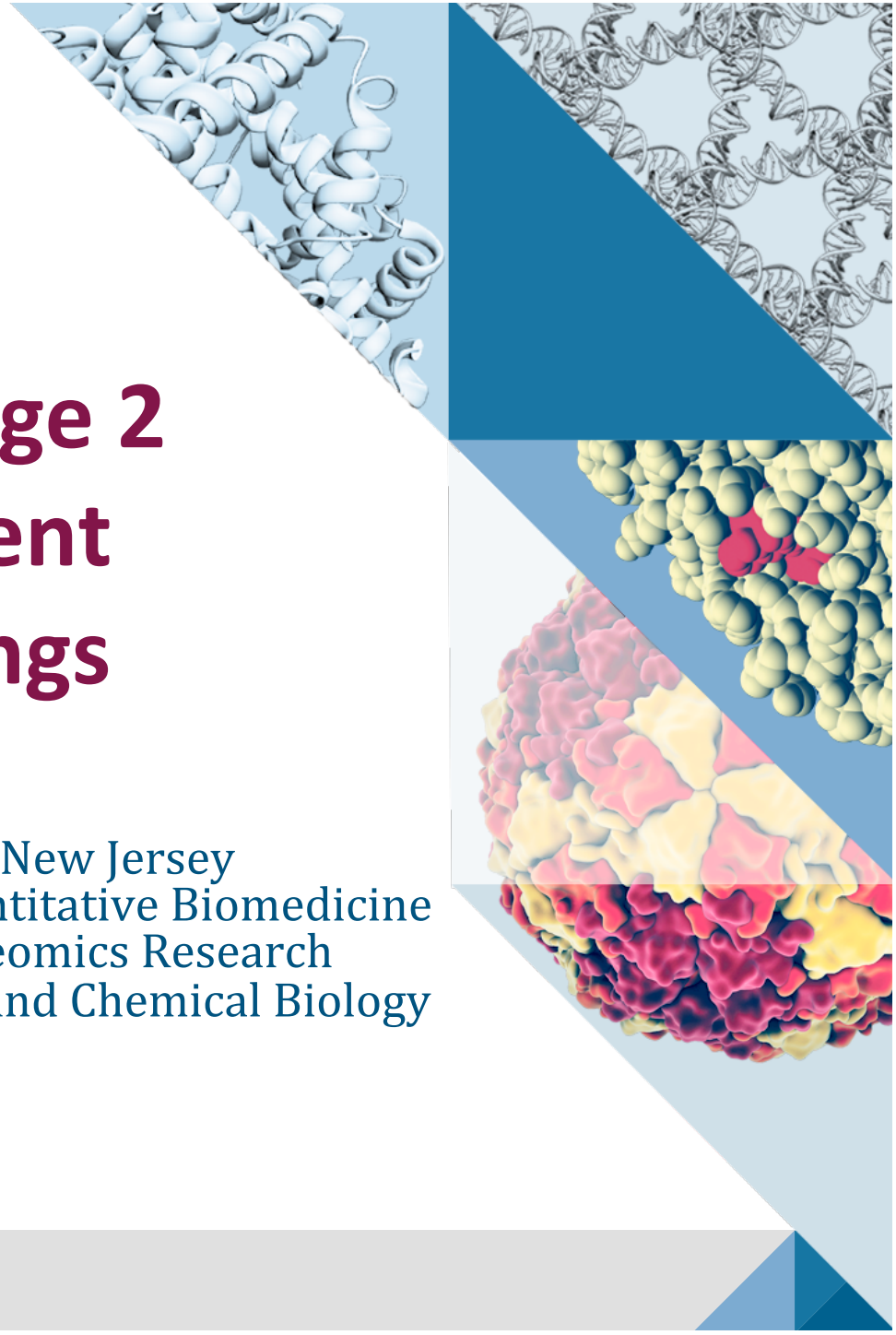
RCSB PDB
PROTEIN DATA BANK



D3R Grand Challenge 2 Structure Refinement Experience/Learnings

Stephen K. Burley, M.D., D.Phil.
Director, RCSB Protein Data Bank
Member, Rutgers Cancer Institute of New Jersey
Founding Director, Institute for Quantitative Biomedicine
Director, Center for Integrative Proteomics Research
Distinguished Professor, Chemistry and Chemical Biology

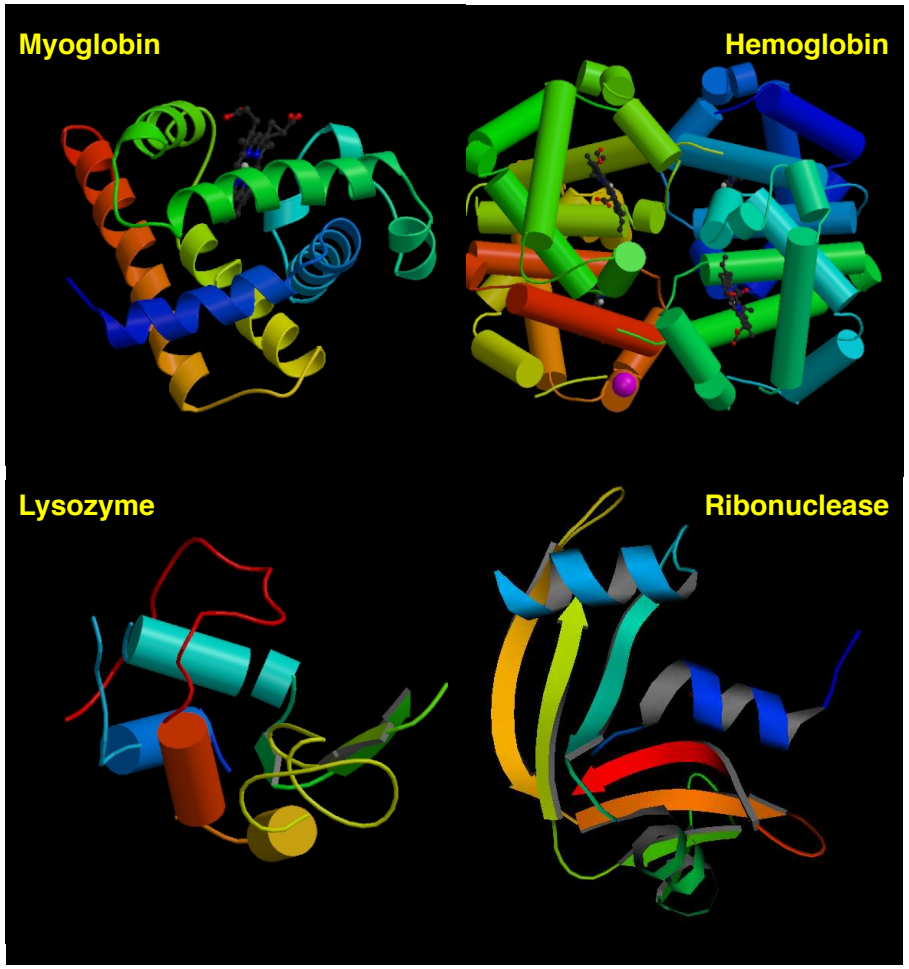
D3R Webinar Meeting
March 27th 2017

Outline

- Protein Data Bank (RCSB PDB, rcsb.org)
- RCSB PDB Group Deposition System
- D3R Structure Refinement Protocol
- Experience/Learnings
- Acknowledgements

Protein Data Bank Archive

- PDB 1st Open Access digital data resource in all of biology
- Founded 1971with 7 X-ray structures
- Today, single global archive
- Managed by wwPDB Partnership(RCSB PDB, PDBe, PDBj, BMRB)



Some of the very first structures in the PDB

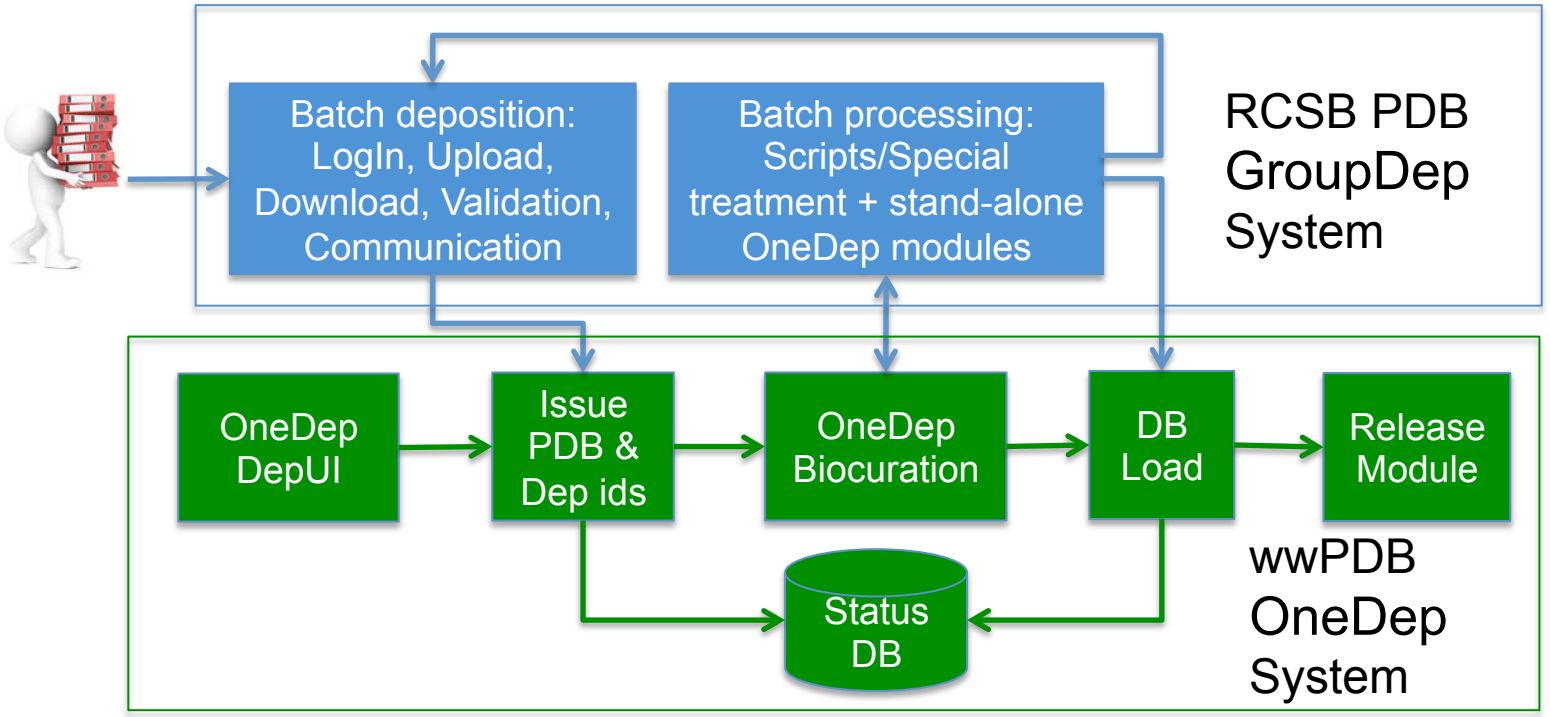
PDB Facts and Figures

- >128,000 Structures Archived since 1971
- ~11,500 New Structures Deposited/Year
- ~30,000 Depositors Worldwide
- >1 Million Unique Users Worldwide/Year
- >550 Million Data Files Downloaded/Year
- ~1.5 Million Data Files Downloaded/Day
- US PDB Headquartered at Rutgers/UCSD
- ~70% of Global PDB Usage through Rutgers/UCSD

RCSB PDB Group Deposition System

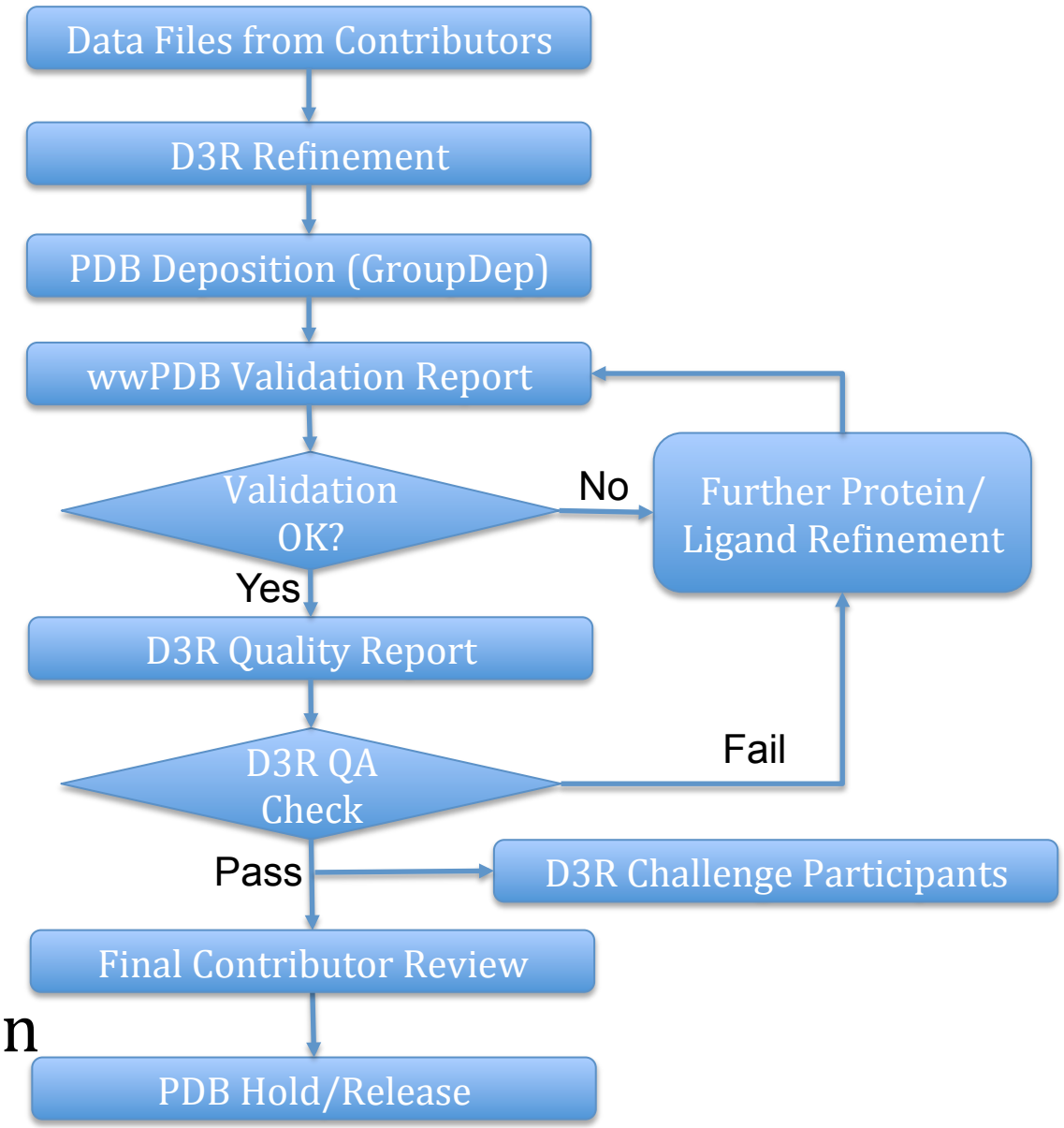
https://deposit-group.rcsb.rutgers.edu/groupdeposit/

- Supports Batch Deposition/Validation of Related Structures
- Originally developed for D3R Blind Challenges
- To date, >1300 structures from Roche, Merck Serono, BMS, SGC, etc., with 10s-100s of structures/group

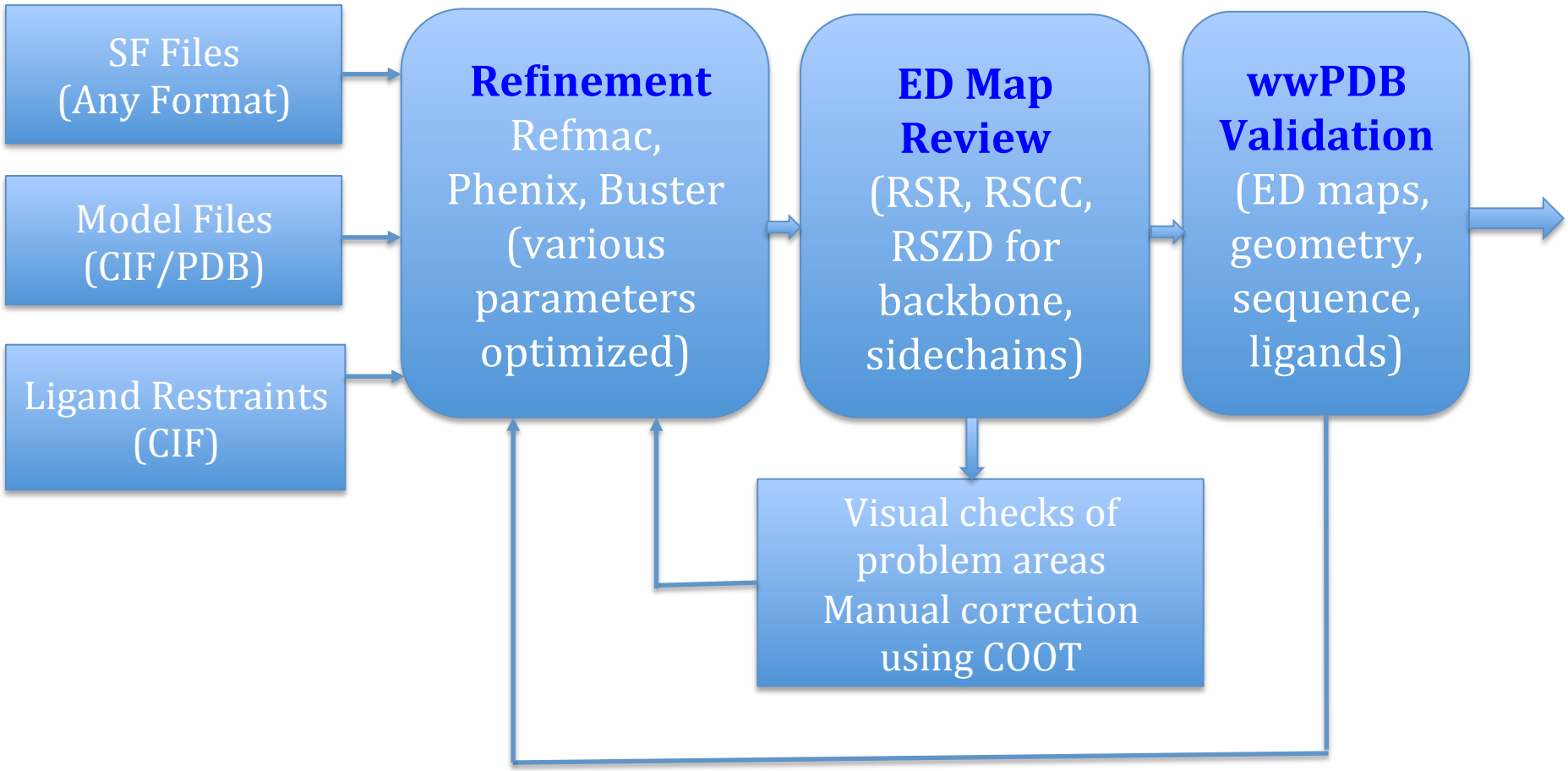


D3R Protocol

- D3R completes refinements
- GroupDep→PDB
- PDB Validates all refined structures
- Contributors and D3R co-author final PDB deposition



D3R Structure Refinement Protocol



Roche FXR Data Set Characteristics

- 36 Co-crystal/1 Apo Structures
  - Resolution Limits: 2.6-1.7Å
  - Asymmetric Unit MWs: 26-110kDa
  - 3 Point Groups/5 Space Groups

- Ligand MWs: 219-564 Da
- Three Ligands: Br atoms
- Four Ligands: Alternative Conformations
- All Ligands: Buried in Binding Pocket

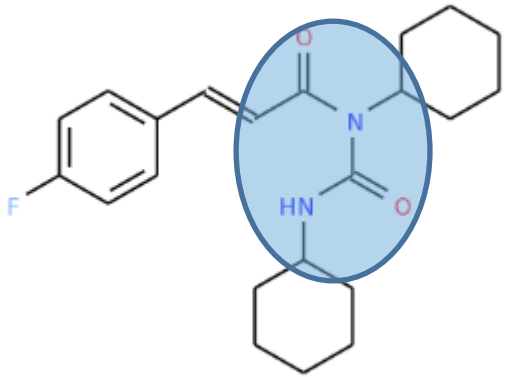
Protein Structure Quality Improved

	Initial Range	Final Range	Average Change
Rfree (%)	23.0-35.0	22.0-31.0	-1.0 (better)
Rwork (%)	17.0-30.0	18.0-25.0	<1.0 (no change)
Rfree – Rwork (%)	4.0-11.0	3.0-10.0	-1.0 (better)
Clash Score	4.38-40.41	0.51-6.90	-15.16 (better)
Ramachandran Outliers (%)	0.0-2.2	0.0-1.3	-0.3 (better)
Sidechain Rotamer Outliers (%)	0.5-24.5	0.5-9.9	-7.9 (better)
% RSRZ Outliers	0.0-29.5	0.0-22.1	-2.2 (better)
Bonds RMSZ	0.3-0.7	0.4-0.9	+0.2 (worse)
Angles RMSZ	0.5-0.8	0.5-0.9	+0.1 (worse)
Overall Quality (%ile of archive, see Shao et al. 2017)	0% - 70%	5% - 80%	+ 25%(better)

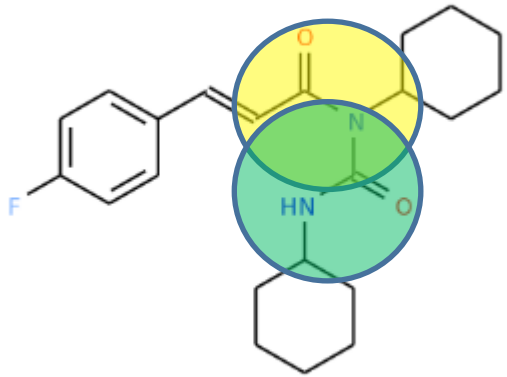
Ligand Structure Quality Improved

	Initial	Final	Average Change
RSR (%)	8.0 – 21.0	7.0 – 19.0	-1.0 (better)
RSCC (%)	82.0 – 98.0	79.0 – 98.0	+1.0 (better)
Bond Lengths Outliers	0 – 17	0 - 10	-3 (better)
Bond Angles Outliers	3 – 24	0 – 7	-7 (better)
Clashes	0-8	0-4	-1 (better)
Bonds RMSZ	0.8 – 2.9	0.5 - 2.9	-0.5 (better)
Angles RMSZ	1.1 - 3.1	0.5 - 2.0	-0.7 (better)
RSZD+(absolute)	0.0 - 2.8	0.0 - 3.4	+0.2 (balanced)
RSZD-(absolute)	0.0 - 4.2	0.0 - 3.3	-0.2 (balanced)

Ligand Geometrical Restraints

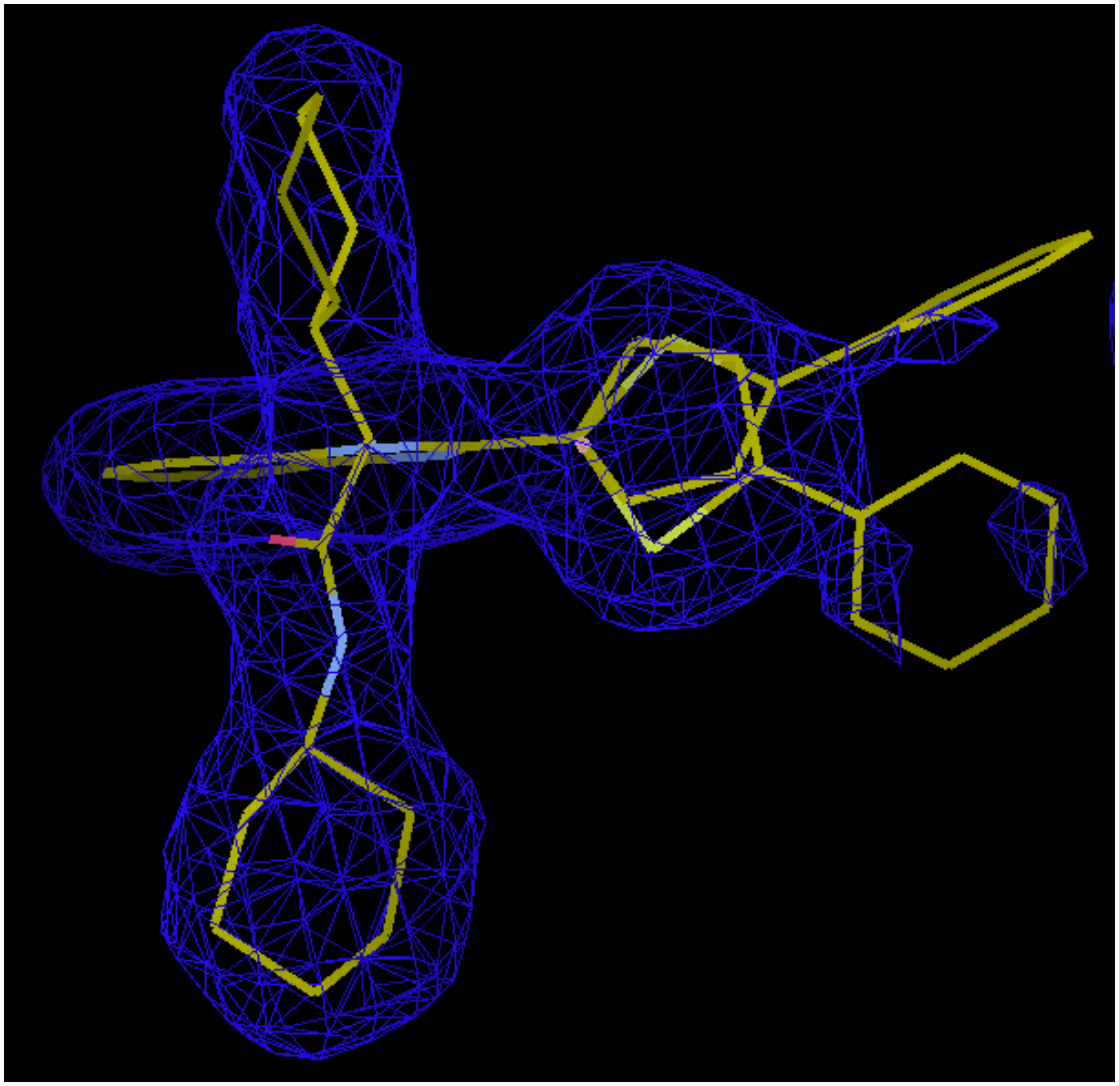


Roche 1ymrc



- Generated ligand restraints with
  - AM1 and RM1 QM optimizers
  - +/- Mogul values
  - Varied planarity settings
- Best results with AM1 + Mogul and tighter planarity, and manual setting of some restraints as needed

Experience/Learnings: Partial Disorder



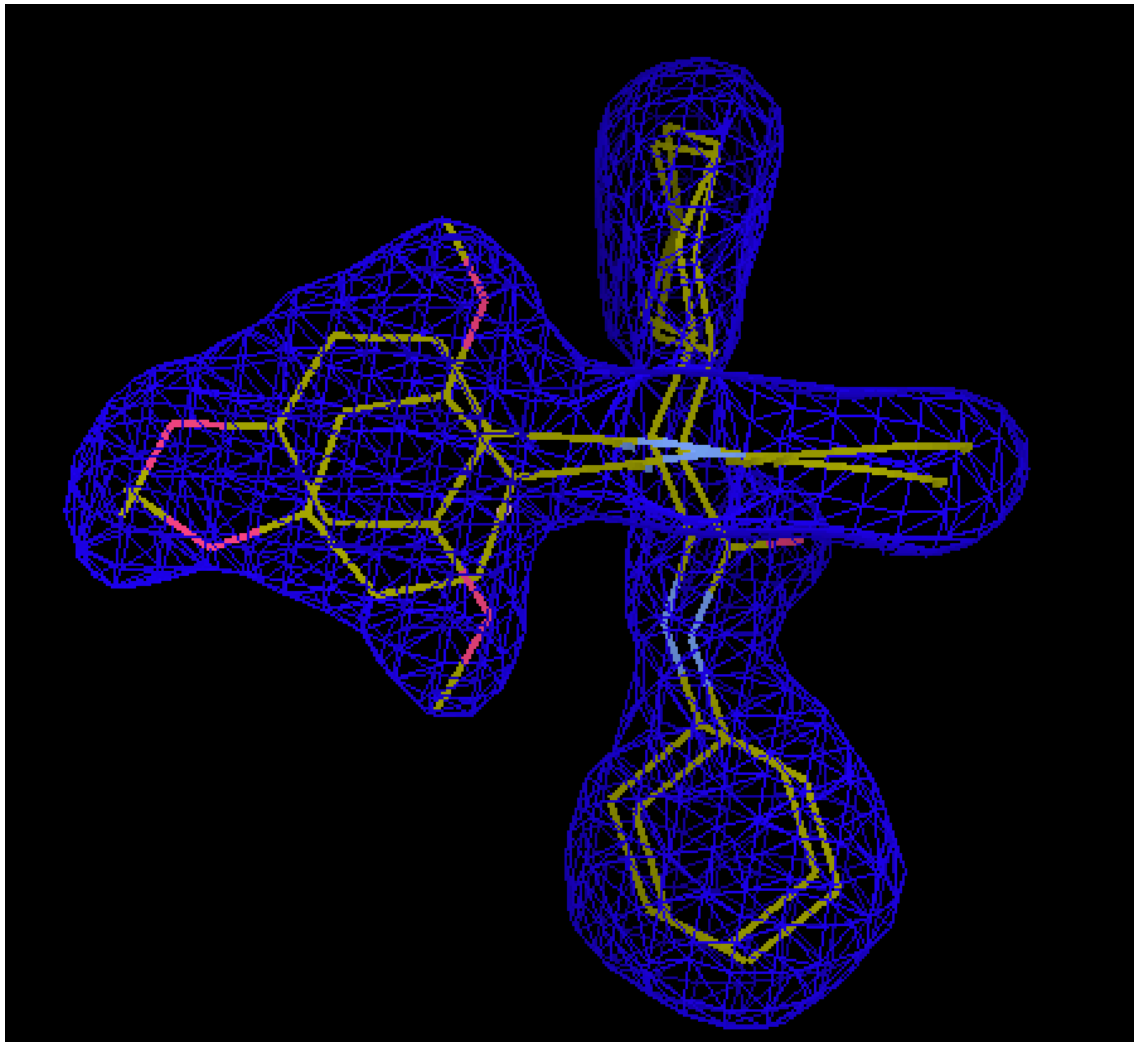
Partial disorder alternative conformers
(1dqzc, 2.0Å)

Thiophene alternative conformations

Cyclohexane disordered

Blue: 2|Fo|-|Fc|/1.0σ

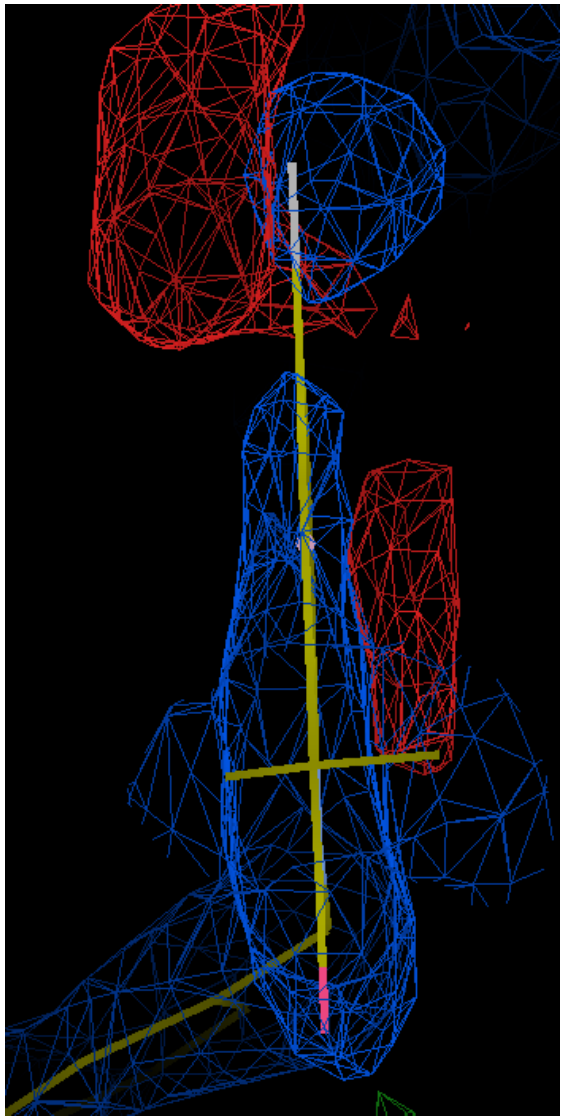
Experience/Learnings: Overall Disorder



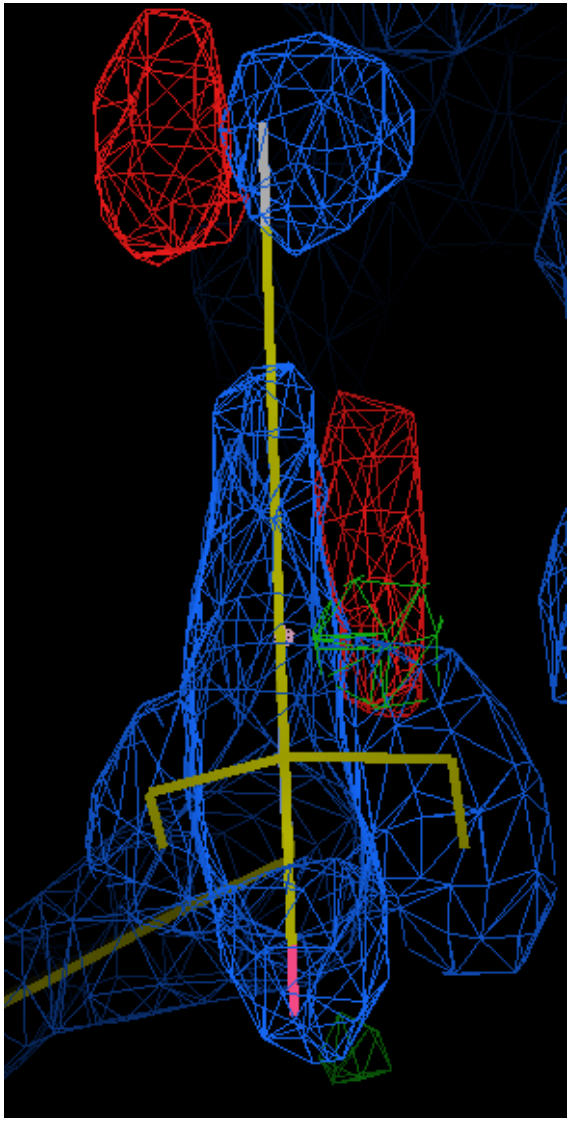
Entire ligand in two orientations
(1txca, 1.9Å)

Blue: 2|Fo|-|Fc|/1.0σ

Experience/Learnings: Loss of Br



Br Occup=1.0



Br Occup=0.5

Ligand with Br
(1xxar, 1.8Å)

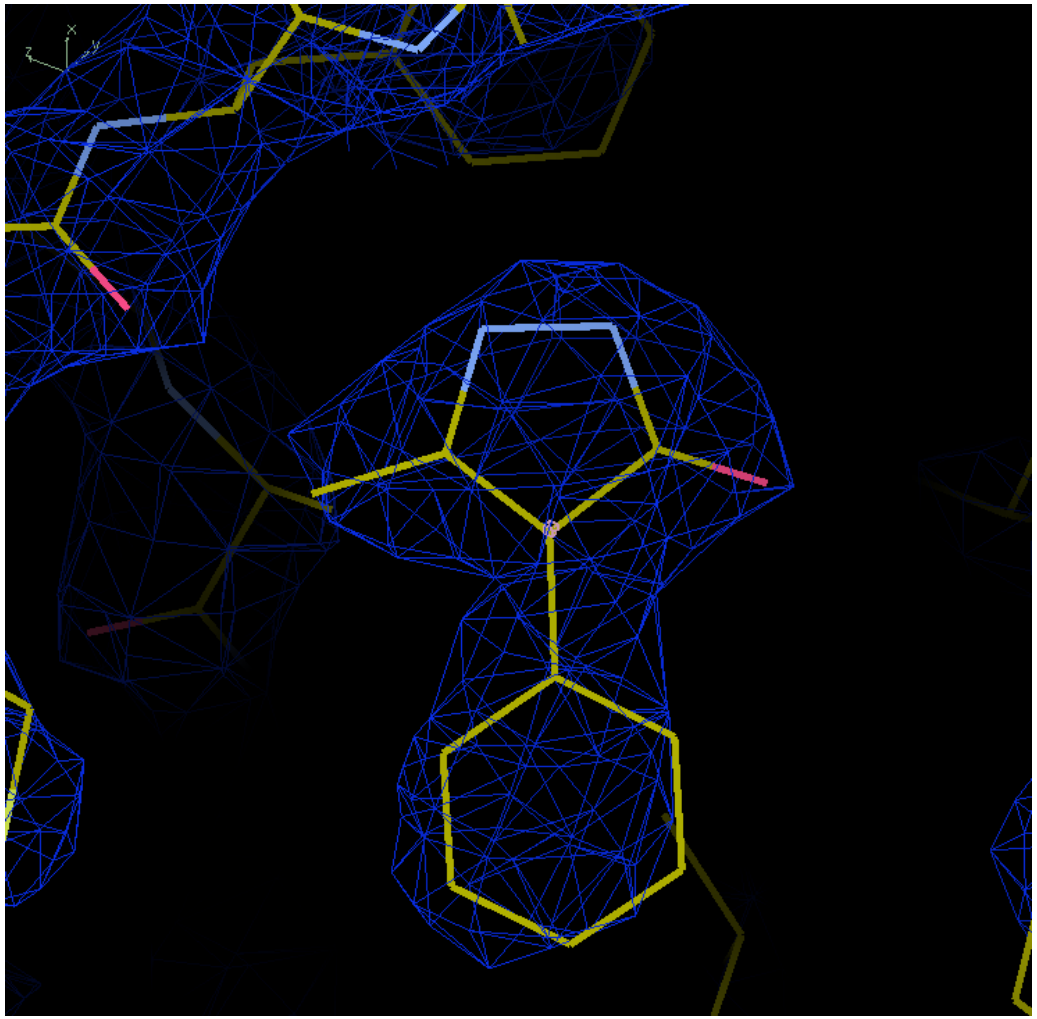
Br-C bond often
broken by X-ray beam

Crystal contains two
Ligand species (+/- Br)

Br Modeled with
reduced Occupancy

Blue: 2|Fo|-|Fc|/1.5σ
Red: -ve |Fo|-|Fc|/3.0σ
Green: +ve |Fo|-|Fc|/3.0σ

Experience/Learnings: Ambiguities



ED Map Ambiguity
(from different study)

Pseudo-symmetric ligands require inspection of binding environment

Wrong Orientation!

Blue: 2|Fo|-|Fc|/1.0σ

Acknowledgements

- Huanwang Yang, Chenghua Shao, Zukang Feng, John Westbrook, Jasmine Young, Ezra Peisach
- Symon Gathiaka, Shuai Liu, Mike Gilson, Rommie Ammaro

The RCSB PDB is funded by a grant from the National Science Foundation, the National Institutes of Health, and the US Department of Energy.