Bioinformatics Applications for the Study of Evolution and Heme Proteins

AP Biology Teachers Workshop
Susan Cates, Ph.D.

www.bioc.rice.edu/~olson/courses.html
Background:

Why do we study heme proteins?

Why do we study evolution?

Heme Proteins

1. Hemoglobin
   oxygen transport (blood)
2. Cytochrome c oxidase
   cellular respiration, ATP synthesis
3. Peroxidases
   reduction of hydrogen peroxide
4. Cyclooxygenases
   formation of inflammatory prostanoids,
   aspirin and ibuprofen inhibit - pain relief
5. Nitric oxide synthases
   NO - cell signaling in nervous, cardiovascular
   and immune systems
6. Cytochrome P450
   metabolism, often have multiple substrates,
   and can catalyze multiple reactions -
   57 human genes for cytochrome P450s
Heme Proteins

Basic Structure of Single Domain Globins

Heme group (red ball and stick prosthetic group)

Heme

prosthetic group that consists of an iron atom contained in a porphyrin ring
Study of Heme Proteins

1. Blood substitutes

2. Historical Precedent

3. Disease

Blood Substitutes

A bloodless revolution

Blood transfusion was once regarded as an acceptable and effective practice to save a patient's life. But the HIV epidemic and concerns that donated blood could be contaminated with HIV and other infectious agents, such as the hepatitis C virus, damaged public confidence in blood transfusions. This only added to the effect of donor supply problems in the USA alone, about 1.5 million patients require some 50 to 60 million blood units per year. As a result, researchers have sought to develop artificial blood that could be used to replace blood lost after trauma and surgery, and treatment is in use in selected localities, as a substitute for whole blood, in cancer and trauma. Because the demand for blood products continues to increase due to more aggressive medical treatment protocols and the expanded health needs of an aging population, there are efforts to develop artificial replacements. The ongoing search for a blood substitute has seen some setbacks and ongoing research, with many in the field attempting to address accidents and trauma that are new to the field of medicine, as well as to treat the limitations of current products in laboratory and clinical settings. However, research into blood substitutes is not easy, and some developers, after several rounds of clinical trials, are now trying to develop artificial blood substitutes that are safer and more effective. In the past, efforts to develop such a product only began in the late 1990s, but since then, developments and advances in the field of artificial blood substitutes have been made. This has led to the development of artificial blood that can be used in clinical settings, as well as in other medical applications. The field of artificial blood substitutes is still in its infancy, but it is expected to continue to grow and evolve as new technologies and methods are developed. doi:10.1038/sj.embor.7400494

Blood Substitutes - A bloodless revolution

Blood Substitutes - A bloodless revolution

Blood Substitutes - A bloodless revolution

Blood Substitutes - A bloodless revolution
Blood Substitutes

Blood Supply Update

10-09-2006

Regional Blood Inventory: Low

<table>
<thead>
<tr>
<th>Blood Type</th>
<th>Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>O+</td>
<td>0.6</td>
</tr>
<tr>
<td>A+</td>
<td>3.2</td>
</tr>
<tr>
<td>B+</td>
<td>1.5</td>
</tr>
<tr>
<td>AB+</td>
<td>2.9</td>
</tr>
<tr>
<td>O-</td>
<td>1.2</td>
</tr>
<tr>
<td>A-</td>
<td>1.2</td>
</tr>
<tr>
<td>B-</td>
<td>0.9</td>
</tr>
<tr>
<td>AB-</td>
<td>2</td>
</tr>
</tbody>
</table>

As of 10-09-2006

American Red Cross
Blood Services
Penn-Jersey Region

Bleeding (Exsanguination, Hemorrhagic Shock) is still the number one cause of death on the battlefield.
Blood Substitutes

Transfused human donor blood carries disease

<p>| Table 3: Testing for the four major TTI's expressed as percentages of total units reported by 38 countries in 2004 |
|-------------------------------------------------|-------------------------------------------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Tested</th>
<th>Untested</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Units</td>
<td>%</td>
<td>Units</td>
</tr>
<tr>
<td>HBV</td>
<td>2185909</td>
<td>98.52%</td>
</tr>
<tr>
<td>HCV</td>
<td>2097201</td>
<td>84.53%</td>
</tr>
<tr>
<td>Syphilis</td>
<td>1796420</td>
<td>80.97%</td>
</tr>
</tbody>
</table>

TTI = transfusion transmissible infection

Advantages of Blood Substitutes:

1. Unlimited availability
2. Universal donor (no typing)
3. Disease-free
4. Longer shelf life
Study of Heme Proteins

1. Blood substitutes

2. Historical Precedent

3. Disease
Historical Precedent

The Nobel Prize in Chemistry 1962

“for their studies of the structures of globular proteins”

Haemoglobin (Hb)
Perutz [Nature, 141 (1938) 523]
&
Myoglobin (Mb)

Max Ferdinand Perutz

1/2 of the prize
United Kingdom

John Cowdery Kendrew

1/2 of the prize
United Kingdom

Historical Precedent

First three dimensional structures of proteins (1957-1960): hemoglobin (Hb) and myoglobin (Mb)
Study of Heme Proteins

1. Blood substitutes

2. Historical Precedent

3. Disease

Hemoglobin related diseases

a. Thalassemia
b. Sickle Cell Anemia
c. Staph Infection - (*Staphylococcus Aureus*)
**Disease**

**Staph Infection - Staphylococcus aureus**

The bacteria produces hemolysins that lyse red blood cells and scavenge the heme from Hb to obtain iron required for growth. It is thought that medicinal bloodletting may have been an attempt to starve pathogenic bacteria of required iron.

---

**Disease**

**Sickle Cell anemia**

Red blood cells containing HbS β subunits

Irreversibly sickled cells

Normal red blood cells with no deformations

http://www.sicklecelldisease.org/research/index.phtml
Hb nomenclature

\[ \text{HbEmbryonic} = \alpha_2 \epsilon_2 \]
\[ \text{HbA} = \alpha_2 \beta_2 \]
\[ \text{HbF} = \alpha_2 \gamma_2 \]
\[ \text{HbS} = \alpha_2 \beta(S)_2 \]

Chromosome 16

\[ \zeta \quad \alpha_1 \quad \alpha_1 \quad \alpha_2 \quad \alpha_2 \]

Chromosome 11

\[ \text{HS} \quad 1-6 \quad \text{HS} \quad 1-6 \]
\[ \epsilon \quad \epsilon \quad \gamma_A \quad \gamma_A \]
\[ \gamma_G \quad \gamma_G \quad \delta \quad \delta \]
\[ \beta \quad \beta_S \quad \beta \quad \beta_S \]

Disease

Sickle Cell anemia

Hb AS is benign (only shortens red cell life time to ~80 days)

Hb SS leads to severe anemia and early death
Disease

**Sickle Cell anemia**

The *Anopheles gambiae* parasite releases the immature parasite which is destroyed by the immune system.

**Infection with Plasmodium falciparum**

HbAS cells only last ~80 days and release the immature parasite which is destroyed by the immune system.

The *Plasmodium* parasite must remain in the red blood cells for more than 90 days to be able to change its cell coat and avoid antibody responses.

---

**Sickle Cell Anemia**

The sickle cell trait is maintained by strong selective pressure because AS individuals are resistant to malaria. It is a classic case of natural selection.
**Thalassemia**

\[ HbA = \alpha_2 \beta_2 \]

**Chromosome 16**

\[ \begin{align*} 
\zeta & \quad \zeta \\
\alpha_1 & \quad \alpha_1 \\
\alpha_2 & \quad \alpha_2 
\end{align*} \]

**Chromosome 11**

\[ \begin{align*} 
\epsilon & \quad \epsilon \\
\gamma_A & \quad \gamma_A \\
\gamma_G & \quad \gamma_G \\
\delta & \quad \delta \\
\beta & \quad \beta 
\end{align*} \]

**Disease**

**Thalassemia**

**Alpha thalassemia**

(i) Silent carrier state - hematologically normal

(ii) Mild alpha-thalassemia - mild anemia

(iii) Hemoglobin H disease - severe anemia

(iv) Hydrops fetalis - usually fatal in utero

**Beta thalassemia**

(i) Thalassemia minor, or thalassemia trait - mild anemia

(ii) Thalassemia intermedia - significant anemia

(iii) Thalassemia major - chronic transfusions required
Background:

Why do we study heme proteins?

Why do we study evolution?

Evolution of Species
Phylogenetic Trees show the relatedness of organisms

Phylogenetic Tree of Life

Common Ancestor (Root of the tree)
Molecular Evolution: relatedness of biological molecules

I. Genes
determined primarily by nucleic acid sequence

II. Proteins
determined primarily by amino acid sequence
Globin phylogenetic trees

Vinogradov et al. (2005) Proc. Natl. Acad. Sci., USA 102, 11385-11389

Multidomain flavohemoglobins (FHbs) & Globin-coupled sensors (GCS)

- Sperm whale Mb
- Human HbA
- Rice non-symbiotic Hb
- E. coli flavoHb
- Cyanobacterial Hb
- Bacillus subtilis Hb
- 2/2 or truncated globins (TrHbs)

3/3 Single domain globins (SDGs or "classical" hemoglobins)

Single domain globin (The classical myoglobin fold)

- Myoglobin (SDgb)
- 3/3 AGH/BEF
- (C and D are small and variable)
The major animal globins participate in O$_2$ storage and transport and provide lessons for designing rHb-based blood substitutes.

Globin phylogenetic trees

Vinogradov et al. (2005) Proc. Natl. Acad. Sci., USA 102, 11385-11389
**Bacillus subtilis Hb**

2/2  BE/GH

(F is variable, A is small)

---

**Globin phylogenetic trees**

Vinogradov et al. (2005) Proc. Natl. Acad. Sci., USA 102, 11385-11389

---

**Multidomain flavohemoglobins (FHbs)**

& **Globin-coupled sensors (GCS)**
Chimeric (bi-domain) globin 3/3 fold
Globin domain

E. coli FlavoHb (Hmp)

FAD

Heme

Reductase domain

VHb (bacterial) (Vitreoscilla stearcorium)

Myoglobin (chordate)

Bioinformatics: Aligning similar sequences using libraries created by sequencing genes from different species

<table>
<thead>
<tr>
<th>Species</th>
<th>F1</th>
<th>F4</th>
<th>F8</th>
<th>helical position</th>
</tr>
</thead>
<tbody>
<tr>
<td>whelk</td>
<td>AKKLSRNHTA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aardvark</td>
<td>IQPLAQSHAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pig</td>
<td>LTPLAQSHAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mouse</td>
<td>IQPLAQSHAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rabbit</td>
<td>IKPLAQSHAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sheep</td>
<td>VKHLAESHAN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bovine</td>
<td>VKHLAESHAN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>horse</td>
<td>LKPLAQSHAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>elephant</td>
<td>IQPLAQSHAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>whale</td>
<td>LKPLAQSHAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dog</td>
<td>LKPLAQSHAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chicken</td>
<td>LKPLAQTHAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>alligator</td>
<td>LKPLAKSHAL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tuna</td>
<td>LKPLANSHT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shark</td>
<td>VKELADTHIN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sea slug</td>
<td>LSQFAKEHVG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>red clam</td>
<td>VEKPAVNHIT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fly larva</td>
<td>VNTIADVPHLR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>roundworm</td>
<td>ARETVDPHLLR</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F-helix sequences
Bioinformatics: Applications

1. Pharmaceutical Design
2. Developmental Biology
3. Protein Engineering

Applications in infectious diseases

Flavohemoglobins (NO dioxygenation)

Defense mechanisms against host macrophages and other sources of NO
Applications in infectious diseases

Macrophage engulfing bacteria or fungi.

Human macrophages up-regulate inducible nitric oxide synthase (iNOS) and produce increased levels of NO after infection.

**Applications in infectious diseases**

Fox, S. I. Human Physiology (7th Edition), Fig. 15.5

**NO is potentially very toxic**

1. NO can inactivate aconitase at levels ≤ 200 nM; shuts down TCA cycle (Gardner et al., 1997 J. Biol. Chem. 272, 25071-76)

2. NO can inhibit cytochrome oxidases at levels ≤ 1 µM (Sarti et al. (M. Brunori), 2000 Biochem. Biophys. Res. Com. 274, 183-7; Tania et al. (R. Poole), 2000, J. Biol. Chem. in press)

3. NO synthesis inhibits respiration in endothelial and smooth muscle cells (Clementi et al. (S. Moncada), 1999, PNAS 96, 1559-1562)

**NO dioxygenase (NOD) activity detoxifies NO (and O₂)**

\[
\text{2Hb(FeII)O}_2 + 2\text{NO} \rightarrow 2\text{NO}_3^- + 2\text{Hb(FeIII)} \\
\text{NADH} \rightarrow \text{NAD} \\
\text{O}_2 \hspace{1cm} \text{cytochrome b₅, or flavoprotein(FAD)} \hspace{1cm} \text{(reductase)}
\]

- FlavoHbs are expressed to detoxify NO, increasing the resistance of pathogenic and symbiotic microorganisms to host defense mechanisms.

(Gardner et al., 1998, PNAS 95, 10378-10383; Stamler, Poole, and others)

(21)
Miranda et al. (2006) J. Biol. Chem. 280, 3674-3576

Sequence comparisons of globins from bacteria to man (weak homologies except in key structural (3°) regions)

- P. catodon SMMβ
  - AL1...B1...M10...C2...CD1...
  - ...E1...E7...E11...E20...J1...F4...F8...G1...

- H. sapiens Hb β
  - AVLSEGEMQVLHFAVKEAVDGAGQDDSLKGLFRSEELLEREDFREK...
  - E52

- A. aeolicus AαTgb
  - MLSEQETIVIKSTPVLKENGTEKAMELLFSKMKFLHEECGELC...
  - E49

- V. stercoraria Hb
  - LDLQCTINIIKATPVLPKLKETTVTTFKLLFAKVRPVEAEQRQESL...
  - E52

- E. coli Fhb
  - MLMTQATQIAATVATPLALTEYGDFIKCFQRMFEMELELKVAYNAMQ...
  - E50

- A. eutrophus Fhb
  - MLTQKYTQIVAVATAPVLIEAKYDFIKCFQRMFEMELELKVAYNAMQ...
  - E48

Applications in infectious diseases

Applications in genetic diseases

Hb AS is benign (only shortens red cell lifetime to ~80 days)

Hb SS leads to severe anemia and early death

Sickle Cell Anemia
Applications in genetic diseases

Thalassemia

Chromosome 16

\[ \begin{align*}
\zeta & \quad \zeta \\
\alpha_1 & \quad \alpha_1 \\
\alpha_2 & \quad \alpha_2
\end{align*} \]

Chromosome 11

\[ \begin{align*}
\text{HS} & \quad \text{HS} \\
1-6 & \quad 1-6 \\
\epsilon & \quad \epsilon \\
\gamma_A & \quad \gamma_A \\
\gamma_G & \quad \gamma_G \\
\delta & \quad \delta \\
\beta & \quad \beta
\end{align*} \]

Thalassemia

Applications in genetic diseases

Alpha thalassemia
(i) Silent carrier state - hematologically normal
1 inactive alpha gene - 3 are normal
(ii) Mild alpha-thalassemia - mild anemia
2 inactive alpha genes - red cells are smaller
(iii) Hemoglobin II disease - severe anemia
3 inactive alpha genes - appear malnourished, enlarged spleens, bony abnormalities
(iv) Hydrops fetalis - usually fatal in utero
4 inactive alpha genes - in utero transfusions followed by life-long, chronic transfusions required to survive

Thalassemia
Applications in genetic diseases

Beta thalassemia
(i) Thalassemia minor, or thalassemia trait - mild anemia
mildly suppressed beta gene - main danger is passing the trait to children with a partner who also has thalassemia trait
(ii) Thalassemia intermedia - significant anemia
suppressed or inactive beta gene - small stature, poor weight gain, poor energy levels, and susceptibility to infection
(iii) Thalassemia major - chronic transfusions required
Severely suppressed or inactive beta gene(s) - chronic blood transfusions are needed

Thalassemia

Applications in Human developmental biology

HbEmbryonic = α2ε2  HbF = α2γ2  HbA = α2β2

Chromosome 16

Chromosome 11
Applications in Human developmental biology

α₂ε₂ Hb embryonic (High O₂ affinity)

α₂γ₂ HbF (moderate O₂ affinity)

α₂β₂ HbA (low O₂ affinity)

ε subunit

γ subunit

β subunit

Chromosome 11

Chromosome 16

α₁, α₂

γΑ, γΓ

δ, δ

β
Hydroxyurea, the first approved drug for the causative treatment of sickle cell anemia was shown to decrease the number and severity of attacks in a study in 1995 (Charache et al) and shown to increase survival time in a study in 2003. This is achieved by reactivating fetal hemoglobin production in place of the hemoglobin S that causes sickle cell anemia.

Applications in protein engineering

Rational design of free Hb, engineered as a blood substitute:

1. Must not scavenge NO - specificity and affinity
2. Must still deliver O2 effectively - specificity and affinity
3. Must not degrade - stability
Acid, base, and isopropanol tests for hemoglobin stability

Embryonic and Fetal Hbs are much more stable (application for blood substitutes)

Applications in protein engineering

1. Engineering more stable Hbs:
   • sequence comparison of more stable Hbs
   • sperm whale Hb
   • fetal Hb

2. Engineering NO, O2 affinities and specificities:
   • primarily by rational design of mutations
   • compare sequences of Hbs with desired properties
Applications in protein engineering

<table>
<thead>
<tr>
<th>BHG2_HUMAN</th>
<th>HGR1_HUMAN</th>
<th>HRH2_HUMAN</th>
<th>HRR_HUMAN</th>
<th>HSD_HUMAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>GQIPEKTDG</td>
<td>GQIPEKTDG</td>
<td>GQIPEKTDG</td>
<td>GQIPEKTDG</td>
<td>GQIPEKTDG</td>
</tr>
<tr>
<td>LTFV</td>
<td>LTFV</td>
<td>LTFV</td>
<td>LTFV</td>
<td>LTFV</td>
</tr>
<tr>
<td>KVATY</td>
<td>KVATY</td>
<td>KVATY</td>
<td>KVATY</td>
<td>KVATY</td>
</tr>
<tr>
<td>TKGRE</td>
<td>TKGRE</td>
<td>TKGRE</td>
<td>TKGRE</td>
<td>TKGRE</td>
</tr>
<tr>
<td>QTVWF</td>
<td>QTVWF</td>
<td>QTVWF</td>
<td>QTVWF</td>
<td>QTVWF</td>
</tr>
<tr>
<td>YTFPQ</td>
<td>YTFPQ</td>
<td>YTFPQ</td>
<td>YTFPQ</td>
<td>YTFPQ</td>
</tr>
<tr>
<td>KVATY</td>
<td>YTVWF</td>
<td>YTVWF</td>
<td>YTVWF</td>
<td>YTVWF</td>
</tr>
<tr>
<td>LTFV</td>
<td>QTVWF</td>
<td>QTVWF</td>
<td>QTVWF</td>
<td>QTVWF</td>
</tr>
<tr>
<td>HRH2_HUMAN</td>
<td>HRH2_HUMAN</td>
<td>HRH2_HUMAN</td>
<td>HRH2_HUMAN</td>
<td>HRH2_HUMAN</td>
</tr>
<tr>
<td>GQIPEKTDG</td>
<td>GQIPEKTDG</td>
<td>GQIPEKTDG</td>
<td>GQIPEKTDG</td>
<td>GQIPEKTDG</td>
</tr>
<tr>
<td>LTFV</td>
<td>LTFV</td>
<td>LTFV</td>
<td>LTFV</td>
<td>LTFV</td>
</tr>
<tr>
<td>KVATY</td>
<td>KVATY</td>
<td>KVATY</td>
<td>KVATY</td>
<td>KVATY</td>
</tr>
<tr>
<td>TKGRE</td>
<td>TKGRE</td>
<td>TKGRE</td>
<td>TKGRE</td>
<td>TKGRE</td>
</tr>
<tr>
<td>QTVWF</td>
<td>QTVWF</td>
<td>QTVWF</td>
<td>QTVWF</td>
<td>QTVWF</td>
</tr>
<tr>
<td>YTFPQ</td>
<td>YTFPQ</td>
<td>YTFPQ</td>
<td>YTFPQ</td>
<td>YTFPQ</td>
</tr>
<tr>
<td>KVATY</td>
<td>YTVWF</td>
<td>YTVWF</td>
<td>YTVWF</td>
<td>YTVWF</td>
</tr>
<tr>
<td>LTFV</td>
<td>QTVWF</td>
<td>QTVWF</td>
<td>QTVWF</td>
<td>QTVWF</td>
</tr>
</tbody>
</table>

Sequence comparison of fetal gamma Hb sequences with adult beta and embryonic epsilon sequences.

---

O3/10/2007