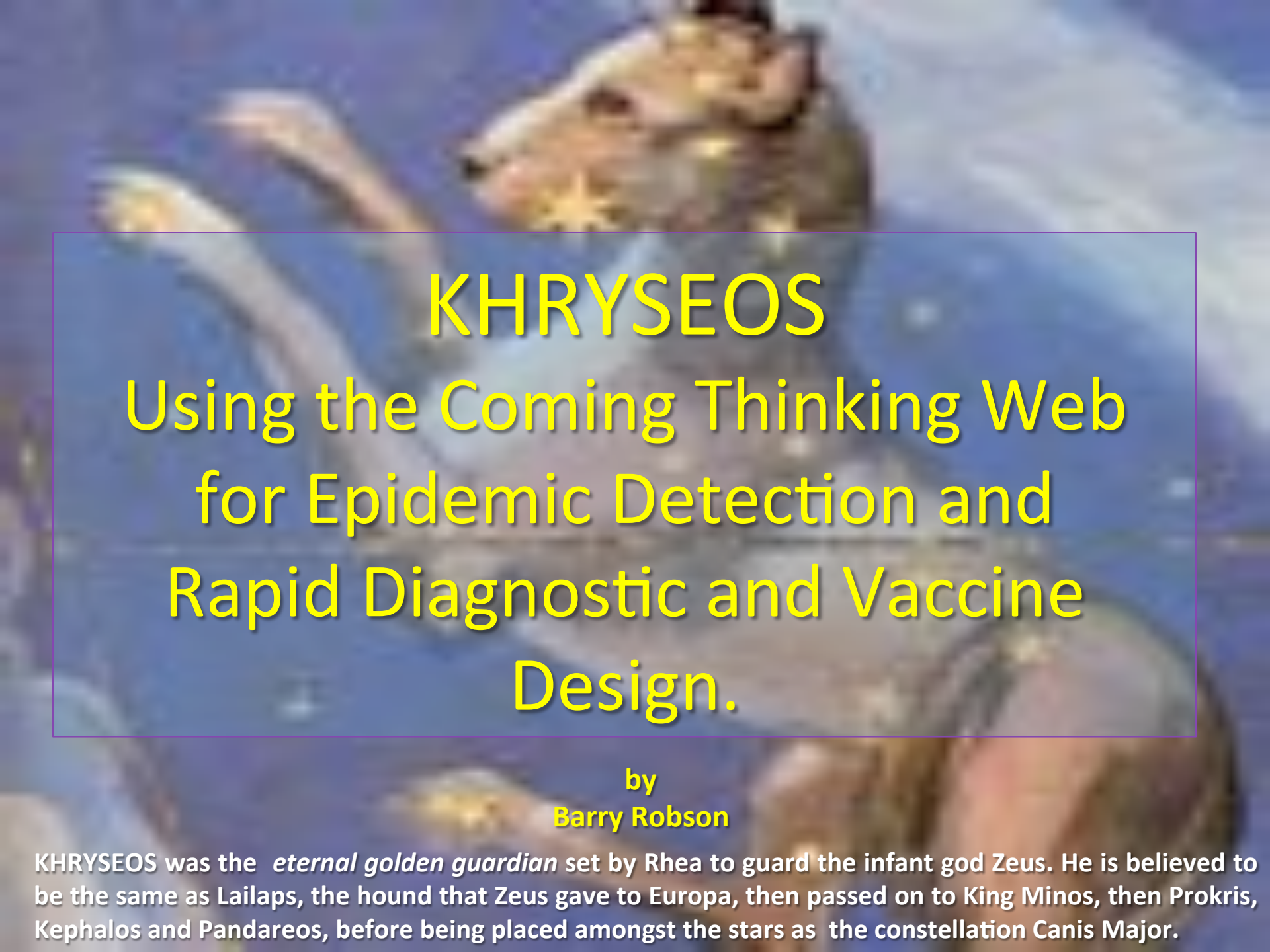


KHRYSEOS
Using the Coming Thinking Web for Epidemic Detection and Rapid Diagnostic and Vaccine Design.

by
Barry Robson

KHRYSEOS was the *eternal golden guardian* set by Rhea to guard the infant god Zeus. He is believed to be the same as Lailaps, the hound that Zeus gave to Europa, then passed on to King Minos, then Prokris, Kephalos and Pandareos, before being placed amongst the stars as the constellation Canis Major.

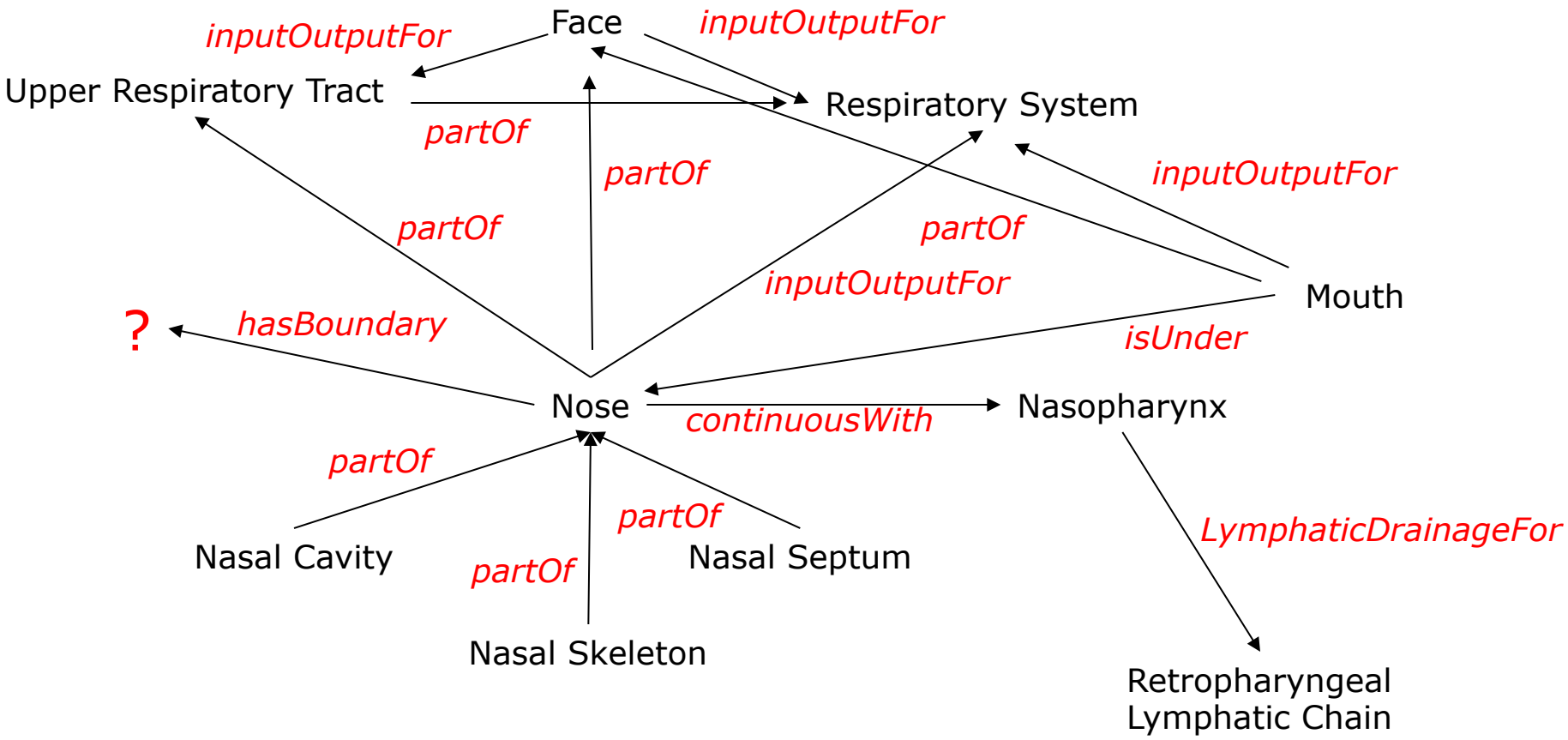
Past and Future Layers of the Internet

- **INTERNET - Connects Computers** - The US Department of Defense awarded contracts as early as the 1960s for packet network systems, including the development of the ARPANET, which would become the first network to use the **Internet Protocol**.
- **World Wide Web 1.0 - Connects Web Pages** - Berners-Lee wrote a proposal in March 1989 for "a large hypertext database with typed **links**" . Although the proposal attracted little interest, Berners-Lee was encouraged by his boss.[] He considered several names, including *Information Mesh*,[7] *The Information Mine* or *Mine of Information*, but settled on *World Wide Web.*
- **World Wide Web 2.0 - Connects People** - It means sites that use technology beyond the static pages of earlier Web sites. Essentially, it connects people by facilitating **social networking**. The term was coined in 1999 by Darcy DiNucci and was popularized by Tim O'Reilly at the O'Reilly Media Web 2.0 conference in 2004.
- **World Wide Web 3.0 – Connects Data and Knowledge - The Semantic Web** is a collaborative movement led by international standards body the World Wide Web Consortium (W3C). The Semantic Web aims at converting the current web, dominated by unstructured and semi-structured documents into a "web of data". The Semantic Web stack on the W3C's Resource Description Framework (RDF).
- **World Wide Web 4.0 - The Thinking Web** - Will organize probabilistic knowledge and reason with it across multiple servers – helps make decisions.

The Medical Semantic Web

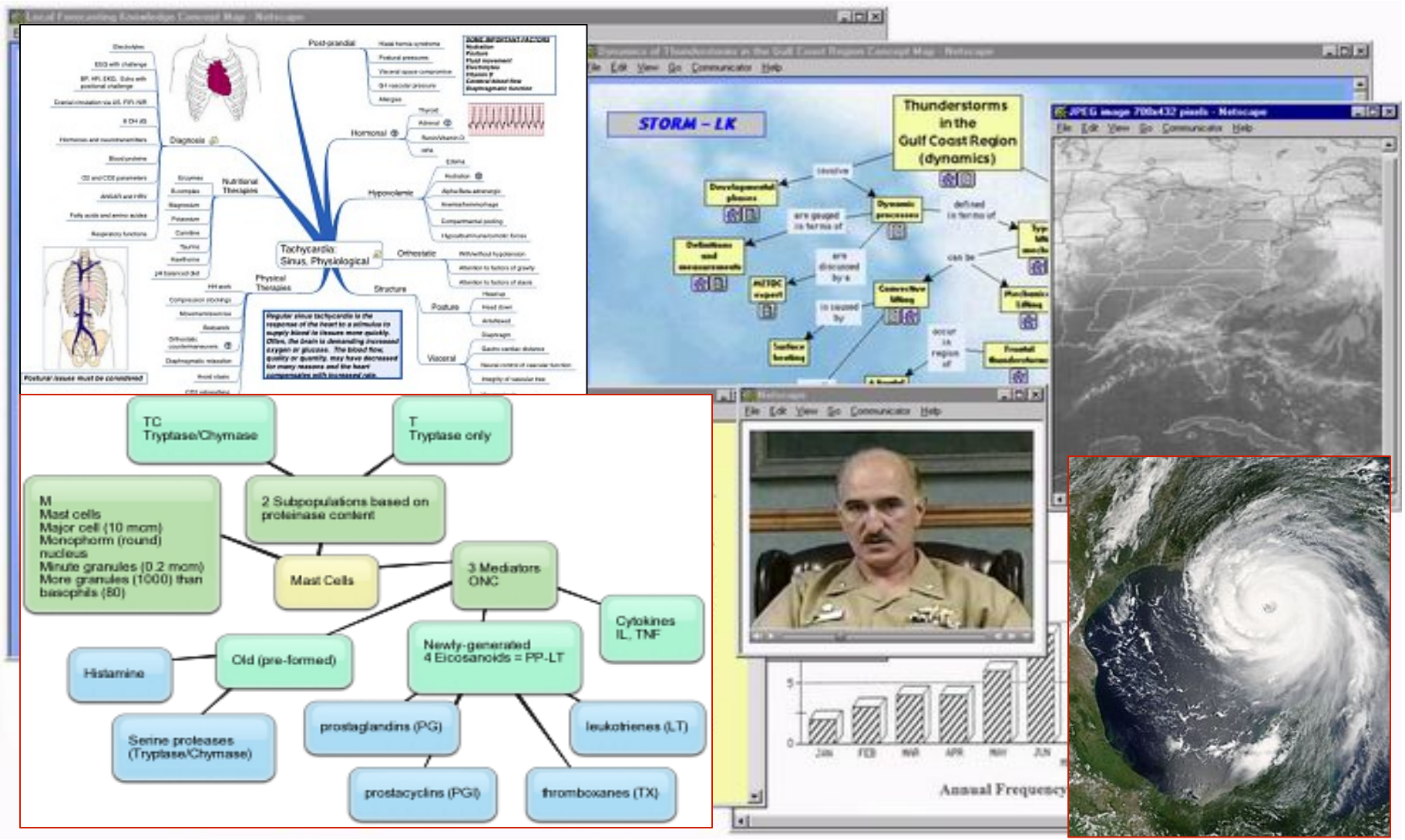
MEDICAL SEMANTIC WEB

W3C Consortium Medical *Semantic Web* Project is Connecting Knowledge and Data on the Internet, in "tags", not just Web Pages

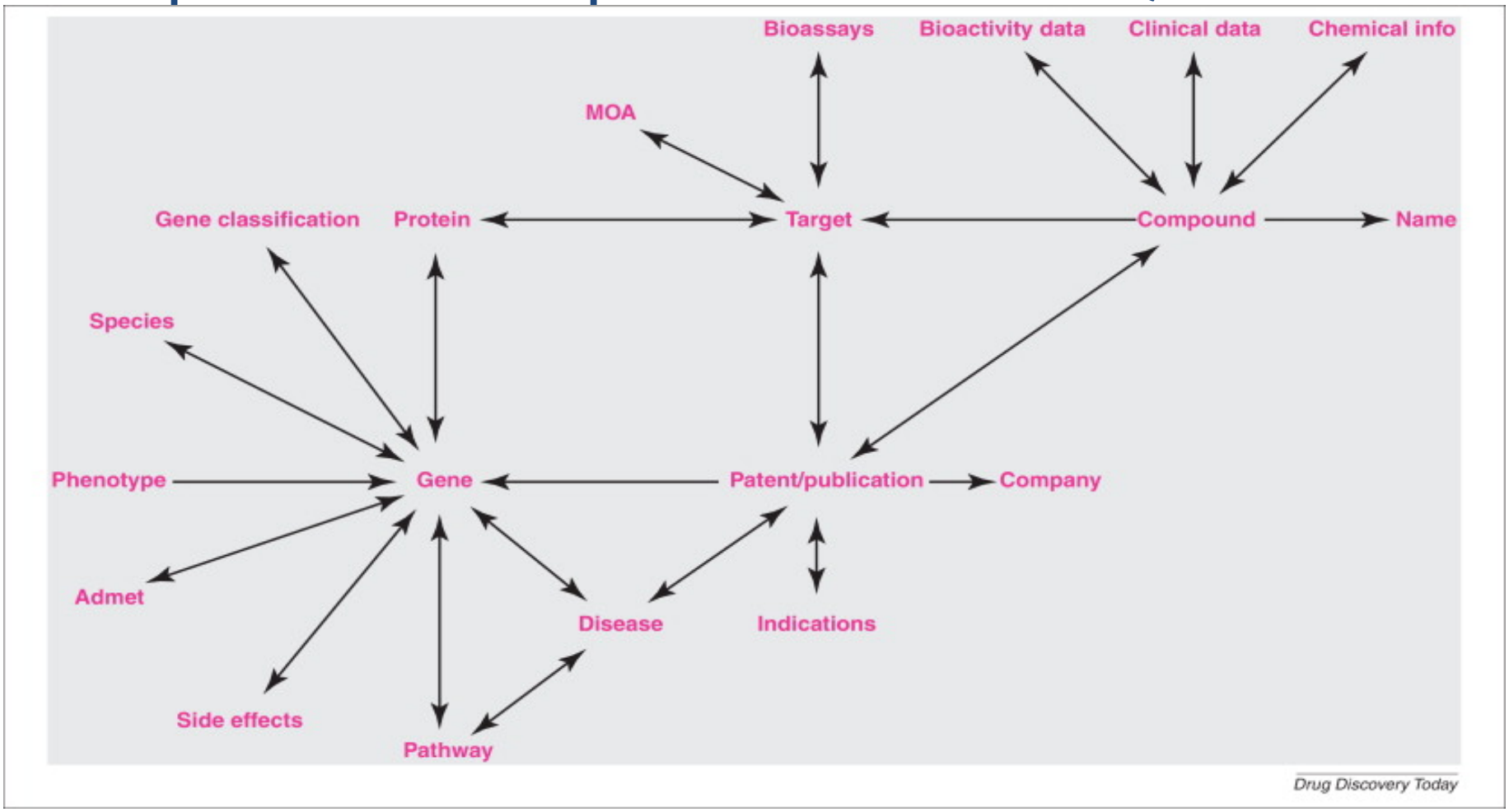


Achille Fokoue, Aditya Kalyanpur, Aaron Kershenbaum, Li Ma, Edith Schonberg, Kavitha Srinivas IBM NY, China Research Lab, Chintan Patel, James Cimino Columbia University Medical Center

Such Knowledge Nets Are Already Increasingly Used To Support Prediction, Risk Assessment, Decision Making



The Network of Associations Implied in the Top Ranked Biopharmaceutical Queries



The network summarizing data associations that are needed to target the top 20 research questions.

K. Azzaoui et. al. (2013) "Scientific competency questions as the basis for semantically enriched open pharmacological space development", Drug Discovery Today, Vol 18, Issues 17–18, Pages 843–852.

THE CALL FOR A UNIVERSAL EXCHANGE LANGUAGE (UEL)

The PCAST Report

REPORT TO THE PRESIDENT REALIZING THE FULL POTENTIAL OF HEALTH INFORMATION TECHNOLOGY TO IMPROVE HEALTHCARE FOR AMERICANS: THE PATH FORWARD

President's Council of Advisors on Science and Technology (December 2010)

The PCAST Concerns

- "In other sectors, universal exchange standards have resulted in new products that knit together fragmented systems into a unified infrastructure."
- "The resulting ' network effect' then increases the value of the infrastructure for all, and spurs rapid adoption."
- "By contrast, health IT has not made this transition."

•"The market for new products and services based on health IT remains relatively small and undeveloped compared with corresponding markets in most other sectors of the economy, and there is little or no network effect to spur adoption."

and so they call for an XML-like "***Universal Exchange Language***"

UEL!

YOSEMITE MANIFESTO
Response to PCAST from the Semantic Web Community

Yosemite Manifesto on RDF as a Universal Healthcare Exchange Language

Position statement from the Workshop on RDF as a Universal Healthcare Exchange Language held at the 2013 Semantic Technology and Business Conference, San Francisco, in response to the President's Council of Advisors on Science and Technology (PCAST) report calling for a universal exchange language for healthcare.

1. **RDF** [the basic link-to-semantic-definition mechanism of the SW] **is the best available candidate for a universal healthcare exchange language.**
2. **Electronic healthcare information should be exchanged in a format that either: (a) is an RDF format directly; or (b) has a standard mapping to RDF.**
3. **Existing standard healthcare vocabularies, data models and exchange languages should be leveraged by defining standard mappings to RDF, and any new standards should have RDF representations.**
4. **Government agencies should mandate or incentivize the use of RDF as a universal healthcare exchange language.**
5. **Exchanged healthcare information should be self-describing, using Linked Data principles, so that each concept URI is de-referenceable to its free and open definition.**

Our Q-UEL Language Predated the Yosemite Manifesto.

Robson, B., Caruso, T, and Balis, U. G. J. (2014) "Suggestions for a Web Based Universal Exchange and Inference Language for Medicine. Continuity of Patient Care with PCAST Disaggregation. *Computers in Biology and Medicine*, 2014 (in press).

Robson, B. (2014) "POPPER, a Simple Programming Language for Probabilistic Semantic Inference in Medicine." *Computers in Biology and Medicine*, 2014 (in press).

Robson, B., Caruso, T, and Balis, U. G. J. (2013)"Suggestions for a Web Based Universal Exchange and Inference Language for Medicine", Computers in Biology and Medicine,1;43(12):2297-310. Epub 2013 Sep 20.Also found in preliminary form, with permission of the Editor-in-Chief, at the US Government S&I website:- http://wiki.siframework.org/file/view/ UELRobson102corrections.pdf/451304614/UELRobson102corrections.pdf

Robson, B. (2014) " Hyperbolic Dirac Nets for Medical Decision Support. Theory, Methods, and Comparison with Bayes Nets" *Computers in Biology and Medicine*, in 2014 Aug;51:183-97.

Robson, B. (2013)"Towards New Tools for Pharmacoepidemiology", Advances in Pharmacoepidemiology and Drug Safety, 1:6, http://dx.doi.org/ 10.4172/2167-1052.100012

Robson, B. (2012) "Towards Automated Reasoning for Drug Discovery and Pharmaceutical Business Intelligence", Pharmaceutical Technology and Drug Research, Pharmaceutical Technology & Drug Research 2012 1: 3 (27 March 2012)

Robson, B., Balis, U. G. J. and Caruso, T. P. (2011)"Considerations for a Universal Exchange Language for Healthcare." In Proceedings of 2011 IEEE 13th International Conference on e-Health Networking, Applications and Services (Healthcom 2011), 173–176. Columbus, MO: IEEE, 2011.

Barry Robson, Ulysses G.J. Balis, and Thomas P. Caruso. Considerations for a Universal Exchange Language for Healthcare, 13th IEEE International Conference on e-Health Networking Applications and Services (IEEE Healthcom '11), Columbia, MO, June 13, 2011, pages 173-176. http://tpcaruso.com/BITT/ Files/HealthCom11-Final_2.pdf

Robson, B, and TP Caruso (2013) "A Universal Exchange Language for Healthcare" MedInfo '13: Proceedings of the 14th World Congress on Medical and Health Informatics, Copenhagen, Denmark, Edited by CU Lehmann, E Ammenwerth, and C Nohr. IOS Press, Washington, DC, USA. http:// quantalsemantics.com/ documents/MedInfo13-RobsonCaruso_V6.pdf; http://ebooks.iospress.nl/publication/ 34165

B. Robson, Rethinking Global Interoperability in Healthcare. Reflections and Experiments of an e-Epidemiologist from Clinical Record to Smart Medical Semantic Web Johns Hopkins Grand Rounds Lectures (last accessed 3/14/2013). http://webcast.jhu.edu/Mediasite/Play/ 80245ac77f9d4fe0a2a2bbf300caa8be1d

Robson, B. (2012) "Schrodinger's Better Patients". http://sils.unc.edu/events/2012/better-patients.

Robson, B. (2009) "Links Between Quantum Physics and Thought" (A. I. Applications in Medicine) , Future of Health Technology Congress, Technology and Informatics, Vol. 149, 157-177 IOS Press

Robson, B. (2009) "Towards Intelligent Internet-Roaming Agents for Mining and Inference from Medical Data", Future of Health Technology Congress, Technology and Informatics, Vol. 149, 157-177 IOS Press

Robson, B. (2007) "Data Mining and Inference Systems for Physician Decision Support in Personalized Medicine" Lecture and Circulated Report at the 1st Annual Total Cancer Care Summit, Bahamas 2007.

Robson, B. (2007) "The New Physician as Unwitting Quantum Mechanic: Is Adapting Dirac's Inference System Best Practice for Personalized Medicine, Genomics and Proteomics?" B. Robson (2007), J. Proteome Res. (A. Chem. Soc.), Vol. 6, No. 8: 3114 – 3126

UEL Challenges, Solutions, and Our Project Status

CHALLENGE	SOLUTION	PROJECT STATUS
AVAILABILITY. Medical records should be rapidly available on the Internet to treat the patient any time, anywhere, but this risks breach of privacy/security.	**Disaggregate** (shred) the record amongst many others to add dimension of "entropy protection" on top of encryption, as proposed by PCAST. Use passwords and digital certificates to reaggregate.	A flexible toolkit provides proof of concept with rapid and scalable solutions.
SECURITY/AUTHORITY/CONSENT. Only the right people should see the data that the patient wants them to see.	Build **fine grained consent** instructions into the basic data combined with use of authority keys. Use the fine grained consent to reproduce separately data visible or lightly encrypted for data mining.	Done, but the consent language continues to be expanded to cover more use cases.
GRANULARITY. The current XML format has spread the record information such that analysts use unstructured data mining techniques to convert to simple Event Attribute Value formats. But that takes time, and there will soon be some 250m constantly changing digital patient records in the USA.	**Rethink data representation. The basic elements of the UEL-based record should essentially be of Event Attribute Value format from the outset.**	**Data attribute specification has been expanded to a metadata ontology language including time stamp and consent description.**
PROBABILITY. Medical data and medical summary rules are fundamentally probabilistic, but there is no agreed Best Practice for managing this nor even of all the essential forms of automated reasoning for	**Rethink the problem based on long proven observation, measurement and probabilistic inference principles used in physics, and build it into medical information representation in a**	**Representation and prototype inference engines use a UEL fundamentally based on Dirac notation**

Q-UEL Objects called tags autosurf and spawn on the Web, gathering knowledge

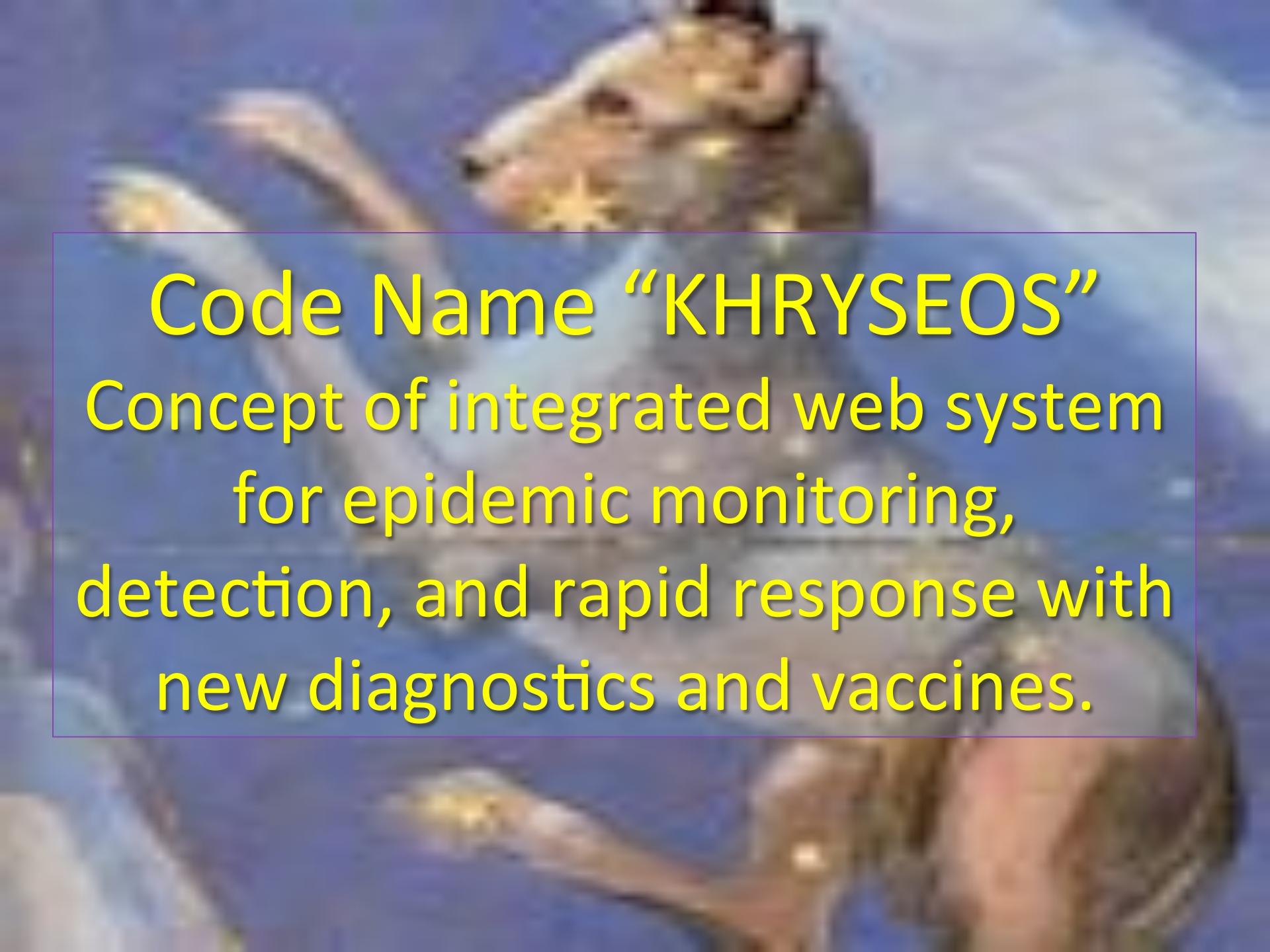
- < Q-UEL-XTRACT-BIOLOGY "`The human _brain |**^is `the center of**| `the human nervous _system [0http://en.wikipedia.org/wiki/Nervous_system]; `The human _brain **|^has `the `same `general _structure as**| `the _brains |**of|** `other mammals [0http://en.wikipedia.org/wiki/Mammal]; `The human _brain **|^is larger than ^expected on `the basis of**| _body _size |**among**| `other primates [0http://en.wikipedia.org/wiki/Primate] [1(0)http://www.ncbi.nlm.nih.gov/pubmed/17148188] [2file:input.txt#cite_note-Brain-num-1]" | **from** | source:='http://en.wikipedia.org/wiki/Human_brain' time:='Wed Oct 3 14:02:19 2012' extract:=0 Q-UEL-XTRACT-BIOLOGY > **(this XTRACT tag is the actual autosurf-and-spawn harvester). Examples of tidied knowledge:-**

- <Q-UEL-MOLECULE Ampicillin | **means**:= www.qexl.org/means_2/ | code:=IUPAC:= '**2S,5R,6R)-6-{[(2R)-2-Amino-2-phenylacetyl]amino}-3,3-dimethyl-7-oxo-4-thia-1-azabicyclo[3.2.0]heptane-2-carboxylic acid' or** code:=SMILES:= O=C(O)[C@@H]2N3C(=O)[C@@H](NC(=O)[C@@H](c1ccccc1)N)[C@H]3SC2(C)C **or** code:=InChI:= InChI=1S/C16H19N3O4S/c1-16(2)11(15(22)23)19-13(21)10(14(19)24-16)18-12(20)9(17)8-6-4-3-5-7-8/h3-7,9-11,14H,17H21-2H3,(H,18,20)(H,22,23)/t9-,10-,11+,14-/m1/s1 **and** 'empirical formula':=C16H19N3O4S and Monoisotopic mass:=349.109619 **and** 'average mass (Da)':= 349.404785 Q-UEL-MOLECULE>

- <Q-UEL-THESAURUS cat | **suggests** | '2. Special Vitality':='366. Animal.':=pbwd:=0.03502 or 'Section II. PRECURSORY CONDITIONS AND OPERATIONS':='455. [The desire of knowledge.] Curiosity.':=pbwd:=0.03226 **or** '(ii) SPECIFIC SOUNDS':='407. [Repeated and protracted sounds.] Roll.':=pbwd:=0.00667 **or** '(ii) SPECIFIC SOUNDS':='412. [Animal sounds.] Ululation.':=pbwd:=0.00632 **or** 'Present Events':='151. Eventuality.':=pbwd:=0.00571 **or** '(iii) PERCEPTIONS OF LIGHT':='441. Vision.':=pbwd:=0.00448 **or** 'SECTION III. ORGANIC MATTER 1. VITALITY 1. Vitality in general':='359. Life.':=pbwd:=0.00392 **or** '3. Fluids in Motion':='348. [Water in motion.] River.':=pbwd:=0.00356 or '3. PROSPECTIVE AFFECTIONS':='864. Caution.':=pbwd:=0.00223 **or** '5. INSTITUTIONS':='975. [Instrument of punishment.] Scourge.':=pbwd:=0.00151 or '3. Contingent Subservience':='668. Warning.':=pbwd:=0.00142 or ……*(etc)*…… Q-UEL-THESAURUS>

Vaccine and Diagnostic Design - Use of knowledge from a medical Semantic Web

- Identify a new strain emerging by unstructured data mining / text analytics of Internet activity.
- Explore to see if new DNA/RNA sequence is available for the new pathogen or new strain of pathogen.
- See if three dimensional structures are available for proteins with related sequences
  - A protein with a detectable sequence homology is almost always homologous in three dimensional structure.
- Search on knowledge base to find facts about
  - symptoms,
  - attack/incidence/prevalence rate, latency/time to symptioms, mortality and fatality rate,
  - preexisting vaccines for the species of pathogen and their limitations,
  - protein sequence variants, three dimensional structure if known,
  - three dimensional structures of related proteins if known,
  - Immunoinformatics, B-eptopicity and T-epitopicity.
- Raise antibodies for continually growing network of biosensors that directly feed into (constantly report to) the Web.
  - "Living in the Connected World. How Global Sensor Networks are Extending the Human Nervous System/ How a Sensor-Filled World Will Change Human Consciousness ", Gershon Dublon and Joseph A. Paradiso (2014) Scientific American, July.


Code Name “KHRYSEOS”
Concept of integrated web system for epidemic monitoring, detection, and rapid response with new diagnostics and vaccines.

Overview of KHRYSEOS

Knowledge in text and images

Unstructured Data mining

WW4 Probabilistic Semantic Web

Structured Data mining

Consented, de-identified patient records

Event Mining

PROBABLISTIC **STATEMENTS** representing information and knowledge

METASTATEMENTS Rules of logic, syllogisms etc. grammar, definitions of vocabulary that manipulate STATEMENTS.

Inference NET

Inference ENGINE

PORTAL as "Expert System" (Decision support System)

Note the important cycle seeking molecular recognition information

Diagnostics/ biosensors

EPITOPES = MOLECULAR RECOGNITION (short peptides from pathogen proteins)

Cartridge vaccines

The web is already used to detect epidemics

Nature **457**, 1012-1014 (19 February 2009)

Detecting influenza epidemics using search engine query data

Jeremy Ginsberg[1], Matthew H. Mohebbi[1], Rajan S. Patel[1], Lynnette Brammer[2], Mark S. Smolinski[1] & Larry Brilliant[1]

- [1] **Google Inc.**, 1600 Amphitheatre Parkway, Mountain View, California 94043, USA
- [2] **Centers for Disease Control and Prevention**, 1600 Clifton Road, NE, Atlanta, Georgia 30333, USA

- "One way to improve early detection is to monitor health-seeking behavior in the form of queries to online search engines, which are submitted by millions of users around the world each day. Here we present a method of analyzing large numbers of Google search queries to track influenza-like illness in a population. "

- "Because the relative frequency of certain queries is highly correlated with the percentage of physician visits in which a patient presents with influenza-like symptoms, we can accurately estimate the current level of weekly influenza activity in each region of the United States, with a reporting lag of about one day. This approach may make it possible to use search queries to detect influenza epidemics in areas with a large population of web search users."

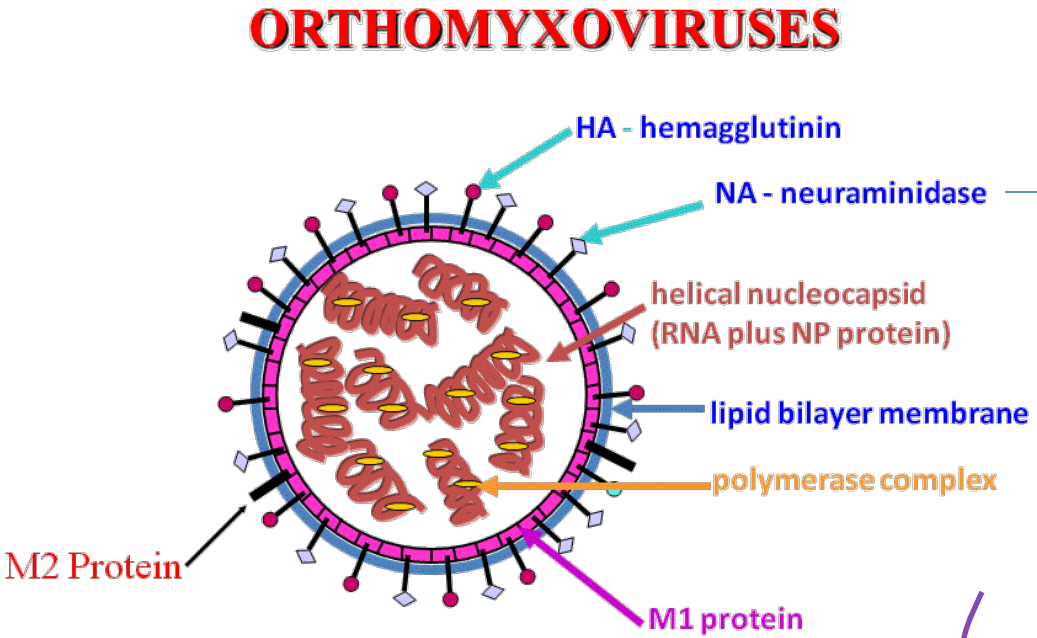
Synthetic (Peptide) Vaccines

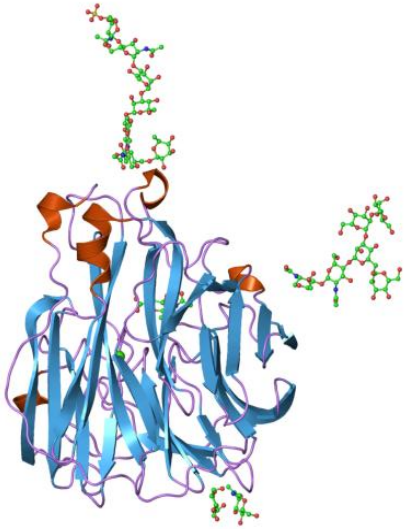
• During the H1N1 outbreak in 2009, vaccines only became available in large quantities after the peak of human infections. This was a learning experience for vaccination companies. Creating vaccines synthetically has the ability to increase the speed of production. This is especially important in the event of a pandemic.

•A **synthetic vaccine** is a vaccine consisting mainly of synthetic peptides, carbohydrates, or antigens, usually linked to a carrier protein to render them immunogenic. They are usually considered to be safer than vaccines from bacterial cultures.

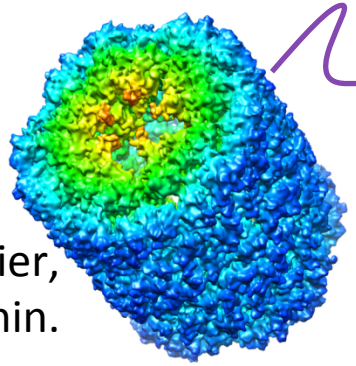
- •The world's first synthetic vaccine was created in 1982 from diphtheria toxin by Louis Chedid (scientist) from the Pasteur Institute and Michael Sela from the Weizmann Institute.
- •In 1986, Manuel Elkin Patarroyo created the SPf66, the first version of a synthetic vaccine for Malaria.
- •Many early vaccines used dead samples of FMDV to inoculate animals, but those early vaccines sometimes caused real outbreaks. Scientists discovered that a vaccine could be made using only a single key protein from the virus. Further, loops from the surface proteins in cloned or synthetic constructs.
- •Novartis Vaccine and Diagnostics, among other companies, developed a synthetic approach that very rapidly generates vaccine viruses from sequence data in order to be able to administer vaccinations early in the pandemic outbreak.

Only small molecular patches, possibly loops that are of 5-10 amino acid residues, matter….





B-epitope



Protein carrier,
e.g. keyhole limpet hemocyanin.

Good for raising antibodies in sheep etc. to use in diagnostics/biosensors. For Vaccines, also attach….

- T-epitopes (cell response, immune memory)
- Molecular adjuvant (e.g. muramyl dipeptide)
- Possibly, anti-inhibitory peptides

Keyhole limpet hemocyanin (KLH)

- Used extensively as a carrier protein in the production of antibodies for research, biotechnology and therapeutic applications.
- Haptens are substances with a low molecular weight such as peptides, small proteins and drug molecules that are generally not immunogenic and require the aid of a carrier protein to stimulate a response from the immune system in the form of antibody production.
- KLH is the most widely employed carrier protein for this purpose. KLH is an effective carrier protein for several reasons.
- Its large size and numerous epitopes generate a substantial immune response, and abundance of lysine residues for coupling haptens allows a high hapten:carrier protein ratio, increasing the likelihood of generating hapten-specific antibodies.
- In addition, because KLH is derived from the limpet, a gastropod, it is phylogenetically distant from mammalian proteins, thus reducing false positives in immunologically-based research techniques in mammalian model organisms, and clinically avoiding autoimmune effects.

Our Peptide Diagnostic and Vaccine Patents (1)

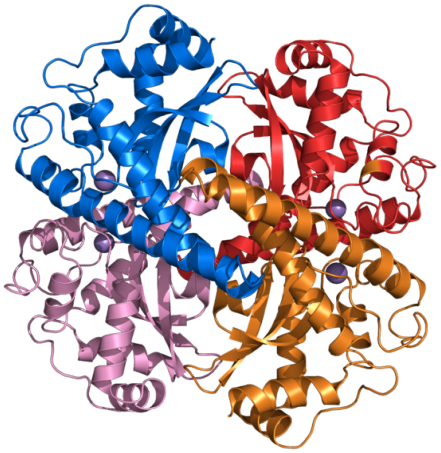
193. Fishleigh, R. V., Robson, B., and Morrison, C. (1987) "Analogues of LHRH United Kingdom: GB 8723072.8.
194. Robson, B. and Fishleigh, R. V. (1988) "Synthetic Peptides Related to HIV-env ProteinsInternational Patent Application No. PCT/GB88/00491, International Publication No: WO88/10267, dated 29th December 1988
195. Morrison, C., Robson, B., and Fishleigh, R. V., (1988) “Improvements in or relating to hormones” EP00293530A2 (12/07/1988)
196. Robson, B. (1989) "Computer Aided Peptide and Protein Engineering", Progress in Clinical and Biological Research, Ed. J. L. Fauchere. Alan R. Liss Inc. New York, 227-23
197. Robson and R.V. Fishleigh, R. V. (1989) EP00298633A2 “Synthetic polypeptides” (01/11/1989)
198. Robson, B. and Fishleigh, R. V. (1989) EP00298633A3 “Synthetic polypeptides” (04/26/1989)
199. Bomford, R., Garnier, J., and Robson, B. (1990) "Computer Aided Peptide and Protein Engineering Software Development", in 'Biotechnology R and D in the EC (BAP) 1985-1989'., (Eds. A Vassarotti and E Magnien). Elsevier, Paris, II, 59-64
200. Fishleigh, R. V., and Robson, B. (1988) "Synthetic Peptides Related to HIV-env Proteins P.C.T. International Patent Application No. PCT/GB88/00491, International Publication No: WO88/10267, dated 29th December 1988. EP00371046A1 “Synthetic Peptides Related To HIV-Env Proteins” (06/06/1990)
201. Morrison, C., Robson, B., and Fishleigh, R. V., (1990) “Improvements in or relating to hormones” EP00293530A (05/30/1990)
202. Fishleigh, R. V. and Robson, B. (1991) "Synthetic Peptides", R. V. Fishleigh, B. Robson. P.C.T. International Patent Application No. PCT/GB91/00392 International Publication No: WO91/13909, dated 19th September 1991
205. Fishleigh, R. V., and Robson, B. "Synthetic Peptides” P.C.T. International Patent Application No. PCT/GB91/00392, WO09113909A1 (09/19/1991)
208. Fishleigh, R. V., Robson, B., (1992) “Synthetic Peptides” EP00519986A1 (12/30/1992)
209. Fishleigh R. V. and Robson, B. (1992) “Analogues of Piscine LHRH” WO09212247A1 (07/23/1992)
210. Robson, B., Mee, R. P., (1993) "Fragments of Prion Protein",.P.C.T. International Patent Application No. PCT/GB92/02246, International Publication No: WO93/11155, dated 10th June 1993
211. Fishleigh, R.V. and Robson, B. (1993) “Synthetic Polypeptides Derived From The HIV Envelope Glycoprotein” P00636145A1 (10/28/1993)
212. Fishleigh, R. V., Robson, B., and Aston, R. (1993) "Synthetic Polypeptides", R. V. Fishleigh, B. Robson and R. Aston, B.P.C.T. International Patent Application No. PCT/GB93/00808 International Publication No: WO93/21218, dated 28th October 1993
213. Fishleigh, R. V., Robson, B., and Aston, R. (1993) "Synthetic Polypeptides", B.P.C.T. International Patent Application No. PCT/GB93/00808
214. Fishleigh, R. V., Robson, B, and Mee, P. (1993) "Fragments of Prion Protein", P.C.T. International Patent Application No. PCT/GB92/02246, International Publication No: WO9311155A1 (06/10/1993), dated 10th June 1993
215. Fishleigh, R. V., Robson, B, and Mee, P. (1993) “Fragments of prion proteins” US Patent US05773572 (06/30/1998)
216. Fishleigh R. V. and Robson, B. (1993) “Analogues of Piscine LHRH” EP00566611A1 (10/27/1993)
217. Fishleigh, R. V. and Robson, B. (1994) “Analogues of Piscine LHRH' R. V. Fishleigh, B. Robson, P.C.T. International Patent Application No. PCT/GB92/0004
218. Fishleigh, R. V. and Robson, B. (1994) “Synthetic Polypeptides”, International Publication No: WO92/12247, dated 23rd July 1994

Our Peptide Diagnostic and Vaccine Patents (2)

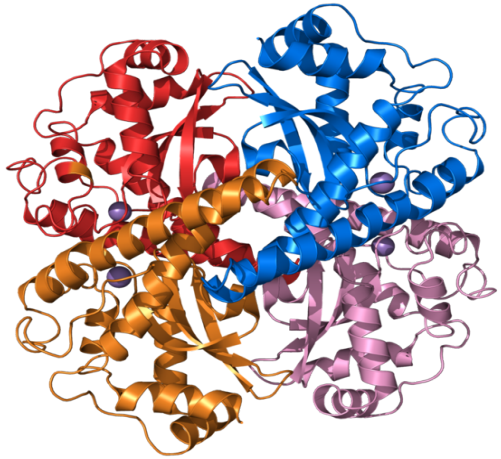
219. Fishleigh, R. V. and Robson, B. (1994) “Synthetic Polypeptides” ZA09302648A 10/15/1994
220. Fishleigh, R. V. and Robson, B. (1994) “Synthetic Polypeptides” CN01090584A 08/10/1994
221. Fishleigh R. V. and Robson, B. (1994) “Analogues of Piscine LHRH” International Patent Application No. PCT/GB92/0004 International Publication No: WO92/12247, dated 23rd July 1994
222. Fishleigh R. V. and Robson, B. (1994) “Analogues of Piscine LHRH” GB02267496B (09/07/1994)
223. Fishleigh R. V. and Robson, B. (1994) “Analogues of Piscine LHRH” AU00652611B2 (09/01/1994)
224. Fishleigh, R .V., Robson, B. and Greaney, P. J. (1994) "Peptides Related to Mycobacterium Bovis Proteins", P.C.T. International Patent Application No. PCT/GB93/01527 International Publication No: WO94/02508, dated 3rd February 1994
225. Fishleigh, R .V., Robson, B. and Greaney, P. J. (1994) "IPNV Vaccine", P.C.T. International Patent Application No. PCT/GB93/01812, International Publication No: WO94/04565, dated 3rd March 1994
226. Fishleigh, R. V., Robson, B., and Greaney, P. J. (1994) “Polypeptide Antigens Of Mycobacterium Bovis” (02/03/1994) WO09402508A2 (02/03/1994) WO09402508A3 (06/09/1994)
227. Fishleigh, R. V., Robson, B., and Greaney, P. J. (1994) P.C.T. International Patent Application No. PCT/GB93/01812, WO09404565A3 (03/03/1994) WO09404565A2 (03/03/1994) WO09404565A3 (06/09/1994)
228. Fishleigh, R. V., Robson, B, and Mee, P. (1995) “Synthetic Polypeptides With At Least One Antigenic Site Of A Prion Protein, Methods Of Their Use And Manufacture, Antibodies Thereto, Compositions Kits NZ00246059A.(08/28/1995)
229. Fishleigh, R. V., Robson, B., and Morrison, C. (1995) “Luteinising Hormone-Releasing Hormone (LHRH) Analogues, DNA Coding For Analogues, And Vaccines” (10/26/1995) NZ00222031A
230. Fishleigh, R,V, Robson, B.. and Aston, R.(1995), ”Synthetic Polypeptides Derived From The HIV Envelope Glycoprotein.” Ep Patent 0,636,145
231. Fishleigh R. V. and Robson, B. (1995) “Analogues of Piscine LHRH” HU00066829A2 (01/30/1995)
232. Fishleigh, R.V. and Robson, B. (1997) “Synthetic Polypeptides Derived From The HIVEnvelope Glycoprotein” (02/01/1995) AU00675053B2 01/23/1997
233. Fishleigh, R.V. and Robson, B. (1997) “Fragments Of Prion Proteins” EP00636145A1 02/01/1995)
238. Fishleigh, R. V., Robson, B, and Mee, P. (1999) “Prion” GR03029740T3 06/30/1999
239. Fishleigh, R. V., Robson, B, and Mee, P. (1999) “Fragmentos De Proteinas De Prion” ES02128362T3 05/16/1999
240. Fishleigh, R. V., Robson, B, and Mee, P. (1999) “Fragmente Von Prion Proteinen” AT00177754E 04/15/1999
241. Fishleigh, R. V., Robson, B, and Mee, P. (1999) “Prion” EP00862447A1 09/09/1998
242. Fishleigh, R. V. and Robson, B. (1999) “LHRH hormones” (04/27/1999) US Patent US05897863

Totally Chemo-Synthetic "Bionanotechnology" Constructs

Normal SOD made of L-amino acids

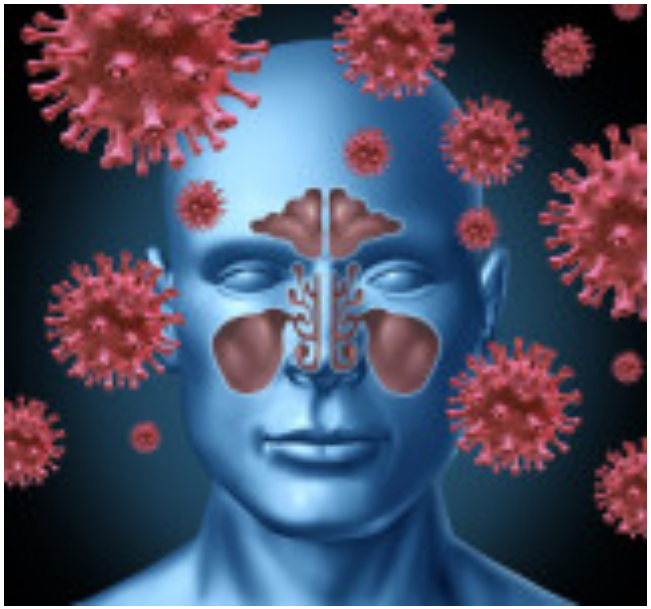


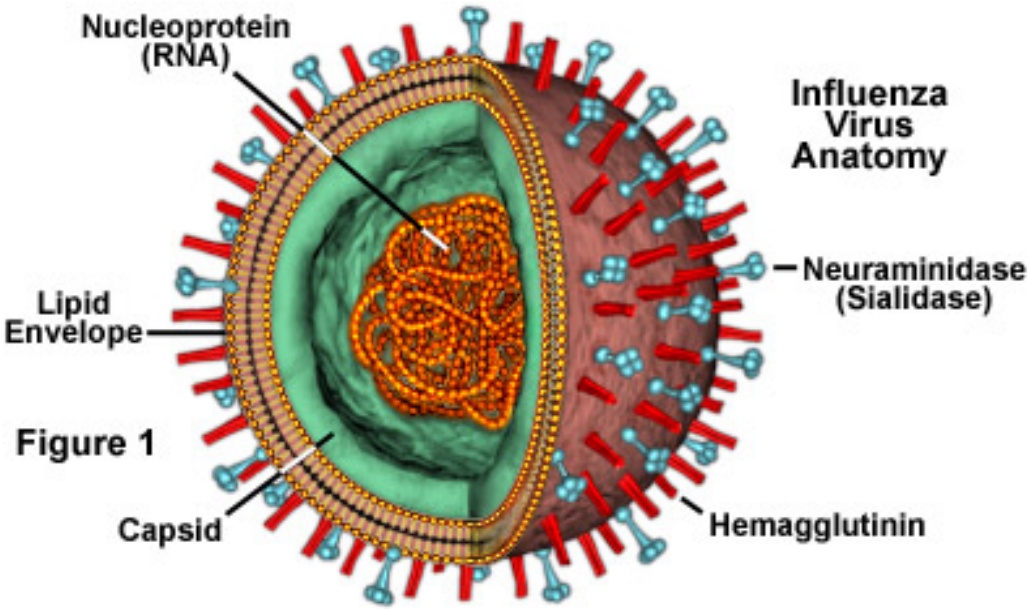
Mirror image SOD made of D-amino acids



- G. M. Figliozzi, M. A. Siani, L. E. Canne, B. Robson, and R. J. Simon (1996)) "Chemical synthesis and activity of D-superoxide dismutase", Protein Science, 5, Suppl.1, 72
- Siani, M. A., Canne, L.E., Figliozzi, G. M., Robson, B. "Total Chemical Synthesis of Proteins including Chemokines and their Analogues, "Chemokines", International Business Communications (1997)
- Robson (1999) "Beyond Proteins" Trends in Biotechnology, 17:311-315
- Canne, L. E., Figliozzi, G. M., Robson, B., Siani, M. A., and Simon, R.J. (1996) "N-Alkoxy amid backbone protection in BOC chemistry : improved synthesis of a 'difficult sequence'", Protein Science, 5, Suppl.1, 72
- Robson, B. (1996) "Doppelganger Proteins as Drug Leads", B. Robson (1996), Nature Biotechnology, 14, 892-893
- Robson, B. (1998) "Generating D-Peptides: Methods And Compositions" EP00862447A1 (09/09/1998) WO09713522A1 (04/17/1997)
- Robson, B. (1997) "Generating D-Peptides: Methods And Compositions" AU07442296A1 04/30/1997
- Robson, B. (1997) "Generating D-Peptides: Methods And Compositions" CA02234723AA 04/17/1997

A Simple Worked Example With Influenza





Use the boolean operators AND, OR, NOT to do complex searches.
Parenthesis will dictate order of evaluation.

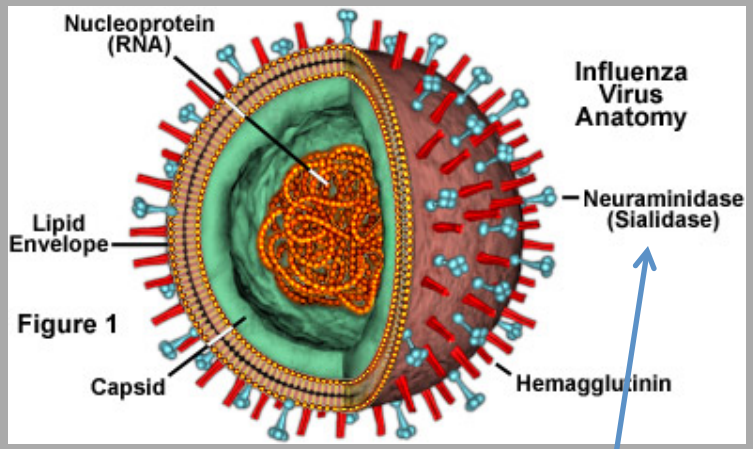
Contains neuraminidase

Hits per page 10 Display mode Full

Search Reset Return Help Report Bugs

Red = Commercial, Green = Public Domain, Blue = Unknown as yet

AAINDEX_1	Amino Acid index database Amino Acid Indices
AAINDEX_2	Amino Acid index database Amino Acid Mutation Matrices
BLOCKS	Multiple alignments of conserved regions in protein families
ECDC	E.Coli Database Collection
EPD	Eukaryotic Promoter Database
GBBCT	Genbank Bacterial Sequences
GBENV	GenBank Environmental Sampling Sequences
GBHTC	GenBank High Throughput cDNA Sequencing Entries
GBINV	GenBank Invertebrate Sequences
GBMAM	GenBank Mammalian Sequences
GBNEW	GenBank Updates
GBPAT	GenBank Patent Sequences
GBPHG	GenBank Phage Sequences
GBPLN	GenBank Plant Sequences incl. fungi and algae
GBPRI	GenBank Primate Sequences
GBREFFUNGI	Genbank Refseq Fungi
GBREFINV	Genbank Refseq Invertebrate
GBREFMAM	Genbank Refseq Mammalian
GBREFMICROBIAL	Genbank Refseq Microbial
GBREFMITO	Genbank Refseq Mitochondrion
GBREFPLANT	Genbank Refseq Plant
GBREFPLASMID	Genbank Refseq Plasmid
GBREFPLASTID	Genbank Refseq Plastid
GBREFPROTOZOA	Genbank Refseq Protozoa
GBREFSEQNEW	Genbank Refseq Updates
GBREFVERT	Genbank Refseq Non Mammalian Vertebrate
GBREFVIRAL	Genbank Refseq Viral



We will target neuraminidase

Using Web Utility “Biology Workbench”.

Enter “neuraminidase” in query box

Click on GBREFVIRAL.
(in general you can select many data bases at the same time).

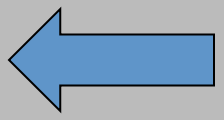
☐ S_pombe_gene	☐ StaurMu50	☐ StaurMu50_gene	☐ StaurN315
☐ StaurN315_gene	☐ StpneR6	☐ StpneR6_gene	☐ StpneTIGR4
☐ StpneTIGR4_gene	☐ Stpyo	☐ Stpyo_gene	☐ Vchol
☐ Vchol_gene			

Databases selected: GBREFVIRAL

Matches (0 to 10) / 44 View Search Results

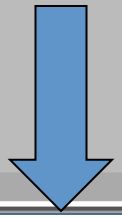
RESULTS OF neuraminidase

Rank	Score	Matching Database Record
0	5	☐ **GBREFVIRAL:8486156** Influenza B virus RNA 6, complete sequence.
1	5	☑ **GBREFVIRAL:8486127** Influenza A virus (A/Puerto Rico/8/34(H1N1)) segment 6, complete
2	4	☐ **GBREFVIRAL:55775697** Simian parainfluenza virus 5, complete genome.
3	4	☐ **GBREFVIRAL:19525721** Human parainfluenza virus 2, complete genome.
4	3	☐ **GBREFVIRAL:9634109** Bovine parainfluenza virus 3, complete genome.
5	3	☐ **GBREFVIRAL:55770820** Simian virus 41, complete genome.
6	3	☐ **GBREFVIRAL:32140165** Influenza A virus (A/Hong Kong/1073/99(H9N2)) segment 6, complete
7	2	☐ **GBREFVIRAL:9695415** Mumps virus, complete genome.
8	2	☐ **GBREFVIRAL:9627219** Sendai virus, complete genome.
9	2	☐ **GBREFVIRAL:83571714** Enterobacteria phage K1E, complete genome.



Suppose you are interested in an H1N1 Strain from Puerto Rico. Click on that one.

Click on “import Sequence”



Search | Next 10 | Show Record(s) | Show Sequence(s) | Import Sequence(s)

Contains neuraminidase

Hits per page 10 **Display mode** Full

Display in Aux. Window **in** Beautified Format Return Reset

SDSC

Influenza A virus (A/Puerto Rico/8/34(H1N1)) segment 6, complete Translatec

```
>8486127 Translated - Frame 1
SESRGLK*IQIRK**PLDQSVW*SD*LA*YCK*GI*SQYGLAIQFKLEVKTILEYATKTSLPIKIAPG*R
TQLQ*Y*PAIHLFVPSVGGLYTAKTIA*ELVPKETFLS*ESPLFHVLTWNAGPFF*PKVPY*MTGIQMGL
LRTEALIGP**AALSVKLRPRTIQDLNRLLGQQVHVMMAWAG*QSEFQVQIMEQWLY*NTTA**LKP*KV
GGRKY*GHKSLNVPV*MVHVLL**LMARVMGWPRTKFSRSKRGRLLNQ*S*MHLILTMRNVPVTLIPAK*
```

```
      S  E  S  R  G  L  K  *  I  Q  I  R  K  *  *  P  L  D  Q  S
1     agcgaaagcaggggtttaaaatgaatccaaatcagaaaataataaccattggatcaatct  60
      V  W  *  S  D  *  L  A  *  Y  C  K  *  G  I  *  S  Q  Y  G
61    gtctggtagtcggactaattagcctaatattgcaaatagggaatataatctcaatatgga  120
      L  A  I  Q  F  K  L  E  V  K  T  I  L  E  Y  A  T  K  T  S
121   ttagccattcaattcaaactggaagtcaaaaccatactggaatatgcaaccaaaacatca  180
      L  P  I  K  I  A  P  G  *  R  T  Q  L  Q  *  Y  *  P  A  I
181   ttacctataaaaatagcacctgggtaaaggacacaacttcagtgatattaaccggcaatt  240
      H  L  F  V  P  S  V  G  G  L  Y  T  A  K  T  I  A  *  E  L
241   catctctttgtcccatccgtgggtgggctatatacagcaaagacaatagcataagaattg  300
      V  P  K  E  T  F  L  S  *  E  S  P  L  F  H  V  L  T  W  N
301   gttccaaaggagacgtttttgtcataagagagccctttatttcatgttctcacttggaat  360
      A  G  P  F  F  *  P  K  V  P  Y  *  M  T  G  I  Q  M  G  L
361   gcaggacctttttctgacccaaggtgccttactgaatgacaggcattcaaatgggactg  420
      L  R  T  E  A  L  I  G  P  *  *  A  A  L  S  V  K  L  R  P
421   ttaaggacagaagcccttatagggccttaatgagctgccctgtcggtgaagctccgtccc  480
      R  T  I  Q  D  L  N  R  L  L  G  Q  Q  V  H  V  M  M  A  W
481   cgtacaattcaagatttgaatcggttgcttggtcagcaagtgcatgtcatgatggcatgg  540
      A  G  *  Q  S  E  F  Q  V  Q  I  M  E  Q  W  L  Y  *  N  T
541   gctggctaacaatcggaatttcaggtccagataatggagcagtggctgtattaaaatataca  600
      T  A  *  *  L  K  P  *  K  V  G  G  R  K  Y  *  G  H  K  S
601   acggcataataactgaaaccataaaaagttggaggaagaaaatattgaggacacaagagt  660
      L  N  V  P  V  *  M  V  H  V  L  L  *  *  L  M  A  R  V  M
661   ctgaatgtgcctgtgtaaatggttcatgttttactataatgactgatggcccgagtgatg  720
      G  W  P  R  T  K  F  S  R  S  K  R  G  R  L  L  N  Q  *  S
721   ggctggcctcgtacaaaattttcaagatcgaaaaggggaaggttactaaatcaatagagt  780
      *  M  H  L  I  L  T  M  R  N  V  P  V  T  L  I  P  A  K  *
781   tgaatgcacctaattctcactatgaggaatgttcctgttaccctgataccggcaaagtga  840
      C  V  C  A  E  T  I  G  M  V  R  T  G  H  G  C  L  S  I  K
841   tgtgtgtgtgcagagacaattggcatggttcgaaccggccatgggtgtctttcgatcaaa  900
      T  W  I  I  K  *  D  T  S  A  V  G  F  S  V  T  T  R  V  P
901   acctggattatcaaataggatacatctgcagtgggggtttcggtgacaacccgcgtccca  960
      K  M  E  Q  A  A  V  V  Q  C  M  L  M  E  Q  T  E  *  R  D
961   aagatggaacaggcagctgtggtccagtgtatgttgatggagcaaacggagtaaagggat  1020
      F  H  I  G  M  V  M  V  F  G  *  E  G  P  K  V  T  V  P  D
1021  tttcatataggtatggtaatggtgtttggataggaaggaccaaaagtcacagttccagac  1080
      M  G  L  R  *  F  G  I  L  M  D  G  Q  R  L  I  V  S  S  L
1081  atgggtttgagatgatttgggatcctaatggatggacagagactgatagtaagttctctg  1140
      *  G  K  M  L  W  Q  *  L  I  G  Q  G  I  A  G  V  S  F  N
1141  tgaggcaagatgttgtggcaatgactgattggtcagggtatagcgggagtttcgttcaac  1200
      I  L  S  *  Q  G  *  T  V  *  G  R  A  S  G  L  N  *  S  G
1201  atcctgagctaacagggctagactgtataaggccgtgcttctgggttgaattaatcaggg  1260
      D  D  L  K  K  K  Q  S  G  L  V  R  A  A  F  L  F  V  A  *
1261  gacgacctaaagaaaaaacaatctggactagtgcgagcagcatttctttttgtggcgtga  1320
      I  V  I  L  *  I  G  L  G  Q  T  V  L  S  C  H  S  P  L  T
1321  atagtgatactgtagattggtcttggccagacggtgctgagttgccattcaccattgaca  1380
      S  S  L  F  K  K  L  L  V  S  T
1381  agtagtctgttcaaaaaactccttgtttctact  1413
```

Translate to amino acid sequence.

Any nucleic acid sequence could be read In 6 reading frames (3 per complementary Nucleic acid strand).

All are done – this shown is reading frame number 1.

Look for the one that has an M (methionine) near the start, and would have the longest ORF (open reading frame). i.e that would not be cut short by a stop codon (blue)

Frame 2, 40 stop codons

Influenza A virus (A/Puerto Rico/8/34(H1N1)) segment 6, complete Translatec

```
-303  ccattaggatcccaaatcatctcaaacccatgtctggaactgtgacttttggtccttcct  -362
      I  Q  T  P  L  P  Y  L  Y  E  N  P  F  T  P  F  A  P  S  T
-363  atccaaacaccattaccatacctatatgaaaatccctttactccgtttgctccatcaaca  -422
      Y  T  G  P  Q  L  P  V  P  S  L  G  R  G  L  S  P  K  T  P
-423  tacactggaccacagctgcctgttccatctttgggacgcgggttgtcaccgaaaacccca  -482
      L  Q  M  Y  P  I  *  *  S  R  F  *  S  K  D  T  H  G  R  F
-483  ctgcagatgtatcctatttgataatccaggttttgatcgaaagacacccatggccggttc  -542
      E  P  C  Q  L  S  L  H  T  H  I  T  L  P  V  S  G  *  Q  E
-543  gaaccatgccaattgtctctgcacacacacatcactttgccggtatcagggtaacaggaa  -602
      H  S  S  *  *  E  L  G  A  F  N  S  I  D  L  V  T  F  P  F
-603  cattcctcatagtgagaattaggtgcattcaactctattgatttagtaaccttccccttt  -662
      S  I  L  K  I  L  Y  E  A  S  P  S  L  G  P  S  V  I  I  V
-663  tcgatcttgaaaatttttgtacgaggccagcccatcactcgggccatcagtcattatagta  -722
      K  H  E  P  F  T  Q  A  H  S  D  S  C  V  L  N  I  F  F  L
-723  aaacatgaaccatttacacaggcacattcagactcttgtgtcctcaatattttcttcctc  -782
      Q  L  F  M  V  S  V  I  M  P  L  Y  F  N  T  A  T  A  P  L
-783  caacttttatggtttcagttattatgccgttgtattttaatacagccactgctccatta  -842
      S  G  P  E  I  P  I  V  S  Q  P  M  P  S  *  H  A  L  A  D
-843  tctggacctgaaattccgattgttagccagcccatgccatcatgacatgcacttgctgac  -902
      Q  A  T  D  S  N  L  E  L  Y  G  D  G  A  S  P  T  G  Q  L
-903  caagcaaccgattcaaatcttgaattgtacggggacggagcttcaccgacagggcagctc  -962
      I  K  A  L  *  G  L  L  S  L  T  V  P  F  E  C  L  S  F  S
-963  attaaggccctataagggcttctgtccttaacagtcccatttgaatgcctgtcattcagt  -1022
       K  A  P  W  V  R  K  K  V  L  H  S  K  *  E  H  E  I  K  G
-1023  aaggcaccttgggtcagaaaaaaggtcctgcattccaagtgagaacatgaaataaagggc  -1082
       S  L  M  T  K  T  S  P  L  E  P  I  L  M  L  L  S  L  L  Y
-1083  tctcttatgacaaaaacgtctccttggaaccaattcttatgctattgtctttgctgtat  -1142
       I  A  H  P  R  M  G  Q  R  D  E  L  P  V  N  I  T  E  V  V
-1143  atagcccacccacggatgggacaaagagatgaattgccggttaatatcactgaagttgtg  -1202
       S  F  T  Q  V  L  F  L  *  V  M  M  F  W  L  H  I  P  V  W
-1203  tcctttacccaggtgctatttttataggtaatgatgttttggttgcatattccagtatgg  -1262
       F  *  L  P  V  *  I  E  W  L  I  H  I  E  I  I  F  P  I  C
-1263  ttttgacttccagtttgaattgaatggctaatccatattgagattatattccctatttgc  -1322
       N  I  R  L  I  S  P  T  T  R  Q  I  D  P  M  V  I  I  F  *
-1323  aatattaggctaattagtccgactaccagacagattgatccaatggttattattttctga  -1382
       F  G  F  I  L  N  P  C  F  R
-1383  tttggattcattttaaaacccctgctttcgct  -1413
```

Actually the Workbench does that for you. The protein sequence from reading frame 3 is best.

Frame 3 [Longest ORF], 0 stop codons

☐ Influenza A virus (A/Puerto Rico/8/34(H1N1)) segment 6, complete Translatec

Now tick this box

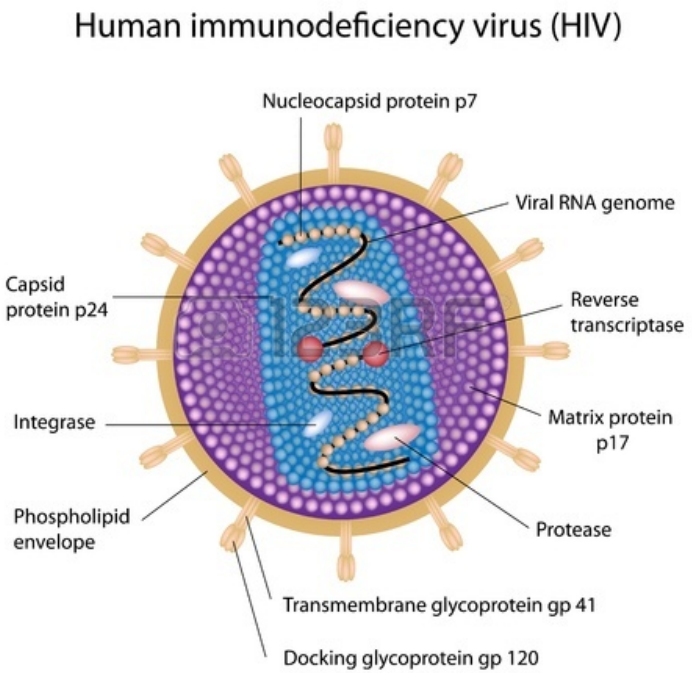
```
>8486127 Translated - Longest ORF [Frame 3]
MNPNQKIITIGSICLVVGLISLILQIGNIISIWISHSIQTGSQNHTGICNQNIITYKNSTWVKDTTSVIL
TGNSSLCPIRGWAIYSKDNSIRIGSKGDVFVIREPFISCSHLECRTFFLTQGALLNDRHSNGTVKDRSPY
RALMSCPVGEAPSPYNSRFESVAWSASACHDGMGWLTIGISGPDNGAVAVLKYNGIITETIKSWRKKILR
TQESECACVNGSCFTIMTDGPSDGLASYKIFKIEKGKVTKSIELNAPNSHYEECSCYPDTGKVMCVCRDN
```

Import Sequence(s) | Return | Help | Report Bugs

SDSC

and click on “Import Sequences”

Now that you have the amino acid sequence, use an old trick …



- Compare the *protein sequences* that are implied by two stains of the virus.
- Differences tell you what segments are likely to be at the surface of the protein and what is like to be buried (as for a B- and a T-epitope respectively), because evolution is fast in surface loops.
- They tell you what is fairly conserved between strains of a pathogen and what varies considerably, an important consideration for diagnostics and vaccines.
- Was quickly applied when the second sequence of an AIDS virus appeared
  - Robson, B., Fishliegh, R. V., and Morrison, C. A. (1987) "Prediction of HIV Vaccine", Nature, 325, 395
  - But changes of regulations (as to testing) delayed development, and entirely new strains, initially HIV2, came along before a single vaccine could arrest the spread.

☐	SWISSPROT	NRAM_IABAN	(P06818) Neuraminidase (EC 3.2.1.18) [Influ...	357	7e-98
☐	SWISSPROT	NRAM_IAKOR	(Q6XUA7) Neuraminidase (EC 3.2.1.18) [Influ...	356	1e-97
☐	SWISSPROT	NRAM_IATOK	(P06820) Neuraminidase (EC 3.2.1.18) [Influ...	355	2e-97
☐	SWISSPROT	NRAM_IAMEB	(Q2VNF0) Neuraminidase (EC 3.2.1.18) [Influ...	355	2e-97
☐	SWISSPROT	NRAM_IAHO1	(Q91MA2) Neuraminidase (EC 3.2.1.18) [Influ...	355	2e-97
☐	SWISSPROT	NRAM_IAAIC	(Q75VQ4) Neuraminidase (EC 3.2.1.18) [Influ...	355	2e-97
☐	SWISSPROT	NRAM_IAZH0	(Q09104) Neuraminidase (EC 3.2.1.18) [Influ...	355	2e-97
☐	SWISSPROT	NRAM_IAMEE	(Q2VND0) Neuraminidase (EC 3.2.1.18) [Influ...	355	3e-97
☐	SWISSPROT	NRAM_IANT6	(P03473) Neuraminidase (EC 3.2.1.18) [Influ...	353	6e-97
☐	SWISSPROT	NRAM_IAEN6	(Q6XTN2) Neuraminidase (EC 3.2.1.18) [Influ...	353	8e-97
☐	SWISSPROT	NRAM_IAME1	(P03471) Neuraminidase (EC 3.2.1.18) [Influ...	353	1e-96
☐	SWISSPROT	NRAM_IAZKA	(Q09106) Neuraminidase (EC 3.2.1.18) [Influ...	350	5e-96
☐	SWISSPROT	NRAM_IAQU7	(Q6XTM7) Neuraminidase (EC 3.2.1.18) [Influ...	350	5e-96
☐	SWISSPROT	NRAM_IAZH4	(Q67344) Neuraminidase (EC 3.2.1.18) (Fragm...	329	1e-89
☐	SWISSPROT	NRAM_INBHK	(P16191) Neuraminidase (EC 3.2.1.18) [Influ...	223	1e-57
☐	SWISSPROT	NRAM_INBOR	(P16201) Neuraminidase (EC 3.2.1.18) [Influ...	223	1e-57
☐	SWISSPROT	NRAM_INBSI	(P16203) Neuraminidase (EC 3.2.1.18) [Influ...	221	4e-57
☑	SWISSPROT	NRAM_INBMD	(P16195) Neuraminidase (EC 3.2.1.18) [Influ...	221	6e-57
☑	SWISSPROT	NRAM_INBYB	(Q90021) Neuraminidase (EC 3.2.1.18) [Influ...	220	1e-56

Use BLAST to find similar sequences.

Select bottom 2 as significant variants to compare

Select sequences and then press 'Import Sequence(s)' to import them to the workbench, or press 'Show Record(s)' to see the database records.

Show Record(s) | Show Sequence(s) | Import Sequence(s) | Return | Help | Report Bugs

>SWISSPROT:NRAM_IAPUE (P03468) Neuraminidase (EC 3.2.1.18)
[Influenza A virus (strain A/Puerto Rico/8/1934 H1N1)]
Length = 454

Score = 960 bits (2481), Expect = 0.0
Identities = 450/454 (99%), Positives = 454/454 (100%)

```
Query: 1   MNPNQKIITIGSICLVVGLISLILQIGNIISIWISHSIQTGSQNHTGICNQNIITYKNST 60
           MNPNQKIITIGSICLVVGLISLILQIGNIISIWISHSIQTGSQNHTGICNQNIITYKNST
Sbjct: 1   MNPNQKIITIGSICLVVGLISLILQIGNIISIWISHSIQTGSQNHTGICNQNIITYKNST 60

Query: 61  WVKDTTSVILTGNSSLCPIRGWAIYSKDNSIRIGSKGDVFVIREPFISCSHLECRTFFLT 120
           WVKDTTSVILTGNSSLCPIRGWAIYSKDNSIRIGSKGDVFVIREPFISCSHLECRTFFLT
Sbjct: 61  WVKDTTSVILTGNSSLCPIRGWAIYSKDNSIRIGSKGDVFVIREPFISCSHLECRTFFLT 120

Query: 121 QGALLNDRHSNGTVKDRSPYRALMSCPVGEAPSPYNSRFESVAWSASACHDGMGWLTIGI 180
           QGALLND+HSNGTVKDRSPYRALMSCPVGEAPSPYNSRFESVAWSASACHDGMGWLTIGI
Sbjct: 121 QGALLNDKHSNGTVKDRSPYRALMSCPVGEAPSPYNSRFESVAWSASACHDGMGWLTIGI 180

Query: 181 SGPDNGAVAVLKYNGIITETIKSWRKKILRTQESECACVNGSCFTIMTDGPSDGLASYKI 240
```

Then click on “Import Sequences”

Sequence alignment

```
Consensus key (see documentation for details)
* - single, fully conserved residue
: - conservation of strong groups
. - conservation of weak groups
  - no consensus


CLUSTAL W (1.81) multiple sequence alignment


NRAM_INBMD                      MLPS--TIQTLTLFLTSGGVLLSLYVSASLSYLLYSDILLKFSPTKRTAP
NRAM_INBYB                      MLPS--TIQTLTLFLTSGGVLLSLYVSASLSYLLYSDILLKFSPTEITAP
8486127_Translated_-_Longest    MNPNQKIITIGSICLVVGLISLILQIGNIISIWISHSIQTGSQNHTGICN
                                * *.   *   :: *. * : * * :.  :*  :  .*    .     .

NRAM_INBMD                      TMSLECVNVSNAQAVNHSATKEMTFLLPEPEWTYPRLSCQGSTFQKALLI
NRAM_INBYB                      KVPLDCANASNVQAVNRSATKGMTLLLSEPEWTYPRLSCQGSTFQKALLI
8486127_Translated_-_Longest    QNIITYKNSTWVKDTTSVILTGNSSLCPIRGWAIYSKDNS----------
                                   :   * : .: ..    .  : * .   *:    . .

NRAM_INBMD                      SPHRFGETRGNSAPLIIREPFVACGPKECRHFALTHYAAQPGGYYNGTRK
NRAM_INBYB                      SPHRFGESRGNSAPLIIREPFIACGPKECKHFALTHYAAQPGGYYNGTRE
8486127_Translated_-_Longest    --IRIG---SKGDVFVIREPFISCSHLECRTFFLTQGALLNDRHSNGTVK
                                  *:*   .:.  ::*****::*.  **: * **: *   . : *** :

NRAM_INBMD                      DRNKLRHLISVKLGKIPTVENSIFHMAAWSGSACHDGREWTYIGVDGPDS
NRAM_INBYB                      DRNKLRHLISVKLGKIPTVENSIFHMAAWSGSACHDGREWTYIGVDGPDS
8486127_Translated_-_Longest    DRSPYRALMSCPVGEAPSPYNSRFESVAWSASACHDGMGWLTIGISGPDN
                                **.  * *:*  :*: *:  ** *. .***.******  *  **:.***.

NRAM_INBMD                      DALIKIKYGEAYTDTYHSYAHNILRTQESACNCIGGDCYLMITDGSASGI
NRAM_INBYB                      NALIKIKYGEAYTDTYHSYANNILRTQESACNCIGGDCYLMITDGSASGI
8486127_Translated_-_Longest    GAVAVLKYNGIITETIKSWRKKILRTQESECACVNGSCFTIMTDGPSDGL
                                .*:  :**.   *:* :*: ::******* * *:.*.*: ::***.:.*:

NRAM_INBMD                      SKCRFLKIREGRIIKEIFPAGRVEHTEECTCGFASNKTIECACRDNSYTA
NRAM_INBYB                      SKCRFLKIREGRIIKEIFPTGRVEHTEECTCGFASNKTIECACRDNSYTA
8486127_Translated_-_Longest    ASYKIFKIEKGKVTKSIELNAPNSHYEECSC-YPDTGKVMCVCRDNWHGS
                                :. :::**.:*:: *.*   .  .* ***:* :... .: *.**** : :

NRAM_INBMD                      KRPFVKLNVETDTAEIRLMCTETYLDTPRPDDGSITGPCESNGDKGLGGI
NRAM_INBYB                      KRPFVKLNVETDTAEIRLMCTETYLDTPRPDDGSITGPCESNGDKGRGGI
8486127_Translated_-_Longest    NRPWVSFDQNLD-YQIGYICSGVFGDNPRPKDG--TGSCGPVYVDGANGV
                                :**:*.:: : *  :*  :*: .: *.***.**  **.* .   .* .*:

NRAM_INBMD                      KGGFVHQRMASKIGRWYSRTMSKTERMGMELYVKYDGDPWTDSDALAPSG
NRAM_INBYB                      KGGFVHQRMASKIGRWYSRTMSKTERMGMELYVKYDGDPWTDSDALAPSG
8486127_Translated_-_Longest    KG-FSYRYGN---GVWIGRTKSHSSRHGFEMIWDPNGWTETDSKFSVRQD
                                ** * ::      * * .** *::.* *:*:  . :* . ***.  . ..

NRAM_INBMD                      VMVSIKEP---GWYSFGFEIKDKKCDVPCIGIEMVHDGGKE--TWHSAAT
NRAM_INBYB                      VMVSMKEP---GWYSFGFEIKDKKCDVPCIGIEMVHDGGKK--TWHSAAT
8486127_Translated_-_Longest    VVAMTDWSGYSGSFVQHPELTGLDCIRPCFWVELIRGRPKEKTIWTSASS
                                *:.  . .   * :    *:.. .*  **: :*:::.  *:   * **::

NRAM_INBMD                      AIYCLMGSGQLLWDTVTGVDMAL----
NRAM_INBYB                      AIYCLMGSGQLLWDTVTGVDMAL----
8486127_Translated_-_Longest    ISFCGVNSDTVDWSWPDGAELPFTIDK
                                  :* :.*. : *.   *.::.:
```

Now use bioinformatics utility CLUSTALW to align sequences.

Black shows highly variable Regions, probably meaning surface Loops.

Influenza A virus (A/Puerto Rico/8/34(H1N1)) segment 6, complete Translated - Longest ORF [Frame 3]

```
>8486127 Translated - Longest ORF [Frame 3]
MNPNQKIITIGSICLVVGLISLILQIGNIISIWISHSIQTGSQNHTGICN
QNIITYKNSTWVKDTTSVILTGNSSLCPIRGWAIYSKDNSIRIGSKGDVF
VIREPFISCSHLECRTFFLTQGALLNDRHSNGTVKDRSPYRALMSCPVGE
APSPYNSRFESVAWSASACHDGMGWLTIGISGPDNGAVAVLKYNGIITET
IKSWRKKILRTQESECACVNGSCFTIMTDGPSDGLASYKIFKIEKGKVTK
SIELNAPNSHYEECSCYPDTGKVMCVCRDNWHGSNRPWVSFDQNLDYQIG
YICSGVFGDNPRPKDGTGSCGPVYVDGANGVKGFSYRYGNGVWIGRTKSH
SSRHGFEMIWDPNGWTETDSKFSVRQDVVAMTDWSGYSGSFVQHPELTGL
DCIRPCFWVELIRGRPKEKTIWTSASSISFCGVNSDTVDWSWPDGAELPF
TIDK
```

LEGEND:

Alpha Helix = H Beta Sheet = E Random Coil = C

[Download Unformatted Results]

Neuraminidase (EC 3.2.1.18) [Influenza B virus (strain B/Yamagata/16/1988)]

```
>NRAM_INBYB
MLPSTIQTLTLFLTSGGVLLSLYVSASLSYLLYSDILLKFSPTEITAPKV
PLDCANASNVQAVNRSATKGMTLLLSEPEWTYPRLSCQGSTFQKALLISP
HRFGESRGNSAPLIIREPFIACGPKECKHFALTHYAAQPGGYYNGTREDR
NKLRHLISVKLGKIPTVENSIFHMAAWSGSACHDGREWTYIGVDGPDSNA
LIKIKYGEAYTDTYHSYANNILRTQESACNCIGGDCYLMITDGSASGISK
CRFLKIREGRIIKEIFPTGRVEHTEECTCGFASNKTIECACRDNSYTAKR
PFVKLNVETDTAEIRLMCTETYLDTPRPDDGSITGPCESNGDKGRGGIKG
GFVHQRMASKIGRWYSRTMSKTERMGMELYVKYDGDPWTDSDALAPSGVM
VSMKEPGWYSFGFEIKDKKCDVPCIGIEMVHDGGKKTWHSAATAIYCLMG
SGQLLWDTVTGVDMAL
```

LEGEND:

Alpha Helix = H Beta Sheet = E Random Coil = C

[Download Unformatted Results]

Neuraminidase (EC 3.2.1.18) [Influenza B virus (strain B/Maryland/1959)]

```
>NRAM_INBMD
MLPSTIQTLTLFLTSGGVLLSLYVSASLSYLLYSDILLKFSPTKRTAPTM
SLECVNVSNAQAVNHSATKEMTFLLPEPEWTYPRLSCQGSTFQKALLISP
HRFGETRGNSAPLIIREPFVACGPKECRHFALTHYAAQPGGYYNGTRKDR
NKLRHLISVKLGKIPTVENSIFHMAAWSGSACHDGREWTYIGVDGPDSDA
LIKIKYGEAYTDTYHSYAHNILRTQESACNCIGGDCYLMITDGSASGISK
CRFLKIREGRIIKEIFPAGRVEHTEECTCGFASNKTIECACRDNSYTAKR
PFVKLNVETDTAEIRLMCTETYLDTPRPDDGSITGPCESNGDKGLGGIKG
GFVHQRMASKIGRWYSRTMSKTERMGMELYVKYDGDPWTDSDALAPSGVM
VSIKEPGWYSFGFEIKDKKCDVPCIGIEMVHDGGKETWHSAATAIYCLMG
SGQLLWDTVTGVDMAL
```

PREDICT LOOPS DIRECTLY.

Use GOR (Garnier-Osguthorpe-Robson) method.

Predict secondary structure of the 3 Sequences, including predictions of loops.

Red alpha–helix, blue=beta sheet, Black – loop or “coil”

When examined, click on “Return”

Influenza A virus (A/Puerto Rico/8/34(H1N1)) segment 6, complete Translated - Longest ORF [Frame 3]

Combining prediction methods.

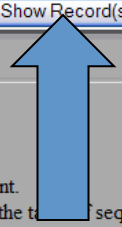
Note approximate but not exact agreement.

Issue of which to use.

JOI is a “joint” or the combined prediction. It tends to make secondary structure too short.

Select sequences and then press 'Import Sequence(s)' to import them to the workbench, or press 'Show Record(s)' to see the database records.

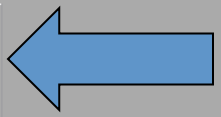
Show Record(s) | Show Sequence(s) | Import Sequence(s) | Return | Help | Report Bugs



Sequences producing significant alignments:

Click on the score value to view the corresponding alignment.
Click on the sequence key in the alignments to get back to the t sequences.

Select	Database	ID	Name	Score	Evalue
☑	PDBSEQRES	1NSD_B	HYDROLASE(O-GLYCOSYL)| (E.C. 3.2.1.18),Neuramin...	783	0.0
☐	PDBSEQRES	1NSD_A	HYDROLASE(O-GLYCOSYL)| (E.C. 3.2.1.18),Neuramin...	783	0.0
☐	PDBSEQRES	1NSC_B	HYDROLASE(O-GLYCOSYL)| (E.C. 3.2.1.18),Neuramin...	783	0.0
☐	PDBSEQRES	1NSC_A	HYDROLASE(O-GLYCOSYL)| (E.C. 3.2.1.18),Neuramin...	783	0.0
☐	PDBSEQRES	1NSB_B	HYDROLASE(O-GLYCOSYL)| (E.C. 3.2.1.18),Neuramin...	783	0.0
☐	PDBSEQRES	1NSB_A	HYDROLASE(O-GLYCOSYL)| (E.C. 3.2.1.18),Neuramin...	783	0.0
☐	PDBSEQRES	1A4Q_B	HYDROLASE| (E.C. 3.2.1.18),biological_unit: tet...	783	0.0
☐	PDBSEQRES	1A4Q_A	HYDROLASE| (E.C. 3.2.1.18),biological_unit: tet...	783	0.0
☐	PDBSEQRES	1A4G_B	HYDROLASE| (E.C. 3.2.1.18),biological_unit: tet...	783	0.0
☐	PDBSEQRES	1A4G_A	HYDROLASE| (E.C. 3.2.1.18),biological_unit: tet...	783	0.0
☐	PDBSEQRES	1INF__	HYDROLASE (O-GLYCOSYL)| (E.C. 3.2.1.18),(sialid...	778	0.0
☐	PDBSEQRES	1IVB__	HYDROLASE (O-GLYCOSYL)| (E.C. 3.2.1.18),Influen...	777	0.0
☐	PDBSEQRES	1INV__	HYDROLASE (O-GLYCOSYL)| (E.C. 3.2.1.18),Influen...	777	0.0
☐	PDBSEQRES	1B9V_A	HYDROLASE| (E.C. 3.2.1.18),(sialidase)	777	0.0
☐	PDBSEQRES	1B9T_A	HYDROLASE| (E.C. 3.2.1.18),(sialidase)	777	0.0
☐	PDBSEQRES	1B9S_A	HYDROLASE| (E.C. 3.2.1.18),(sialidase)	777	0.0
☐	PDBSEQRES	1VCJ_A	HYDROLASE| (E.C. 3.2.1.18),fragment: catalytic ...	775	0.0
☐	PDBSEQRES	2HU4_H	HYDROLASE|Mutant	221	1e-57
☐	PDBSEQRES	2HU4_G	HYDROLASE|Mutant	221	1e-57
☐	PDBSEQRES	2HU4_F	HYDROLASE|Mutant	221	1e-57
☐	PDBSEQRES	2HU4_E	HYDROLASE|Mutant	221	1e-57
☐	PDBSEQRES	2HU4_D	HYDROLASE|Mutant	221	1e-57
☐	PDBSEQRES	2HU4_C	HYDROLASE|Mutant	221	1e-57
☐	PDBSEQRES	2HU4_B	HYDROLASE|Mutant	221	1e-57
☐	PDBSEQRES	2HU4_A	HYDROLASE|Mutant	221	1e-57



Hunt for similar proteins of known 3D structure.

Select and tick a protein with zero Difference Evalue.

Then click on “Show Record”

- Home
- Getting Started
- Download Files
- Deposit and Validate
- Structural Genomics
- Dictionaries & File Formats
- Software Tools
- General Education
- Site Tutorials
- BioSync
- General Information
- Acknowledgements
- Frequently Asked Questions
- Report Bugs/Comments

Quick Tips :

Click here to see structures released this week.

ALERT: Our data files are changing soon. Please see http://www.wwpdb.org for more details.

Help | Structure Summary | Biology & Chemistry | Materials & Methods | Sequence Details | Geometry

1nsd

DOI 10.2210/pdb1nsd/pdb

Red - Derived Information

Title	INFLUENZA B VIRUS NEURAMINIDASE CAN SYNTHESIZE ITS OWN INHIBITOR
Authors	Burmeister, W.P., Ruigrok, R.W.H., Cusack, S.
Primary Citation	**Burmeister, W.P., Henrissat, B., Bosso, C., Cusack, S., Ruigrok, R.W.** Influenza B virus neuraminidase can synthesize its own inhibitor. *Structure v1 pp.19-26 , 1993* [Abstract]
History	Deposition 1993-05-24 Release 1993-10-31
Experimental Method	Type X-RAY DIFFRACTION Data N/A

Parameters	Resolution[Å]	R-Value	R-Free	Space Group
	1.80	0.169 (obs.)	n/a	P 3_1 2 1

Unit Cell						
Length [Å]	a	88.90	b	88.90	c	222.80
Angles [°]	alpha	90.00	beta	90.00	gamma	120.00

Molecular Description Asymmetric Unit

Polymer: 1 Molecule: NEURAMINIDASE Chains: A,B EC no.: 3.2.1.18

Classification Hydrolase(o Glycosyl)

Source

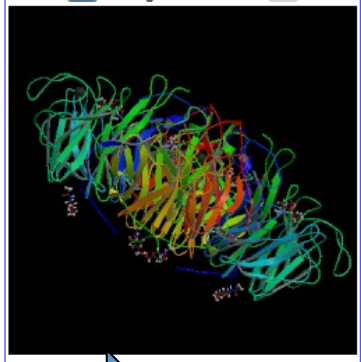
Polymer: 1 Scientific Name: Synthetic construct

Chemical Component

Identifier	Name	Formula	Drug Similarity	Hapten Similarity	Ligand Structure	Ligand Interaction
NAG	N-ACETYL-D-GLUCOSAMINE	$C_8H_{15}NO_6$			[View]	[View]

Images and Visualization

<< Biological Molecule >>



Display Options

KiNG
Jmol
WebMol
MBT SimpleViewer*
MBT Protein Workshop
QuickPDB
All Images

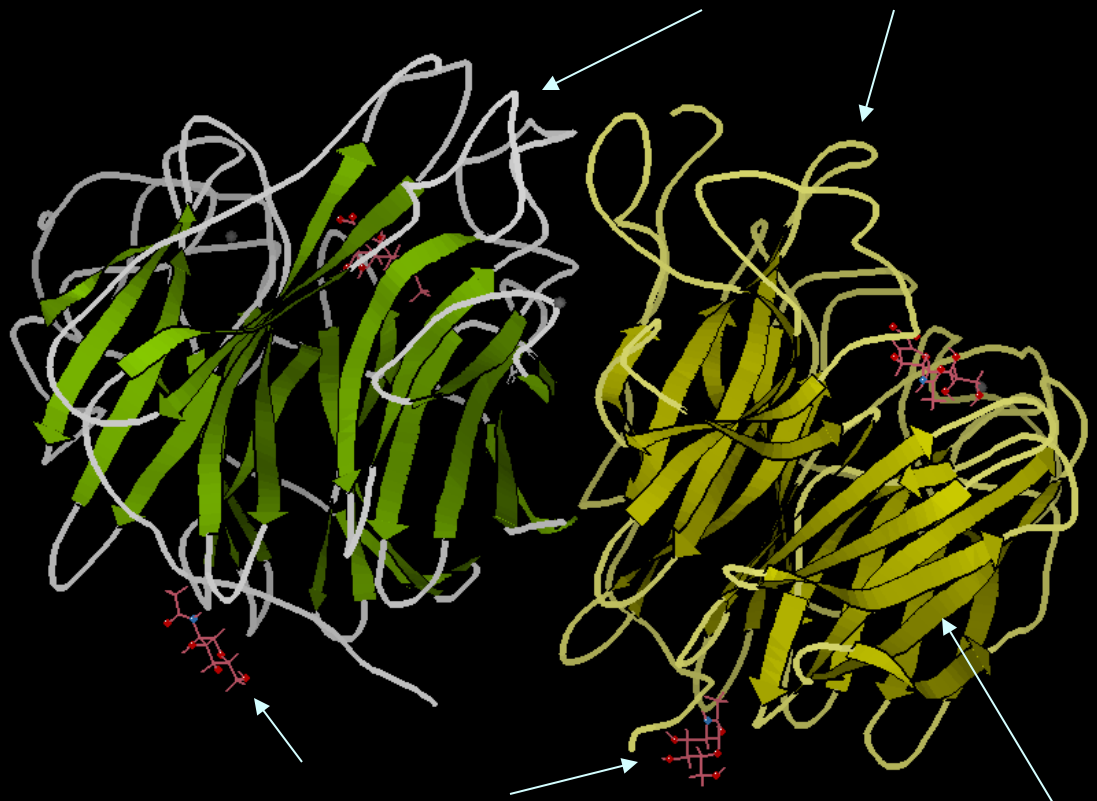
* Capable of displaying biological molecules.

Select a viewer program that suits you, say KiNG

KiNG 1.39

File Edit Views Display Tools Help

Surface loops accept mutations easily?
(because the fewer internal interactions do not mess up structure?)



Notice bound sugars

Core is mainly pleated sheet

Kinemage #1

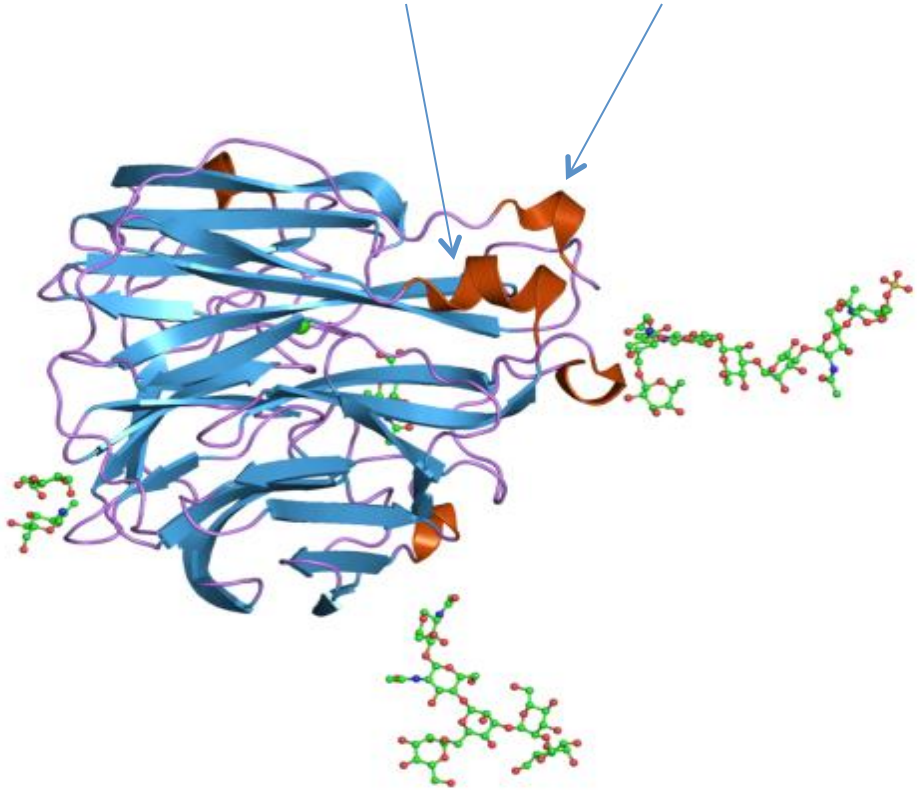
- 1NSDa
- 1NSDb
- 1NSD
- hets
- atoms
- ribbon
- coil
- beta

Zoom
Clipping
Pick center
Markers
Show text
Show hierarchy

start 54% 12:08 PM

The elements of a-helix are somewhat distorted and show up in that “viewer” as coil/loop, which is unfortunate.

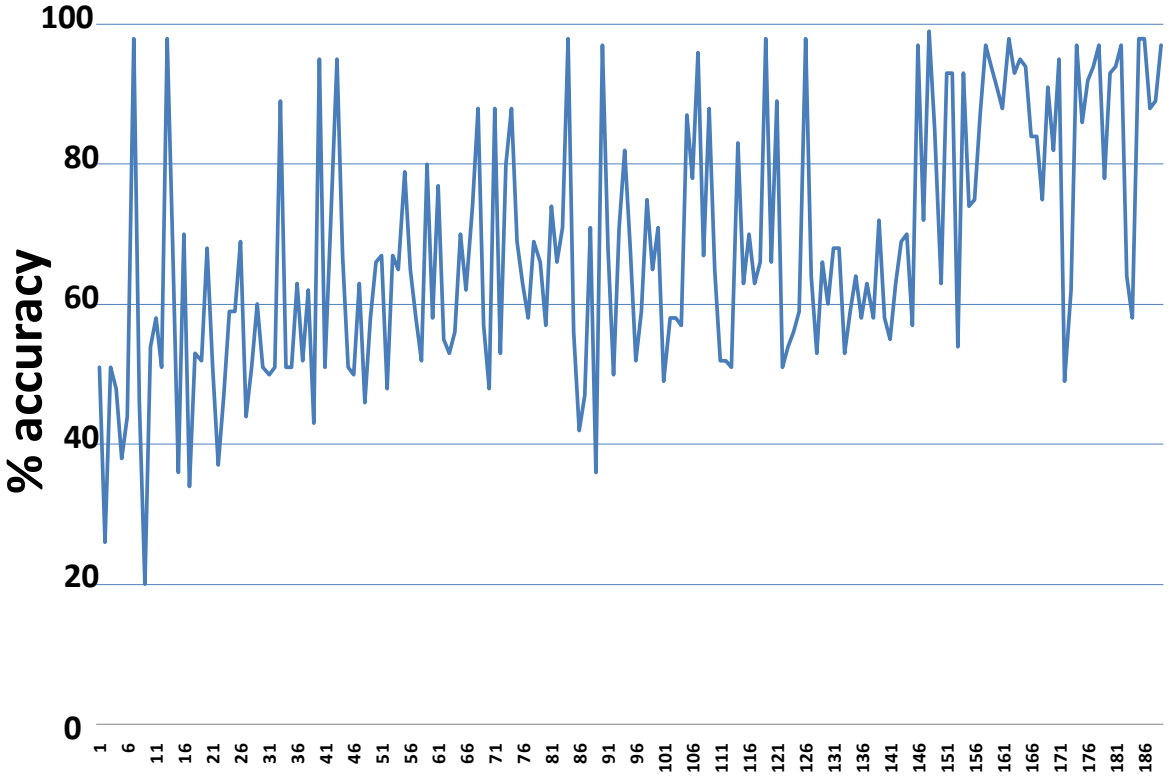
They are clearer in this representation…



Learning to Predict Protein Structure from probabilistic Semantic Web rules about amino acid sequence and 3D structure relationships

Increase in accuracy with size of data set.

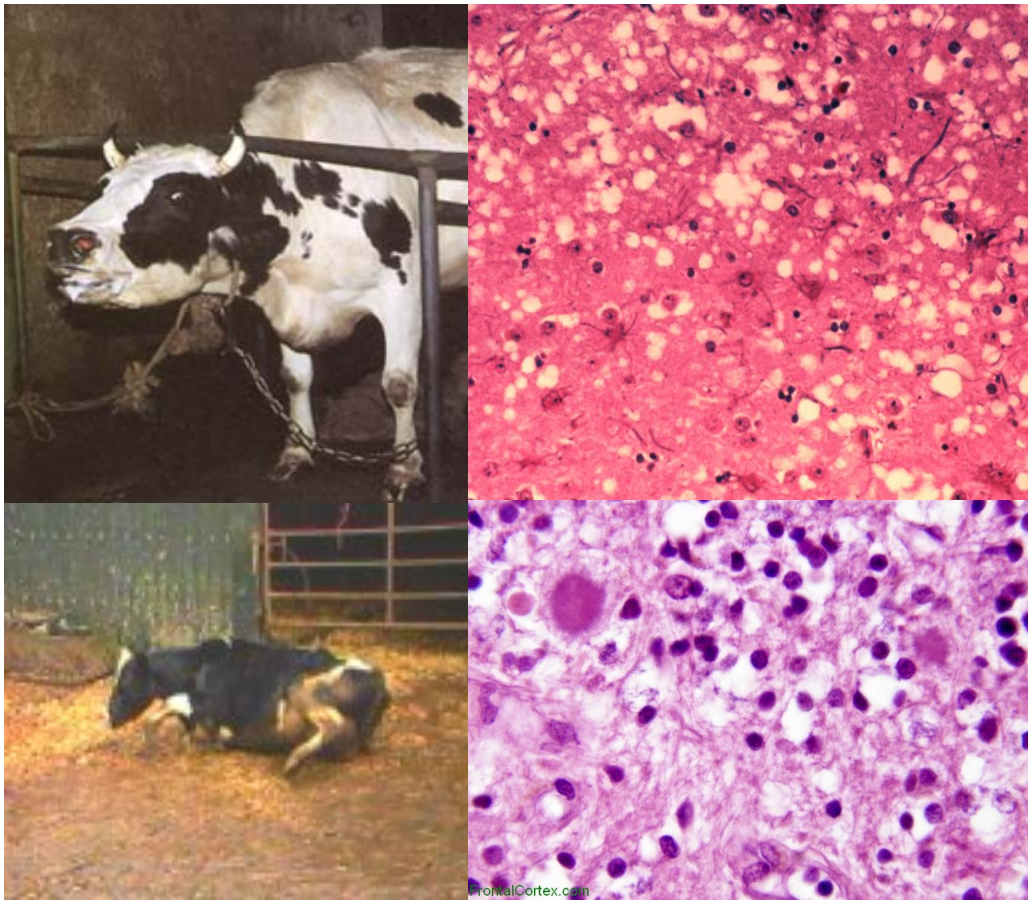
Each predicted protein is not included in the data set, *but is picked arbitrarily without regard to homology* with others except for ignoring those likely to be homologous by name.



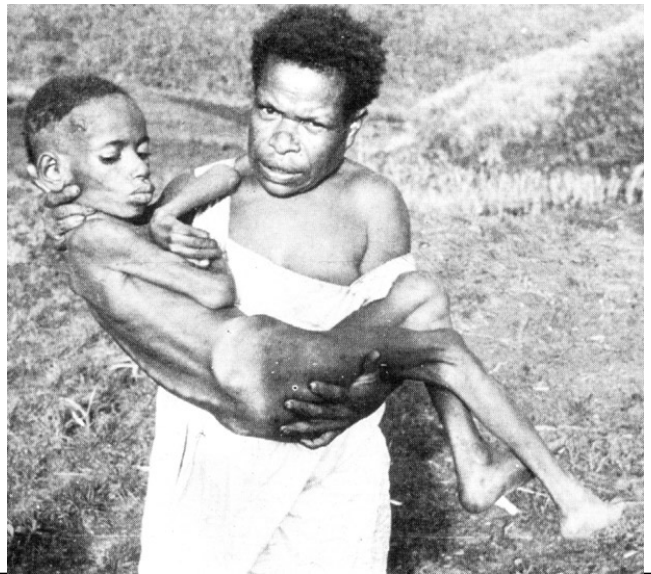
Robson, B. (2014) " Hyperbolic Dirac Nets for Medical Decision Support. Theory, Methods, and Comparison with Bayes Nets" *Computers in Biology and Medicine*, in 2014 Aug;51:183-97.

Proof of Concept: Earlier Similar Work on Bovine Spongiform Encephalopathy Diagnostic

(a tough one – we have to distinguish the malign form from the same natural protein in the brain/meat)



Prion Diseases

Infected species	Disease
sheep, goat	Scrapie
cattle	Bovine spongiform encephalopathy (BSE), mad cow disease
mink	Transmissible mink encephalopathy (TME)
white-tailed deer, elk, mule deer, moose	Chronic wasting disease (CWD)
cat	Feline spongiform encephalopathy (FSE)
nyala, oryx, greater kudu	Exotic ungulate encephalopathy (EUE)
ostrich	Spongiform encephalopathy (Has not been shown to be transmissible.)
human 	Creutzfeldt–Jakob disease (CJD)
	Iatrogenic Creutzfeldt–Jakob disease (iCJD)
	Variant Creutzfeldt–Jakob disease (vCJD)
	Familial Creutzfeldt–Jakob disease (fCJD)
	Sporadic Creutzfeldt–Jakob disease (sCJD)
	Gerstmann–Sträussler–Scheinker syndrome (GSS)[
	Fatal familial insomnia (FFI)
	Kuru

PROMETHEUS EXPERT SYSTEM

- In 1987, the Internet was just being born. Around then CERN had just begun installation and operation of TCP/IP to interconnect its major internal computer systems, workstations, PCs and an accelerator control system.

- PROMETHEUS by another form and name was the conceptual descendant of a prototype Expert System at the University of Manchester that used communication between computers to share knowledge and design biotechnological compounds.

  - "Proteus was formed in 1987 on the basis of the value of Prometheus and its potential in drug discovery applications." (It went to the London Stock Exchange as Proteus International plc in 1990)

  - "Initially it was intended that Prometheus be sold as software to other companies but it was soon decided that it made more commercial sense to keep Prometheus as a proprietary piece of software and to apply the software to the discovery of drugs and vaccines, for exploitation with major pharmaceutical companies."

  - **"Proteus has used the PROMETHEUS software to identify regions on PrP which can be used to raise antibodies that, as part of an overall diagnostic test protocol, can differentiate between normal PrP and the abnormally folded form ...."**

An Early Role of PROMETHEUS in BSE Science

- Collaborators were doubting that prion protein was the cause of Scrapies (sheep) and BSE (cows).
- It seemed unlikely that a protein without nucleic acid could be a pathogen.
- Knock-out mice without prion protein seemed to be fine…..except for a few hints of learning difficulties (and rather like an Alzheimer's patient in a nursing home!). So the protein seemed to serve no useful function vital to life, suggesting that it was somewhat unimportant.
- PROMETHUES predicted that the level of prion protein conservation between species was high enough that it had an important function.
  - It modeled a system of communication between prion with a dual receptor and transmitter ligand role in its lifetime that caused synapses to make contact between nerve cells involved in the learning process.
  - It suggested how this could go wrong, leading to the disease state.

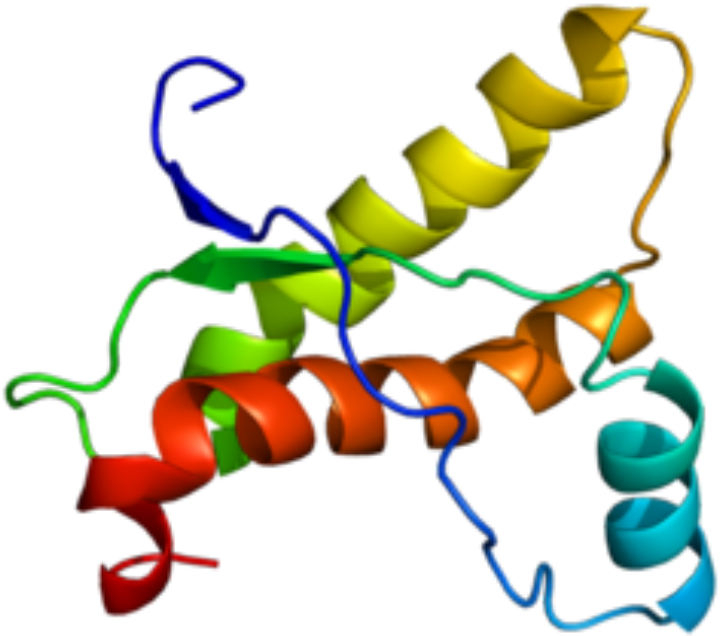
Natural Endogenous Cellular Prion

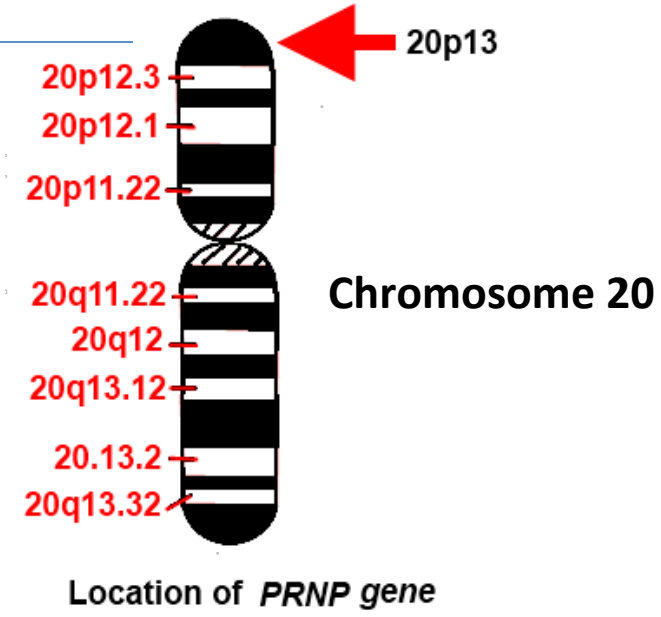
• **Major prion protein** (**PrP**, for **pr**ion **p**rotein or **p**rotease-**r**esistant **p**rotein), also known as **CD230** (cluster of differentiation 230), is encoded by the ***PRNP*** gene (PRioN Protein)

•Researchers subsequently commonly proposed roles for PrP^{c} in cell signaling or in the formation of synapses, and that it may have a memory role similar to the Alzheimer's precursor protein, but in higher centers of consciousness.

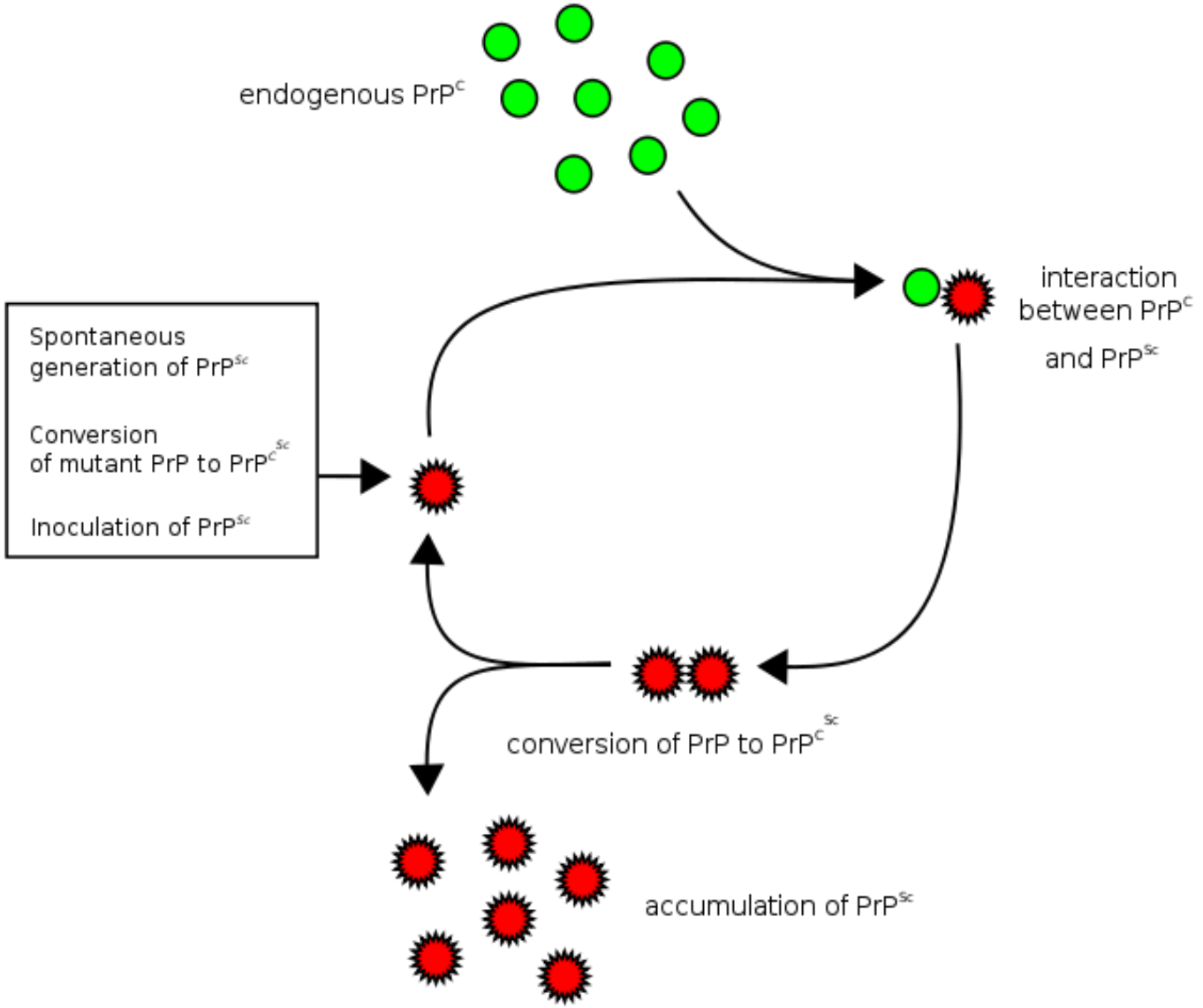
•PrP^{C} attaches to the outer surface of the cell membrane by a glycosylphosphatidylinositol anchor at its C-terminal Ser231.

•Prion protein contains 5 amino-terminal octapeptide repeats with sequence PHGGGWGQ. This is thought to generate a copper-binding domain

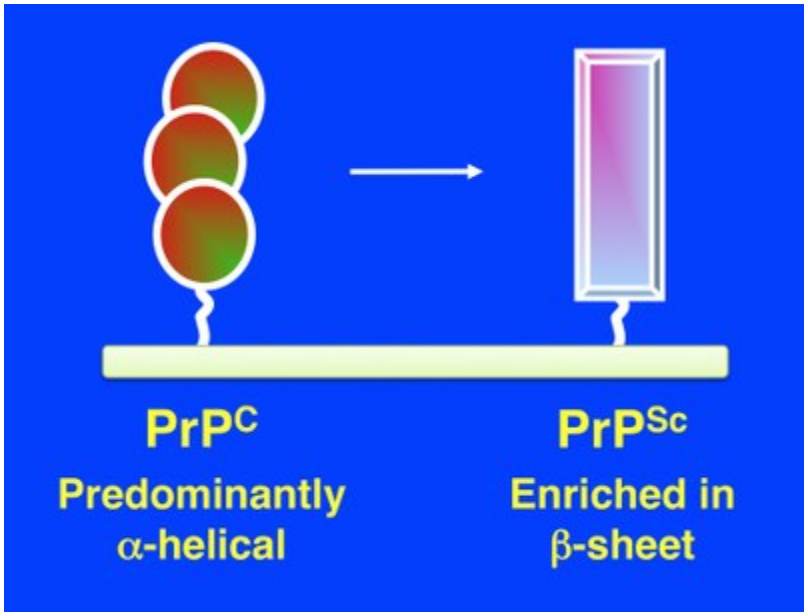


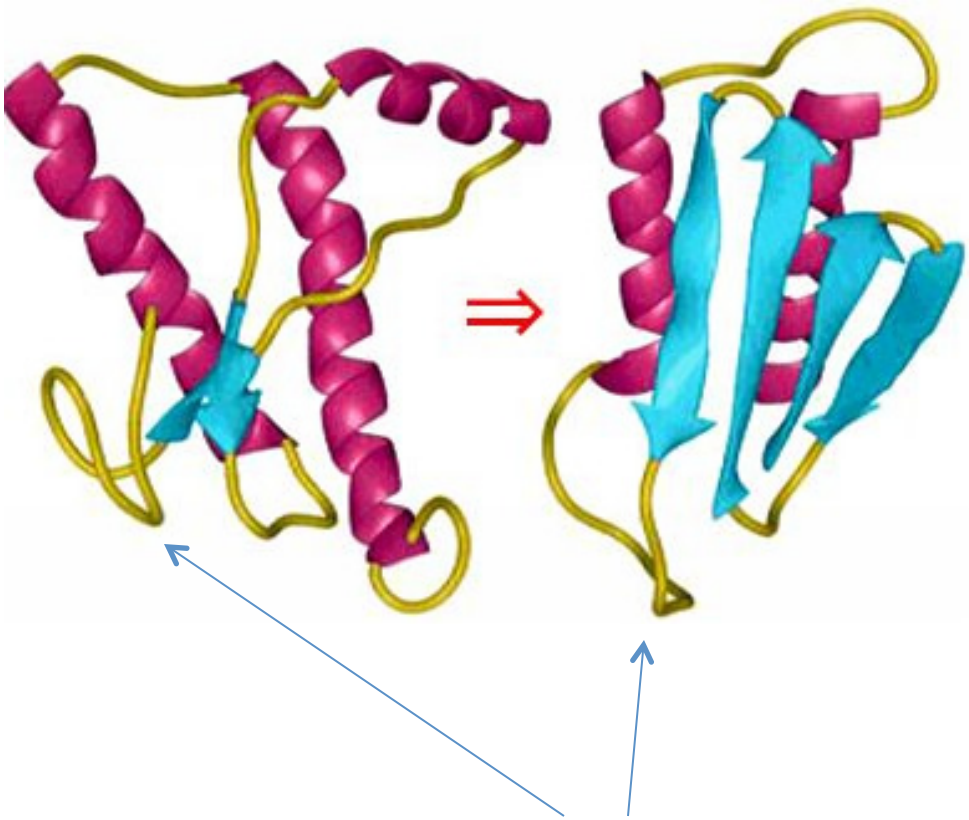


Coercion of Endogenous Prion



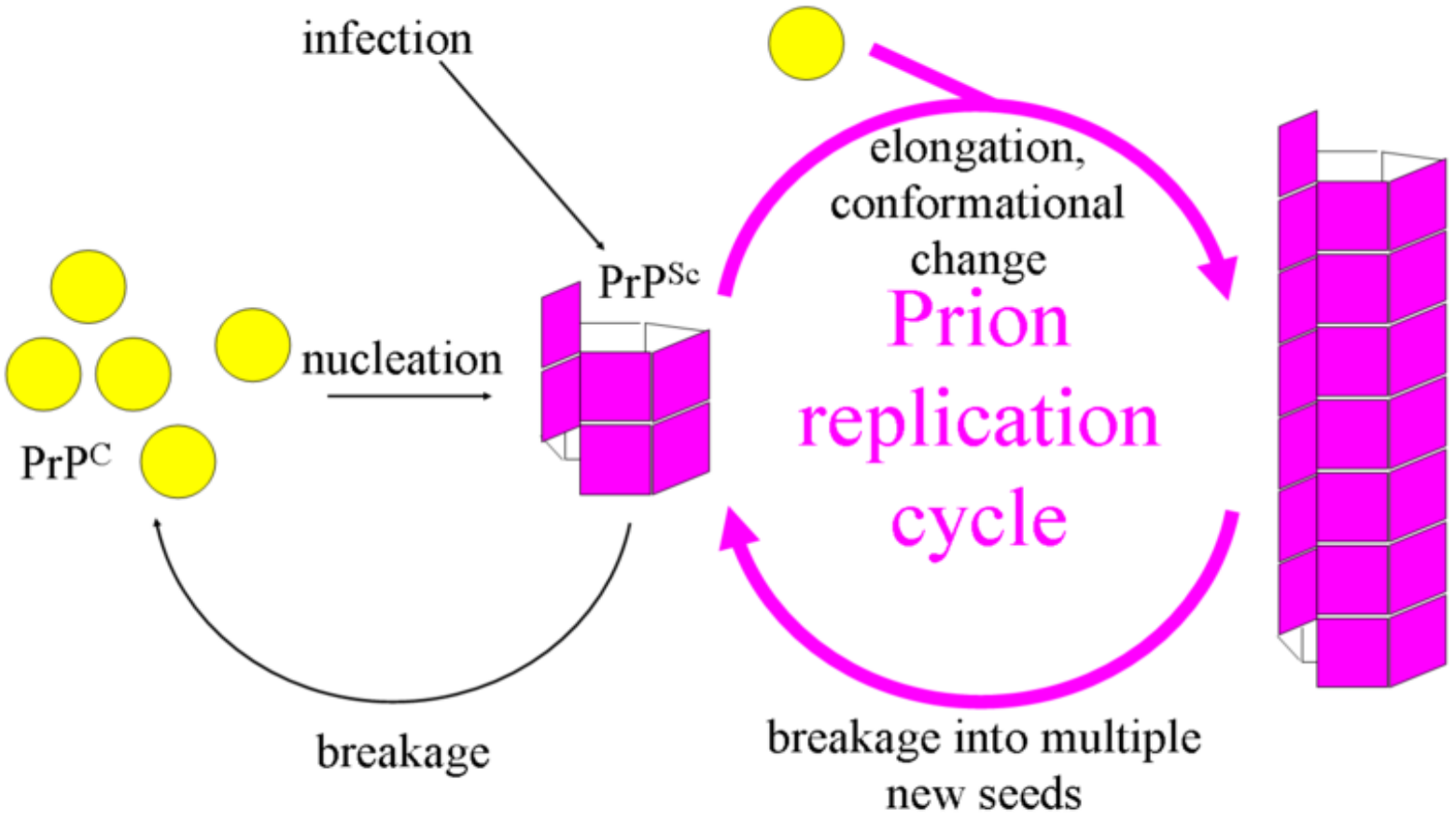
Coercion Involves α → β Transition





This region will become the one of interest as an epitope.

Coercion of Endogenous Prion



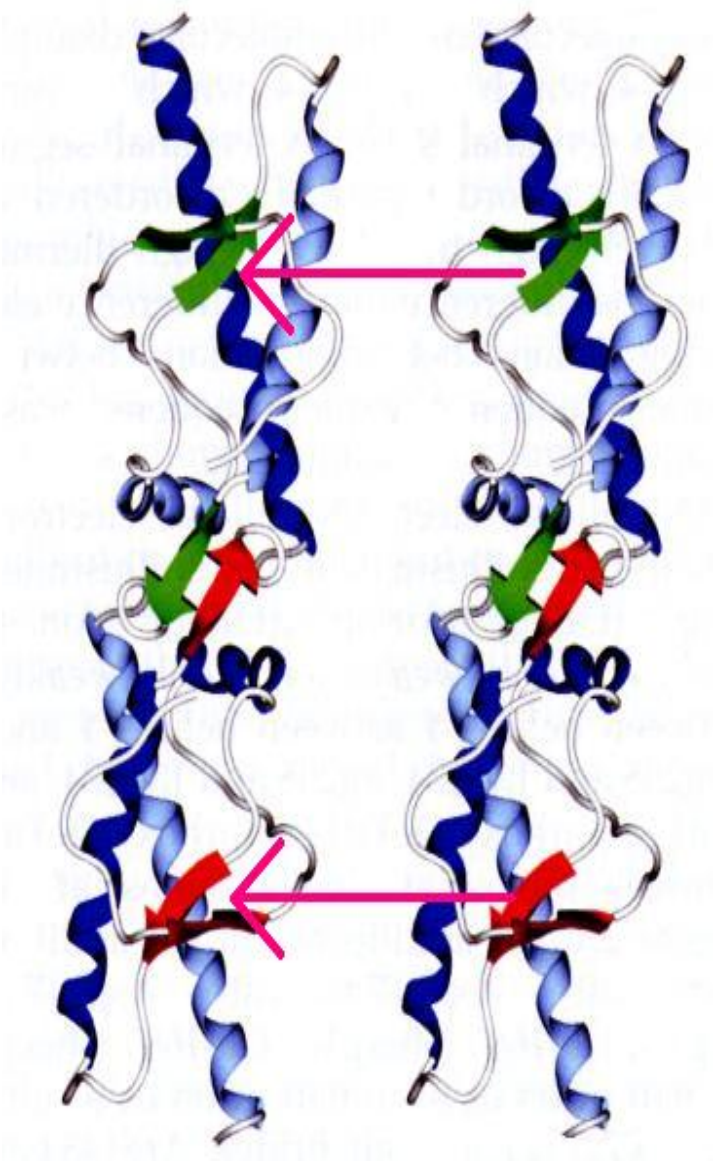
Conformational Change to Malign Form

•Much later than our diagnostic patent, Haire et al. (2004) proposed a mechanism for the oligomerization of normal prion protein rather similar to PROMETHEUS predictions.

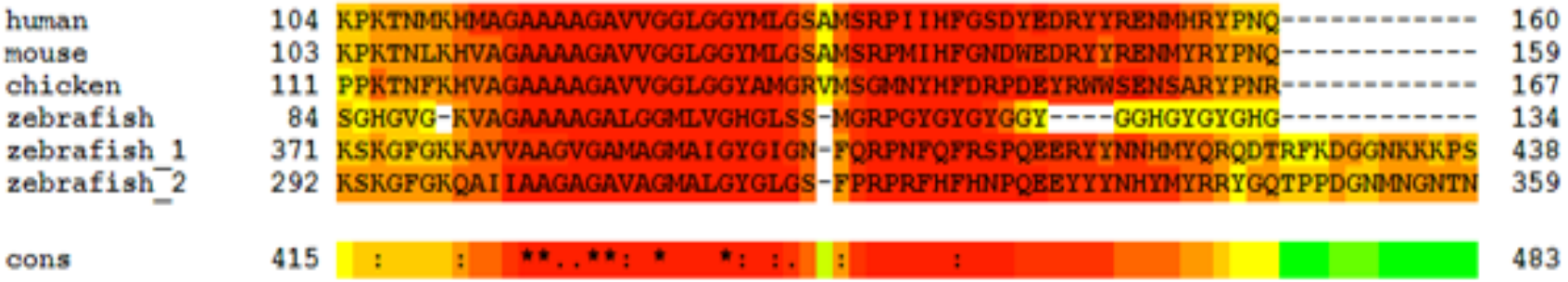
•The primary conformational change is the destruction of the cysteine bond that bridges H2 and H3.

•Haire et al. hypothesized that after disruption of this disulphide bond, H3 is exchanged between two PrP molecules, followed by the annealing of the disulphide bonds between the cysteines of the newly aligned PrPs.

•This is just the beginning of oligomerization. Once a dimer is formed, two dimers can aggregate to form a tetramer, and so on. The dimers combine by 'stacking' themselves upon each other. During this process, the anti-parallel beta sheet pairs are transformed into 4 stranded intermolecular beta-sheets



A Suspicious Region



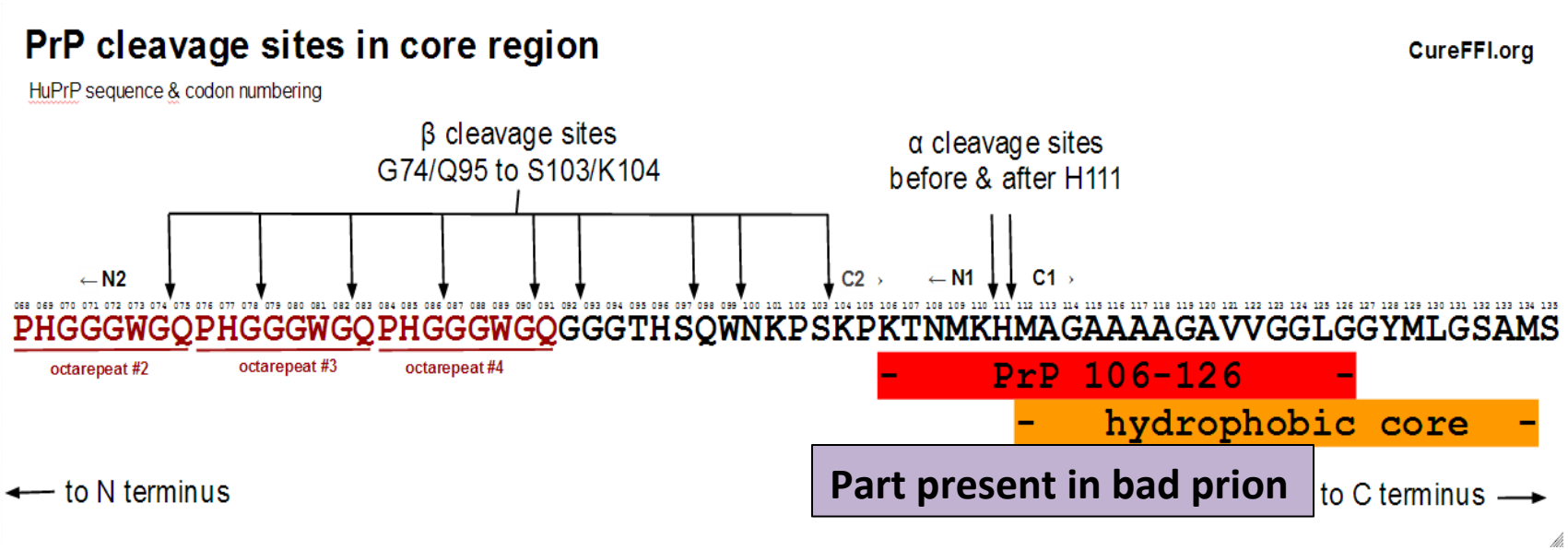
•If you zoom out from Mammalia and look at other vertebrates, there is only one large part of PrP that is really conserved even between us, chickens and zebrafish.

•This stretch runs from human amino acids 104 – 160, with particularly strong conservation at 112 – 150 or so.

•The first half of this tract, approximately 112 – 134, is the so-called “hydrophobic core”, rich in A, M, L and V that gets embedded in the membrane in ^{Ctm}PrP and ^{Ntm}PrP [reviewed in Harris 2003)

•Mutations that increase the hydrophobicity of this stretch cause GSS and tend to increase the fraction of PrP that ends up with a ^{Ctm}PrP transmembrane topology [Hegde 1998]

Cleavage Site Tends to Differ in Malign Form



•**Alpha cleavage** cleaves through the core region of amino acids 106-126. This is thought to render the remaining C1 fragment incapable of converting to PrP^{Sc}. While a few other studies have suggested that C1 makes cells or animals more sensitive to pro-apoptotic stimuli (which is bad), Westergard observed no neurological symptoms in C1 mice even with ~7x expression levels. Overall, evidence seems to point to **alpha cleavage being a good thing from the standpoint of prion disease.**

•**Beta cleavage**, by contrast, leaves PrP 106-126 intact, and when a C2 fragment starting at amino acid 73 is expressed alone, it is still capable of supporting prion disease [Fischer 1996]. This, together with the observation that C2 is more abundant in prion-diseased brains than healthy ones [Chen 1995], casts **beta cleavage as a bad thing from the standpoint of prion disease.**

Initial Prion Diagnostic Patent

- **Fragments of prion proteins**
- **Application number:** 20030199013
- **Abstract:** Synthetic polypeptides having at least one antigenic site of a prior protein, methods for their use and manufacture, antibodies raised against such polypeptides and diagnostic kits containing these polypeptides or antibodies.
- **Type:** Application
- **Filed:** April 5, 2002
- **Issued:** October 23, 2003
- **Assignee:** Proteus Molecular Design Limited
- **Inventors:** Robert Vincent Fishleigh, Barry Robson, Roger Paul Mee

- With regard to region A, our invention provides a synthetic peptide sequence according to general formula (I): 1 Seq. I.D. No: 52 X-(R1-Lys-His-R2)-Ala-Gly-Ala-Ala-Ala-R3-Gly-Ala- Val-Val-Gly-Gly-Leu-Gly-Gly-Tyr-Met-Leu-Gly-Ser- Ala-Met-Ser-(Arg-Pro-R4-R5)-Y (I)

Later Prion Diagnostic Product

- **Prion-specific polyclonal antibodies**
- **Patent number:** 7777011
- **Abstract:** Synthetic polypeptides having at least one antigenic site of a prior protein, methods for their use and manufacture, antibodies raised against such polypeptides and diagnostic kits containing these polypeptides or antibodies.
- **Type:** Grant
- **Filed:** April 5, 2002
- **Issued:** August 17, 2010
- **Assignee:** Protherics Medicines Development Limited
- **Inventors:** Robert Vincent Fishleigh, Barry Robson, Roger Paul Mee

"Synthetic polypeptides having at least one antigenic site of a prior protein, methods for their use and manufacture, antibodies raised against such polypeptides and diagnostic kits containing these polypeptides or antibodies."

- Seq. I.D. No: 52 (I) X-(R_1-Lys-His-R_2)-Ala-Gly-Ala-Ala-Ala-R_3-Gly-Ala- Val-Val-Gly-Gly-Leu-Gly-Gly-Tyr-Met-Leu-Gly-Ser- Ala-Met-Ser-(Arg-Pro-R_4-R_5)-Y
  wherein R_1 is an amino acid residue selected from Met, Leu and Phe;
  - R_2 is either Met or Val; R_3 is Ala or is absent; R_4 and R_5 are independently an amino acid residue selected from Leu, Ile and Met; one or more residues within brackets may be present or absent with the proviso that if they are present they are attached to the rest of the peptide in sequence; and X and Y may each independently be absent or independently be one or more additional amino acid residues.
- It will be apparent for example that the residues at the N-terminal of the sequence may be present as "R_2"- or "His-R_2-," or "Lys-His-R_2-" or "R_1-Lys-His-R_2-." Similarly, the preferable residues at the C-terminal may be present as "-Arg", or "-Arg-Pro," or "-Arg-Pro-R_4," or "-Arg-Pro-R_2-R_5."
- Preferably, R_1, if present, is Met, R_3, is Ala and R_5, if present, is Ile. Also, if R_4 is Met then R_4, if present, is Ile. Below are preferred sequences (Seq. I.D. No: 1 and Seq. I.D. No: 2) of formula I relating to bovine and ovine and to human prion proteins respectively:
- Seq. I.D. No: 1 X-(Met-Lys-His-Val)-Ala-Gly-Ala-Ala-Ala-Ala-Gly- Ala-Val-Val-Gly-Gly-Leu-Gly-Gly-Tyr-Met-Leu-Gly- Ser-Ala-Met-Ser-(Arg-Pro-Leu-Ile)-Y; and Seq. I.D. No: 2 X-(Met-Lys-His-Met)-Ala-Gly-Ala-Ala-Ala-Ala-Gly- Ala-Val-Val-Gly-Gly-Leu-Gly-Gly-Tyr-Met-Leu-Gly- Ser-Ala-Met-Ser-(Arg-Pro-Ile-Ile)-Y.
- A particularly preferred sequence according to formula I is Seq. I.D. No: 51:

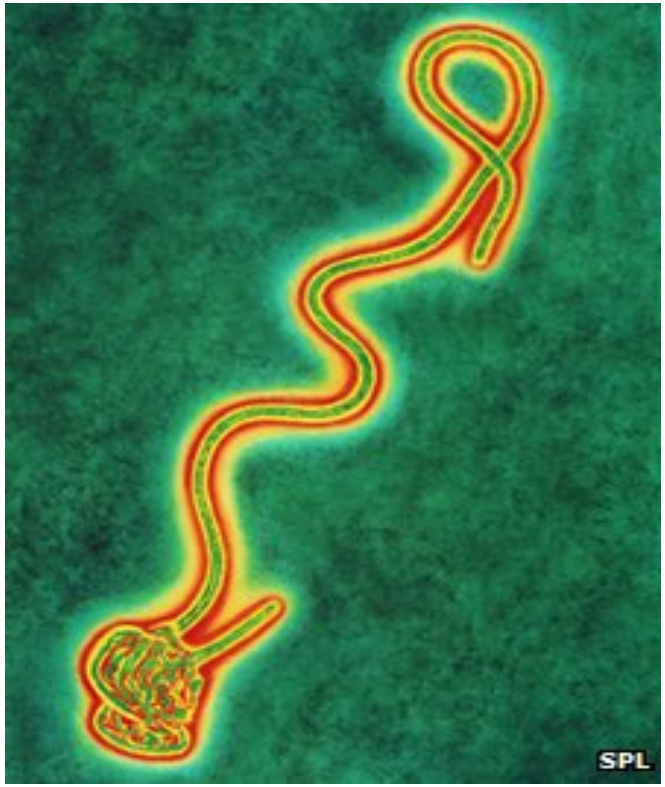
Lys-His-Met-Ala-Gly-Ala-Ala-Ala-Ala-Gly-Ala-Val- Val-Gly-Gly-Leu-Gly-Gly-Tyr-Met-Leu-Gly-Ser-Ala- Met-Ser-Arg-Gly-Cys.

"PROTHERICS FILES PATENT FOR HUMAN APPLICATIONS OF EXISTING PRION RECOGNITION TECHNOLOGY"

"The Motley Fool – The Worlds Greatest Investment Community" (Feb 5 2004)

- Protherics PLC today announces that it has filed a UK priority patent application in the area of diagnostic testing for Transmissible Spongiform Encephalopathies (TSE) in human tissues.
- Enfer Scientific Limited (" Enfer ") has the exclusive, worldwide rights to exploit Protherics antibody based TSE detection technology in animal applications.
- Protherics receives royalties from Enfer on the worldwide sales of its post mortem TSE test, including BSE, developed using Protherics' proprietary antibody technology.
- Enfer markets its test in Ireland and has licensed the distribution rights for this test for the rest of the world to Abbott Laboratories (" Abbott").
- Following the recent outbreak of BSE in the United States, Abbott has filed a Veterinary Biological Product License Application with the U.S. Department of Agriculture (USDA) to market and distribute Enfer's BSE test in the US.

A Final Word…



Similar Routes May Work for Ebola Vaccine

- **Survivors have high antibody titres.**

- **Development of a preventive vaccine for Ebola virus infection in primates.**
  - Sullivan NJ, Sanchez A, Rollin PE, Yang ZY, Nabel GJ.
  - Nature. 2000 Nov 30;408(6812):605-9.
  - A combination of DNA immunization and boosting with adenoviral vectors **that encode viral proteins** generated cellular and humoral immunity in cynomolgus macaques. … findings demonstrate that it is possible to develop a preventive vaccine against Ebola virus infection in primates.

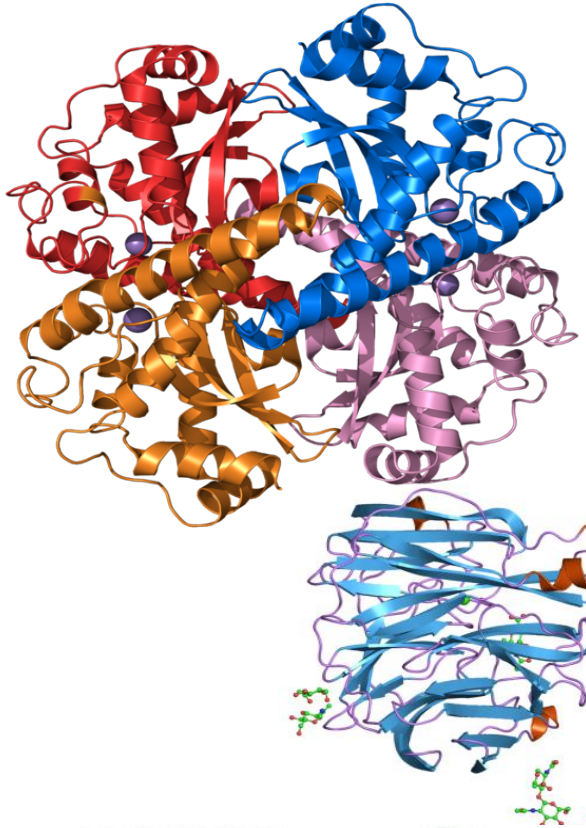
- **A nonreplicating subunit vaccine protects mice against lethal Ebola virus challenge.**
  - Waranyoo Phoolcharoen, John M. Dye, Jacquelyn Kilbourne Khanrat Piensook, William D. Pratt, Charles J. Arntzen, Qiang Chen, Hugh S. Mason, Melissa M. Herbst-Kralovetz
  - PNAS vol. 108, 51 20695–20700
  - Survival after vaccination with EIC [Ebola Immune Complex] plus [polyinosinic:polycytidylic acid (PIC, a Toll-like receptor 3 agonist] was statistically equivalent to that achieved with an alternative viral vector vaccine candidate reported in the literature.
  - Because nonreplicating subunit vaccines offer the possibility of formulation for cost-effective, long-term storage in biothreat reduction repositories, is an attractive option for public health defense measures.

Problems, but a plausible protein target…

- **Steric Shielding of Surface Epitopes and Impaired Immune Recognition Induced by the Ebola Virus Glycoprotein**
  - Joseph R. Francica, Angel Varela-Rohena, Andrew Medvec, Gabriela Plesa, James L. Riley, Paul Bates
  - P.L.O.S., September 09, 2010 DOI: 10.1371/journal.ppat.1001098

- **Conserved proline-rich region of Ebola virus matrix protein VP40 is essential for plasma membrane targeting and virus-like particle release.**
  - Reynard O1, Nemirov K, Page A, Mateo M, Raoul H, Weissenhorn W, Volchkov VE.
  - J Infect Dis. 2011 Nov;204 Suppl 3:S884-91. doi: 10.1093/infdis/jir359.

    - The matrix protein VP40 is essential for Ebola virus (EBOV) and Marburg virus assembly and budding at the plasma membrane.
    - In this study we have investigated the effect of single amino acid substitutions in a conserved proline-rich region of the EBOV VP40 located in the carboxy-terminal part of the protein.
    - We demonstrate that substitutions within this region result in an alteration of intracellular VP40 localization and also cause a reduction or a complete block of virus-like particle budding, a benchmark of VP40 function.



THANK YOU!

Any Questions?