Protein Ontology (PRO) workshop 2016

PRO workshop: Proteoforms in disease

- When and Where
- Mission and workshop objectives
- Agenda
  - Day 1 (Conference Room 1300)
  - Day 2 (Conference Room 4300)
- Participants
- Area information
  - Hotels
  - Airports
  - Parking and Local transportation

When and Where

**Dates:** November 14-15, 2016

**Venue:** 3300 Whitehaven NW., Harris Building, Georgetown University, Washington DC 20007. See in map

This is a close meeting, attendance is by invitation.

Mission and workshop objectives

PRO mission is to enable protein-disease understanding by presenting knowledge about proteins in biological context richly and accurately for both human understanding and computational reasoning.

**Workshop objectives**

1. Creation and implementation of strategies to enable researchers to curate PRO branches in collaboration with PRO experts
2. Identification of relationships and biological contexts (e.g., kinase-substrate, proteoform-disease) to describe relations of proteins to disease
3. Curation of terms needed to address use cases
4. Construction of semantic queries to address use cases

In this workshop we have identified two areas in need of community collaboration/input

1. **PTM/variants and disease.** One of the aims in PRO is to provide comprehensive coverage of proteoforms in their biological context. More specifically, create a comprehensive representation of known proteoforms that integrates proteomic evidence of clinical variants and PTMs with rich relationships and biological contexts (e.g., kinase-substrate, proteoform-disease). To work on this topic ProKinO, MGI, NeXtProt and Reactome groups were recruited.

2. **Epitopes, antibodies, proteoforms and disease.** The immunology community is keenly interested in ontological terms for key immune-related proteins such as cell surface markers, cytokines, transcription factors, and the highly variable antigen receptors. In some cases, these requests highlight the need for development of the PRO framework, e.g., to allow the proper mapping between the antibody and the set of recognized proteins. To work on this topic, the Immport and IEDB groups were recruited.

Agenda

**Link to the presentations:** https://drive.google.com/open?id=0B4QyLB3JI6qRR05pQUgxdEthQzQ

**Day 1 (Conference Room 1300)**

8:30-9:00 Breakfast

9:00-9:15 PRO mission and workshop objectives (Cathy Wu)

9:15-9:35 PRO Ontological Framework (Darren Natale)

9:35-10:45 Accessing PRO scientific content: this session will introduce some resources to access PRO data

- 9:35-10:00 PRO website [includes overview of the website and demonstration of scientific content] (Cecilia Arighi)
- 10:00-10:25 SPARQL queries [includes demonstration of federated query using UniProt as example] (Chuming Chen)
- 10:25-10:45 iPTMnet [includes how PRO is consumed by iPTMnet and enriches its content] (Karen Ross)

10:45-11:00 COFFEE BREAK
11:00-12:30 Project’s introductions: 15 min presentations from each resource to provide a general information, followed by workshop-related questions about the intended outcome of the workshop and how PRO could assist to achieve such outcome.

- 11:00-11:15 ProKinO (Kannan Natarajan)
- 11:15-11:30 NeXIProt (Pascale Gaudet)
- 11:30-11:45 Reactome (Peter D'Eustachio)
- 11:45-12:00 MGI (Cynthia Smith)
- 12:00-12:15 Immport (Alex Diehl)
- 12:15-12:30 IEDB (Randi Vita)

12:30-1:30 LUNCH

1:30-2:00 Intro to use cases. This section will layout the most relevant issues and use cases to be discussed in different working sessions. By the end of the day, each group should give a presentation of the major outcomes of the discussion. (Cecilia Arighi)

2:00-3:00 Session 1-General session. This section intends to get (hopefully) a consensus opinion on common topics of interest, so we can provide our recommendation to the ontology community. Moderators should have a proposal ready to be discussed.

- 2:00-2:20 Connecting proteoforms/complexes and disease. How to relate protein and complexes to disease. Moderators: Judith Blake, Peter D'Eustacchio, Cynthia Smith
- 2:40-3:00 Representation of variants and isoforms. How to relate variants and isoforms. Moderators: Karen Ross, Harold Drabkin

3:00-3:15 COFFEE BREAK

3:15-4:45 Parallel group sessions. For the proposed questions/use cases the discussion should include: who has the data and how to connect to such data, and a proposal of federated queries/scientific questions from each group.

Group 1- PTM/variants and disease.

Moderator: Judith Blake

Participants: Karen Ross, Cathy Wu, Cynthia Smith, Harold Drabkin, Pascale Gaudet, Kannan Natarajan, Liang-Chin Huang and Peter D'Eustacchio

- PTM-variants and disease
- What are variants affecting phosphorylation site (gain or loss)?
- What proteoforms and complexes would be affected?
- What pathways would be affected?
- Connected to what disease?
- What are variants affecting kinase function?
- What targets would be affected?
- What pathways would be affected?
- Connected to what disease?
- Isoforms-variants and connection to disease in mouse and human
- Connecting mouse and human isoforms and variants to allelic information in MGI, disease models and human disease
- Connecting functional annotation of variants

Group 2- Epitopes/antibodies and disease

Moderator: Cecilia Arighi

Participants: Darren Natale, Alex Diehl and Randi Vita

- MHC related proteins
- Epitopes/proteins and antibodies in autoimmune disease
- Cytokines

Group 3 – Semantic integration

Moderator: Chuming Chen

Participants: Alan Ruttenberg, Krys Kochut

4:45-5:30 Wrap up session day 1

- 4:45-4:55 Presentation group 1
- 4:55-5:05 Presentation group 2
- 5:05-5:15 Presentation group 3
- 5:15-5:30 General Discussion that should lead to action items and regrouping for day 2

6:30 Working Dinner at Cafe Divan, 1834 Wisconsin Ave, NW. Washington, DC 20007

Day 2 (Conference Room 4300)
8:30-9:00 Breakfast

9:00-10:30 Group session (cont). For the proposed questions/use cases the discussion should now model different cases. If possible, apply to real case.

Group 1- PTM/variants and disease. (cont)

Moderator: Cathy Wu

Participants: TBD after day 1

- PTM-variants and disease
- Isoforms-variants and connection to disease in mouse and human

Group 2- epitopes/antibodies and disease (cont.)

Moderator: Cecilia Arighi

Participants: TBD after day 1

- MHC related proteins
- Epitopes/proteins and antibodies in autoimmune disease
- Cytokines

10:30-10:45 COFFEE BREAK

10:45-12:30 Breakout sessions to streamline data integration

1. ProkinO ontology development as an extension of PRO (Alan Ruttenberg, Darren Natale and Kannan’s group)
2. NeXtProt and IEDB modified peptides and epitopes, respectively, and link to proteins in iPTMnet and PRO (Cathy Wu, Randi Vita and Pascale Gaudet)
3. MGI-PRO establish workflow for variants in MGI to PRO IDs and annotation. Modeling with Noctua (Cecilia Arighi and MGI group)

12:30-1:30 LUNCH

1:30-2:30 Presentation of breakout session

- 1:30-1:45 ProkinO ontology development as an extension of PRO
- 1:45-2:00 NeXtProt and IEDB modified peptides and epitopes, respectively, and link to proteins in iPTMnet and PRO
- 2:00-2:15 MGI-PRO establish workflow for variants in MGI to PRO IDs and annotation. Modeling with Noctua?
- 2:15-2:30 General discussion

2:30-2:45 COFFEE BREAK

2:45-4:00 Wrap up use case sessions and next steps

- Review of meeting objectives and their outcome
- Planning for manuscripts
  - Manuscript 1-proteoforms/complexes->noctua modeling
  - Manuscript 2-epitopes/antibodies/proteoforms
- Next steps

End of workshop

Participants

On Site:

<table>
<thead>
<tr>
<th>Entity</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>EIDB</td>
<td>Randi Vita</td>
</tr>
<tr>
<td>ProKinO</td>
<td>Kannan Natarajan</td>
</tr>
<tr>
<td>ProKinO</td>
<td>Liang-Chin Huang</td>
</tr>
<tr>
<td>ProKinO</td>
<td>Krys Kochut</td>
</tr>
<tr>
<td>ProKinO</td>
<td>Daniel McSkimming</td>
</tr>
<tr>
<td>JAX</td>
<td>Harold Drabkin</td>
</tr>
<tr>
<td>JAX</td>
<td>Judith Blake</td>
</tr>
<tr>
<td>JAX</td>
<td>Cynthia Smith</td>
</tr>
<tr>
<td>Reactome</td>
<td>Peter D'Eustacchio</td>
</tr>
</tbody>
</table>
Area information

Hotels
Here are some hotels close to the venue

**The Georgetown Inn**
1310 Wisconsin Avenue, NW (near N Street, NW)
Washington, DC 20007
1-866-308-8883
http://www.georgetowninn.com/

**Georgetown Hotel and Conference Center**
3800 Reservoir Road, NW
Washington, DC 20057
1-888-902-1606
http://www.acc-guhotelandconferencecenter.com/

**Kimpton Glover Park Hotel**
2505 Wisconsin Avenue, NW
Washington, DC 20007
1-877-219-2970
http://www.gloverparkhotel.com/

For additional information please contact Sandra Smith

Airports
The closest airport is *Ronald Reagan Washington National - DCA* (directions)
Distance: approx. 7.0 mi
Transportation information: http://www.flyreagan.com/dca/parking-transportation

* Dulles International Airport - IAD* (directions)
Distance: approx. 37.0 mi
Transportation information: http://www.flydulles.com/iad/parking-transportation

**Parking and Local transportation**

If you plan to drive please let us know in advance so we can arrange parking at the Harris Building garage.

Georgetown University has a free shuttle that connects the Main Campus with university offices on Wisconsin Avenue and the Georgetown Safeway. The shuttle runs Monday through Friday and stops at 3300 Whitehaven Street. To check schedule: https://transportation.georgetown.edu/guts/wisconsin-avenue

There are many public buses along Wisconsin Ave. **30s series buses (30N, 30S, 31, 33)** - connects with Friendship Heights and Tenleytown-AU metro stations (Red Line), as well as Foggy Bottom and Farragut West metro stations (Blue, Orange, and Silver Lines).

Metro information is available at http://www.wmata.com/