Quantifying Radiation Damage in Real Space

Markus Gerstel, C. M. Deane, E. F. Garman

Systems Biology Doctoral Training Centre
Department of Biochemistry
University of Oxford
Radiation Damage
Global Radiation Damage - Diffraction Pattern

RhoGDI EA, DLS I03, 100K, 22.10.2012

Gerstel et al. (2013)
Specific Radiation Damage

Weik et al. (2002) TcAChE; 10 MGy/dataset; densities from datasets 1 and 9, 1.5σ, 100 K
Specific Radiation Damage

Weik et al. (2002) TcAChE; 10 MGy/dataset; densities from datasets 1 and 9, 1.5σ, 100 K
Specific Damage Progression

Order of Specific Damage Progression:
- Reduction of metallo-centres
- Disulphide bond elongation and breakage
- GLU/ASP decarboxylation
- TYR –OH
- MET S–C
Specific Radiation Damage – Who cares?

- Specific Radiation Damage
- Structural Changes
- Your data look good, but they are wrong.
Specific Radiation Damage – Who cares?

- Specific Radiation Damage
- Structural Changes

Your data look good, but they are wrong.
Specific Radiation Damage – Who cares?

- Specific Radiation Damage
- Structural Changes

Your data look good, but they are wrong.
Radiation damage misidentified as reaction intermediate
(Matsui et al., 2002; bR)

Reaction intermediate missed due to radiation damage
(Kort et al., 2004; pyp)
Progression of Structural Changes

Order of Structural Changes:
- Reduction of metallo-centres
- Disulphide bond elongation and breakage
- GLU/ASP decarboxylation
- TYR –OH
- MET S–C

Preferential Damage: Order within groups
But why?
\textbf{B}_{\text{Damage}}
Idea

- Statistics
- Use PDB structures
- Identify specific damage in a single dataset
- Locations with increased uncertainty
- Occupancy / atomic B factor
### Occupancy is not refined

<table>
<thead>
<tr>
<th>Occupancy</th>
<th>Atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{occ} &lt; 0 )</td>
<td>&lt; 0.01% (5 in 3hkw)</td>
</tr>
<tr>
<td>( \text{occ} = 0 )</td>
<td>0.09%</td>
</tr>
<tr>
<td>( 0 &lt; \text{occ} &lt; 0.5 )</td>
<td>0.67%</td>
</tr>
<tr>
<td>( \text{occ} = 0.5 )</td>
<td>2.30%</td>
</tr>
<tr>
<td>( 0.5 &lt; \text{occ} &lt; 1 )</td>
<td>0.71%</td>
</tr>
<tr>
<td>( \text{occ} = 1 )</td>
<td>96.23%</td>
</tr>
<tr>
<td>( \text{occ} &gt; 1 )</td>
<td>&lt; 0.01% (5 in 1ie0 and 1wul)</td>
</tr>
</tbody>
</table>
## Occupancy is not refined

<table>
<thead>
<tr>
<th>Occupancy</th>
<th>Atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>$occ &lt; 0$</td>
<td>&lt; 0.01% (5 in 3hkw)</td>
</tr>
<tr>
<td>$occ = 0$</td>
<td>0.09%</td>
</tr>
<tr>
<td>$0 &lt; occ &lt; 0.5$</td>
<td>0.67%</td>
</tr>
<tr>
<td>$occ = 0.5$</td>
<td>2.30%</td>
</tr>
<tr>
<td>$0.5 &lt; occ &lt; 1$</td>
<td>0.71%</td>
</tr>
<tr>
<td>$occ = 1$</td>
<td><strong>96.23%</strong></td>
</tr>
<tr>
<td>$occ &gt; 1$</td>
<td>&lt; 0.01% (5 in 1ie0 and 1wul)</td>
</tr>
</tbody>
</table>

Gerstel, Garman, Deane (2014)
Atomic B factor

- Isotropic vs. Anisotropic (35%)
- Can be refined per atom, per side-/mainchain, per residue, ...
- Correlated with packing density (Weiss, 2007)

Correct for packing density?
Atomic B factor

- Isotropic vs. Anisotropic (35%)
- Can be refined per atom, per side-/mainchain, per residue, ...
- Correlated with packing density (Weiss, 2007)
- Correct for packing density?
$B_{\text{Damage}}(a) := \frac{B(a)}{\overline{B(S_{P(a)})}}$

Simply put:

- Take the atomic $B$ factor
- Divide by the average $B$ factor of all atoms within the same protein structure with a similar packing density.
- We used Atomic Contact Numbers (ACN; Weiss, 2007).
\[ B_{\text{Damage}}(a) := \frac{B(a)}{\overline{B}(S_{P(a)})} \]

Simply put:

- Take the atomic B factor
- Divide by the average B factor of all atoms within the same protein structure with a similar packing density.
- We used Atomic Contact Numbers (ACN; Weiss, 2007).
$B_{Damage}(a) := \frac{B(a)}{\overline{B}(S_{P(a)})}$

Simply put:

- Take the atomic B factor
- Divide by the average B factor of all atoms within the same protein structure with a similar packing density.
- We used Atomic Contact Numbers (ACN; Weiss, 2007).
\[ B_{\text{Damage}}(a) := \frac{B(a)}{B(S_{P(a)})} \]

Simply put:

- Take the atomic B factor
- Divide by the average B factor of all atoms within the same protein structure with a similar packing density.
- We used Atomic Contact Numbers (ACN; Weiss, 2007).
$B_{Damage}(a) := \frac{B(a)}{B(S_{P(a)})}$

Simply put:

- Take the atomic B factor
- Divide by the average B factor of all atoms within the same protein structure with a similar packing density.
- We used Atomic Contact Numbers (ACN; Weiss, 2007).
Results – Test set
## Nanao et al. (2005) dataset

6 proteins before and after X-ray burn

<table>
<thead>
<tr>
<th>Protein</th>
<th>Elastase</th>
<th>Insulin</th>
<th>Lysozyme</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>before</td>
<td>after</td>
<td>before</td>
</tr>
<tr>
<td>Dose per data set</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X-ray burn (kGy)</td>
<td>80</td>
<td>80</td>
<td>2000</td>
</tr>
<tr>
<td># CYS/ASP/GLU</td>
<td>8 / 8 / 4</td>
<td>6 / – / 4</td>
<td>8 / 7 / 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein</th>
<th>Ribonuclease A</th>
<th>Thaumatin</th>
<th>Trypsin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>before</td>
<td>after</td>
<td>before</td>
</tr>
<tr>
<td>Dose per data set</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X-ray burn (kGy)</td>
<td>3000</td>
<td>3000</td>
<td>800</td>
</tr>
<tr>
<td># CYS/ASP/GLU</td>
<td>8 / 5 / 5</td>
<td>16 / 12 / 6</td>
<td>12 / 6 / 4</td>
</tr>
</tbody>
</table>
Atomic B factor vs $B_{\text{Damage}}$ (Nanao et al., 2005)

- **atomic B factor, entire structure**
  - low-dose datasets
  - high-dose datasets
  - overlap

- **$B_{\text{Damage}}$, entire structure**
  - low-dose datasets
  - high-dose datasets
  - overlap

Normalized density vs $B_{\text{Damage}}$ for low-dose and high-dose datasets, with overlap regions.

Gerstel, Garman, Deane (2014)
Atomic B factor vs $B_{\text{Damage}}$ (Nanao et al., 2005)

atomic B factor, CYS sulfurs

- low-dose datasets
- high-dose datasets
- overlap

$B_{\text{Damage}}$, CYS sulfurs

- low-dose datasets
- high-dose datasets
- overlap

Normalized density

Gerstel, Garman, Deane (2014)
GLU Oε: B factor (Nanao et al., 2005)

![Box plot of atomic B factors for Lysozyme and Ribonuclease A, showing low-dose and high-dose datasets.](Image)

- **Lysozyme**
  - Low-dose dataset: Approximately 10-15 Å
  - High-dose dataset: Approximately 15-20 Å

- **Ribonuclease A**
  - Low-dose dataset: Approximately 10-15 Å
  - High-dose dataset: Approximately 15-20 Å

Gerstel, Garman, Deane (2014)
GLU $\epsilon$: $B_{\text{Damage}}$ (Nanao et al., 2005)

Gerstel, Garman, Deane (2014)
$\Delta B_{\text{Damage}}$ between low-/high-dose datasets (Nanao)

- **Elastase** (2 / 2.2 MGy)
- **Trypsin** (2 / 2.8 MGy)
- **Lysozyme** (2 / 3.2 MGy)
- **Thaumatin** (2 / 3.6 MGy)
- **Insulin** (2 / 6.0 MGy)
- **Ribonuclease A** (4 / 10 MGy)

Gerstel, Garman, Deane (2014)
Results – PDB analysis
PDB Analysis

2,704 protein structures
database size: 373 GB

Gerstel, Garman, Deane (2014)
Only $\pi$-helix secondary structure shows significant differences.

But: rarest secondary structure as assigned by STRIDE (0.016%).
Disulphide Bond Types

Spiral

Hook

Staple

Azimi et al. (2011)
Staple disulphide bonds show different behaviour.

Includes allostERIC (protein function regulating), but not bulk of structural or catalytic disulphide bonds

(Schmidt et al., 2006)

Not visible with B factor

Gerstel, Garman, Deane (2014)
Solvent Accessibility of GLU O\(\epsilon\) (Nanao et al., 2005)

Positive correlation between solvent accessibility and \(B_{\text{Damage}}\)

Gerstel, Garman, Deane (2014)
## Solvent Accessibility correlations (PDB)

<table>
<thead>
<tr>
<th>Atom</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARG N$\epsilon$</td>
<td>0.574*</td>
</tr>
<tr>
<td>ARG C$\zeta$</td>
<td>0.615*</td>
</tr>
<tr>
<td>ARG N$\eta$</td>
<td>0.643*</td>
</tr>
<tr>
<td>CYS S$\gamma$</td>
<td>0.286*</td>
</tr>
<tr>
<td>MET S$\delta$</td>
<td>0.438*</td>
</tr>
<tr>
<td>MET C$\epsilon$</td>
<td>0.338*</td>
</tr>
<tr>
<td>GLN N$\epsilon$</td>
<td>0.432*</td>
</tr>
<tr>
<td>GLN O$\epsilon$</td>
<td>0.393*</td>
</tr>
<tr>
<td>GLU O$\epsilon$</td>
<td>0.309*</td>
</tr>
</tbody>
</table>

Slope of simple linear regression model fitted using solvent accessibility as explanatory and $B_{\text{Damage}}$ as dependent variable.

Starred values significant with $p < 0.05$. 

Gerstel, Garman, Deane (2014)
Summary

- $B_{\text{Damage}}$ can be used to identify affected sites
- $B_{\text{Damage}}$ unlocks the PDB for radiation damage research
- Secondary structure is not correlated with susceptibility
- Disulphide bond type is
- Solvent accessibility matters
Thank you

Elspeth Garman, Charlotte Deane, Raimond Ravelli, Philip Hogg, Isaac Turner
EPSRC; IUCr; SysBio DTC, St Hilda’s College, and Biochemistry Oxford
Dose numbers – Where do they come from?

SF-38      RADDOSE-3D

Sunday, August 10  15:15–16:00  519b
Electron Densities
Electron Densities

Back to the Roots

Weik et al. (2002) TcAChE; 10 MGy/dataset; densities from datasets 1 and 9, 1.5σ, 100 K
Electron Densities

Aftershock

Observing Electron Density Progression

Gerstel et al. (2013)
Looking at the Numbers...

Electron densities in fixed real space volumes

- A231
- A254
- A265
- A402
- A521
- A67
- A94

Gerstel et al. (2013)
Sensitivity, Quantified

Electron Densities in fixed real space volumes

- A231
- A254
- A265
- A402
- A521
- A67
- A94

Residue Half life

<table>
<thead>
<tr>
<th>Residue</th>
<th>Half life</th>
</tr>
</thead>
<tbody>
<tr>
<td>A231</td>
<td>6.6</td>
</tr>
<tr>
<td>A254</td>
<td>2.5</td>
</tr>
<tr>
<td>A265</td>
<td>1.8</td>
</tr>
<tr>
<td>A402</td>
<td>3.8</td>
</tr>
<tr>
<td>A521</td>
<td>2.0</td>
</tr>
<tr>
<td>A67</td>
<td>2.5</td>
</tr>
<tr>
<td>A94</td>
<td>4.4</td>
</tr>
</tbody>
</table>

Gerstel et al. (2013)
Sensitivity, Quantified

Electron densities in fixed real space volumes

Residue Half life

A231 6.6
A254 2.5
A265 1.8
A402 3.8
A521 2.0
A67 2.5
A94 4.4

dataset (Weik et al., 2000)

Gerstel et al. (2013)
## Sensitivity, Quantified

<table>
<thead>
<tr>
<th>S–S bond</th>
<th>Half lifes</th>
<th>Mean half life</th>
</tr>
</thead>
<tbody>
<tr>
<td>A231</td>
<td>6.6</td>
<td>6.6</td>
</tr>
<tr>
<td>A254–A265</td>
<td>2.5 1.8</td>
<td>2.15</td>
</tr>
<tr>
<td>A402–A521</td>
<td>3.8 2.0</td>
<td>2.9</td>
</tr>
<tr>
<td>A67–A94</td>
<td>2.5 4.4</td>
<td>3.45</td>
</tr>
</tbody>
</table>
RhoGDI mutants

Longenecker et al. (2001)


Ooi/ACN Distributions on the PDB dataset

**Distribution of Ooi\(_8\) numbers**

- Number of residues vs. Ooi number of residue; 8 Å sphere

**Distribution of Ooi\(_{14}\) numbers**

- Number of residues vs. Ooi number of residue; 14 Å sphere

**Distribution of ACN\(_8\) numbers**

- Number of atoms vs. Atomic contact number of atom; 8 Å sphere

**Distribution of ACN\(_{14}\) numbers**

- Number of atoms vs. Atomic contact number of atom; 14 Å sphere

Gerstel, Garman, Deane (2014)
B factor vs $B_{\text{Damage}}$ on ASN/ASP Oδ

![Graphs showing B factor vs $B_{\text{Damage}}$ for various proteins: Elastase, Insulin, Lysozyme, Ribonuclease A, Thaumatin, and Trypsin.](image)

- **Elastase**
  - ASN
  - ASP

- **Insulin**
  - ASN
  - ASP

- **Lysozyme**
  - ASN
  - ASP

- **Ribonuclease A**
  - ASN
  - ASP

- **Thaumatin**
  - ASN
  - ASP

- **Trypsin**
  - ASN
  - ASP

**Gerstel, Garman, Deane (2014)**

**Burn / total dose [MGy]:**
- 2 / 2.2
- 2 / 6
- 2 / 3.2
- 4 / 10
- 2 / 3.6
- 2 / 2.8