

1
2
3
4
5
6
7
8
9
10 **Analysis of global and site-specific radiation**
11 **damage in cryo-EM**
12
13
14
15
16
17
18
19
20

21 **Johan Hattne¹, Dan Shi¹, Calina Glynn², Chih-Te Zee², Marcus Gallagher-Jones²,**
22 **Michael W. Martynowycz¹, Jose A. Rodriguez² & Tamir Gonen^{1,3,4*}**
23
24
25
26
27

28 ¹Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA.
29

30 ²Department of Chemistry and Biochemistry, UCLA-DOE Institute, University of
31 California, Los Angeles, Los Angeles, CA 90095, USA.
32

33 ³ Howard Hughes Medical Institute, University of California, Los Angeles, Los Angeles
34 CA 90095 USA.
35

36 ⁴ Departments of Physiology and Biological Chemistry, David Geffen School of Medicine,
37 University of California, Los Angeles, Los Angeles CA 90095 USA.
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54

55 *To whom correspondence should be addressed: T.G: tgonen@ucla.edu
56
57
58

59 Short title: Radiation damage in cryo-EM
60
61
62
63
64
65

Summary

Micro-crystal electron diffraction (MicroED) is an emerging method in cryo-EM for structure determination using nanocrystals. It has been used to solve structures of a diverse set of biomolecules and materials, in some cases to sub-atomic resolution. However, little is known about the damaging effects of the electron beam on samples during such measurements. We assess global and site-specific damage from electron radiation on nanocrystals of proteinase K and of a prion hepta-peptide and find that the dynamics of electron-induced damage follow well-established trends observed in X-ray crystallography. Metal ions are perturbed, disulfide bonds are broken, and acidic side chains are decarboxylated while the diffracted intensities decay exponentially with increasing exposure. A better understanding of radiation damage in MicroED improves our assessment and processing of all types of cryo-EM data.

Keywords

electron cryo microscopy (cryoEM); MicroED; radiation damage; Electron diffraction; Electron crystallography

Introduction

Structure determination relies on interpreting the outcome of interactions of a beam of quanta with a sample. Most quanta pass through samples without any interaction at all (Henderson, 1995). Other quanta scatter either elastically, whereby they interact with the sample without losing any energy, or inelastically by depositing part of their energy to the sample. In a conventional diffraction measurement, the information that can be gained increases with the number of elastically scattered quanta. Inelastic scattering events manifest as damage, introduce noise, and ultimately limit the signal that can be extracted from the sample.

The success of crystallographic structure determination ultimately depends on the ratio of elastic to inelastic scattering events (Nave & Hill, 2005). While this ratio is greater in electron diffraction than in X-ray diffraction (Henderson, 1995), a single incident electron carries sufficient energy to knock out several electrons from an atom in an inelastic scattering event. These ejected, secondary electrons and their associated Auger electrons are mobile even at 77 K (Jones et al., 1987), and can, depending on the chemical composition of the sample and its surrounding mother liquor, cause additional ionization and excitation events in the crystal (Garman, 2010). Additional damage from thermal diffusion of atomic and molecular radicals produced by these ionization events is curbed by keeping the sample at cryogenic temperatures, where diffusion is limited (Henderson & Unwin, 1975; Hayward & Glaeser, 1979; Uyeda et al., 1980; Jeng & Chiu, 1984). The absorbed energy, which is related to radiation damage, depends on the chemical composition of the crystal and the medium in which it is embedded, the temperature at which the measurement is performed, and the energy of the incident electrons.

1
2
3
4 As damage accumulates, its effects become apparent at specific sites as well as throughout
5 the entire crystal. As crystalline order deteriorates, the fraction of unit cells contributing to
6 crystalline diffraction decreases, the *B*-factor generally increases (Kmetko et al., 2006),
7 and the observed unit cell volume may increase as the lattice expands (Ravelli et al., 2002).
8 Consequently, the signal of the obtained diffraction pattern, which varies with the square
9 of the number of scattering unit cells, decreases. To some extent, damage can be
10 compensated for by appropriate scaling procedures (Diederichs, 2006), and may be further
11 mitigated in nanocrystals as the probability that secondary electrons escape before causing
12 further damage is higher than in large crystals (Nave & Hill, 2005; Sanishvili et al., 2011).
13
14

15
16 Site-specific damage may be observed if the impact of radiation damage on the crystal is
17 not uniformly distributed but more selective towards certain moieties (Weik et al., 2000).
18 Unlike global damage, site-specific damage can only be assessed once the dataset has been
19 integrated and phased, and a real-space density map is calculated. Site-specific damage
20 becomes apparent when certain bonds are more susceptible to damage than others, and may
21 remain invisible if it only occurs in a small fraction of the unit cells or if it is masked by
22 phases calculated from an undamaged model.
23
24

25
26 Exposure of the sample to the electron beam results in immediate damage even at cryogenic
27 temperatures as in electron cryo-electron microscopy (cryo-EM). Low-dose procedures
28 helped minimize the exposure to the sample prior to data recording (Uyeda et al., 1980).
29 Early studies indicated that organic samples deteriorate 4–5× faster at room temperature
30 compared with cryogenic temperatures (Unwin & Henderson, 1975; Hayward & Glaeser,
31 1979; Jeng & Chiu, 1984). In these studies, either two-dimensional crystals and/or three-
32 dimensional crystals were used in electron diffraction to look at the overall decay of
33 reflections following exposure to the electron beam. These studies indicated that atomic
34 resolution information (defined as better than 3 Å) was lost after exposure of the sample to
35 only 3 e⁻ Å⁻². In 2013, a new method for cryo-EM was unveiled and termed MicroED, or
36 3D electron crystallography of microscopic crystals (Shi et al., 2013; Nannenga et al.,
37 2014b). With continuous rotation MicroED, which is the preferred method of data
38 collection in MicroED, significant loss of diffraction intensity was observed at resolutions
39 better than ~3 Å when only ~3 e⁻ Å⁻² were used, consistent with past studies (Hayward &
40 Glaeser, 1979; Jeng & Chiu, 1984; Baker et al., 2010).
41
42
43
44

45
46 Owing to the strong interaction of electrons with matter, and the fact that only diffraction
47 data are collected (no imaging), high-resolution structures can be determined by MicroED
48 from three-dimensional nanocrystals with significantly less total electron exposure than
49 what is normally used for other cryo-EM modalities. Recent MicroED experiments
50 demonstrated that complete data sets could be collected from a single nanocrystal using a
51 total exposure of less than ~1–2 e⁻ Å⁻² (Nannenga et al., 2014a; de la Cruz et al., 2017),
52 making it possible to design new and improved experiments to test for beam induced
53 damage to the specimen with increasing exposure to the electron beam.
54
55

56
57 In this study, we set out to determine the damaging effects of electron radiation using
58 MicroED and nanocrystals of a well characterized sample, proteinase K, and a short hepta-
59 peptide with a bound metal. MicroED data were collected using exposures of 0.0017–
60
61
62
63
64
65

1
2
3
4 0.007 e⁻ Å⁻² s⁻¹. Such low exposure allowed us to repeatedly measure the same wedge of
5 reciprocal space from the same crystal and to compare the data obtained over increasing
6 exposure as the experiment progressed. The data was also sufficiently complete to allow
7 us to investigate the effects not only in reciprocal space but also in real space, such that
8 both global and site-specific damage can be observed. The data indicate that beam damage
9 is a limiting factor in high-resolution cryo-EM methods, and the results have implications
10 for all EM methods that use considerably higher electron exposures for imaging.
11
12
13
14

15 Results

16 Indicators of global damage

17
18 For both the globular proteinase K and hepta-peptide samples, the overall weakening of
19 the diffraction spots resulting from the loss of crystalline order can be modeled by
20 exponential decay as a function of absorbed dose (Blake & Phillips, 1962; Liebschner et
21 al., 2015) (Figure 1). We note that this model appears to systematically underestimate the
22 intensities of the weakest reflections at high exposures for all samples. This observation
23 does not necessarily invalidate the model: weak reflections are difficult to measure
24 accurately due to noise; to the extent these reflections can be measured at all; they are not
25 discernible by eye. Profiles derived from stronger surrounding reflections are likely to
26 overestimate their intensities, and outlier rejection due to, for example, ill-fitting
27 background models introducing further bias toward more intense reflections.
28
29
30
31

32
33 After an exposure of 1 e⁻ Å⁻² the average intensity of all observed reflections in
34 proteinase K has decreased to 73% of its extrapolated value at zero dose (Figure 1). After
35 an additional 1.6 e⁻ Å⁻², the high-resolution limit of the data has dropped from 1.7 Å to
36 1.9 Å (Table 1). Similar trends can be seen in the peptide images at high and low exposure
37 rates where the intensities have dropped to 61% and 65%, respectively, by 1 e⁻ Å⁻².
38 However, the highest-resolution reflections in the peptide data are much stronger than those
39 in proteinase K and the effect of exposure on optical resolution is consequently not as
40 pronounced. For both peptide data sets, the optical resolution decreases to ~1.1 Å by 2 e⁻
41 Å⁻² (Table 2, Table 3).
42
43
44

45 The highest resolution reflections, those providing atomic resolution, are most sensitive to
46 disruption of crystalline order. Consequently, as a result of subtle long-range changes in
47 the lattice, weak, high-resolution reflections fade into the background faster than strong,
48 low-resolution reflections (Blake & Phillips, 1962; Howells et al., 2009), and the highest
49 observable resolution decreases with dose (Table 1, Table 2, Table 3). In conjunction, fine
50 features in real space density maps disappear at a higher rate than the overall molecular
51 envelope. Generally, high resolution information, defined as better than 2 Å, was
52 significantly decayed when exposures greater than ~3 e⁻ Å⁻² were used. Beyond ~4 e⁻ Å⁻²,
53 the reflections at a resolution finer than 2 Å have dropped to 10% of their extrapolated
54 value at zero dose in diffraction patterns of proteinase K. The corresponding exposure for
55 the peptide at high and low dose rates are similar at 2.7 e⁻ Å⁻² and 3.2 e⁻ Å⁻², respectively.
56
57
58
59
60
61
62
63
64
65

1
2
3
4 As an additional proxy for global damage by electrons we use changes in relative B -factor,
5 B_{rel} (Kmetko et al., 2006), calculated over the reflections in the resolution range common
6 to all datasets of a given sample and exposure rate. For all but the poorest diffracting
7 crystals, B_{rel} increases monotonically with absorbed dose (Figure 2). We also find that the
8 unit cell volume, although a much less reliable indicator of global radiation damage,
9 generally increases with exposure (Murray & Garman, 2002). In our measurements, the
10 unit cell volumes of proteinase K and of the peptide increased by a modest 1.8% from 452
11 to 460 nm³ and 1.15 to 1.17 nm³, respectively.
12
13

14 **Site-specific damage**

15
16 Localized chemical changes within the macromolecule can be analyzed by observing
17 changes in the density attributed to specific atoms in real space. We observe site-specific
18 damage even at exposures as small as 0.1 e⁻ Å⁻². For example, the presence of positive mF_o -
19 DF_c difference density is detected around the sulfur atoms of the disulfide bonds in
20 proteinase K, indicating that the disulfide bridge was breaking apart in a significant fraction
21 of the unit cells (Helliwell, 1988) even at total exposure <0.9 e⁻ Å⁻² (Figure 3). At this
22 exposure, the overall diffraction intensity was reduced to 75% of its extrapolated value at
23 zero dose, indicating that 86% of the unit cells are still diffracting to high resolution (Blake
24 & Phillips, 1962). As the exposure increases, the positive difference density is replaced by
25 negative difference density in the location of the bond and the $2mF_o$ - DF_c density
26 progressively weakens.
27
28
29
30

31 We also observe site-specific damage in the decarboxylation of the acidic side chains
32 (Figure 3). Like the disulfide bonds, these moieties have been observed to be particularly
33 sensitive to radiation in both X-ray crystallography (Weik et al., 2000) and single-particle
34 cryo-EM (Bartesaghi et al., 2014; Barad et al., 2015). In proteinase K, the density around
35 the side chains of glutamate and aspartate begins to deteriorate starting at a total exposure
36 of ~2 e⁻ Å⁻² and completely disappears after approximately 5 e⁻ Å⁻² (Figure 3).
37
38
39

40 The peptide unit cell contains a zinc atom as well as an acetate and three modeled water
41 molecules, all of which display significant signs of site-specific radiation damage at
42 exposures >0.8 e⁻ Å⁻² (Figure 4). For the bound zinc, the radiation damage is primarily
43 modeled using atomic displacement parameters (ADP). Unlike occupancies, which model
44 large-scale discrete disorder and were fixed at unity in all models, the ADP describe
45 harmonic vibrations around the mean position of the atoms. While the density around the
46 zinc atom in the model remains positive even at the highest exposure, its ADP steadily
47 increases with exposure (Figure 4). The displacement begins with as little total exposure
48 as 0.2 e⁻ Å⁻², when the ADP of the zinc is 1.3× higher than the average ADP in the model.
49
50
51

52 Site specific damage was further assessed in real space using RIDL (Bury et al., 2015),
53 which calculates the maximum density loss D_{loss} for each atom in the model. This provides
54 a means to objectively establish the sensitivity of the different amino acids to electron
55 radiation. Generally, the results from MicroED confirm the order and appearance of site-
56 specific damage observed in X-ray crystallography: metals are significantly more prone to
57 damage and glutamate, aspartate, and cysteine residues accumulate damage even at very
58 low exposures to the electron beam.
59
60
61
62
63
64
65

Discussion

Past studies of radiation damage in cryo-EM described the exposure-dependent decay of diffraction intensities up to and including 3 Å resolution (Henderson & Unwin, 1975; Hayward & Glaeser, 1979; Jeng & Chiu, 1984; Baker et al., 2010). No data was analyzed at resolutions better than 3 Å likely because such data was not recorded at the time. Unfortunately, even a recent study that reported a 2.6 Å resolution single particle reconstruction included an analysis of beam induced damage only to lower resolution (Grant & Grigorieff, 2015). Moreover, while the early experiments analyzed the decay of diffraction intensities, little or no analysis of the effects of exposure on real space was described. This is likely because at low resolutions, the effects of specific damage in real space are difficult to characterize to any degree of accuracy.

We performed a deep analysis of the effects of electron beam radiation damage on biological samples at resolutions better than 1 Å using MicroED (Shi et al., 2013; Nannenga et al., 2014b). Using ultra-low exposures allowed us to collect sufficient data from several crystals to allow structure determination. Each crystal was then sequentially exposed to the electron beam and additional structures were determined, from the very same crystals, at increasing levels of total exposure. In this way, we could investigate the effects of exposure on both reciprocal and real space informing us on both global and site-specific damage to the sample. With this approach, we could follow the trends of beam induced damage in biological matter at very high resolutions (better than 1 Å) in a way not previously possible.

Our real-space analysis shows that site-specific damage is apparent at high resolution even with exposures less than $1 \text{ e}^- \text{ \AA}^{-2}$; these exposures are well below those currently used in other cryoEM modalities, for example imaging in single particle EM. This analysis therefore holds important implications for all cryo-EM methods, particularly single particle EM as recently resolutions that approach those commonly observed in crystallography have been reported (Merk et al., 2016).

Imaging in single particle cryo-EM allows the determination of protein structures from a collection of thousands of projection images of individual particles oriented randomly in vitrified ice. When a sufficiently fast camera is available, single-particle cryo-EM data is often collected as a movie. The exposure delivered to the sample reflects a tradeoff between contrast and loss of high-resolution information to radiation damage. Using catalase crystals, it was previously suggested that the optimal tradeoff between signal and damage was $\sim 20 \text{ e}^- \text{ \AA}^{-2}$ for the target resolution of 20 Å, while for 3 Å resolution it was recommended that $10 \text{ e}^- \text{ \AA}^{-2}$ be used (Baker et al., 2010). The total exposure, typically more than $20 \text{ e}^- \text{ \AA}^{-2}$, is fractionated over a sequence of short exposures. This allows individual frames to be corrected for specimen drift and beam-induced movement prior to averaging, while the first frame is usually discarded during processing. The last frames aid alignment, but are excluded from the final average because they contribute little high-resolution information. The average of the first 2 to ~ 5 frames from the movie therefore reflect a superposition of the same particles exposed to electron doses in the range of $2 \text{ e}^- \text{ \AA}^{-2}$ to $10 \text{ e}^- \text{ \AA}^{-2}$ and these

1
2
3
4 are used for the final reconstruction when combined with data from thousands of other
5 particles.
6

7
8 Since the damage mechanisms in all of these EM methods originate from the same
9 phenomena, it is likely that the effects of randomly distributed damage events are washed
10 out during the immense averaging and the use of methods to exclude certain particles from
11 the final reconstruction in single particle EM. However, site-specific damage to acidic side
12 chains has already been observed at 3.2 Å in $10\text{ e}^- \text{Å}^{-2}$ exposures (Bartesaghi et al.,
13 2014). We surmise that the underlying damage to the sample will limit the attainable
14 resolution in single-particle cryo-EM at the doses currently used (Grant & Grigorieff, 2015;
15 Merk et al., 2016) and that with such exposure levels the collected data is of mostly
16 damaged particles. When known structures are reconstructed by single particle EM, even
17 if the resulting map is noisy because of damage, it can still be readily interpreted because
18 the correct answer is available. But for novel structures, where the correct structure is
19 unknown, building structures de novo with noisy maps of damaged protein is very
20 challenging and could be prohibitive.
21
22
23
24

25 Quantification of radiation damage and estimates of crystal lifetime under irradiation not
26 only depend on the sample (e.g. the number of scattering unit cells and their size,
27 composition and thickness of the surrounding mother liquor and the embedding vitrified
28 ice) and the measurement setup (e.g. quanta, dose, and temperature) but also on how data
29 were processed and analyzed. Measures such as the upper resolution limit, B factors, and
30 unit cell volume are often the result of some optimization procedure and may be affected
31 by factors other than the actual damage to the crystal (Kmetko et al., 2006). This is reflected
32 in the literature by the wide spread of dose limits. The D_{50} value of $2.2\text{ e}^- \text{Å}^{-2}$ calculated
33 from reflections of proteinase K in the 21.0–1.7 Å interval is consistent with past
34 measurements using electron diffraction from two-dimensional crystals (Stark et al., 1996)
35 as well as three-dimensional crystals (Unwin & Henderson, 1975; Jeng & Chiu, 1984;
36 Baker et al., 2010). When D_{50} was calculated from the reflections in the 14.0–1.7 Å interval
37 of the much smaller hepta-petide, its value was $2.0\text{ e}^- \text{Å}^{-2}$ and $2.2\text{ e}^- \text{Å}^{-2}$ for high and low
38 exposures, respectively, very close to the value obtained from proteinase K. In line with
39 previous studies in synchrotron X-ray crystallography at comparable flux densities and
40 temperatures, we do not see any effects from the dose rate on the observed global damage
41 (Holton, 2009; Warkentin et al., 2013).
42
43
44
45
46

47 Given that the absorbed dose will ultimately limit the amount of meaningful data that can
48 be extracted from a sample, data collection in the face of radiation damage may be viewed
49 as an optimization problem. The more electrons are delivered to the sample in cryo-EM the
50 stronger the signal, but then noise and damage accumulate and the high-resolution
51 information suffers. Where each sample need only be exposed once, e.g. single-particle
52 cryo-EM or serial femtosecond crystallography (Schlichting, 2015), the exposure can be
53 tuned to maximize the signal-to-noise ratio. When pictures are recorded in cryo-EM a
54 minimum level of exposure is required so that sufficient signal is recorded on the camera
55 to facilitate phase contrast and faithful reconstructions. This need for phase contrast makes
56 it hard to lower the total dose in these experiments where it was recommended to use
57
58
59
60
61
62
63
64
65

1
2
3
4 $10 \text{ e}^- \text{ \AA}^{-2}$ for optimal tradeoff between signal and noise for the target resolution of only 3 \AA
5 (Baker et al., 2010).
6

7
8 In MicroED, where phases are lost and only amplitudes are recorded, the minimum
9 exposure necessary for recording the signal is significantly lower than in single particle
10 EM. This allows data collection at extremely low exposures $<0.01 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$ and entire data
11 sets to be collected from a total exposure less than a single electron per \AA^2 at which point
12 atomic resolution information can be preserved. Our estimated D_{50} rate of approximately
13 $2 \text{ e}^- \text{ \AA}^{-2}$ sets an upper target for single particle cryo-EM experiments seeking to maximize
14 resolution. This is particularly challenging but with increasingly sensitive cameras and
15 better algorithms to allow the use of the first few frames of the recorded movies one might
16 be able to achieve such a feat.
17
18
19
20
21

22 **Materials and methods**

23
24 Frozen-hydrated crystals of proteinase K and a 7-residue peptide were prepared as
25 previously described (Shi et al., 2016; Martynowycz et al., 2017), except the blotting force
26 and time of the FEI (now Thermal Fischer) Vitrobot Mark IV were optimized to position
27 24 and 12 s, respectively, at an environment humidity of 30% for proteinase K and 20
28 and 20 s for GSNQNNF. Electron diffraction datasets from separate crystals were collected
29 using an FEI Tecnai F20 transmission electron microscope operated at 200 kV, with the
30 objective aperture fully open to evenly illuminate an area extending beyond the sample and
31 setting the selected area aperture to closely match the size of the crystal. A frozen-
32 hydrated grid was loaded onto a Gatan 626 cryo-holder and transferred to the
33 microscope, where the specimen temperature was maintained at $\sim 100 \text{ K}$. All diffraction
34 images were acquired using a TVIPS TemCam-F416 CMOS camera and corrected to
35 account for negative pixel values (Hattne et al., 2016) prior to further processing.
36
37
38
39

40 **Proteinase K data collection and processing**

41
42 For each crystal of thickness 200–400 nm, the same 23° wedge was repeatedly collected
43 up to five times by continuously rotating the stage from -12° to $+11^\circ$ (-38° to -15° for
44 crystal 3) off its untilted orientation at a constant rate of $0.089^\circ \text{ s}^{-1}$ (Nannenga et al., 2014b).
45 The rate of electron exposure was adjusted to $0.007 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$, calibrated using a Faraday
46 cage. The individual datasets, each consisting of 49–50 frames with exposure time 5.1 s
47 were recorded at a camera length of 1.2 m, corresponding to a virtual detector distance of
48 2.2 m.
49
50

51
52 Proteinase K data were indexed and integrated in $P4_32_12$ using MOSFLM (Leslie &
53 Powell, 2007). Wedges from six different crystals were merged by the order in which they
54 were collected using AIMLESS (Evans & Murshudov, 2013), and the set of free reflections
55 was copied from the molecular replacement search model, PDB entry 5i9s. Neither of these
56 crystals yield a complete dataset on their own, but since proteinase K does not exhibit any
57 pronounced preferred orientation, this produced five reasonably complete datasets, each
58 comprised of frames with a similar degree of exposure. The choice between intensities
59 derived from summation integration and profile fitting was left up to the optimization
60
61
62
63
64
65

1
2
3
4 algorithm implemented in AIMLESS; in all cases this resulted in profile-fitted intensities
5 being used. Relative B -factors were calculated between merged single-crystal datasets with
6 SCALEIT from the CCP4 suite (Howell & Smith, 1992; Winn et al., 2011).
7

8
9 To quantify these global effects of radiation damage, integrated intensities were averaged
10 for each diffraction image. Averages from different crystals were scaled by a single factor
11 in the range [0.10, 1.0] and simultaneously fit to a common function on the form $A \times \exp(-$
12 $B \times x)$ using the BFGS minimizer implemented in scipy (Oliphant, 2007). The mean
13 effective resolution was calculated by EFRESOL (Urzhumtseva et al., 2013) and used as
14 an objective high-resolution cutoff for all datasets.
15
16

17
18 The first dataset was phased by molecular replacement using MOLREP (Vagin &
19 Teplyakov, 1997) from PDB entry 5i9s, and the solution was reused for all subsequent
20 datasets. Water molecules and ions were excluded from the refined structure: while these
21 improve the quality of the model at high resolution, they are difficult to reliably model
22 once damage degraded the quality of the data. This model was also used to calculate
23 solvent-accessible areas with AREAIMOL (Winn et al., 2011).
24
25

26 All models were refined with REFMAC (Murshudov et al., 2011), with electron scattering
27 factors calculated using the Mott–Bethe formula. The occupancies were set to unity for all
28 atoms and no alternate conformations were used to model partial damage to specific sites
29 of the molecule. Further processing and refinement statistics are given in Table 1.
30
31

32 **GSNQNNF data collection and processing**

33
34 Crystals of the hepta-peptide GSNQNNF that were 100–500 nm thick, were tilted over
35 $\sim 60^\circ$ at a three-fold higher rotation rate (0.3°s^{-1}) than was used for proteinase K and up to
36 12 sweeps were collected from each crystal. To probe the effect of dose rate on radiation
37 damage, peptide data were collected at $0.0028 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$ and $0.0017 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$. These rates
38 were tuned to maximize the number of sweeps collected from an individual crystal. Single
39 crystal datasets comprised of approximately 100 images were collected with an exposure
40 time of 2.1 s and camera length 0.73 m which corresponds to an effective sample to
41 detector distance of 1.2 m. Because two orders of magnitude fewer reflections are typically
42 observed on a diffraction pattern from short segments like GSNQNNF than from
43 proteinase K, intensities were averaged for each dataset instead of for each frame when
44 estimating the effects of global damage on the hepta-peptide.
45
46
47

48
49 The datasets were indexed and integrated in $P1$ with XDS (Kabsch, 2010b) and an
50 isomorphous subset of 10 and 12 crystals for the peptide at $0.0028 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$ and
51 $0.0017 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$, respectively, was scaled and merged with XSCALE (Kabsch, 2010a).
52 Phases for the GSNQNNF data were determined *ab initio* by direct methods from the first
53 collected data set using SHELXD (Sheldrick, 2008). XDSCONV (Kabsch, 2010b) was
54 used on this dataset to assign a free set of reflections, which was subsequently reused for
55 all later peptide datasets. A ligated acetate, three water molecules, and a single zinc atom
56 were included with the GSNQNNF model, because they constitute a significant fraction of
57 the unit cell contents, and all atoms were fixed at full occupancy. Otherwise data collection
58
59
60
61
62
63
64
65

1
2
3
4 and processing were performed as detailed for proteinase K; statistics for the datasets at
5 the high and low dose rates are given in Table 2 and Table 3, respectively.
6
7

8 9 **Acknowledgements**

10
11 The Gonen laboratory is supported by funds from the Howard Hughes Medical Institute.
12 J.A.R. is partly supported by NSF Grant No. DMR 1548924, DOE Grant DE-FC02-
13 02ER63421, and as a Searle Scholar and a Beckman Young Investigator.
14
15
16
17

18 19 **Author Contributions**

20
21 DS, JAR, and TG designed the experiment; DS, CG, and CZ collected data; JH, CZ, MG-J,
22 and MWM analyzed data; JH, DS, JAR, and TG wrote the manuscript with input from
23 all authors.
24
25
26
27

28 29 **Declaration of Interests**

30
31 The authors declare no competing interests.
32
33
34
35

36 37 **References**

- 38 Baker, L. A., Smith, E. A., Bueler, S. A., & Rubinstein, J. L. (2010). *J. Struct. Biol.* **169**,
39 431–437.
40
41
42 Barad, B. A., Echols, N., Wang, R. Y. R., Cheng, Y., Dimaio, F., Adams, P. D., & Fraser, J. S.
43 (2015). *Nat. Methods.* **12**, 943–946.
44
45
46 Bartesaghi, A., Matthies, D., Banerjee, S., Merk, A., & Subramaniam, S. (2014). *Proc.*
47 *Natl. Acad. Sci.* **111**, 11709–11714.
48
49
50
51 Blake, C. C. F. & Phillips, D. C. (1962). 183–191.
52
53
54 Bury, C., Garman, E. F., Ginn, H. M., Ravelli, R. B. G., Carmichael, I., Kneale, G., &
55 McGeehan, J. E. (2015). *J. Synchrotron Radiat.* **22**, 213–224.
56
57
58 Diederichs, K. (2006). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62**, 96–101.
59
60 Evans, P. R. & Murshudov, G. N. (2013). *Acta Crystallogr. D. Biol. Crystallogr.* **69**, 1204–
61
62
63
64
65

1
2
3
4 1214.
5

6 Garman, E. F. (2010). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 339–351.
7

8 Grant, T. & Grigorieff, N. (2015). *Elife.* **4**, e06980.
9

10 Hattne, J., Shi, D., de la Cruz, M. J., Reyes, F. E., & Gonen, T. (2016). *J. Appl. Crystallogr.*
11 **49**, 1029–1034.
12

13 Hayward, S. B. & Glaeser, R. M. (1979). *Ultramicroscopy.* **4**, 201–210.
14

15 Helliwell, J. R. (1988). *J. Cryst. Growth.* **90**, 259–272.
16

17 Henderson, R. (1995). *Q. Rev. Biophys.* **28**, 171.
18

19 Henderson, R. & Unwin, P. N. (1975). *Nature.* **257**, 28–32.
20

21 Holton, J. M. (2009). *J. Synchrotron Radiat.* **16**, 133–142.
22

23 Howell, P. L. & Smith, G. D. (1992). *J. Appl. Crystallogr.* **25**, 81–86.
24

25 Howells, M. R., Beetz, T., Chapman, H. N., Cui, C., Holton, J. M., Jacobsen, C. J., Kirz, J.,
26 Lima, E., Marchesini, S., Miao, H., Sayre, D., Shapiro, D. A., Spence, J. C. H., & Starodub,
27 D. (2009). *J. Electron Spectros. Relat. Phenomena.* **170**, 4–12.
28

29 Hryc, C. F., Chen, D.-H., Afonine, P. V., Jakana, J., Wang, Z., Haase-Pettingell, C., Jiang, W.,
30 Adams, P. D., King, J. A., Schmid, M. F., & Chiu, W. (2017). *Proc. Natl. Acad. Sci.* **114**,
31 3103–3108.
32

33 Jeng, T.-W. & Chiu, W. (1984). *J. Microsc.* **136**, 35–44.
34

35 Jones, G. D. D., Lea, J. S., Symons, M. C. R., & Taiwo, F. a (1987). *Nature.* **330**, 772–773.
36

37 Kabsch, W. (2010a). *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 133–144.
38

39 Kabsch, W. (2010b). *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132.
40

41 Kmetko, J., Husseini, N. S., Naides, M., Kalinin, Y., & Thorne, R. E. (2006). *Acta*
42 *Crystallogr. D. Biol. Crystallogr.* **62**, 1030–1038.
43

44 de la Cruz, M. J., Hattne, J., Shi, D., Seidler, P., Rodriguez, J., Reyes, F. E., Sawaya, M. R.,
45 Cascio, D., Weiss, S. C., Kim, S. K., Hinck, C. S., Hinck, A. P., Calero, G., Eisenberg, D., &
46 Gonen, T. (2017). *Nat. Methods.* **14**, 399–402.
47
48
49
50
51
52
53
54
55

1
2
3
4 Leslie, A. G. W. & Powell, H. R. (2007). *Evolving Methods for Macromolecular*
5
6 *Crystallography* Dordrecht: Springer Netherlands.

7
8 Liebschner, D., Rosenbaum, G., Dauter, M., & Dauter, Z. (2015). *Acta Crystallogr. Sect.*
9
10 *D Biol. Crystallogr.* **71**, 772–778.

11
12 Martynowycz, M., Glynn, C., Miao, J., de la Cruz, M. J., Hattne, J., Shi, D., Cascio, D.,
13
14 Rodriguez, J., & Gonen, T. (2017). *bioRxiv*.

15
16 Merk, A., Bartesaghi, A., Banerjee, S., Falconieri, V., Rao, P., Davis, M. I., Pragani, R.,
17
18 Boxer, M. B., Earl, L. A., Milne, J. L. S., & Subramaniam, S. (2016). *Cell.* **165**, 1698–1707.

19
20 Mitsuoka, K., Hirai, T., Murata, K., Miyazawa, A., Kidera, A., Kimura, Y., & Fujiyoshi, Y.
21
22 (1999). *J. Mol. Biol.* **286**, 861–882.

23
24 Murray, J. & Garman, E. (2002). *J. Synchrotron Radiat.* **9**, 347–354.

25
26 Murshudov, G. N., Skubák, P., Lebedev, A. a, Pannu, N. S., Steiner, R. a, Nicholls, R. a,
27
28 Winn, M. D., Long, F., & Vagin, A. a (2011). *Acta Crystallogr. D. Biol. Crystallogr.* **67**,
29
30 355–367.

31
32 Nannenga, B. L., Shi, D., Hattne, J., Reyes, F. E., & Gonen, T. (2014a). *Elife.* **3**, 1–11.

33
34 Nannenga, B. L., Shi, D., Leslie, A. G. W., & Gonen, T. (2014b). *Nat. Methods.* **11**, 927–
35
36 931.

37
38 Nave, C. & Hill, M. A. (2005). *J. Synchrotron Radiat.* **12**, 299–303.

39
40 Oliphant, T. E. (2007). *Comput. Sci. Eng.* **9**, 10–20.

41
42 Ravelli, R. B. G., Theveneau, P., McSweeney, S., & Caffrey, M. (2002). *J. Synchrotron*
43
44 *Radiat.* **9**, 355–360.

45
46 Sanishvili, R., Yoder, D. W., Pothineni, S. B., Rosenbaum, G., Xu, S., Vogt, S., Stepanov, S.,
47
48 Makarov, O. A., Corcoran, S., Benn, R., Nagarajan, V., Smith, J. L., & Fischetti, R. F.
49
50 (2011). *Proc. Natl. Acad. Sci.* **108**, 6127–6132.

51
52 Schlichting, I. (2015). *IUCrj.* **2**, 246–255.

53
54 Sheldrick, G. M. (2008). *Acta Crystallogr. A.* **64**, 112–122.

- 1
2
3
4 Shi, D., Nannenga, B. L., Iadanza, M. G., & Gonen, T. (2013). *Elife*. **2**, e01345.
5
6 Shi, D., Nannenga, B. L., de la Cruz, M. J., Liu, J., Sawtelle, S., Calero, G., Reyes, F. E.,
7 Hattne, J., & Gonen, T. (2016). *Nat. Protoc.* **11**, 895–904.
8
9 Stark, H., Zemlin, F., & Boettcher, C. (1996). *Ultramicroscopy*. **63**, 75–79.
10
11 Unwin, P. N. T. & Henderson, R. (1975). *J. Mol. Biol.* **94**, 425–440.
12
13 Urzhumtseva, L. M., Klaholz, B., & Urzhumtsev, A. G. (2013). *Acta Crystallogr. D. Biol.*
14 *Crystallogr.* **69**, 1921–1934.
15
16 Uyeda, N., Kobayashi, T., Ishizuka, K., & Fujiyoshi, Y. (1980). *Nature*. **285**, 95–97.
17
18 Vagin, A. & Teplyakov, A. (1997). *J. Appl. Crystallogr.* **30**, 1022–1025.
19
20 Warkentin, M., Hopkins, J. B., Badeau, R., Mulichak, A. M., Keefe, L. J., & Thorne, R. E.
21 (2013). *J. Synchrotron Radiat.* **20**, 7–13.
22
23 Weik, M., Ravelli, R. B. G., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P.,
24 Silman, I., Kroon, J., & Sussman, J. L. (2000). *Proc. Natl. Acad. Sci.* **97**, 623–628.
25
26 Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan,
27 R. M., Krissinel, E. B. E. B., McCoy, A. J., McNicholas, S. J., Murshudov, G. N., Pannu, N. S.,
28 Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. A., Wilson, K. S., & Leslie, A. G. W.
29 (2011). *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 235–242.
30
31 Yonekura, K., Kato, K., Ogasawara, M., Tomita, M., & Toyoshima, C. (2015). *Proc. Natl.*
32 *Acad. Sci.* **112**, 3368–3373.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figures and Tables

Figure 1: Exposure dependency of the mean intensity of the unscaled integrated reflections (a) proteinase K, (b) the hepta-peptide, GSNQNNF, recorded at a exposure rate of $0.0028 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$, and (c) GSNQNNF at $0.0017 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$. Spots at high resolution fade significantly faster than spots at low resolution. For proteinase K, the electron exposure at which the mean intensity across the entire recorded resolution range is reduced to 70%, 50%, and e^{-1} of its extrapolated value at zero dose is estimated to be $D_{70} = 1.1 \text{ e}^- \text{ \AA}^{-2}$, $D_{50} = 2.2 \text{ e}^- \text{ \AA}^{-2}$, and $D_e = 3.1 \text{ e}^- \text{ \AA}^{-2}$. The corresponding exposures for the $0.0028 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$ and $0.0017 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$ peptide datasets are $D_{70} = 0.73 \text{ e}^- \text{ \AA}^{-2}$, $D_{50} = 1.4 \text{ e}^- \text{ \AA}^{-2}$, $D_e = 2.1 \text{ e}^- \text{ \AA}^{-2}$ and $D_{70} = 0.83 \text{ e}^- \text{ \AA}^{-2}$, $D_{50} = 1.6 \text{ e}^- \text{ \AA}^{-2}$, $D_e = 2.3 \text{ e}^- \text{ \AA}^{-2}$, respectively.

Figure 2: Exposure dependency of the unit cell volume, V_{uc} , and relative B -factor, B_{rel} V_{uc} and B_{rel} were averaged across all the crystals at each exposure. For B_{rel} only reflections in a sufficiently large resolution range common to all datasets were considered (20.8–3.20 Å for proteinase K; 14.0–1.2 Å for GSNQNNF).

Figure 3: Disulfide bond breakage and decarboxylation of acidic side chains indicate site-specific radiation damage in proteinase K $2mF_o-DF_c$ maps (blue meshes) are contoured at 1.5σ above the mean, mF_o-DF_c difference densities (green/red meshes) are contoured at $\pm 3 \sigma$ above/below the mean. Maps up to and including those calculated at $4.3 \text{ e}^- \text{ \AA}^{-2}$ use data extending to 2.2 Å; the two maps at the highest exposure only use reflections up to 3.2 Å. Densities are carved to 2 Å around the selected atoms. All figures were generated using PyMol (Schrödinger LLC, 2014).

Figure 4: Exposure dependency on the hepta-peptide density $2mF_o-DF_c$ (blue meshes, countered at 1.5σ above the mean) and mF_o-DF_c density (red/green meshes, contoured at $\pm 3 \sigma$ above/below the mean) from the GSNQNNF hepta-peptide at $0.0017 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$. The maps at electron exposures up to and including $1.5 \text{ e}^- \text{ \AA}^{-2}$ are calculated using all observed reflections to 1.01 Å; remaining maps use reflections up to 1.45 Å. The atomic displacement parameter of the Zn atom, ADP_{Zn} , (purple sphere to the left), increases steadily over the course of exposure.

Figure 5: Accumulated density loss Density loss in arbitrary units for all the amino acids, ligands, and ions present in the refined models of (a) proteinase K and (b) the hepta-peptide. The entities are sorted in the approximate order of initial damage onset. The number in parentheses denotes the occurrence of the respective amino acid in the structure. Only reflections in the resolution range common to all datasets are considered (20.8–3.20 Å for proteinase K; 14.0–1.45 Å for GSNQNNF).

Table 1: Processing and refinement statistics for proteinase K. The exposure time of a multi-crystal dataset is denoted $\langle t_{\text{exp}} \rangle$; it is defined as the mean cumulative irradiated time of all the frames in the dataset. Numbers in parentheses refer to the highest-resolution shell for either merging or refinement. Owing to varying response to radiation, damage-induced non-isomorphism causes the number of datasets that can be merged to decrease at the higher exposures.

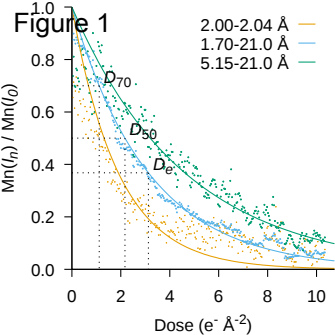
Proteinase K	Set 1	Set 2	Set 3	Set 4	Set 5
Number of crystals	6	6	6	5	4
$\langle t_{\text{exp}} \rangle$ (s)	123.1	370.0	617.3	862.6	1107.9
Resolution (Å)	20.74–1.71 (1.74–1.71)	20.78–1.91 (2.05–1.91)	20.81–2.19 (2.27–2.19)	20.87–2.80 (2.95–2.80)	2.92–3.19 (3.46–3.19)
Completeness (%)	93.4 (71.9)	96.9 (97.0)	94.3 (94.6)	88.4 (87.7)	76.8 (77.7)
Multiplicity	6.1 (4.9)	6.2 (6.3)	5.7 (5.8)	5.6 (5.8)	3.3 (3.4)
$CC_{1/2}$	0.950 (0.171)	0.947 (0.135)	0.925 (0.097)	0.908 (0.146)	0.831 (0.184)
$R_{\text{work}} / R_{\text{free}}$	22.20 / 25.35	21.91 / 25.97	24.44 / 30.49	23.82 / 32.48	21.01 / 32.39

Table 2: Processing and refinement statistics for the hepta-peptide GSNQNNF at $0.0028 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$. Rows as per Table 1.

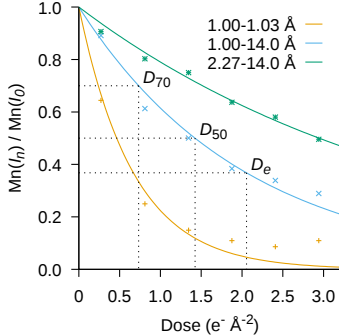
GSNQNNF ($0.0028 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$)	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6
Number of crystals	10	10	10	10	3	3
$\langle t_{\text{exp}} \rangle$ (s)	96.7	289.1	480.1	670.1	860.5	1050.5
Resolution (Å)	13.95–1.01 (1.03–1.01)	13.92–1.01 (1.03–1.01)	13.94–1.01 (1.03–1.01)	14.01–1.15 (1.19–1.15)	13.85–1.26 (1.48–1.26)	6.53–1.37 (1.48–1.37)
Completeness (%)	80.7 (72.5)	81.6 (71.9)	82.4 (68.9)	83.4 (86.1)	56.1 (56.6)	63.5 (61.1)
Multiplicity	6.9 (6.2)	6.9 (6.5)	6.6 (5.4)	6.5 (6.6)	3.3 (3.3)	2.7 (2.8)
$CC_{1/2}$	98.8 (91.1)	98.9 (91.5)	98.6 (60.0)	98.9 (79.1)	96.1 (44.6)	3.3 (14.2)
$R_{\text{work}} / R_{\text{free}}$	17.49 / 15.18	18.88 / 18.18	24.68 / 22.82	26.36 / 29.87	36.54 / 32.75	36.45 / 43.44

Table 3: Processing and refinement statistics for the hepta-peptide GSNQNNF at $0.0017 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$. Rows as per Table 1.

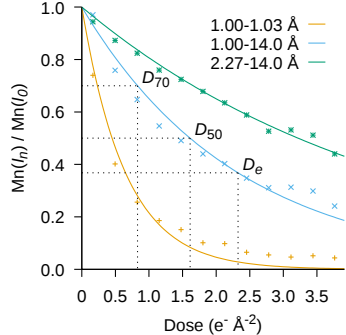
GSNQNNF ($0.0017 \text{ e}^- \text{ \AA}^{-2} \text{ s}^{-1}$)	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10	Set 11	Set 12
Number of crystals	12	12	12	12	5	12	11	12	12	11	11	11
$\langle t_{\text{exp}} \rangle$ (s)	97.2	291.3	484.6	677.1	869.0	1060.6	1251.0	1442.9	1636.2	1832.2	2024.3	2214.0
Resolution (Å)	11.48–1.01 (1.03–1.01)	13.97–1.01 (1.03–1.01)	13.89–1.01 (1.03–1.01)	13.89–1.02 (1.03–1.02)	13.86–1.01 (1.03–1.01)	13.86–1.01 (1.03–1.01)	13.70–1.13 (1.15–1.13)	13.76–1.16 (1.20–1.16)	13.71–1.21 (1.26–1.21)	13.72–1.31 (1.38–1.31)	13.67–1.46 (1.59–1.46)	13.68–1.45 (1.57–1.45)
Completeness (%)	79.4 (72.7)	79.2 (71.4)	79.0 (70.8)	76.3 (64.4)	74.3 (67.0)	77.8 (64.0)	74.4 (76.8)	78.6 (81.2)	80.8 (81.0)	79.4 (79.9)	77.5 (78.5)	79.8 (67.6)
Multiplicity	8.6 (7.5)	8.4 (6.5)	8.5 (7.0)	8.5 (7.2)	3.7 (3.3)	8.1 (5.1)	7.5 (7.7)	7.8 (8.3)	7.7 (6.8)	7.4 (6.7)	7.7 (7.9)	7.6 (7.9)
$CC_{1/2}$	98.3 (92.0)	98.2 (92.8)	98.0 (83.3)	98.4 (66.5)	98.2 (37.3)	95.7 (44.3)	86.2 (70.1)	98.2 (75.0)	89.5 (69.9)	95.8 (60.1)	93.9 (31.9)	89.9 (71.1)
$R_{\text{work}} / R_{\text{free}}$	17.92 / 16.78	18.15 / 16.80	18.67 / 17.69	19.82 / 20.49	23.60 / 24.47	24.55 / 26.72	21.49 / 22.13	24.70 / 30.84	27.49 / 31.61	24.14 / 33.40	23.22 / 35.72	23.84 / 32.74



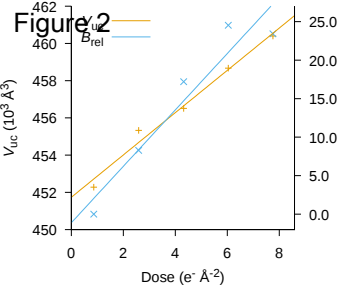
(a)



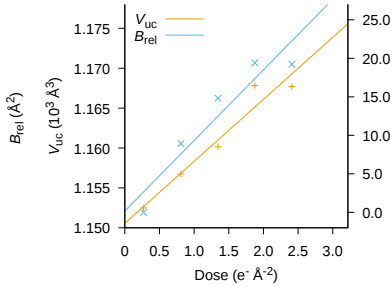
(b)



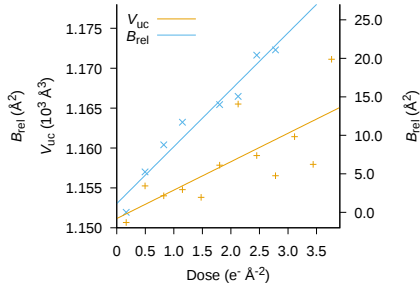
(c)



(a)



(b)



(c)

Figure 3 $0.86 e^{-\text{\AA}^{-2}}$

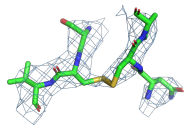
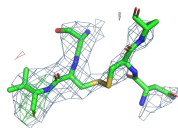
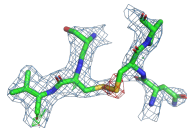
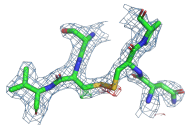
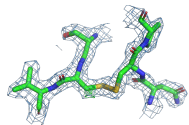
$2.59 e^{-\text{\AA}^{-2}}$

$4.32 e^{-\text{\AA}^{-2}}$

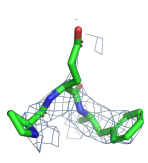
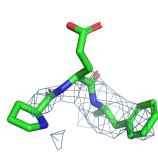
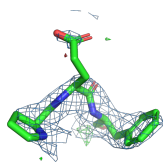
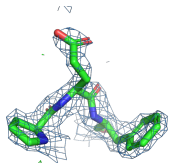
$5.78 e^{-\text{\AA}^{-2}}$

$7.75 e^{-\text{\AA}^{-2}}$

Cys 139-Cys 228



Glu 153



Asp 170

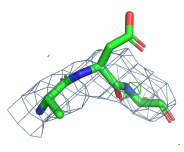
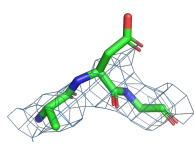
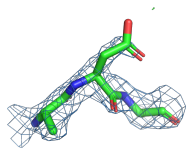
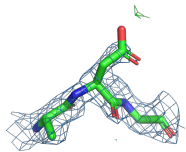
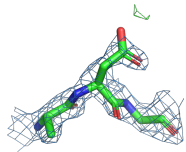
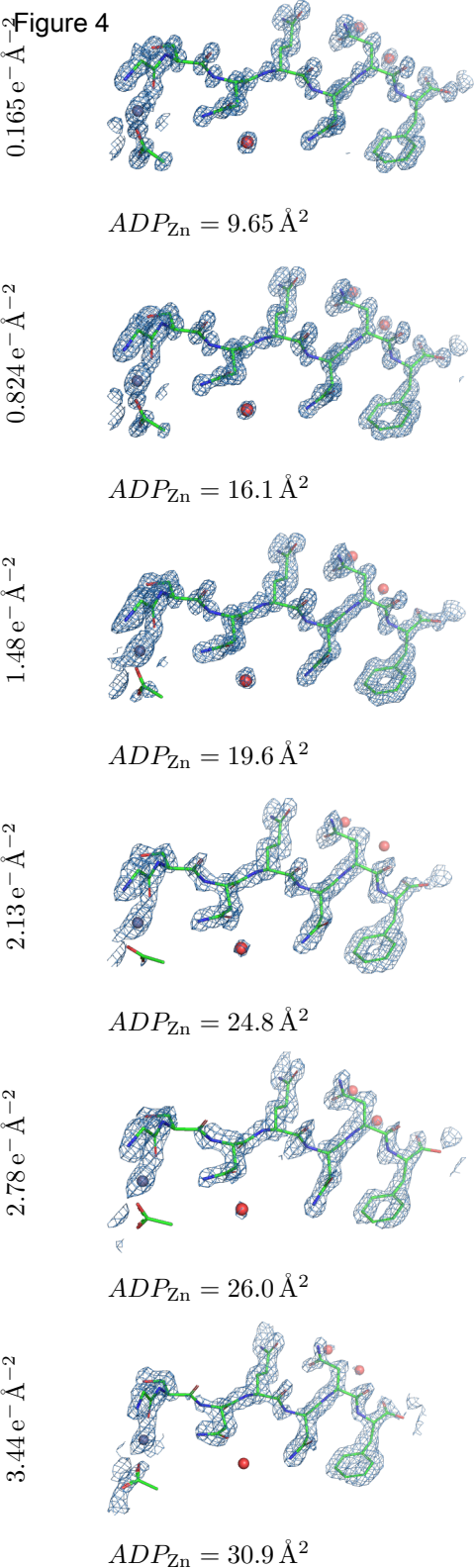
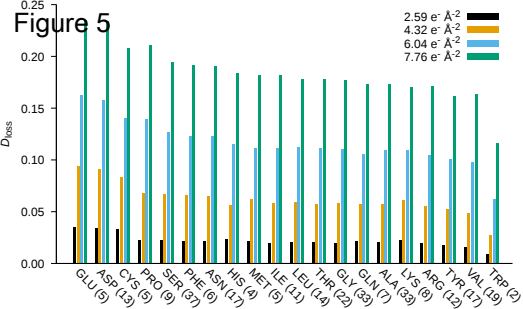
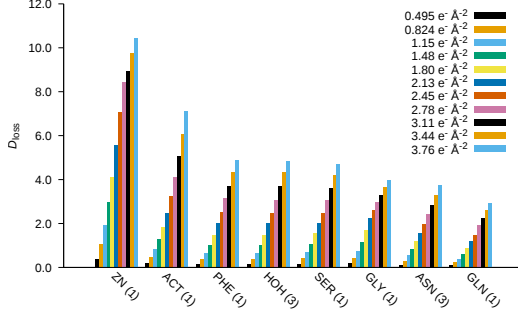


Figure 4





(a)



(b)