Experimental Phasing with CCP4

Experimental Phasing		
ata Preparation		
Itomated Search & Phasing		
SHELX C/D/E pipeline		
Crank		
autoSHARP		
eavy Atom Location		
asing & Refinement		
sualisation		
ilities		
		H
	Experimental Phasing ata Preparation Momated Search & Phasing SHELX C/D/E pipeline Crank autoSMAMA eavy Atom Location tasing & Refinement isualisation tilities	ata Preparation ata Preparation atomated Search & Phasing SHELX C/D/E pipeline Crank autostion eavy Atom Location basing & Refinement isualisation tilities

- Speaker: Charles Ballard
- Program Authors:
 - Raj Pannu, Pavol Skubak
 - Randy Read, Airlie McCoy
 - Rudolf de Graaff, Jan Abrahams
 - Kevin Cowtan
 - Garib Murshudov
 - Phil Evans
 - (George Sheldrick)
- "Every program is becoming an experimental phasing program" -Kevin Cowtan



Data and Resolution



- Spacegroup
 - Translation symmetry
 - Fixed origin
 - No special positions
 - P21 21 21 ideal
- Redundancy
 - K. Diederichs, Acta D66 (2010 733.
- Scala
 - Completeness
 - <|>/<sig|>
 - Inter dataset anomalous correlation MAD (> 0.3)
 - Intra dataset anomalous correction -SAD



Data and Resolution II



- <Dano>/<sigDano>
 - > 1.2
 - Output by AFRO
 - Also SHELXC via crank.
- George Sheldrick's rule of thumb
 - Hi Res + 0.5
 - 3.0 A



Running CRANK http://www.bfsc.leidenuniv.nl/software/crank/

C CRANK				
	Help			
Title				
MTZ in PROJECT - Bro	owse View			
MTZ out PROJECT Bro	wse View			
PDB out PROJECT - Bro	owse View			
Input Intensities - Setup experiment SAD -				
Input protein sequence 🔳				
SEQ in PROJECT - Bro	wse View			
DNA/RNA present 🔟				
Input R-free flag 🔟				
FreeR set will consist of 0.05 fraction of the data.				
Crystal # 1	—			
Native _ Substructure atom Number of substructure atoms per monomer				
Dataset : 1 Type SAD — Anomalous				
Data collected at CuKa wavelength				
IP+ SIGIP+				
IP- SIGIP-	_			
Derived parameters				
Experimental Pipeline				
Start the pipeline with Substructure detection 🖃 and end with Model building	-			
Pipeline : CUSTOM -				
Substructure detection: AFRO/CRUNCH2 -				
Substructure refinement: BP3 -				
Hand determination: SOLOMON - Density modification: SOLOMON				
Model building: BUCCANEER -				
Display individual program options 🔟				
Run - Save or Restore -	Close			

- Takes Merged data
- Experiment type
- Atom types plus f' and f''
- Customize pipeline
- Heart of CRANK
 - AFRO
 - CRUNCH2
 - BP3



Flow inside CRANK



AFRO: Multivariate SAD equation for F_A estimation

 $\mathsf{E}(\mathsf{|F_A|;|F^{+}|,|F^{-}|) =$

 $\iiint F_A |P(|F_A|, \alpha_A, |F^*|, \alpha^*, |F^-|, \alpha^-) d|F_A |d\alpha_A d\alpha^* d\alpha^-$

 $\iiint P(|F_A|, \alpha_A, |F^*|, \alpha^*, |F^*|, \alpha^-) d|F_A| d\alpha_A d\alpha^+ d\alpha^-$

- Direct methods very sensitive to FA values
- The multivariate FA estimation leads to more substructures being determined than
 ΔF = ||F+|-|F-||



Crunch2: substructure detection

- Algebraic approach based on rank reduction of Karle/Hauptman matrices
- Includes higher order collections of reflections, not just triplets
- de Graff et al. (2001) Acta Cryst. D57, 1857-1862.
- Standalone via crank interface



Scoring the substucture solution

- CRUNCH2: FOM > 1.0
 SHELXD: CC_{weak} > 30%
- Both conservative criteria.
 Solutions may be missed

Crunch2 Version 1.3.0 Run at 11:40:13 on 7/04/2009 Finished with: Normal termination

Result:

Bp3 Luzzati parameter indicates solution Total number of trials: 1.



BP3 can be run in "PHAS" mode to verify



BP3: substructure refinement

- Multivariate substructure refinement
 - Includes effect of model and measurement errors
- SAD, MAD, SIR(AS) and MIR(AS)
- Refines atomic and error parameters
- Outputs H-L coeffs, FOM and PHIB
- Normal and fast phasing (PHAS)



Solution validation

• FOM

SAD - Anomalous Luzatti Parameter

- > 0.7 strong solution
- 0.3 unlikely
- Build

Bp3 version 1.01 Run at 11:40:33 on 7/ 4/2009 Finished with: Normal termination

Result:

The overall FOM is 0.457 and the average anomalous Luzzati error is 0.818 to 2.73 resolution



[Show logfile summary] [Show full logfile]



Density Modification

Map improvement

- SOLOMON, (SHELXE)
- PIRATE statistical in recipriocol space
- PARROT MLHL target function, pairwise NCS
- SAD Hand selection
 - SOLOMON/SHELXE:
 - contract local map density
 - higher better

• PARROT:

sigmaa value – higher better (DM:Real Space Free R, lower better)

Tolerences about 0.05





Refinement - Refmac

- Relevant refinement targets
 - SAD (and SIRAS coming)
 - Greatly improves maps
 - Skubak et al (2004) Acta D
 - MLHL
 - Prior phase information used indirectly in the form of H-L coefficients
 - Assume prior phase information is independent of refined phases
 - Generally applicable
 - Pannu et al (1998) Acta D



Model Building - Buccaneer

 Statistical model building. Likelihood function based on conserved density

HIS

MET

- Requires phase probabilities
- Finding Growing: Cα environment



AI A

Sequencing: Cβ environment

CYS



THR

GerE SAD – a challenging example solved by default

- Example data with CCP4. Originally MAD+native
- 2.7Å SAD peak data with 12 Se (F`-3.6, F" 6.0)
- Early Crank version failed
- Crank v1.3 builds 70% by default
- Crank v1.4 builds 93%, and 70% at inflection point.

Phaser



- Multivariate SAD target
- Refines coordinates, temperature, occupancy
- Iterative atom location
 - Improves substructure and sensitivity
- Maps
 - Electron density
 - Log likelihood gradient (by atom type)



Phaser – log likelihood gradient



- Peak height dependent on f'/f''' ration
 - May distinguish atom types
 - Can be fooled if normal scatter present at site
- Very sensitive to detail
 - Image shows anisotropic





MR-SAD



- Input: partial structure/MR solution
- Possible atom types





Better maps



Other Programs

- Acorn
 - Direct method based density modification
 - Free lunch
- OASIS
 - SAD ambuguity breaking
- Venerable
 - RANTAN, MLphare
- External
 - SHELX, HKL2MAP
 - Autosharp

