Large Scale Data Analysis for Serial Femtosecond X-ray Crystallography

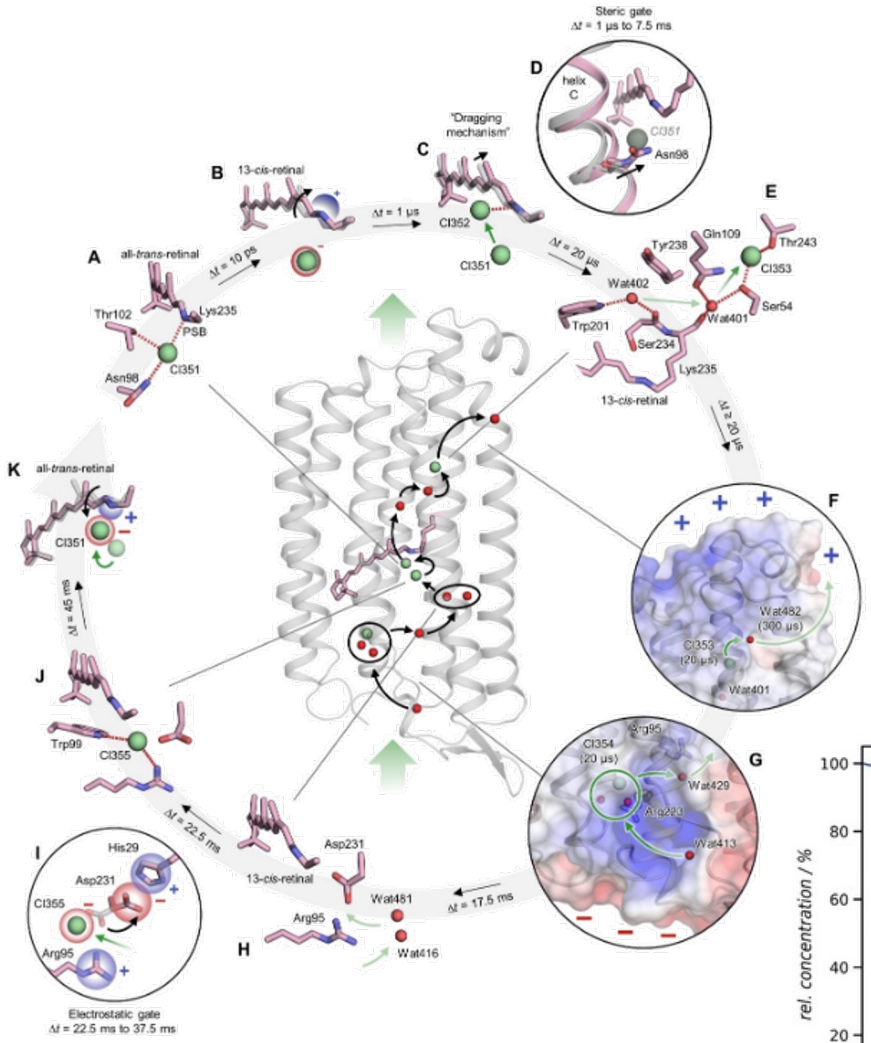
A Facility Perspective

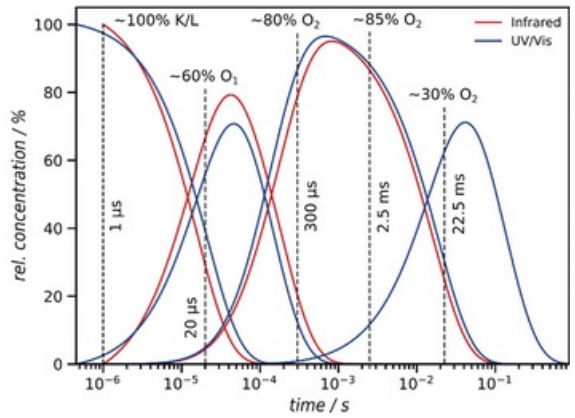
Sandra Mous

SLAC NATIONAL ACCELERATOR LABORATORY

Stanford University | U.S. DEPARTMENT OF ENERGY

Recording molecular movies of charge transport across the membrane





- 6 XFEL structures for a “movie”
- 100’000 indexed diffraction patterns per structure
- 20% hit rate, with 80% indexing rate of hits
  - 16% overall indexing rate
- At 120 Hz, 1.5 hour per structure

How HE will be transformative

- Same amount of data could potentially be collected within 20 seconds
- Disentangle mixture of intermediates
- Record real movies

Realistic view of the experiment

Ideally

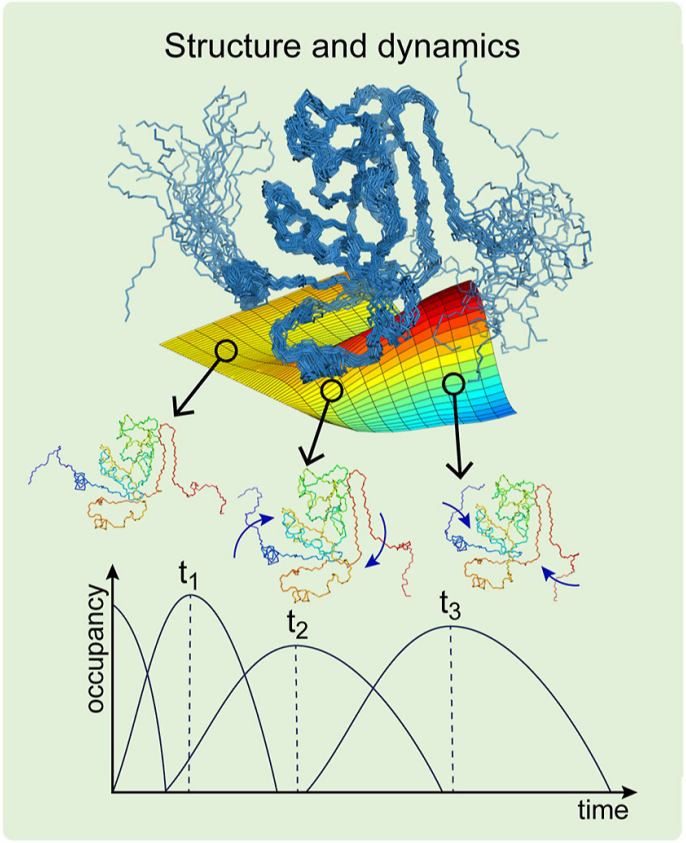
- 1.5 hour per structure (120 Hz)
- 20 seconds per structure (30 kHz)

Actually

- **6 hours** per structure (120 Hz)
  - Beamline alignment: pointing and focusing
  - Laser tuning, e.g. tweaking compressor of Ti:Sapph system, power measurement
  - Laser alignment, spatial overlap
  - Sample exchange
  - Lag in data analysis and decision making, overcollection

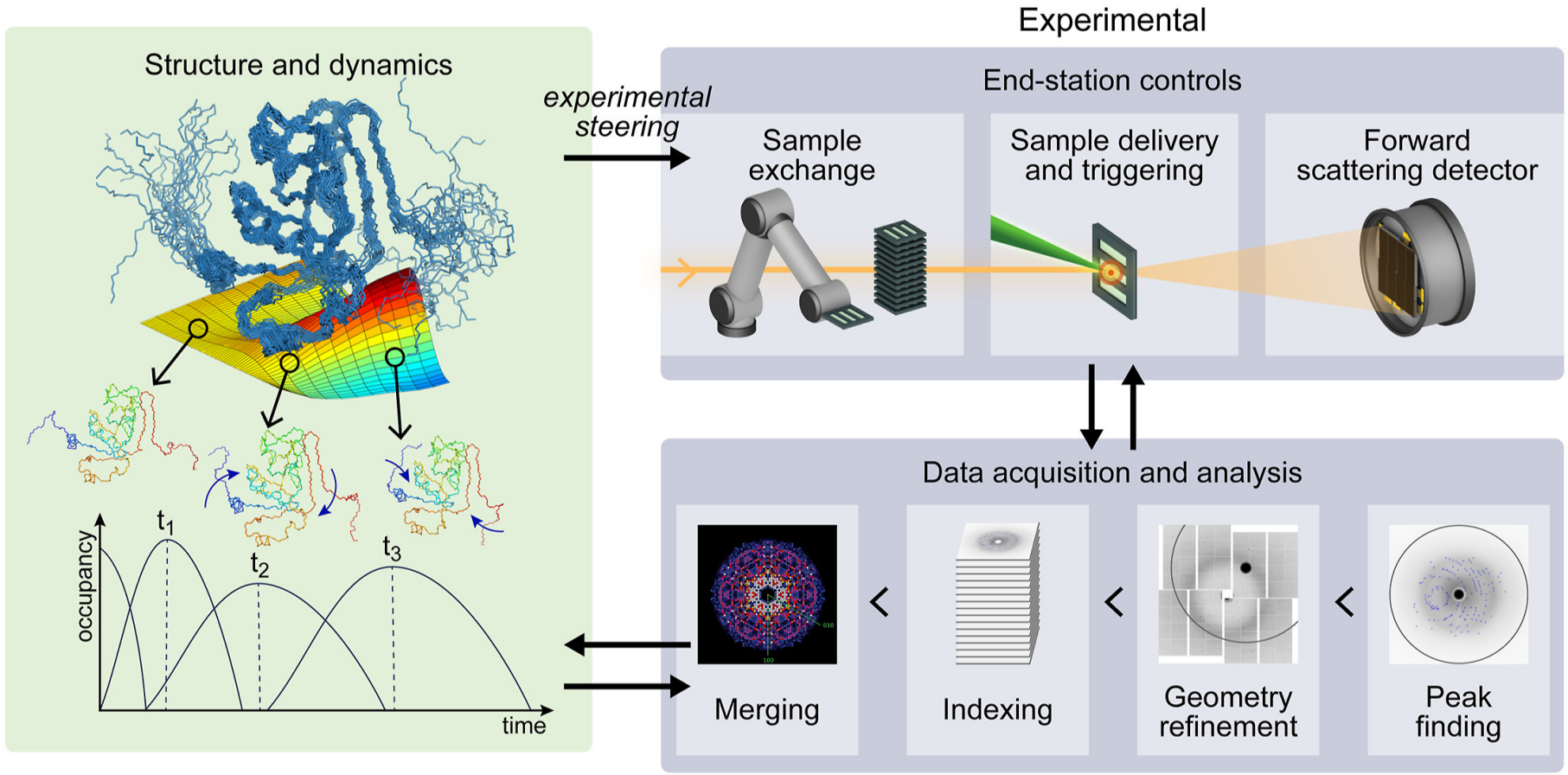
With current systems, higher X-ray repetition rates do not proportionally shorten data collection time!

Coupling petascale data generators to exascale computing



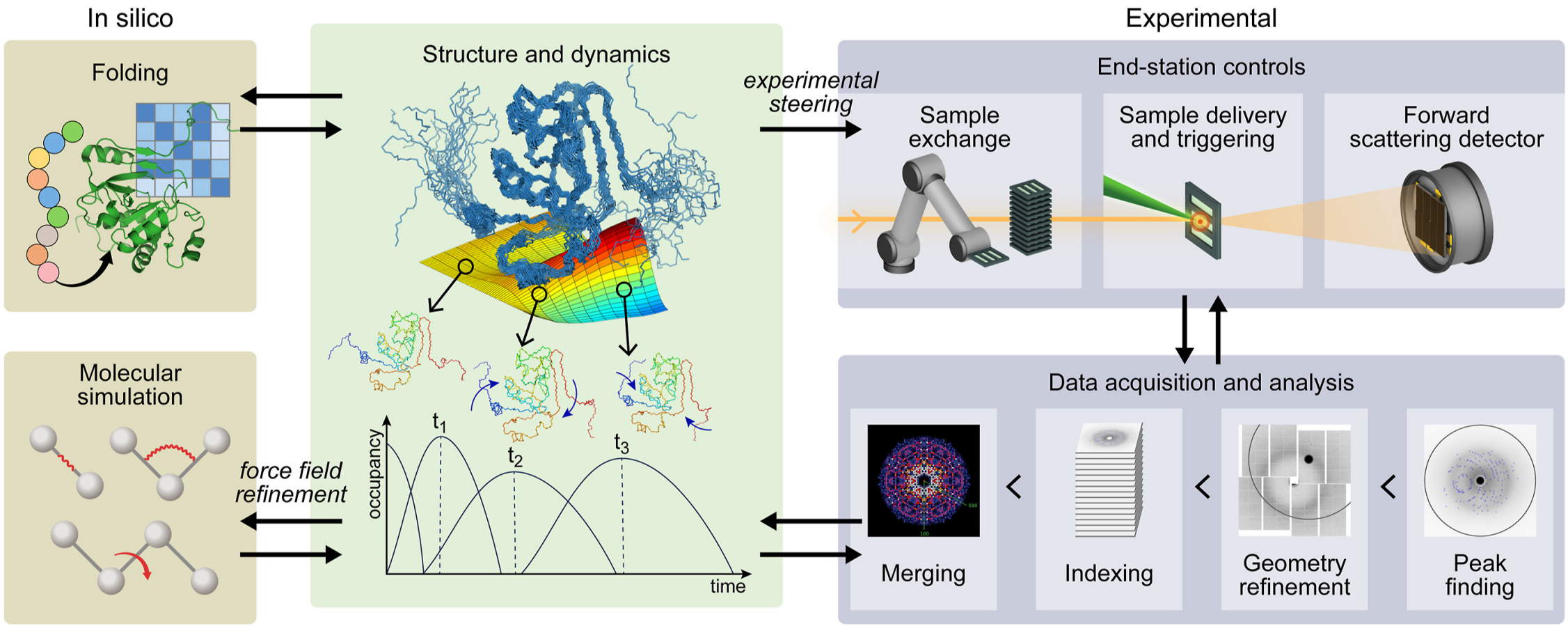
"Structural biology in the age of X-ray free-electron lasers and exascale computing" (2024) *Curr Opinion Struc Biol* **86**: 102808.

Coupling petascale data generators to exascale computing



"Structural biology in the age of X-ray free-electron lasers and exascale computing" (2024) *Curr Opinion Struc Biol* **86**: 102808.

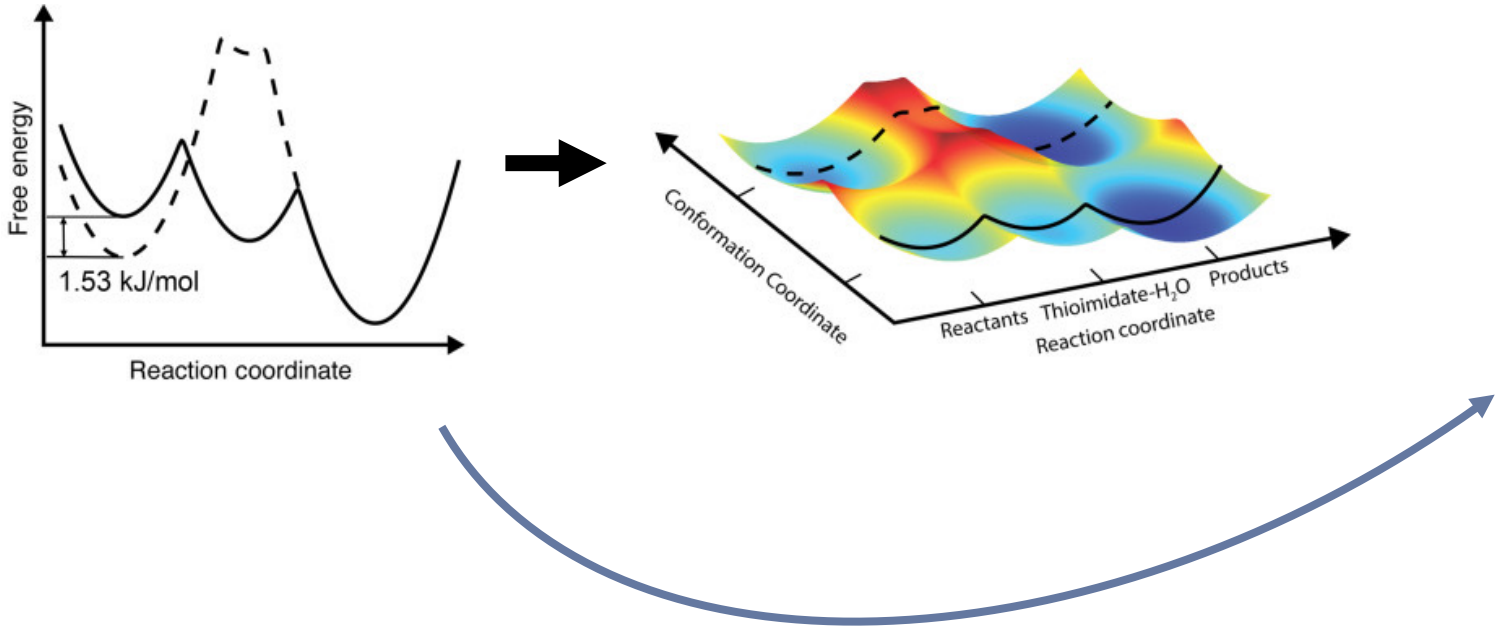
Coupling petascale data generators to exascale computing



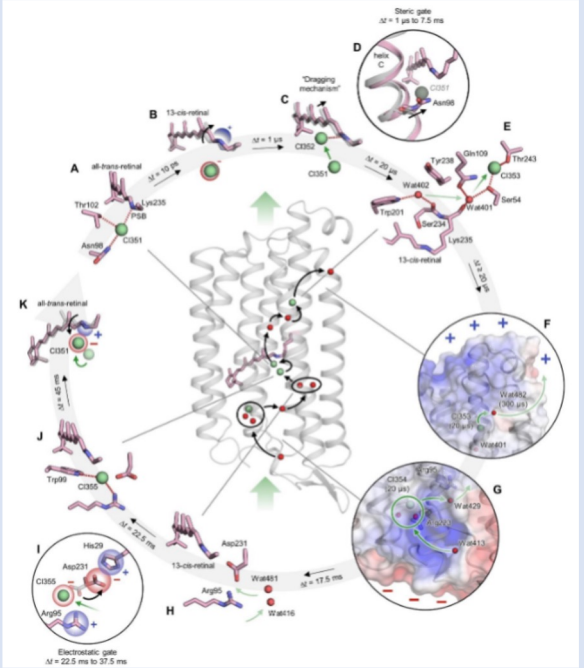
"Structural biology in the age of X-ray free-electron lasers and exascale computing" (2024) *Curr Opinion Struc Biol* **86**: 102808.

Opportunities at high repetition rate XFELs

From reaction coordinate to energy landscape

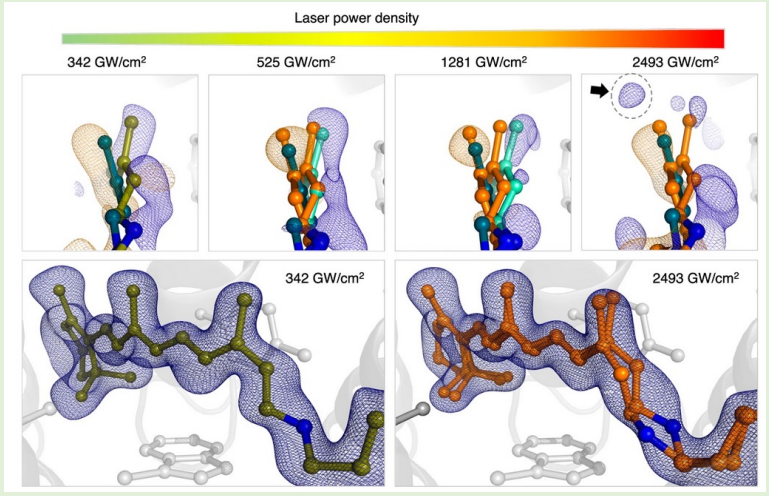


Classically: time-resolved optical pump-probe



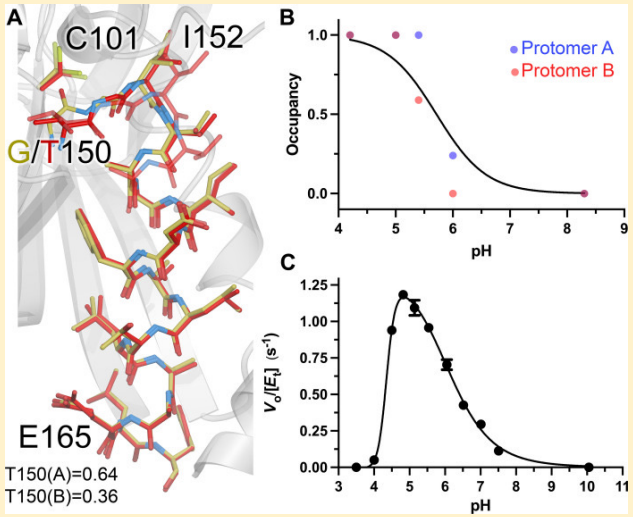
Opportunities at high repetition rate XFELs

Laser parameters
Power density



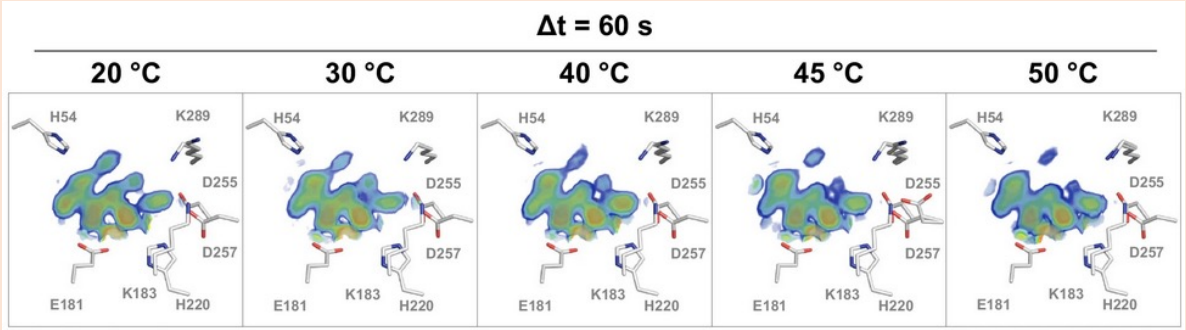
Bertrand *et al.*
Nat Comm (2024)

pH
Protonation state



Schmidt *et al.*
Sci Adv (2024)

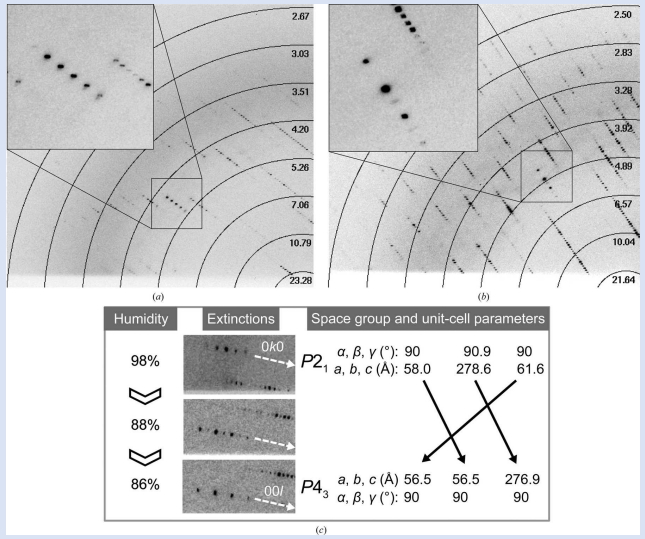
Temperature



Schulz *et al.*
Nat Comm (2025)

Opportunities at high repetition rate XFELs

Humidity Hydration

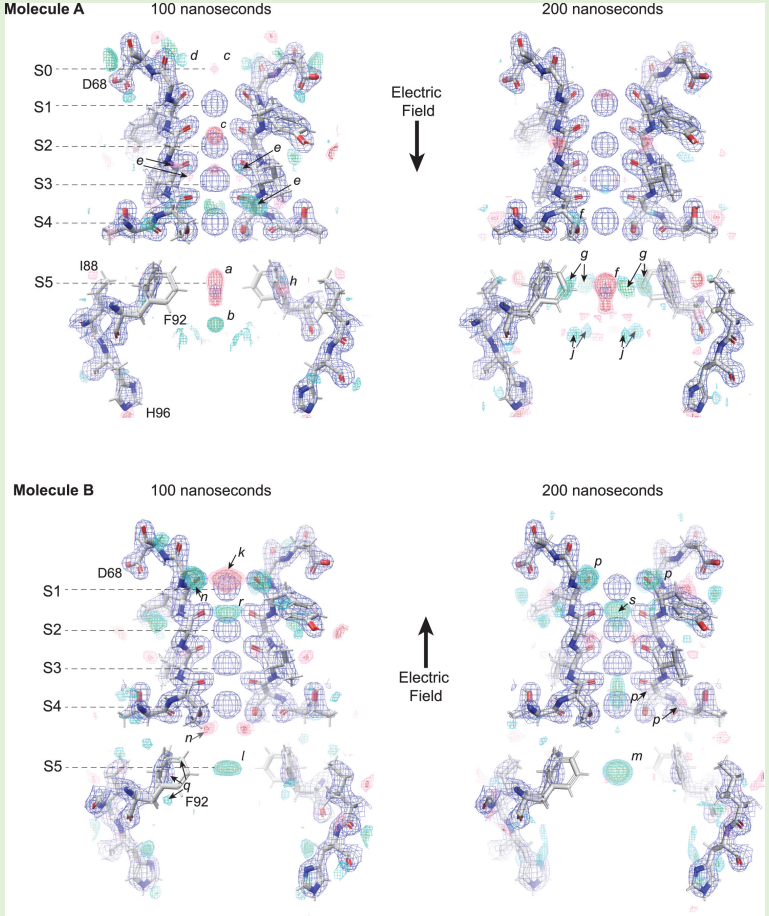


Klingl *et al.*
Acta D (2015)

- The future of XFEL crystallography is
  - High-throughput
  - Multi-state
  - Multi-condition
- Need for more bandwidth, compute and storage
- Need for more monitoring and capturing metadata

Electric field

Lee *et al.*
Cell (2025)



Thank you