Modeling Protein Degradation Processes and the Development of Rational Approaches to Stabilization

A New Strategy of Molecular QbD

Naresh Chennamsetty¹, Vladimir Voynov¹, Veysel Kayser¹, Curt Schneider¹, Diwakar Shulka¹, Bernhard Helk² and Bernhardt L. Trout¹

1) MIT  2) Novartis Pharma AG
New Strategic Approach

- Incorporate developability and manufacturability early.
- Incorporate QbD.
- Reduce overall time from discovery to market launch.
- Molecular QbD presents a new strategic option.
Trout Research Group
Research Areas in Trout Group
Molecular QbD

- Formulation and Stabilization of Biotherapeutics.
  - Aggregation
  - Oxidation
  - Deamidation
  - Hydrolysis

- Crystallization and New Technologies for the Manufacturing of Small Molecular Pharmaceuticals.

- Major Initiatives
  - Novartis-MIT Center for Continuous Manufacturing
  - Singapore-MIT Program on Chemical and Pharmaceutical Engineering
Objective for today:

Identify major problems that you face, and determine how we might be able to help.
Molecular QbD for Therapeutic Antibody Stabilization
Aggregation of Therapeutic proteins: E.g. Therapeutic Antibodies

- Antibody is a large glyco-protein (~ 1300 residues, 150kDa)
- Therapeutic antibodies are used in the treatment of cancer, Rheumatoid Arthritis, etc.
Therapeutic protein sales are growing fast

- Antibody sales are growing at a fast pace
- The sales could reach $56 billion by 2012, a compound annual growth rate (CAGR) of 13%
Problems: Antibody Aggregation

- Therapeutic antibodies aggregate during manufacture and storage

  Aggregation
  Antibodies
  Aggregates are
  - Inactive against disease
  - Immunogenic
  - Can have serious side effects
Why do antibodies aggregate?

- What regions are aggregation prone?
- Can we modify these aggregation prone regions to enhance stability?

Molecular Simulations + Experiment → Molecular level detail on aggregating regions → Validation
Overview of methodology

Unstable antibody

Molecular simulation of a single antibody

SAP Technology*

Identify aggregating regions

Stable antibody

Mutate predicted aggregating regions

* SAP (Spatial Aggregation Propensity) technology developed in this project
Simulation methodology

- Detailed atomistic model for antibody
- Explicit atomistic model for water

- CHARMM force field\textsuperscript{1} for protein, TIP3P water model\textsuperscript{2}
- CHARMM\textsuperscript{3} and NAMD\textsuperscript{4} simulation packages
- Simulations in the NPT ensemble at 300K and 1atm
- Ewald summation for electrostatics
- Supercomputer resources from NCSA

For unknown X-ray structures:

Homology modeling with canonical structures

- Validation: Structure obtained by homology modeling matches very well with the X-ray structure

Full antibody simulation

- Full antibody constructed from fragments using another antibody, 1HZH, as template
- Simulated using supercomputer
- First full MAb simulation in the literature
Fc fragment simulation

- Significant fluctuations in protein and sugar groups
- These fluctuations could dynamically expose buried hydrophobic residues
SAP tool applied after simulation

Unstable antibody

Molecular simulation of a single antibody

SAP Technology

SAP to Identify aggregating regions
Spatial–Aggregation–Propensity (SAP)

SAP finds the dynamically exposed hydrophobic patches on the protein surface.

\[
\text{SAP} = \frac{\sum_{\text{Residues within } R} \left( \frac{\text{SAA of side chain atoms}}{\text{SAA of middle residue in extended tripeptide 'Ala-X-Ala'}} \right) \times \text{Residue Hydrophobicity}}{\sum_{\text{Simulation Average}} \text{Residue fraction exposed}}
\]

SAA of middle residue in extended tripeptide ‘Ala -X- Ala’

SAA = Solvent Accessible Area

SAP mapped onto antibody structure

- RED regions are highly hydrophobic dynamically exposed patches
- BLUE regions are highly hydrophilic dynamically exposed patches
Mutation of SAP predicted aggregation prone regions

- 5 sites with high SAP values selected for mutations
- These sites are mutated to more hydrophilic residues
Stability analysis of mutants by SEC-HPLC

- All 8 mutants lead to increase in monomers (decrease in aggregates)
- This validates SAP predictions
DSC analysis of mutants

- The mutants have higher melting transition for the $C_H^2$ domain
- This indicates increased stability of the mutants
SAP predicts the aggregation prone region of Hemoglobin S

* PNAS (2009)
Can SAP predict protein binding regions?

SAP Technology

Predicts aggregating regions

Predicts binding regions?
SAP predicts protein binding regions as well

- Using simple hydrophobicity would be difficult to predict binding regions
- High SAP regions correlate well with protein binding regions
SAP predicts protein binding regions of antibody

- High SAP regions correlate well with protein binding regions
SAP predicts binding regions of EGFR

- Using simple hydrophobicity would be difficult to predict binding regions
- High SAP regions correlate well with protein binding regions
SAP predicts binding regions of EGFR

- High SAP regions correlate well with protein binding regions
More SAP applications

SAP Technology

Predicts aggregating regions

Predicts binding regions

Developability ranking?
(Aggregation propensity ranking)
SAP for developability ranking

Protein sequences
1. VTEQDSS...
2. DSKQDSS...
3. IKDQDSK...
4. KDPPDSK...

Structures from homology modeling

SAP analysis

Total SAP

Least stable
Most stable

1 2 3 4 mAbs

SAP scale

+0.5
-0.5

Variable region

mAb1

mAb2

mAb3

mAb4

SAP will be optimized for developability ranking
Current drug discovery process

Discovery → Development

Trial and error methods
Expensive and time consuming

Stable mAbs
Unstable mAbs

Improved drug discovery process using SAP

Discovery → SAP screening → Development

SAP stabilization

Stable mAbs
More SAP applications

SAP Technology

Predicts aggregating regions

Predicts binding regions

Developability ranking

Find stable payload conjugation sites?
SAP improves the determination of sites for payload conjugation

To obtain conjugation sites that yield stable monomers, these sites should be partially exposed and away from high-SAP regions.
Summary:
Developed the SAP tool to aid in discovery-commercialization.
Macroscopic modelling

Macroscopic modeling and mathematical connection between long-term and short-term stability tests

Need model of aggregation for a given temperature and the temperature dependence of the rate constants
Monomer loss kinetics: Examples of 1st-, 2nd-order fits

1st order:
\[ \frac{dM}{dt} = -k_{obs}M \]

2nd order:
\[ \frac{dM}{dt} = -k_{obs}M^2 \]

Full function:
\[ \frac{dM}{dt} = -nk_nM^n - k_gMC \]
Temperature dependence: Kinetics are Non-Arrhenius

\[ k = A \exp\left(-\frac{E_a}{RT}\right) \]
\[ \ln k = \ln A - \frac{E_a}{RT} \]

A: pre-exponential coef.
E\(_a\): activation energy
R: gas constant
T: temperature

Need Non-Arrhenius model
VFT method
(Vogel, Fulcher, Tammann)

- Where $T_o$ is a reference temperature at which the relaxation time relevant to molecular displacements becomes infinite, i.e. where the entropy changes suddenly

- Liu et al. found that $T_o = T_m$ for H exchange rates (DNA melting T)

- Can we also use VTF for highly non-Arrhenius behaving aqueous protein samples?

- We have found a similar trend for MAB2 but a higher T for MAB1

Arrhenius:

$$k = A \exp\left(-\frac{E_a}{RT}\right)$$

VFT:

$$k = A \exp\left(\frac{B}{(T - T_o)}\right)$$

Angell et al., J. Appl. Phys., Vol. 88, No. 6, 15 September 2000
Prediction of MAB1 aggregation with the model fitted to short term data

2\textsuperscript{nd} order reaction

LT: long-term data
PD: predicted kinetics
Prediction of MAB2 aggregation with the model fitted to short term data

2\textsuperscript{nd} order reaction

LT: long-term data
PD: predicted kinetics
### Long-Term vs Predicted

- **Time in months**
- **LT: long-term data**
- **2nd order reaction fit**

<table>
<thead>
<tr>
<th></th>
<th>MAB1</th>
<th>MAB2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>5C</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>time</td>
<td>2nd</td>
<td>LT</td>
</tr>
<tr>
<td>0</td>
<td>99.2</td>
<td>99.2</td>
</tr>
<tr>
<td>3</td>
<td>99.19</td>
<td>99.3</td>
</tr>
<tr>
<td>6</td>
<td>99.17</td>
<td>99.3</td>
</tr>
<tr>
<td>12</td>
<td>99.13</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>99.05</td>
<td>99.2</td>
</tr>
<tr>
<td><strong>12C</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>time</td>
<td>2nd</td>
<td>LT</td>
</tr>
<tr>
<td>0</td>
<td>99.2</td>
<td>99.2</td>
</tr>
<tr>
<td>3</td>
<td>99.17</td>
<td>99.5</td>
</tr>
<tr>
<td>6</td>
<td>99.13</td>
<td>99.5</td>
</tr>
<tr>
<td>12</td>
<td>99.05</td>
<td>99.2</td>
</tr>
<tr>
<td></td>
<td>98.14</td>
<td>98.5</td>
</tr>
<tr>
<td><strong>25C</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>time</td>
<td>2nd</td>
<td>LT</td>
</tr>
<tr>
<td>0</td>
<td>99.2</td>
<td>99.2</td>
</tr>
<tr>
<td>3</td>
<td>99.02</td>
<td>99</td>
</tr>
<tr>
<td>6</td>
<td>98.84</td>
<td>98.6</td>
</tr>
<tr>
<td>12</td>
<td>98.47</td>
<td>97.6</td>
</tr>
<tr>
<td></td>
<td>98.8</td>
<td>97.6</td>
</tr>
<tr>
<td><strong>40C</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>time</td>
<td>2nd</td>
<td>LT</td>
</tr>
<tr>
<td>0</td>
<td>99.2</td>
<td>99.2</td>
</tr>
<tr>
<td>1</td>
<td>98.2</td>
<td>98.2</td>
</tr>
<tr>
<td>3</td>
<td>96.4</td>
<td>98.6</td>
</tr>
<tr>
<td>6</td>
<td>93.81</td>
<td>93.5</td>
</tr>
</tbody>
</table>
Molecular QbD for the Design of Protein-Cosolute Interactions
Starting Point: Arginine

- Well known stabilizer.
- Action of stabilization unknown.
- Seems to interact net neutrally with biomolecules.
- Can we better understand arginine and develop better additives?
Proposed mechanism by which arginine inhibits aggregation

Enhanced Crowding

Encounter Complex

Small Aggregate

Native protein

Partially unfolded Intermediate

$\frac{1}{2} A_2$

$\frac{1}{2} A_2$

Free Energy

Reaction Co-ordinate
Preferential Interaction Coefficients

- Measures the degree of attraction or repulsion of cosolutes to proteins.
- Positive values mean that cosolutes are attracted to the protein. (e.g. Gnd, urea)
- Negative values mean that cosolutes are repelled from the protein. (e.g. sucrose, mannitol)
- Indicates the degree to which an additive stabilizes the folded state of a protein.
Arginine has a concentration dependent preferential interaction.

VPO Technique:

$$\Gamma_{XP} \equiv \left( \frac{\partial m_x}{\partial m_p} \right)_{T,P,\mu_x} = -\left( \frac{\partial \mu_p}{\partial \mu_x} \right)_{T,P,m_p}$$

$$\Gamma_{XP} \approx \frac{m_x}{m_p} \left( 1 - \frac{(\partial Osm/\partial m_x)_{m_p}}{(\partial Osm/\partial m_x)_{m_p=0}} \right)$$

Computational Methodology

**MD simulation of aqueous arginine solutions**
- Temperature: 278-368 K
- Concentration: 0.25-2.75 molal

**MD simulation of protein in aqueous arginine solution**
- Protein: α-Chymotripsinogen A, Lysozyme
- Temperature: 298 K
Computing Preferential Interaction Coefficients

Preferential Interaction coefficient ($\Gamma_{23}$): excess number of additive molecules in local domain

$$\Gamma_{23} = \left\langle n_3^{II} - n_1^{II} \left( \frac{n_3}{n_1} \right) \right\rangle$$

$$= \rho_3(\infty) \int (g_3(r) - g_1(r))dV$$

1-Water 2-protein 3-additive

(Scatchard notation)

$$(g_3(r) - g_1(r)) = 0$$

$$\Gamma_{23}(r,t) = n_3(r,t) - n_1(r,t) \left( \frac{n_3 - n_3(r,t)}{n_1 - n_1(r,t)} \right)$$

$n_3$ total number of cosolvent molecules

$n_1$ total number of water molecules

---

**Graph**

- Water
- Additive

$r (\text{Å})$

- ArgHCl
- Glycerol
- Urea
experimentally preferential interaction data is only available up to 0.7 molal.
Interactions in aqueous arginine solutions

- Arginine tends to form clusters via hydrogen bonding and Gdn Stacking
Interactions between arginine and A protein

- Arginine interacts with charged and aromatic residues.
- Interaction with aromatic residues could stabilize unfolded intermediates.
- Clustering in arginine solution leads to enhanced crowding.

Contact coefficient = local/bulk concentration
We have created novel compounds that:
- Solvate proteins much like water
- Have little influence on the folding equilibrium
- Specifically inhibit protein association

We call such excipients “neutral crowders”.

\[ \Gamma_{XP} \approx 0 \]
\[ \delta \Delta G_u^\circ \approx 0 \]

Aggregation: High Temperature

Protein: 10 mg/mL α-Chymotrypsinogen A
Temperature: 52.5 °C
Buffer: 20 mM Sodium Citrate, pH 5

aCtn Monomer Concentration (C/C₀)

Time (Hours)

Compound A (0.2 M)
Arginine HCl (0.67 M)
No Additive
Aggregation: Body Temperature

Protein: 10 mg/mL α-Chymotrypsinogen A
Temperature: 37 °C
Buffer: 20 mM Sodium Citrate, pH 5

Graph showing the effect of different compounds on the aggregation of α-Chymotrypsinogen A over time.
Aggregation Rates vs. Concentration

Protein: 10 mg/mL α-Chymotrypsinogen A
Temperature: 52.5 °C
Buffer: 20 mM Sodium Citrate, pH 5
Other Excipients

Effect of Additives at the Same Osmotic Pressure as Blood
(Osm = 289 mmol/kg)

Additive (Isotonic Concentration)

Sucrose 283 mM
ArgHCl 170 mM
NaH2PO4 212 mM
Na2SO4 142 mM
Compound C 197 mM
Compound A 79 mM
Compound B 138 mM

Rate Reduction Factor = \( \frac{k}{k_0} \)

T = 52.5 °C, 20 mM Sodium Citrate, pH 5
Shelf Life Predictions (5% Loss)

<table>
<thead>
<tr>
<th>k/k₀ (aCgn Aggregation)</th>
<th>10 mg/mL aCgn</th>
<th>40 mg/mL aCgn</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>37 °C</td>
<td>45 °C</td>
</tr>
<tr>
<td>Compound A</td>
<td>4.6%</td>
<td>6.1%</td>
</tr>
<tr>
<td>Compound B</td>
<td>1.5%</td>
<td>2.5%</td>
</tr>
<tr>
<td>ArgHCl</td>
<td>23.7%</td>
<td>27.0%</td>
</tr>
</tbody>
</table>

Aggregation suppression is fairly constant at various temperatures and concentrations.

Shelf Life extended from a few days to several months.

### t₉₅:alpha-Chymotrypsinogen A Monomer Loss (5%)

<table>
<thead>
<tr>
<th>T (°C)</th>
<th>No Additive</th>
<th>Arginine k/k₀ = 0.25</th>
<th>Compound B k/k₀ = 0.025</th>
</tr>
</thead>
<tbody>
<tr>
<td>52.5</td>
<td>2 Minutes</td>
<td>8 Minutes</td>
<td>1.3 Hours</td>
</tr>
<tr>
<td>45</td>
<td>2.1 Hours</td>
<td>8.4 Hours</td>
<td>3.5 Days</td>
</tr>
<tr>
<td>37</td>
<td>8.6 Hours</td>
<td>1.4 Days</td>
<td>14 Days</td>
</tr>
<tr>
<td>25*</td>
<td>3.4 Days</td>
<td>12 Days</td>
<td>5 Months</td>
</tr>
</tbody>
</table>

10 mg/mL aCgn, 20 mM Sodium Citrate, pH 5
*Predicted Value (Arrhenius Plot of Low Temperature Data)
Summary of Molecular Simulation Approaches for Cosolutes

- Gain Mechanistic Understanding
- Allow Rational Design
  - E.g. additives
  - Buffers
A very brief summary of the oxidation of therapeutic antibodies
Correlation between WCN and the Relative Rates of Oxidation

Solves puzzle of 0 SAA from X-ray structure.
Expected Structural Effect

Fit to a two-state protein unfolding model $N \leftrightarrow U$

Gibbs-Helmholtz equation

$$\frac{\Delta G_{\text{unfolding}}}{RT} = \frac{\Delta H_{\text{(unf)}}^0}{RT_m} - \frac{\Delta H_{\text{(unf)}}^0}{RT} - \frac{\Delta C_p(T - T_m)}{RT} + \frac{\Delta C_p}{R} \log \frac{T}{T_m}$$

Equilibrium denaturation monitored by fluorescence

$\Delta C_p = 0$ or Arrhenius-like

$\Delta C_p \neq 0$

Non-Arrhenius-like structural change

Possibly non-Arrhenius-like oxidation kinetics!
Extrapolation Analysis
On the physical basis when there is no structural effect

Expect a non-Arrhenius behavior connecting high T and low T regions
A Phenomenological Model

One of the several models we developed

More buried met
More exposed met
Met in peptides

Structural effect is an activated process
**Expression for Rate Constant**

Use equilibrium condition, mass balance and kinetic expressions

\[
k_{\text{apparent}} = \frac{1}{[O]_0} \frac{A e^{\Delta E^\dagger/RT}}{1 + \frac{c^\phi}{[O]_0} e^{\Delta G^{(l)}/RT}}
\]

- **A** pre-factor
- **[O]_0** initial oxidant concentration
- **c^\phi** standard concentration
- **ΔE^\dagger** intrinsic reaction free energy barrier
- **ΔG^{(l)}** Gibbs free energy change of local structural change

Classification of structural effect

- No structural dependence when \( \Delta G^{(l)} \ll RT \)
- Local structural dependence when \( \Delta G^{(l)} \approx RT \)
- Global structural dependence when \( \Delta G^{(l)} \gg RT \)

Can be simplified into Arrhenius equation

Only when temperature is near the local \( T_m \), structural effect results in non-Arrhenius...
Conclusions

- New Strategic Approach: Molecular QbD for Integration of Discovery, Development, and Manufacturing. Objective: reduce overall time from Discovery to Market Delivery.

- Areas of Impact:
  - Discovery
    - Developability/Manufacturability
      - Aggregation
      - Oxidation
      - Deamidation
      - Fragmentation
    - Payload Conjugation
  - Development
Conclusions

- Areas of Impact:
  - Discovery
    - Developability/Manufacturability
      - Aggregation
      - Oxidation
      - Deamidation
      - Fragmentation
    - Payload Conjugation
  - Development
    - Formulation
    - Stability modeling