Supplementary Materials for
Atomically Resolved Trajectories of a Bimolecular Photochemical Reaction

Authors: R. Patrick Xian[1,2†], Stuart A. Hayes[1,2†], Gastón Corthey[2,3] Carole A. Morrison[4], Alexander Marx[2], Hazem Daoud[1], Cheng Lu[1] and R. J. Dwayne Miller[1*]

Affiliations:
[1] Departments of Chemistry and Physics, University of Toronto, 80 St. George Street, Toronto M5S 3H6, Canada.
[2] Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, 22761 Hamburg, Germany.
[3] Instituto de Nanosistemas, Escuela de Bio y Nanotecnologías, Universidad Nacional de San Martín, Av. 25 de Mayo 1021, San Martín, Buenos Aires 1650, Argentina.
[4] School of Chemistry and EaStCHEM Research School, University of Edinburgh, The King's Buildings, West Mains Road, Edinburgh EH9 3JJ, UK.
[†] These authors contributed equally to the work.
[*] dmiller@lphys.chem.utoronto.ca

Contents

S1 Optimisation of measurement conditions for femtosecond electron diffraction

Due to the sensitivity of the sample to environmental conditions such as vacuum, temperature and laser excitation parameters, before comprehensively studying its time-resolved dynamics, we carefully investigated the effect of these and chose the key parameters to balance sample reversibility and the signal-to-noise ratio (SNR).

Vacuum stability. Tests of sample conditions were carried out with an electron gun with radiofrequency (RF) compression at the University of Toronto (*Gao et al.,* 2013) as well as a compact gun at the Max Planck Institute for the Structure and Dynamics of Matter in Hamburg (*Ishikawa et al.* 2015). It was found that the sample quality degrades significantly over the course of a day at room temperature when placed inside the vacuum chamber (background residual pressure at ~ 10^{-7} mbar), as shown in Fig. S1a-b. However, mild cooling to 270 K can completely stop the vacuum decay (see Fig. S1c-d). During the down time between consecutive days of measurements, the sample temperature was kept at this temperature using a Peltier thermoelectric device sandwiched between the sample holder and the manipulator arm. The configuration is shown in Fig. S1e.

Measurement temperature. To the best of our knowledge, there are no recorded phase transitions for $TBAI_3$ between 77 K and 300 K. This is also supported by our observation that the temperature only modifies the diffraction intensity to some extent. Eventually, we chose 173 K as the measurement temperature, regulated by a PID controller (LakeShore) to stabilize the temperature within 0.1 °C.

Laser repetition rate. We used a custom-designed chopper blade to lower the repetition rate of the laser from 1kHz to 125 Hz, the same as that documented in the previous optical study. This repetition rate strikes a good balance between measurement speed and sample decay rates. Even though the sample response still shows slow optical damage over time (see Fig. S2a), the number of tolerable laser shots (and therefore diffraction patterns) is sufficient for producing high-quality data for structural inversion. The measurement duration of each dataset at a sample location typically lasts between 8-10 hours.

Fluence dependence and sample reversibility. A series of measurements at fixed pump-probe delays were carried out with pump fluences ranging from 0.21 to 0.43 mJ/cm^2. The results of selected diffraction peaks (see Fig. S2b) show a linear dependence of the dynamics, measured at the probe delay of ~ 1 ps, on pump fluence. Although higher fluence leads to higher intensity changes, they generally lead to faster sample decay, therefore, we used a fluence of ~0.25 mJ/cm^2 for the time-resolved experiments.

Polarization dependence. A number of measurements were conducted at a fixed time delay with different linear pump polarization angles to optimize the observed signal (amplitude of change in diffraction intensities). The general characteristics of these measurements (see Fig. S3) show a maximal change when the laser is polarised aligned along the long axis of the sample, parallel to the iodine chains.

S2 Data processing details

After collecting several datasets and performing exploratory data analysis, two datasets were chosen for fitting a series of atomic coordinates. These were collected at two orientations which were determined as described in the methods as (fractional coordinates):

- Beam direction of sample 1: [1, 0.0636, -0.3907]
- Beam direction of sample 2: [1, -0.1244, 0.5336]

The processing of the raw image data involved the following steps:

- **Dark-field correction** to account for the bias on the CCD detector, dark current accumulated during acquisition and any stray background light from the room. During the course of an experiment, the lighting in the room was kept low and unchanged.
- **Flat-field correction** for non-uniformity of the scintillator and optical coupling to the CCD. This was calibrated by scattering the electron beam from a 2 um thick film of aluminium in transition geometry with the magnetic lens turned off.
- **Beam-stop masking**
- **Outlier rejection** - This was based on the sum of pixel intensities for each image and was necessary due to occasional discharges from the electron gun. These could be easily identified by calculating the median and inter-quartile range (IQR) for a subset of images and rejecting those exceeding the median by more than 3*IQR.
- **Radial background subtraction** - A radial background was calculated using the average of all the ground-state patterns for a particular time point. This was determined by calculating the 5th percentile of pixel intensities at each integer number of pixels from the centre of the pattern and applying a minimal smoothing to remove noise. The same radial background subtraction was applied to both the ground-state and excited-state patterns for a given time point to prevent any uncertainty in the background influencing the absolute changes in Bragg peaks.
- **Peak finding and ROI (region of interest) creation**. The first ground-state image in the series was used to locate the most prominent peaks and define ROIs for extracting intensities for each reflection. ROIs were defined as circular spots with a single diameter for all peaks in a data set that approximately matched the observed peak diameter.
- **Data reduction**. ROI Intensities for each image with and without excitation were calculated and from these a mean intensity and a standard error were calculated for a given time point, yielding matrices of intensities, $\mathbf{I}^{\text{pump-on}}$ and $\mathbf{I}^{\text{pump-off}}$, and standard errors expressed as relative to the absolute values, $\sigma^{\text{pump-on}}$ and $\sigma^{\text{pump-off}}$. From these, the intensity ratios, $\mathbf{R}^{\text{obs}}$ were calculated

as defined in the methods. The mean peak uncertainties were calculated using the following equation:

$$\sigma_i = \frac{1}{n_t} \sum_{t=1}^{n_t} \sqrt{(\sigma_{i,t}^{\text{pump-on}})^2 + (\sigma_{i,t}^{\text{pump-off}})^2}$$

FED data would often also require subtraction of pump laser background and correction for beam-pointing instabilities, but in this case, the aluminium coating on the scintillator effectively blocked all of the laser light at 400 nm, so no scattered light from the pump laser could be detected. The position of the diffraction pattern on the detector did not move significantly from image to image or over the course of an experiment, so no correction was required for beam pointing instabilities.

Both datasets spanned a time range of 10 ps and used equal laser fluence of 0.25 mJ/cm^2 for the excitation, but were acquired with slightly different time steps of 120 and 150 fs respectively. To adjust for this mismatch in temporal sampling, dataset 2 was interpolated to 120 fs and the mean uncertainties were scaled by a factor of sqrt(150/120) to account for the artificial increase in the number of data points. The ends of the datasets that did not overlap were truncated, so that the combined dataset spanned a time window of 8.64 ps. Any peaks that differed in intensity for the ground state by over 50% from the theoretical values were removed from further analysis, resulting in a total of 260 peaks at 73 time points.

S3 Hyperparameter tuning

Hyperparameter tuning was performed using five-fold cross-validation where the data was first partitioned into two equally sized sets according to the RMS signal-to-noise ratios (SNRs) for each peak, defined as the root-mean-square intensity change divided by the uncertainty, with one for high and another for low SNR. Each of these sets were then randomly divided into five folds with equal sizes and the two sets with high and low SNRs were recombined to form five folds. Five splits were then created, each containing one fold as test data and four folds of training data. The model was fitted to the training data and the root-mean-squared error (RMSE) was calculated on both the training and test data for a range of hyperparameter values. Each split was treated independently throughout the fitting procedure, resulting in five sets of coordinates at each step. The hyperparameter values that were considered optimal were those for the lowest test error based on the average of all five splits.

In step 1, λ_1 , λ_2 and λ_{bonds} were treated independently and the RMSE values are displayed in figure S4. In step 2, λ_1 and λ_2 were set to be equal and λ_{bonds} was independent as shown in figure S5. In step 3, the number of refinement parameters increased dramatically, so only the positional restraint was tested, with λ_{bonds} fixed at a value of 10, resulting in an optimal value of $\lambda_1 = \lambda_2 = 15$ as shown in figure S6. The RMSE values for each fold and step of the refinement are displayed in figure S7. As a reference, an RMSE of 1 is the estimated uncertainty in the raw diffraction data although the experimental uncertainty appears to be slightly overestimated since the

data at negative time points has an RMSE of approximately 0.7, likely due to a correlation of errors between the signals with and without excitation, while the data reduction step treated these independently. The RMSE for step 1 was calculated both before and after projecting the SVD components into the time domain. In the first case, using the to facilitate comparison with steps 2 and 3 and shows remarkably good agreement given that assumptions made. It is visible that while the training error decreases between steps 1, 2 and 3, the test error is fairly constant. This indicates that, in this case, fitting the SVD components results in a remarkably accurate structures, which is surprising since diffraction intensities are not linear functions of the atomic positions. Both methods of selecting the hyperparameters produced similar results in terms of the structural dynamics. Estimates of the uncertainties of inter-atomic distances are shown in figure 3 of the main text are the standard deviation for these parameters for the five $\mathbf{x}^{S3}$ trajectories obtained during cross validation.

S4 Theoretical structures and supplementary discussion

The resulting geometries for the Chain-A and -B triplet surface calculation are shown in Figure S8. The theoretical structures support the conclusion from FED and transient absorption that I_4^- is indeed a stable species at least in the context of this crystalline environment and the agreement seems reasonable given the assumptions made in modelling both the quantum chemistry and the electron diffraction data. There remains some discrepancies in the exact bond lengths that could be attributed to any of a number of differences when comparing theory and experiment, and we list these here to make readers aware of the underlying assumptions and also scope for future studies.

The first difference that should be noted is that for frequencies faster than the instrument response, FED samples the averaged structure of vibrationally excited species, which may still be far from equilibrium. These need not be centred at the minimum of the excited state potential energy surface, since it may take time for the primary reaction to induce structural rearrangements of the neighbouring species. In contrast, the theoretical structure is the minimum on the potential energy surface. Due to the limited information, we adopted the simplest model for describing the FED experiment that made sense of the data and focuses on the primary reaction rather than the extended chain. However, it is possible that the pathway back to the ground state proceeds via more delocalised structural changes, where the distinction between individual molecules becomes more blurred. To check this it would be desirable to use time-resolved crystallography in the picosecond time domain with fewer time points but a complete sampling of reciprocal space.

The second noteworthy difference is the appearance of I_5 units in the calculation. This could be an artefact of the supercell size, and the convergence could be checked using larger supercells. Transient spectroscopy (*Xian et al.,* 2017) also showed clear signs of I_2^- vibrational dynamics, so there cannot be I_5^{2-} present on the timescale of these experiments, but could form af-

ter dissipation of the energy in I_2^- into the lattice. I_5^- is also known to be stable in the ground state, as is I_4^{2-} which can also be considered an [$I^- \cdot I2 \cdot I^-$] adduct (see for example *Svensson et al.,* 2003), thus a plausible route back to the ground state is via
$I_3^- + I_2^- + I_4^-$ (triplet) → $I_5^{2-} + I_4^-$ (triplet) → $I_5^- + I_4^{2-}$ (singlet).

Supplementary Figures

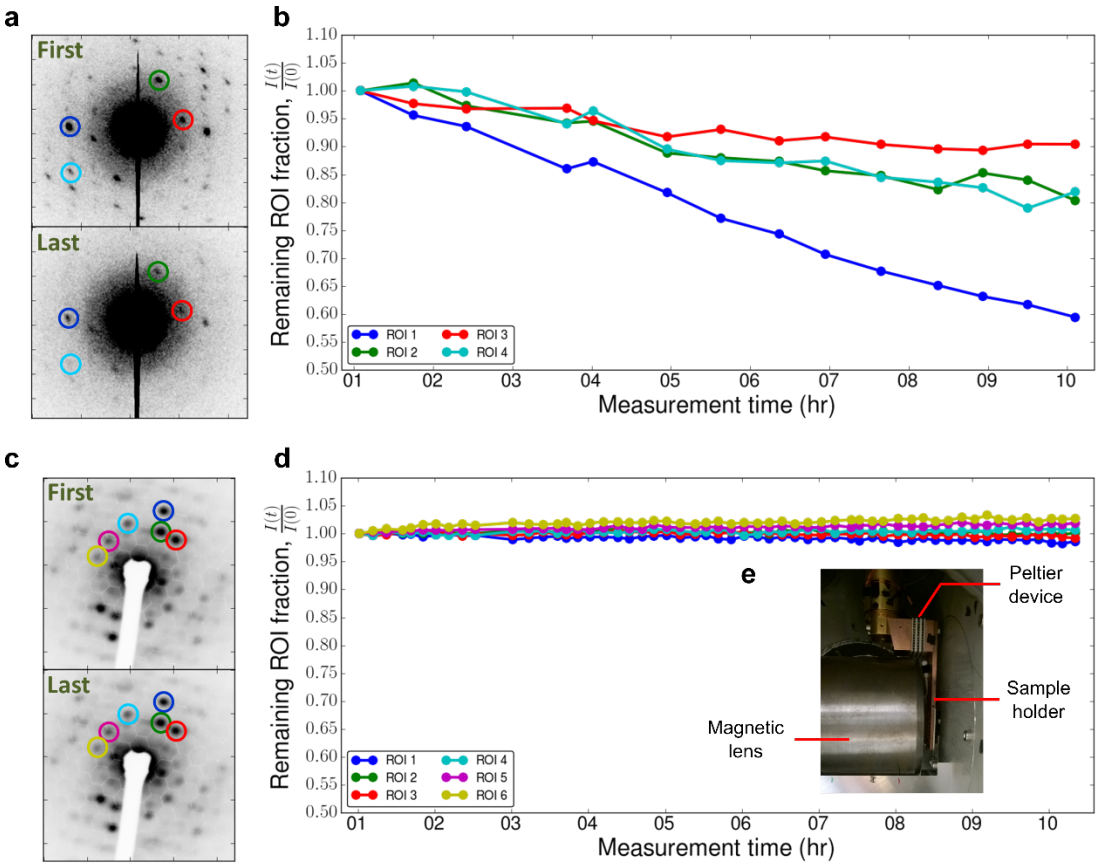

Figure S1 Temperature stabilization and $TBAI_3$ sample vacuum stability. The results in **a**,**b** were measured in Toronto with an RF electron gun but without temperature stabilization, while the results in **c**,**d** were measured in Hamburg with a DC electron gun and in-vacuum temperature stabilization with a Peltier device glued between the manipulator and the sample stage. The hardware arrangement is depicted in **e**.

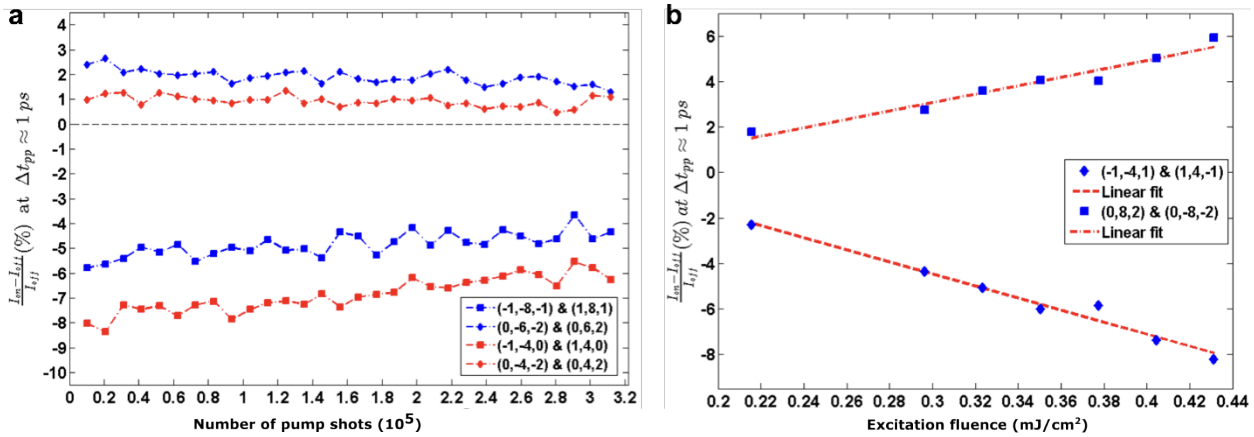

Figure S2 Fluence dependence and reversibility. **a**, Fluence dependence of two selected Friedel pairs and their respective linear fits. **b**, Decay of relative changes averaged over selected Friedel pairs. In both sets of measurements, the pump-probe delays are fixed at ~ 1 ps.

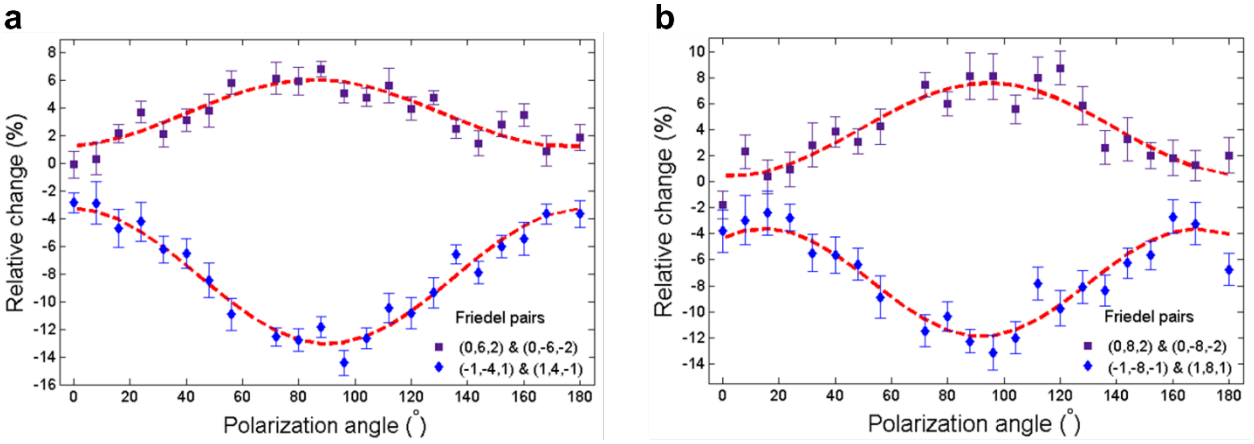

Figure S3 Polarization dependence of electron diffraction intensity changes. The measurements were conducted at a fixed time delay of ~1 ps after time zero. The polarization dependence of four Friedel pairs is shown in **a** and **b**. Their signals reach a maximum deviation at the same polarization angle.

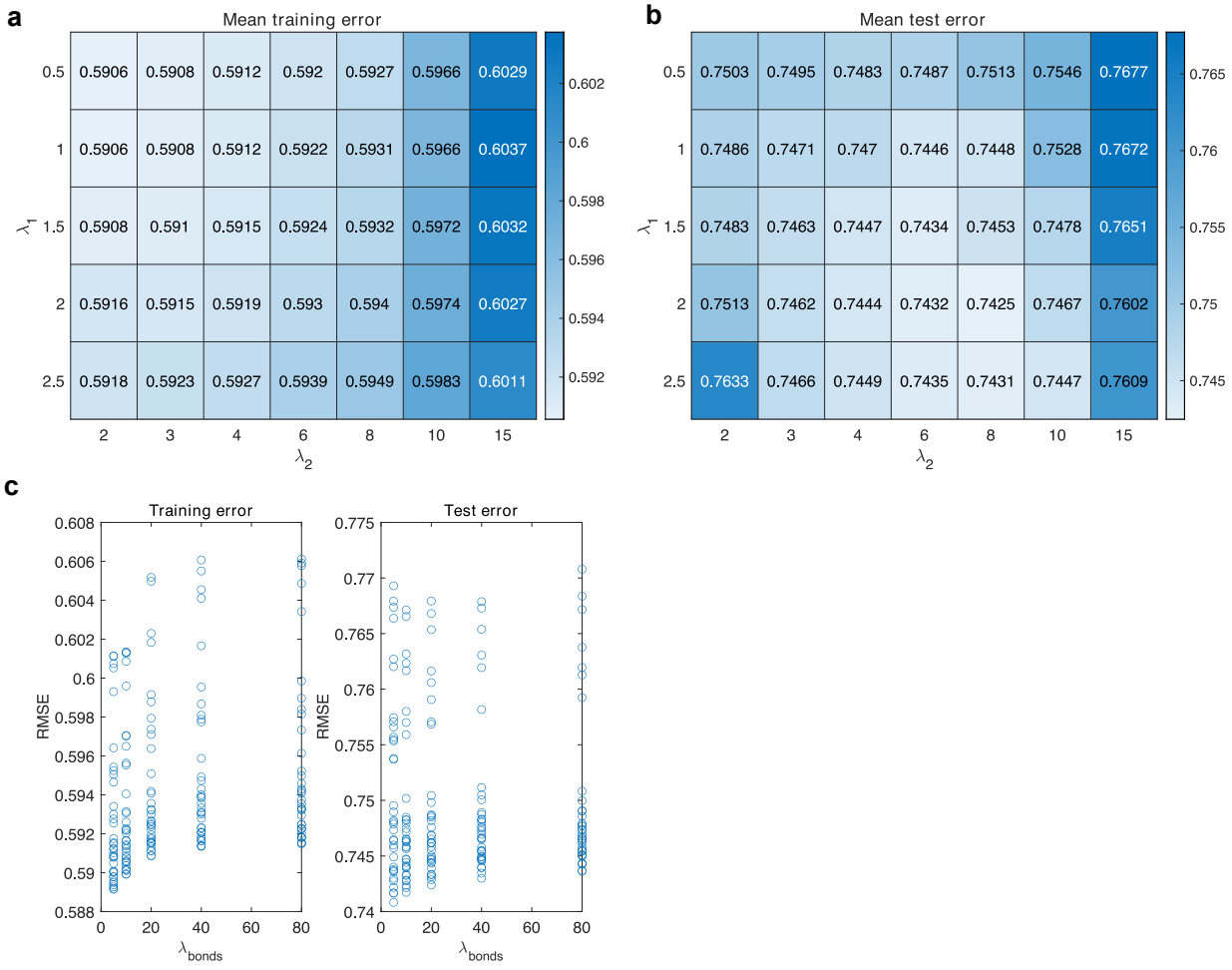

Figure S4 Hyperparameter testing for step 1. The RMSE for the (**a**) training and (**b**) test sets averaged over all folds. These are based on the fits to only the first three SVD components. As expected, weaker restraints allow a better fit to the training data, but the test set shows the optimum values to be $\lambda_1 = 2$ and $\lambda_2 = 8$. (**c**) Dependence of RMSE on λ_{bonds}.

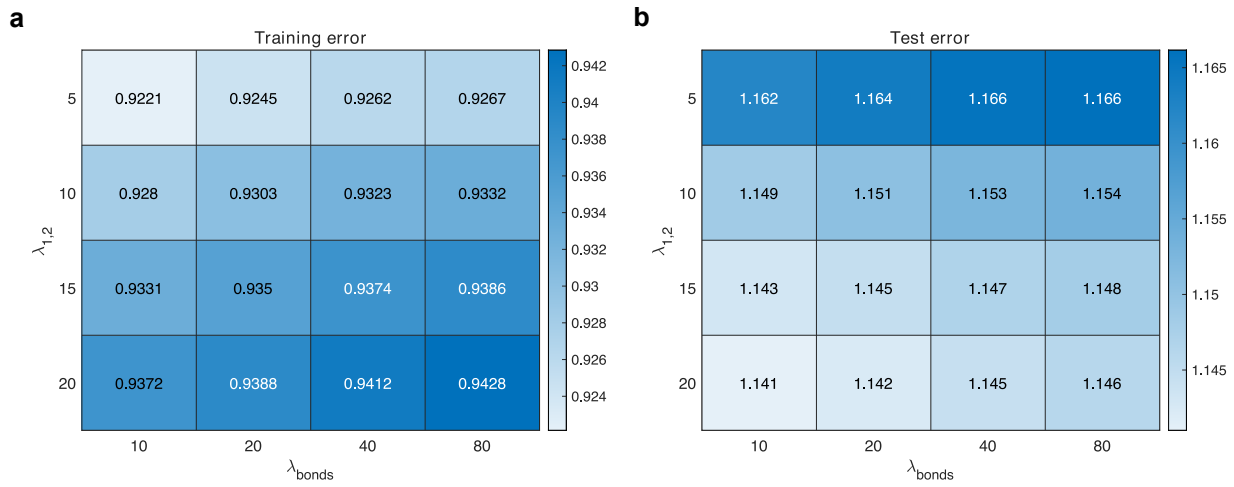

Figure S5 Hyperparameter testing for step 2. The RMSE for the (**a**) training and (**b**) test sets averaged over all folds. The restraint strength on the atomic positions (λ_1 and λ_2) were set to be equal.

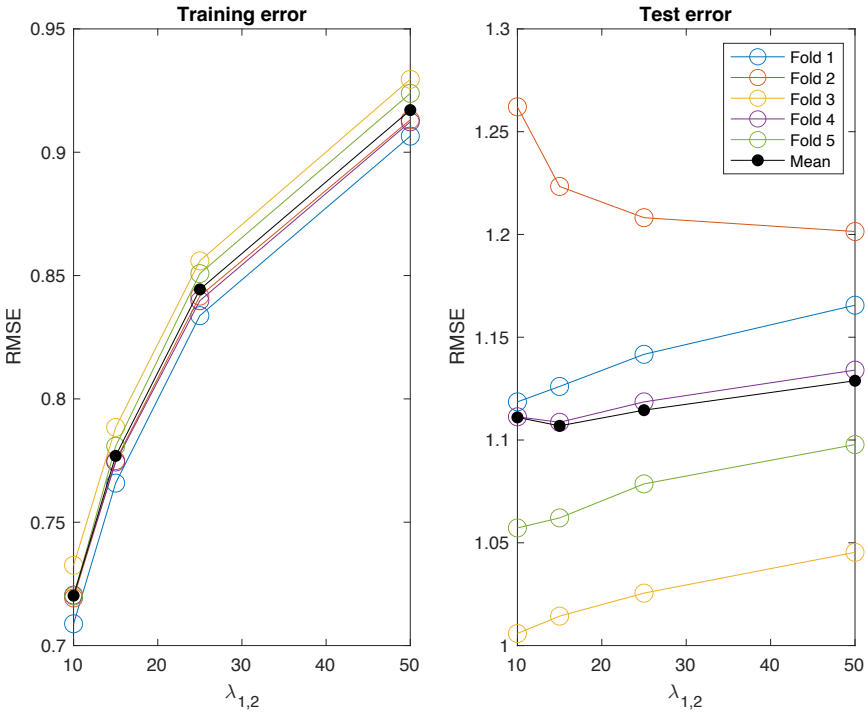

Figure S6 Hyperparameter testing for step 3. The restraint strength on primary and secondary atomic positions, λ_1 and λ_2 were set to be equal. The output from step 1, $\mathbf{x}^{S1}$, was used as the prior estimate, and the output from step 2 was used as the starting point. λ_{bonds} was fixed at a value of 10.

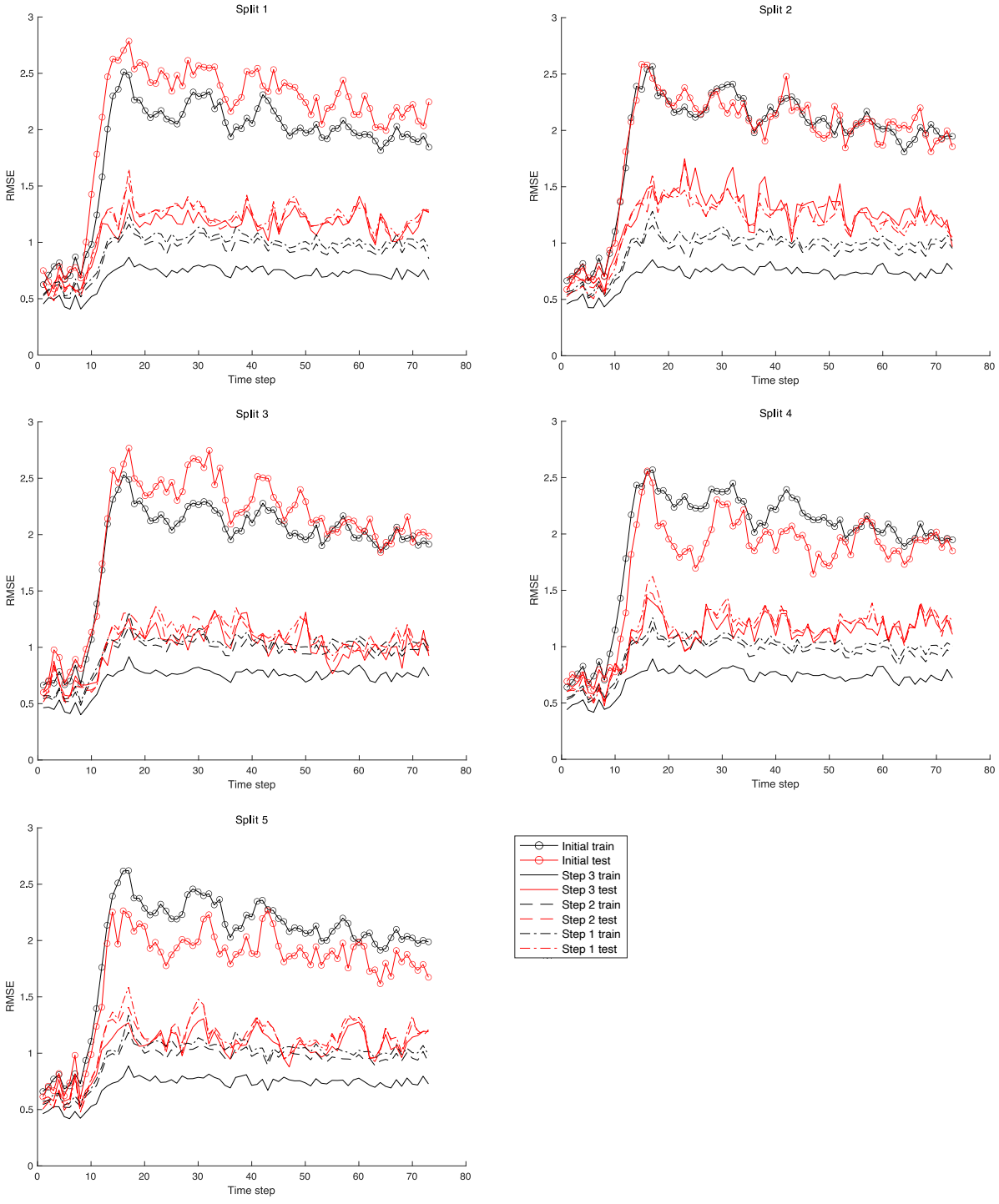

Figure S7 Root-mean-square errors (RMSE). The RMSE for each train-test split and step of the refinement plotted as a function of the time variable, t. The training data is shown in black and test data is plotted in red.

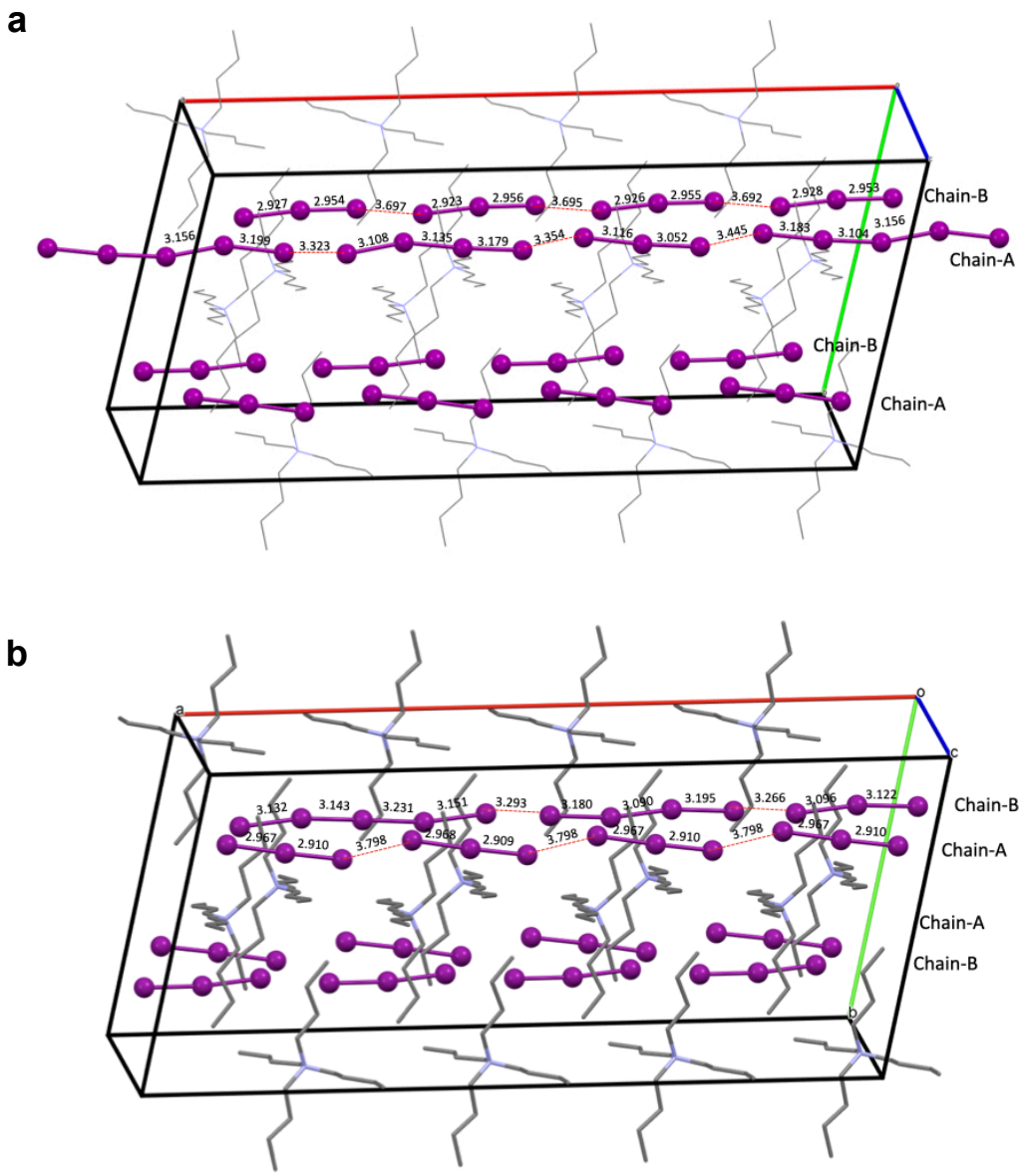

Figure S8 Theoretical metastable structures following dissociation. $TBAI_3$ is represented as a 4×1×1 supercell on the triplet potential energy surface. The starting points were created by elongating an I-I bond in either (**a**) chain A or in (**b**) chain B. Bond lengths are shown in Ångstroms.