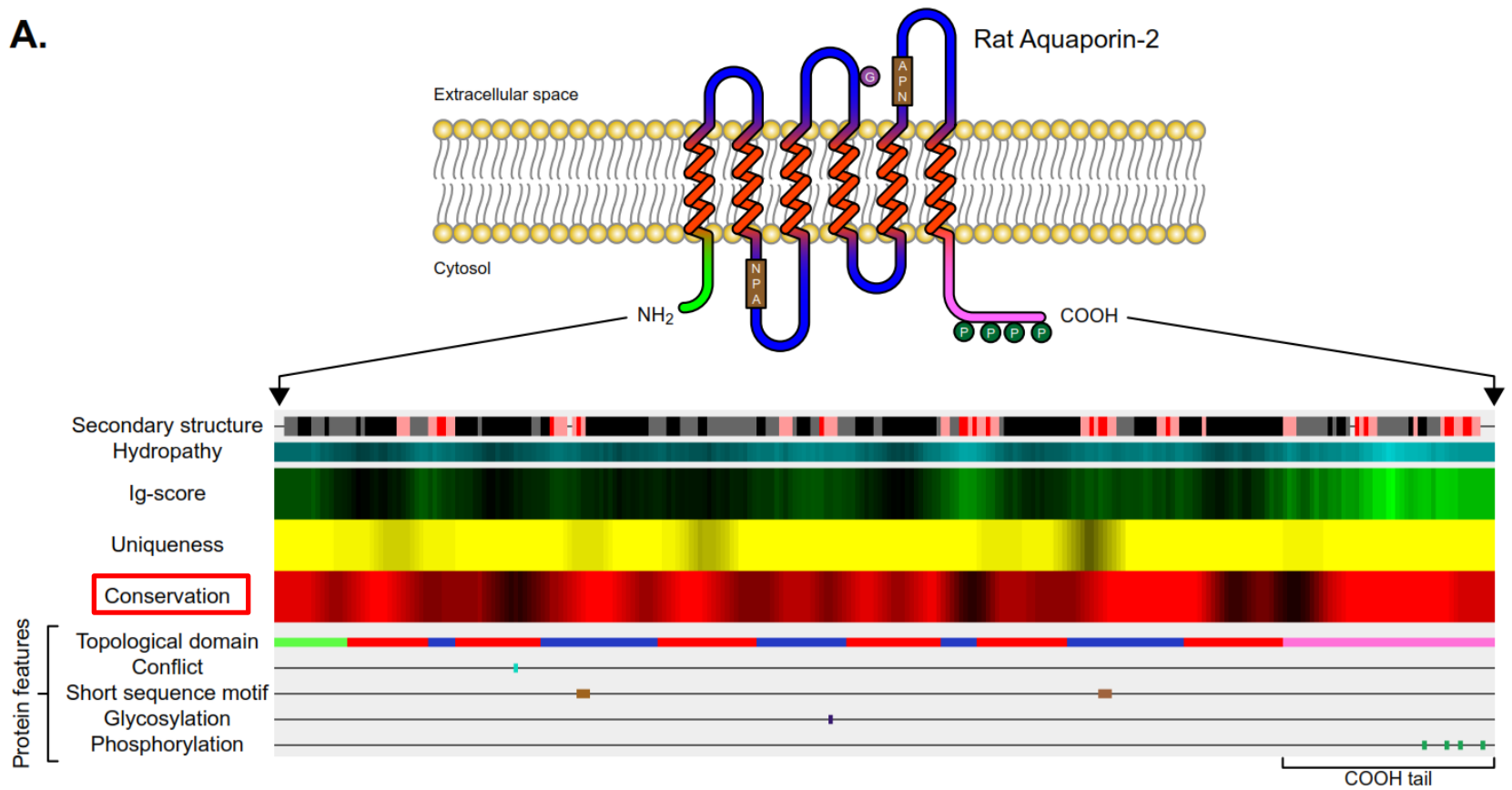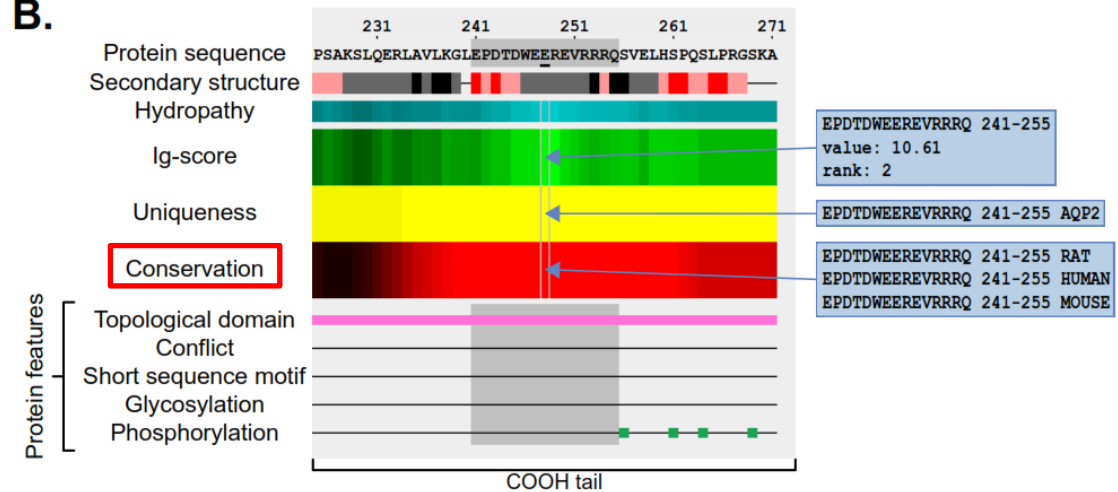
Immunogenicity score (Ig-score)
- Ig-score predicts immunogenicity of a peptide (ability of peptide to provoke an immune response).

Uniqueness score
- Uniqueness score predicts specificity (to a targeted protein) of an antibody produced by a particular peptide.
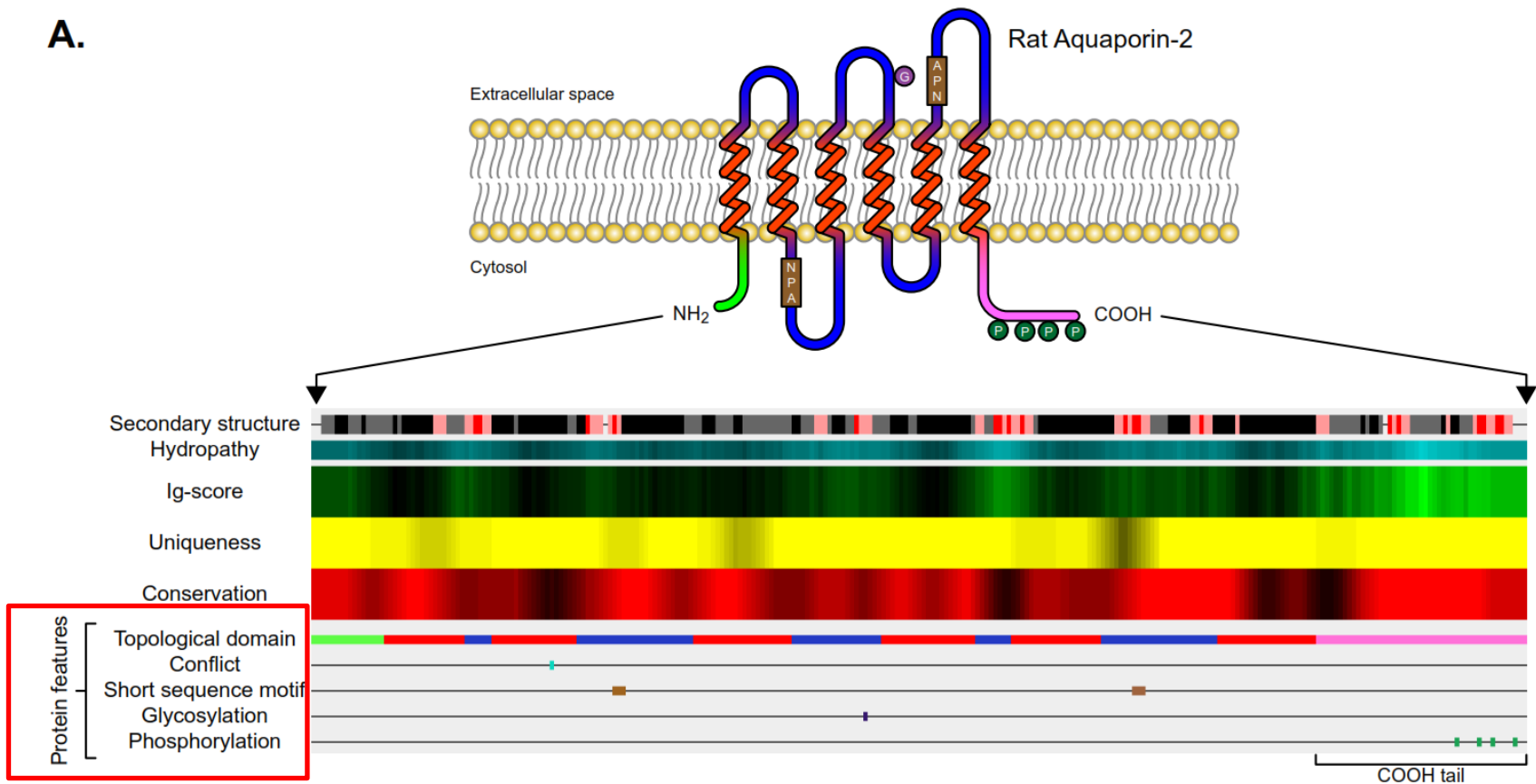
Conservation score
- Conservation score predicts likelihood of multi-species recognition of an antibody produced by a particular peptide.
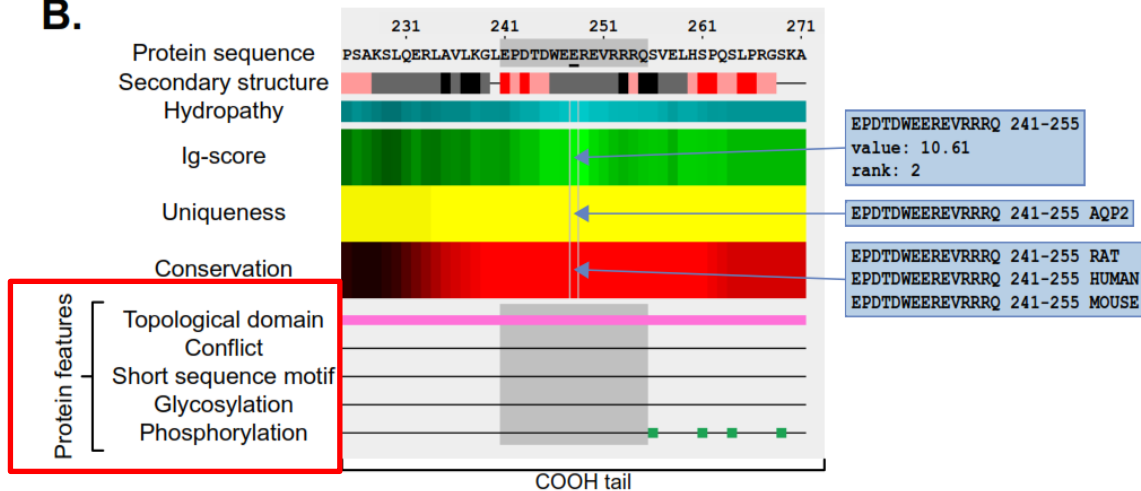
Protein features
- Describes regions or sites of interest in the protein sequence such as:
  - post-translational modifications
  - binding sites
  - enzyme active sites
  - local secondary structure
  - sequence conflicts
  - other characteristics reported in literatures

National **Heart**
Lung and Blood Institute

A.

Rat Aquaporin-2

Extracellular space

Cytosol

NH₂

COOH

Secondary structure
Hydropathy
Ig-score
Uniqueness
Conservation

Protein features
  Topological domain
  Conflict
  Short sequence motif
  Glycosylation
  Phosphorylation

COOH tail

B.

Protein sequence
PSAKSLQERLAVLKGLEPDTDWEEREVRRRQSVELHSPQSLPRGSKA

Secondary structure
Hydropathy
Ig-score
Uniqueness
Conservation

Protein features
  Topological domain
  Conflict
  Short sequence motif
  Glycosylation
  Phosphorylation

EPDTDWEEREVRRRQ 241-255
value: 10.61
rank: 2

EPDTDWEEREVRRRQ 241-255 AQP2

EPDTDWEEREVRRRQ 241-255 RAT
EPDTDWEEREVRRRQ 241-255 HUMAN
EPDTDWEEREVRRRQ 241-255 MOUSE

COOH tail

Methods implemented by *AbDesigner*

Immunogenicity score (Ig-score)

- Kyte-Doolittle hydropathy

  [Kyte J, Doolittle RF. J Mol Biol. 1982 May 5;157(1):105-32]

- Modified Chou-Fasman secondary structure

  [Chou PY, Fasman GD. Biochemistry. 1974 Jan 15;13(2):222-45]

Uniqueness score

- Pattern matching with other proteins in the same species
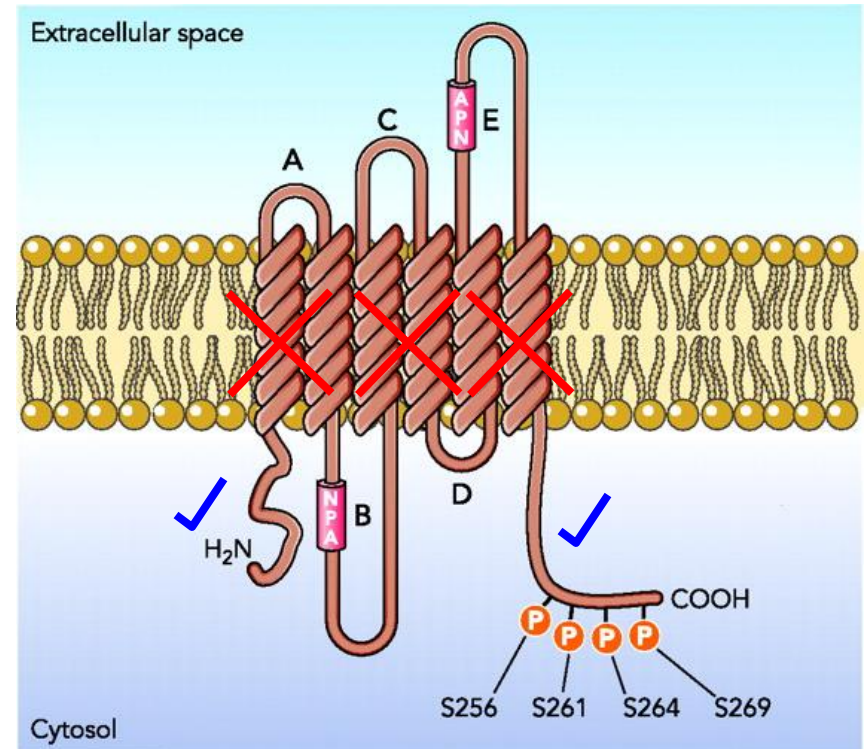
Conservation: conservation score

- Pattern matching with orthologous proteins in the other species

Protein features

- Data extraction from Swiss-Prot database

National **Heart**
Lung and Blood Institute

Immunogenicity Score

Hydropathy * P-Turn * tail bonus

Possible number of epitopes per peptide

1 P D T D W E E R E V R R R Q S 15

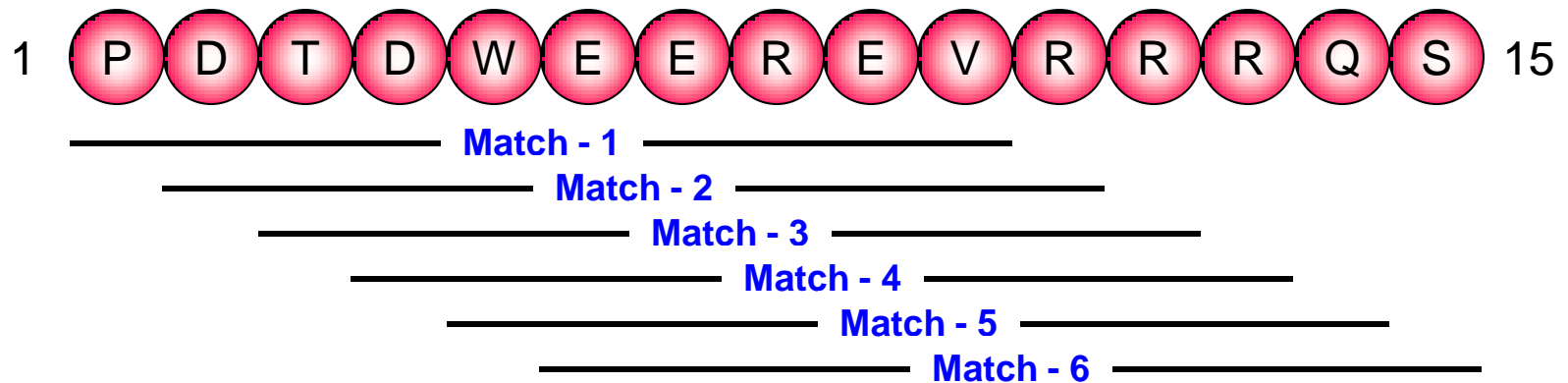Choice 1 P D T D W E E R E V

Choice 2 D T D W E E R E V R

Choice 3 T D W E E R E V R R

Choice 4 D W E E R E V R R R

Choice 5 W E E R E V R R R Q

Choice 6 E E R E V R R R Q S

National **Heart**
Lung and Blood Institute

Tandem pattern matching

1 P D T D W E E R E V R R R Q S 15

Match - 1

Match - 2

Match - 3

Match - 4

Match - 5

Match - 6

Identical Matches - 1

Identical Matches - 2

Identical Matches - 3

Identical Matches - 4

Identical Matches - 5

Identical Matches - 6

" Uniqueness Score"
Pattern matching with other proteins in the same species

"Conservation score"
Pattern matching with the same protein in other species

National **Heart**
Lung and Blood Institute

Comparison of *AbDesigner* to the other existing solutions

	AbDesigner	DNASTAR Protean	Gene Inspector	GenScript	Antigen Profiler	Abie Pro 3.0
Open-source	yes	no	no	no	no	no
Interactive GUI	yes	yes	not tested	no	no	yes
Web application	yes	no	no	no	no	yes
- hydropathy	K-D	K-D, H-W	K-D, H-W	yes	K-D, H-W	K-D, H-W
- secondary structure	C-F	C-F, G-R	C-F, G-R	yes	no	no
- surface probability	no	E	yes	J, E	no	no
- integrated score	Ig-score	J-W	no	J-W	yes	no
Specificity assessment	Uniqueness score	no	no	no	yes, manually?	yes
Multi-species recognition assessment	Conservation score	no	no	no	yes, manually?	no
Protein features	NCBI	NCBI	Prosite Motif Search	no	yes	no

Abbreviations: K-D = Kyte-Doolittle, H-W = Hopp-Woods, C-F = Chou-Fasman, G-R = Garnier-Robson, J = Janin, E = Emini, J-W = Jameson-Wolf

National **Heart** **Lung and Blood** Institute
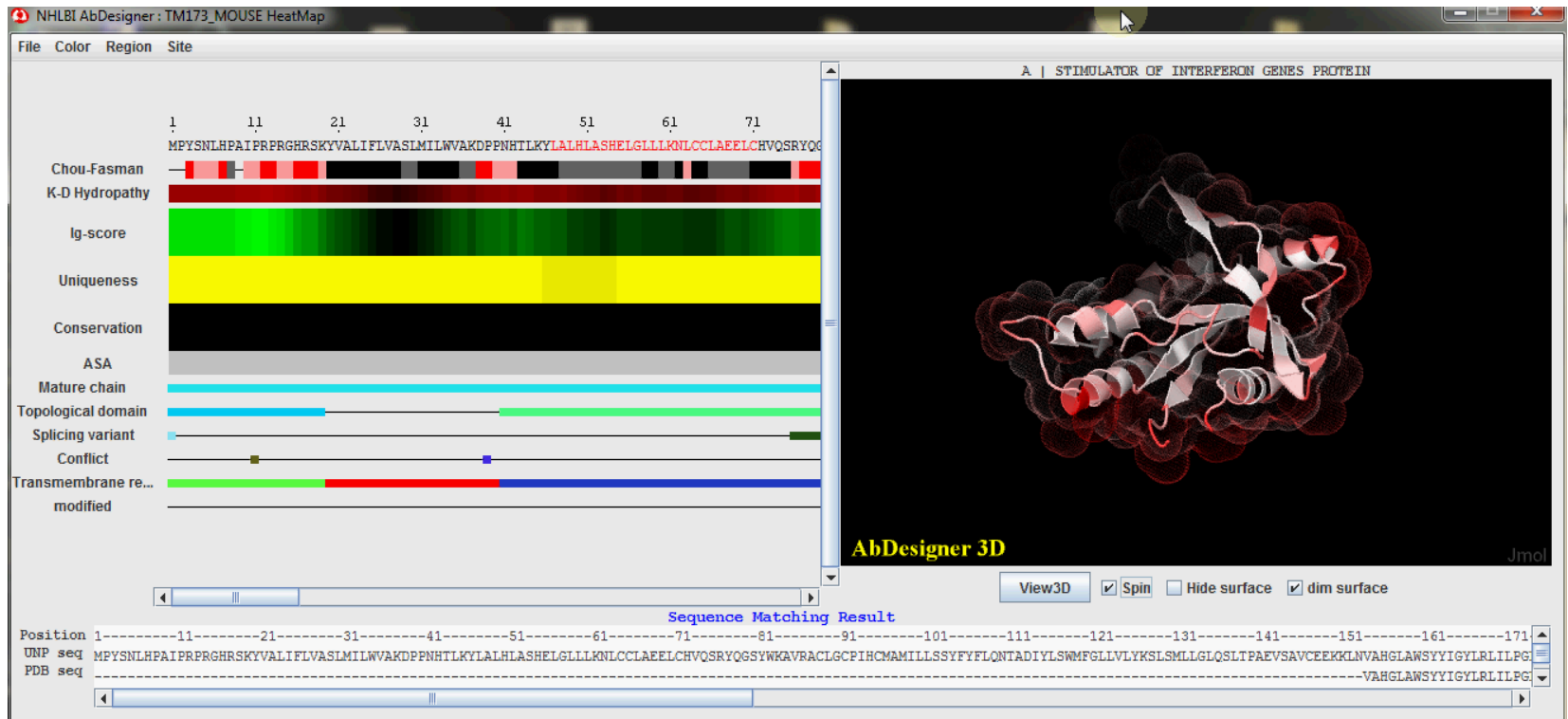
AbDesigner-3D

AbDesigner-3D
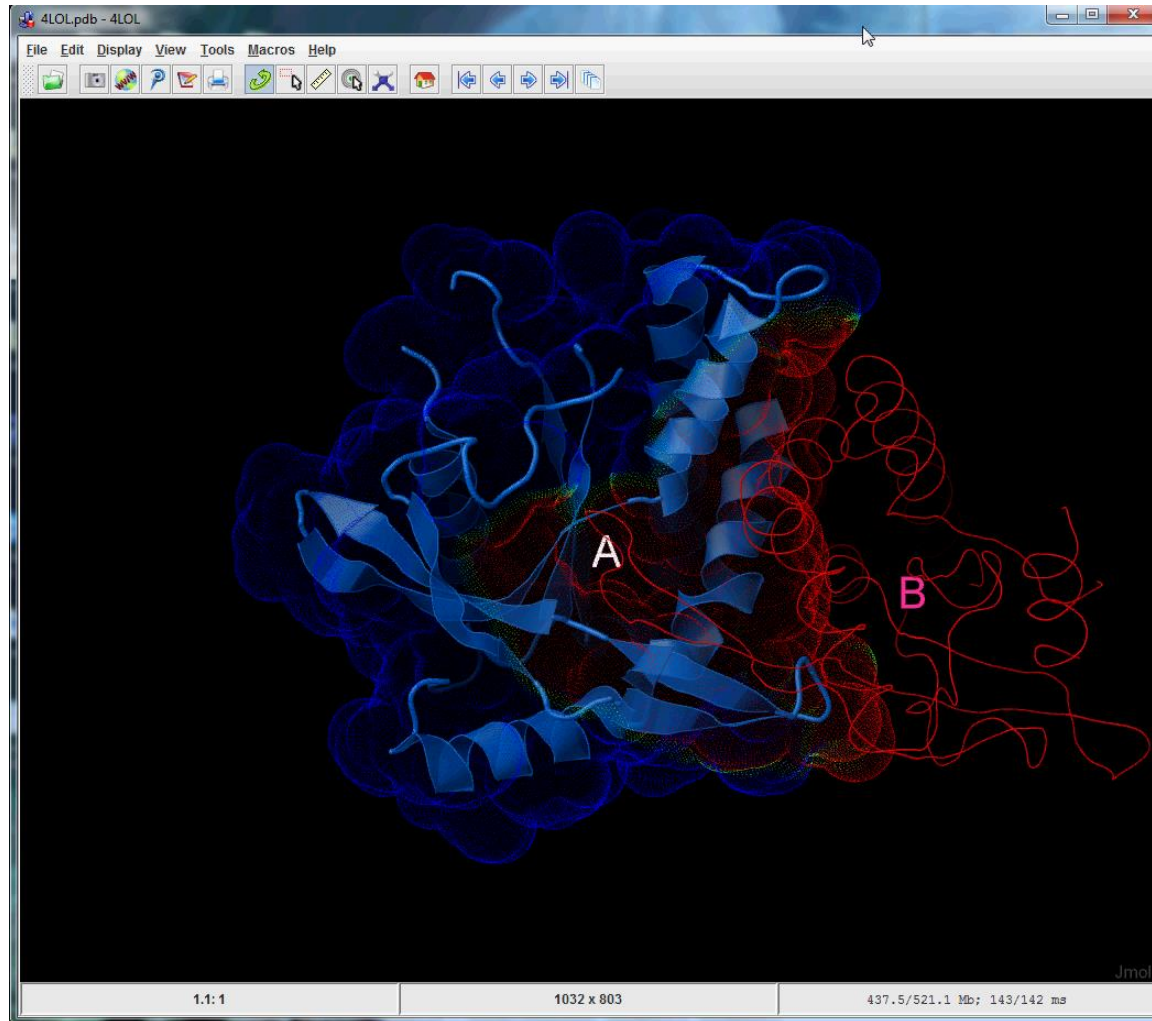
Speculate on additional research uses, commercial potential and/or public health benefits.

AbDesigner can be used for large-scale design of peptide-directed antibodies against whole proteome of organisms which facilitates studies at systems biology level

National **Heart Lung and Blood** Institute

Thank you