

# ***SnB: A Direct Methods Program***

---

***SnB***

**Russ Miller**

**Hauptman-Woodward Med. Res. Inst.**

**Principal Contributors:**

**C.-S. Chang**

**G.T. DeTitta**

**S.M. Gallo**

**H.A. Hauptman**

**H.G. Khalak**

**D.A. Langs**

**R. Miller**

**S. Potter**

**C.M. Weeks**

**Partial funding from NIH and NSF.**

# Outline of Talk

---

*SnB*

- ◆ **Direct Methods**
- ◆ ***Shake-and-Bake: The Method***
  - **Cyclical Dual Space Procedure**
  - **The Minimal Function**
- ◆ ***SnB: The Program***
- ◆ **Applications of *SnB***
- ◆ ***SnB* version 2.0**

# The Phase Problem

---

*SnB*

- ◆ Experiment yields:
  - reflections
  - associated intensities
- ◆ *Phase angles are lost in experiment.*
- ◆ Underlying atomic arrangement is related to the reflections by a 3-D Fourier transform.
- ◆ ***Phase Problem:*** determine the set of phases corresponding to the reflections.

# Data Structures

*SnB*

## Reflections

	$h$	$k$	$l$	$ E $	$\phi$
1					
2					
3					
$M$					

Reciprocal Space

3-D Fourier  
Transform

## Atoms

	$x$	$y$	$z$
1			
2			
3			
$A$			

Real Space

# Direct Methods

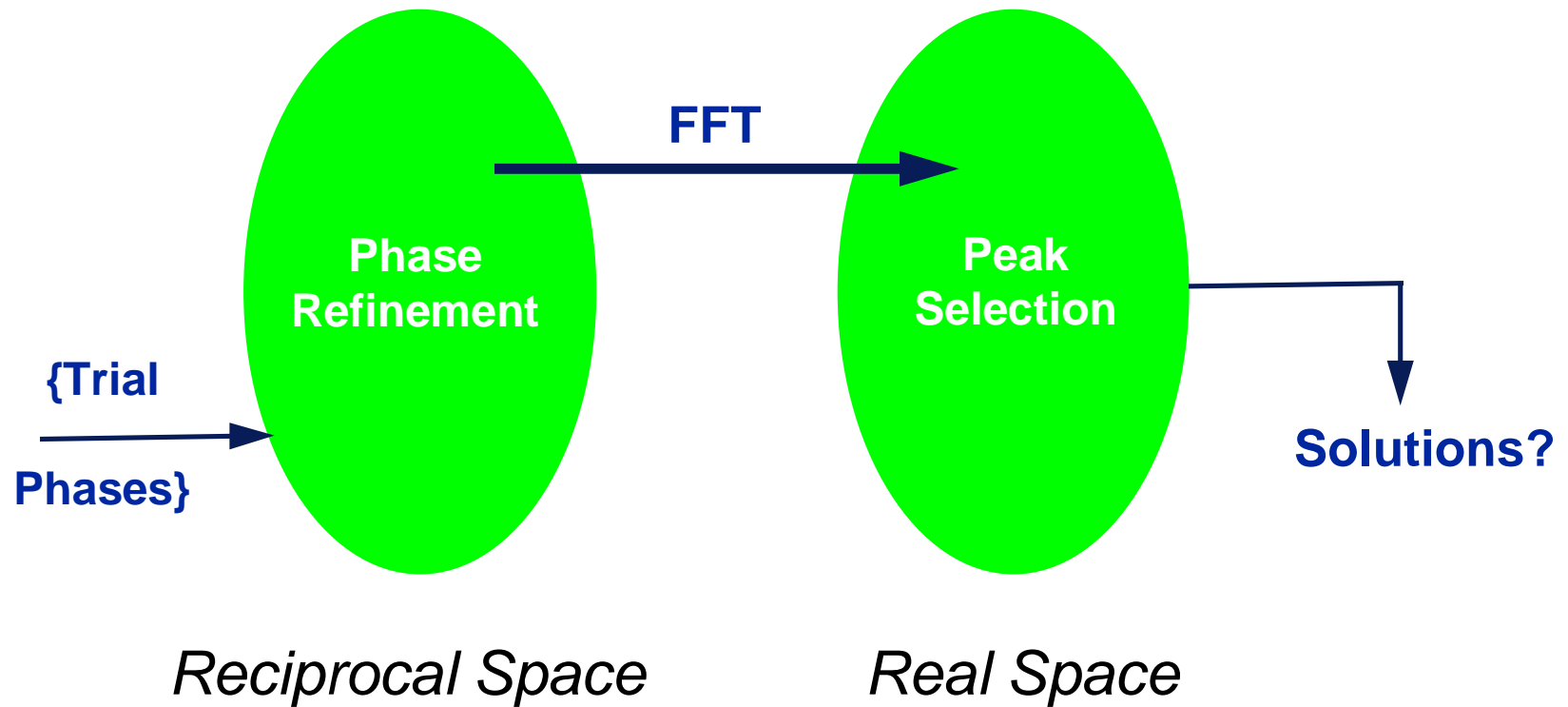
---

*SnB*

- ◆ *Direct Methods* use probabilistic theories to exploit linear relationships among phases.
- ◆ Resolution of 1.2Å or better.
- ◆ Routinely applied to structures with 150 or fewer atoms.
- ◆ Standard packages:
  - SHELX
  - teXsan
  - SIR92/96

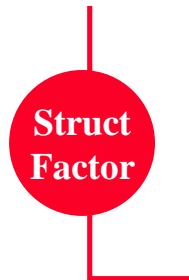
# Conventional Direct Methods

*SnB*



# *Shake-and-Bake*

{Trial  
Structures}

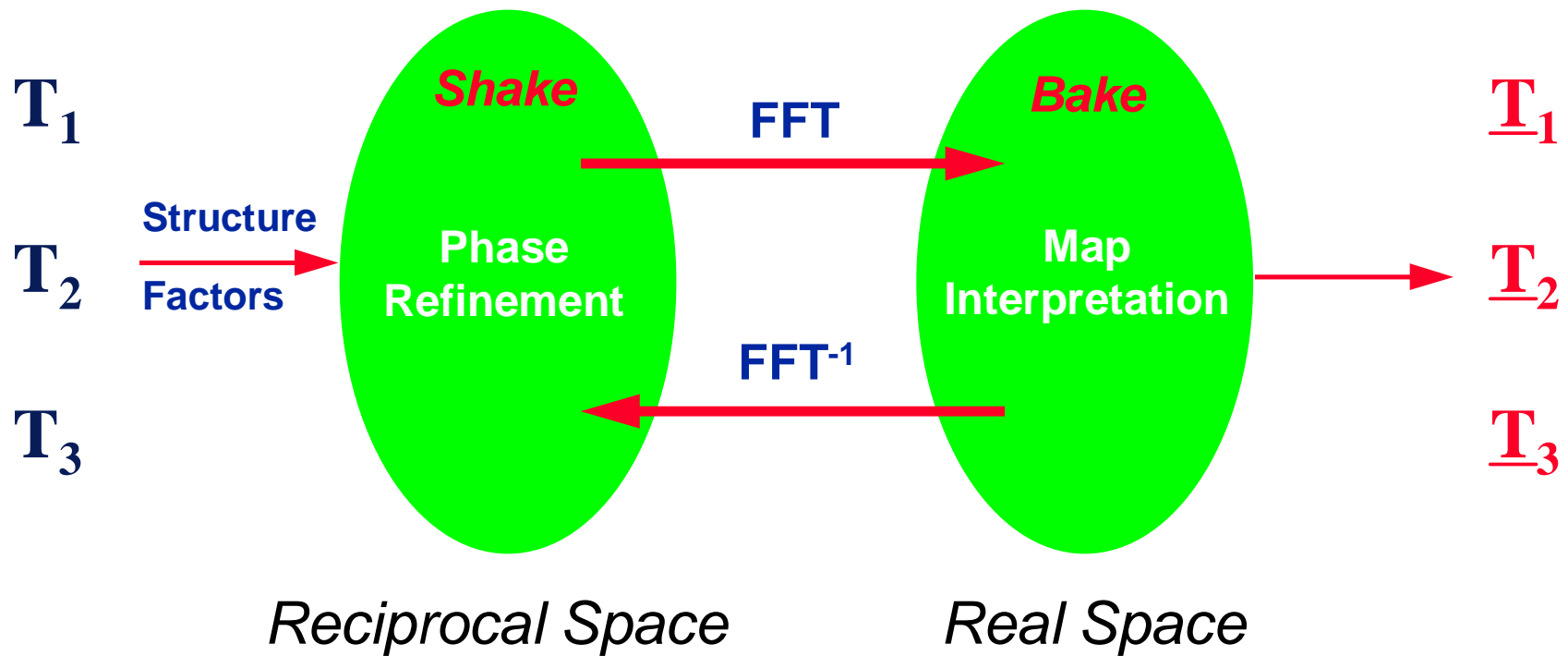


FFT<sup>-1</sup>



# Shake-and-Bake

*SnB*





# *Shake-and-Bake*

---

*SnB*

- ◆ **Direct Methods Optimization Technique**
- ◆ **Multiple Trial Structures**
- ◆ **Real Space  $\Leftrightarrow$  Reciprocal Space**
- ◆ **Phase Refinement Techniques**
  - **Parameter Shift (Minimal Function)**
  - **Tangent Formula**
- ◆ **Minimal Function as FOM**

# The Minimal Function

*SnB*

$$R = \frac{\sum_T W_T (\cos \phi_T - est_T)^2}{\sum_T W_T}$$

Triple:  $\phi_T = \phi_h + \phi_k + \phi_{-h-k}$

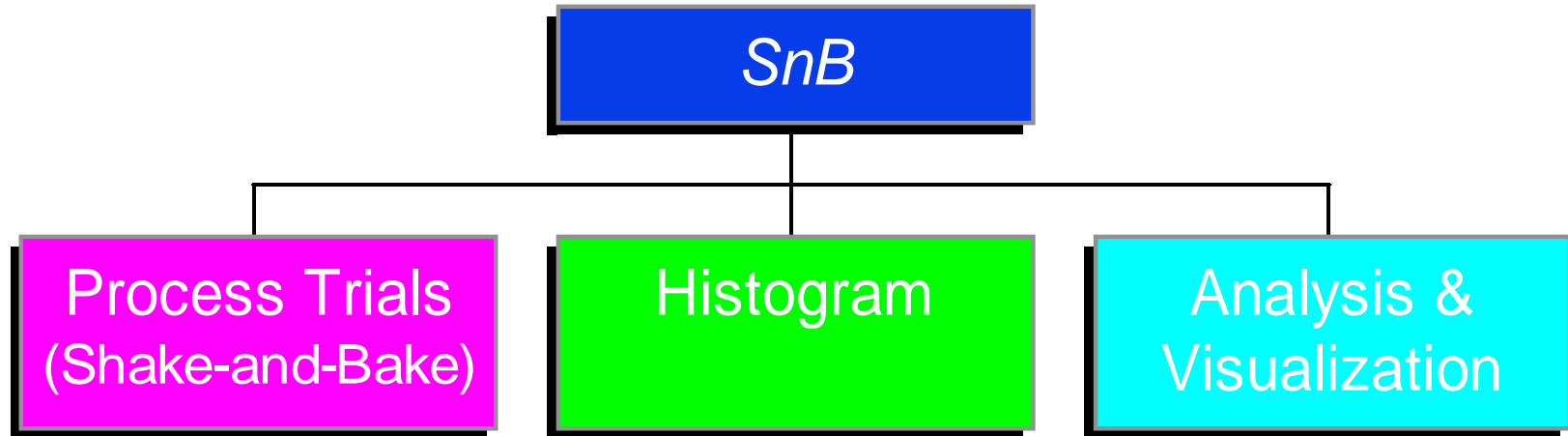
$$W_T = \left( \frac{2}{N^{1/2}} \right) |E_h E_k E_{-h-k}|$$

$est_T$  is the known expected value of  $\cos \phi_T$

# Structure of *SnB*

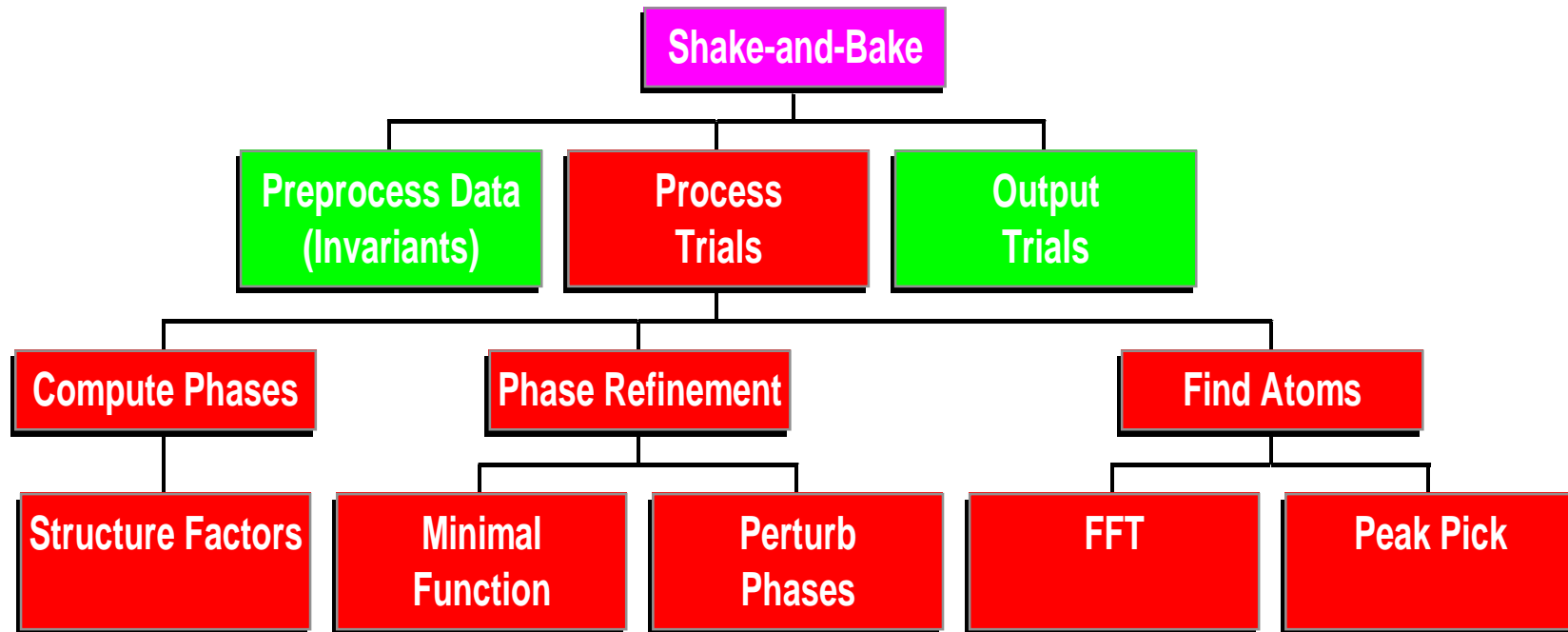
---

*SnB*



# Shake-and-Bake

*SnB*



# *SnB* Parameters

The logo for SnB, featuring the letters 'SnB' in a stylized, italicized font. The 'S' is red, the 'n' is grey, and the 'B' is red. The letters are slightly offset and shadowed to create a 3D effect.

---

	<b>Default</b>	<b>Ph8755</b>	<b>ToxII</b>
<b>Atoms (a.u.)</b>	<b>n</b>	<b>74</b>	<b>508</b>
<b>Phases</b>	<b>8n - 10n</b>	<b>740</b>	<b>5,000</b>
<b>Triples</b>	<b>70n - 100n</b>	<b>7,400</b>	<b>50,000</b>
<b>Cycles (PS)</b>	<b>n/2</b>	<b>40</b>	<b>255</b>
<b>Peaks recycled</b>	<b>0.8n - n</b>	<b>74</b>	<b>400</b>
<b>E-Fourier Steps</b>	<b>2</b>	<b>2</b>	<b>5</b>

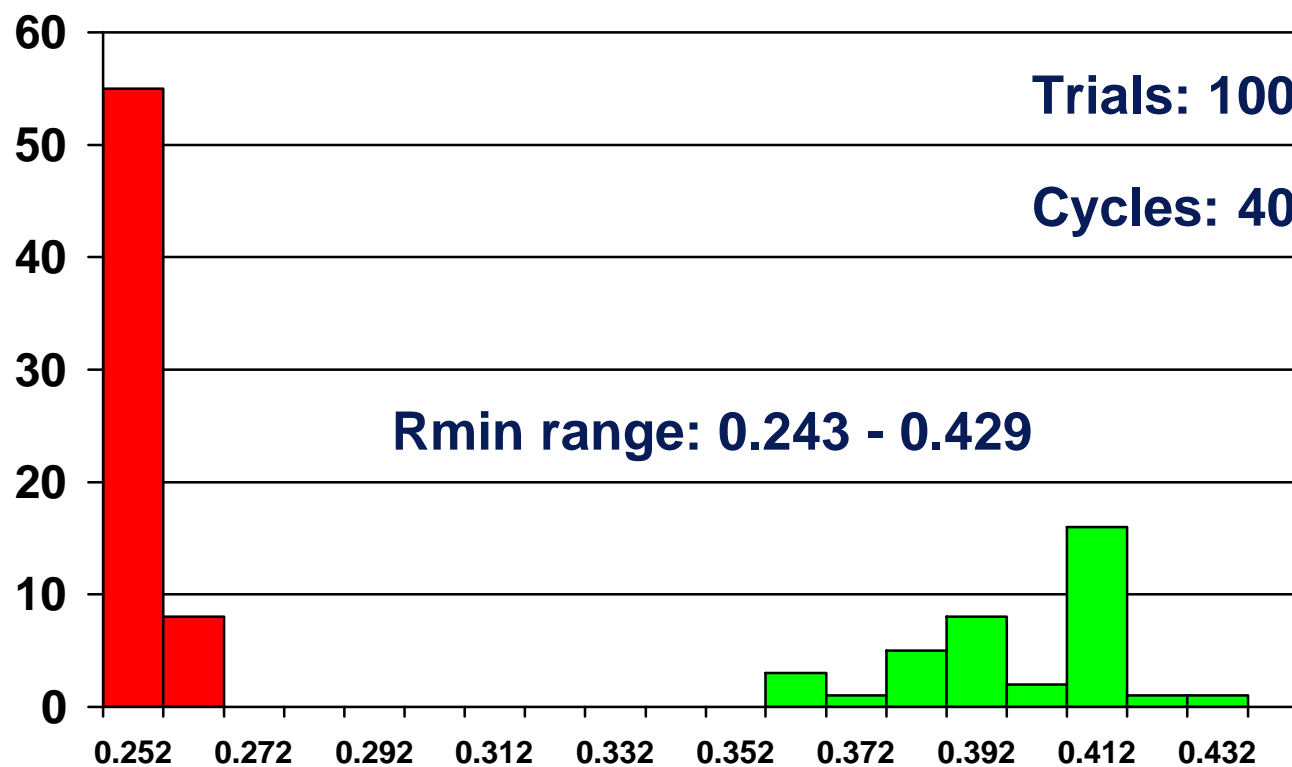
---

# Sample Histogram (Ph8755)

*SnB*

Atoms: 74  
Space Group: P1

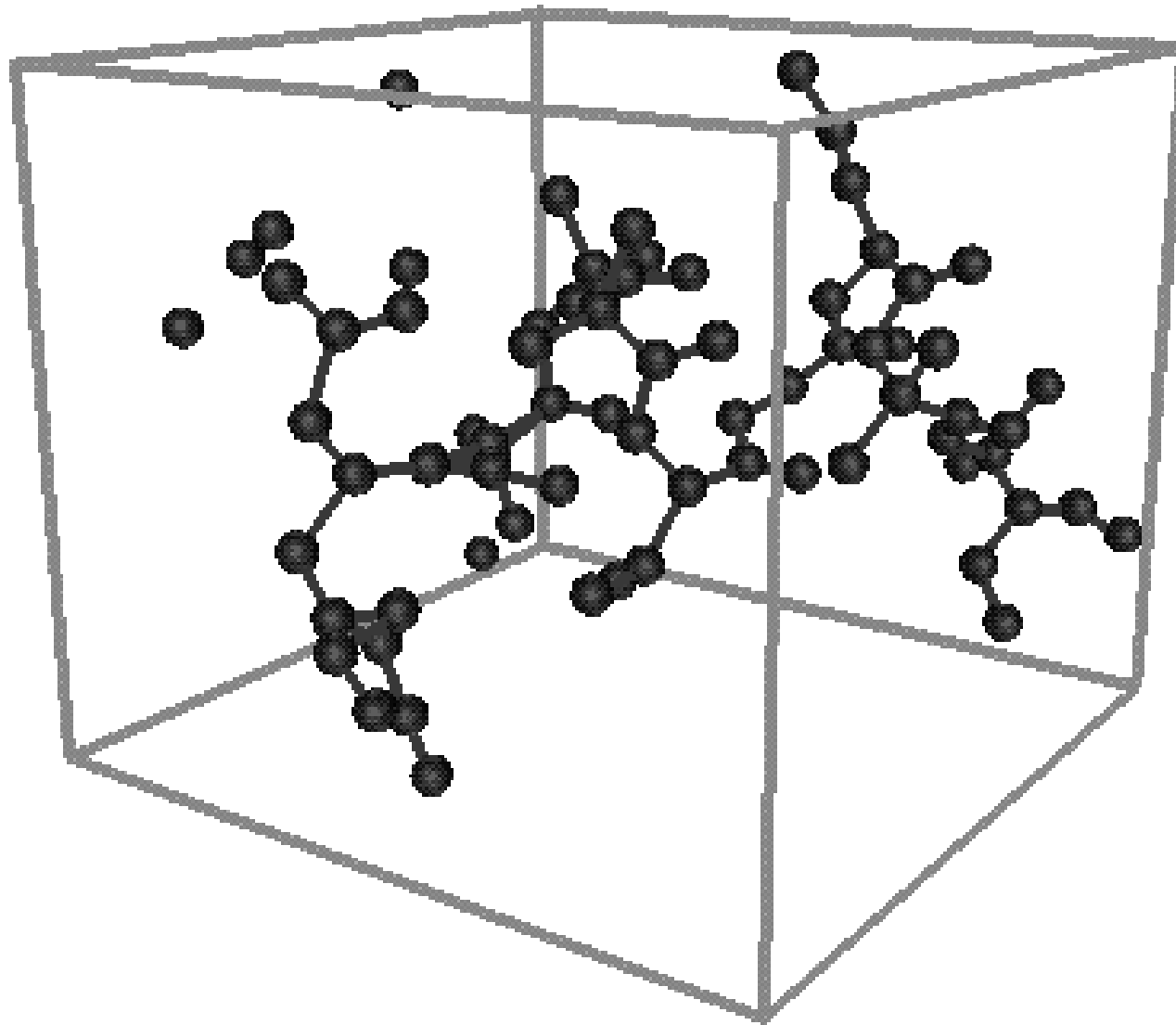
Phases: 740  
Triples: 7,400



# Visualization in *SnB* (Ph8755)

*Geomview*: Geometry Center, U. Minn.

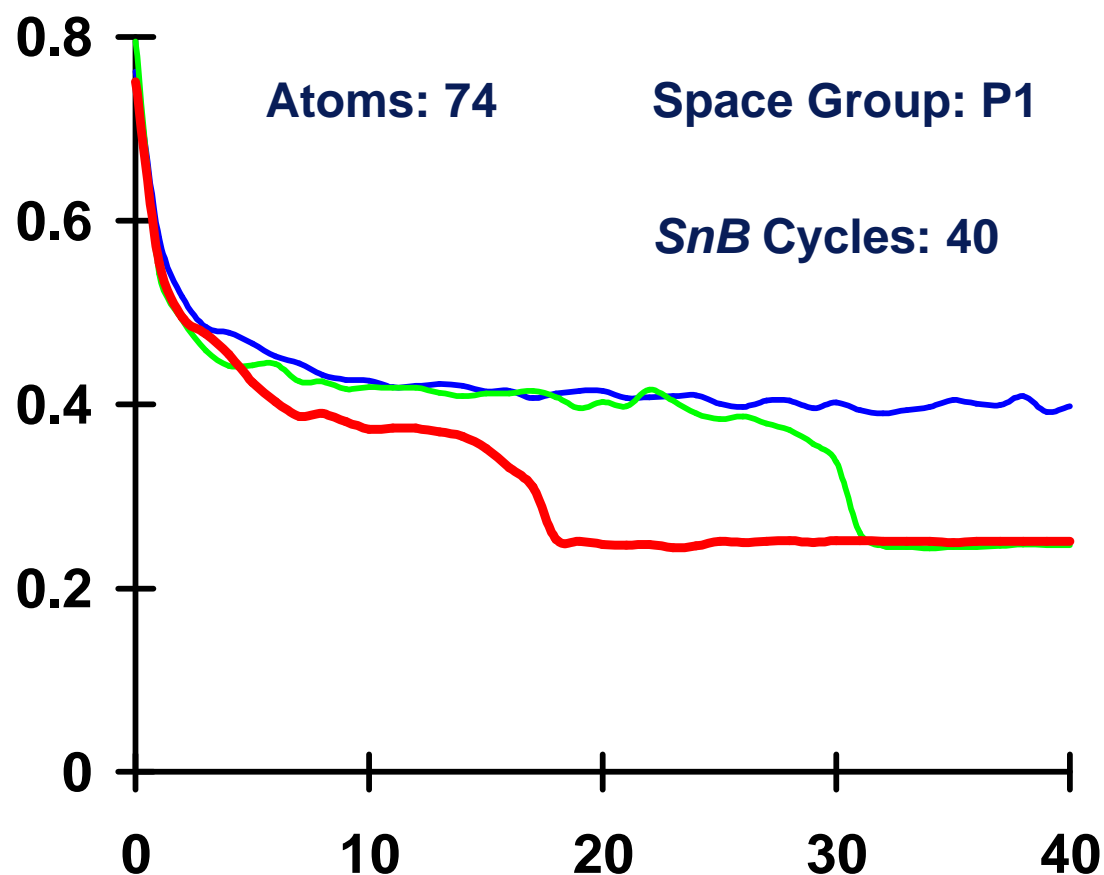
*SnB*



# The Minimal Function is Diagnostic

Structure: Ph8755

*SnB*





# ToxII: *SnB* Histogram

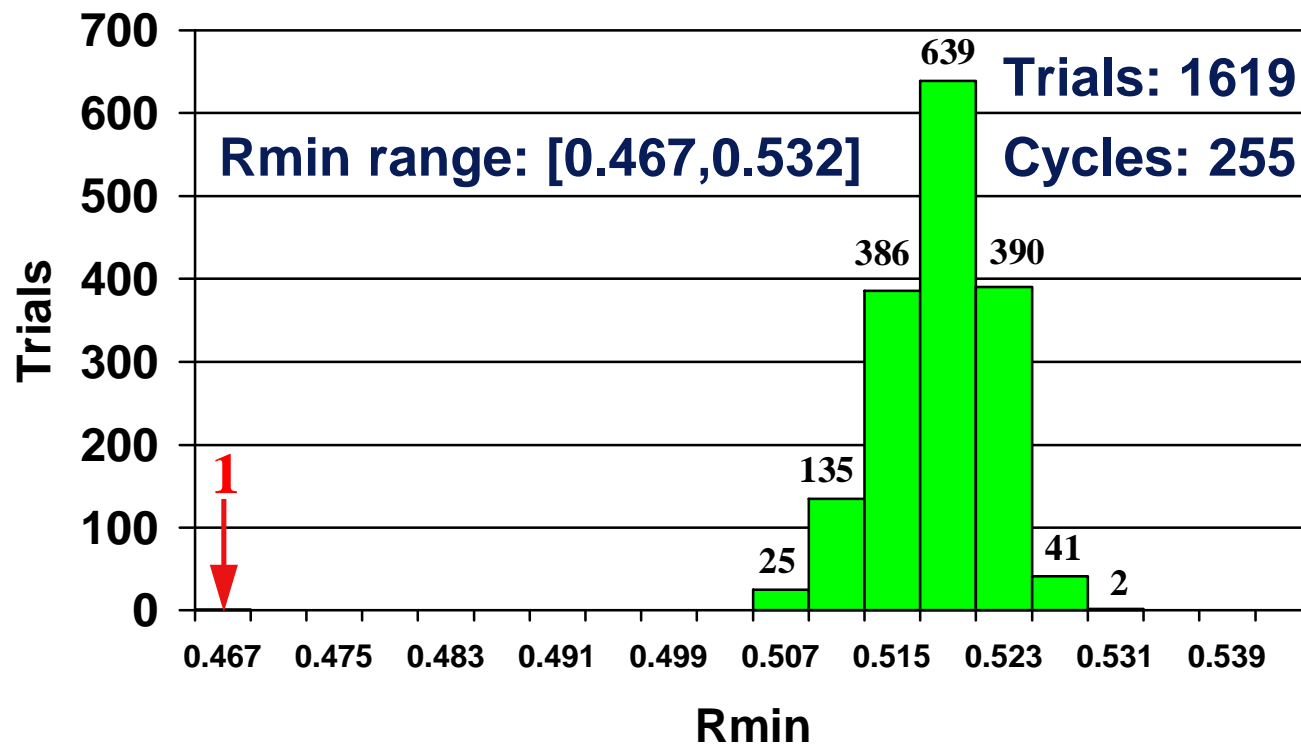
*SnB*

Atoms: 500

Phases: 5,000

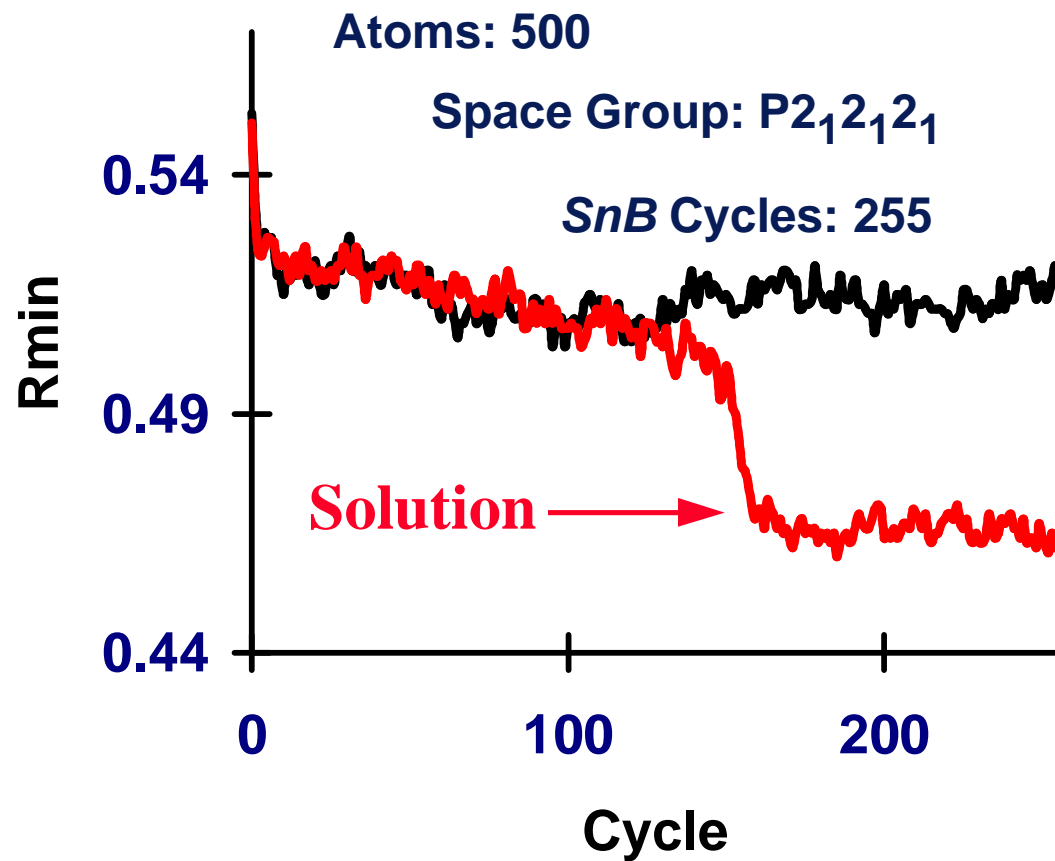
Space Group:  $P2_12_12_1$

Triples: 50,000



# Tox II: Trace of *SnB* Solution

*SnB*



# Some *SnB* Solutions



STRUCTURE	LOCATION	ATOMS	SPACE GROUP
Scripps	Scripps	144	P1
Riboflavin tetrabutryate	MSC	180	P1
Vancomycin	Penn	255	P <sub>4,2,2</sub>
Er-1 pheromone	UCLA	302	C2
Gramicidin A	HWI	317	P <sub>2,2,2</sub> <sub>1</sub>
Crambin	HWI	400	P <sub>2</sub> <sub>1</sub>
Alpha-1 peptide	OCI/U. of T.	450	P1
Rubredoxin	HWI	500	P <sub>2</sub> <sub>1</sub>
ToxII	HWI	630	P <sub>2,2,2</sub> <sub>1</sub>

# Additional *SnB* Solutions

The logo for SnB, featuring the letters 'SnB' in a stylized, italicized font. The 'S' and 'n' are red, and the 'B' is blue. The letters are slightly offset and shadowed to create a 3D effect.

<b>STRUCTURE</b>	<b>LOCATION</b>	<b>ATOMS</b>	<b>RES</b>	<b>SPACE GROUP</b>
<b><math>\alpha</math> -conotoxin PnIA</b>	<b>Queensland</b>	<b>110</b>	<b>1.1Å</b>	<b>P2<sub>1</sub></b>
<b><math>\alpha</math> -conotoxin PnIB</b>	<b>Queensland</b>	<b>111</b>	<b>1.1Å</b>	<b>C2</b>
<b><math>\alpha</math> -conotoxin G1</b>	<b>Queensland...</b>	<b>117</b>	<b>1.1Å</b>	

# Notes on Structures

**SnB**

<b>STRUCTURE</b>	<b>ATOMS</b>	<b>SPACE GROUP</b>	<b>SUCCESS RATE</b>
Ternatin	110	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	2%
Scripps	144	P1	
Riboflavin tetrabutyrates (MSC)	180	P1	(8 years with traditional methods)
Er-1 pheromone	302	C2	
Gramicidin A	317	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	0.3%
Crambin	400	P2 <sub>1</sub>	4%
Alpha-1 peptide	408	P1	5%
Rubredoxin	500	P2 <sub>1</sub>	2.7%
ToxII	630	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	1/1619

# Computing Platforms

---

*SnB*

## ◆ Unix Workstations

- ❑ SGI, Sun, DEC/Alpha
- ❑ IBM RS6000
- ❑ Wintel/Linux

## ◆ Parallel Computers

- ❑ Thinking Machines Corp. CM-5 (NCSA)
  - ❑ Cray T3D (PSC)
  - ❑ IBM SP2 (CTC)
- ## ◆ Cray C90 (PSC)

# SnB Sites



---

Abbott Labs	Clemson	IVIC (Venezuela)	Pittsburgh	Windsor
Adelaide	CUHK (Hong Kong)	Keene State	Roche	Yale
Ames	Dalhousie	Kentucky, Louisville	Scripps	York
Ariad	Delaware	Loughborough	Shell	Zurich
ANU	Dow	McMaster	Southampton	Kodak
Berkeley	Dupong	Michigan State	South Carolina	LLNL
Birmingham	Duquesne	Minnesota	St. Andrews	SmithKline Beecham
Boehringer Ingelheim	East Anglia	MIT	Stanford	NIH
Boston College	Genetics Institute	NIU - Illinois	Sydney	Cornell
Brandeis	Griffith	Northwestern	SYSNU (Taiwan)	UCSF
BYUH	Guelph	Oak Ridge (ORNL)	UBC	Imerial College (London)
Calgary	Helsinki	Pavia	UCLA	Caltech
Cambridge	Hoffman Laroche	Univ. Penn.	UGA	Madras
Chinoin (Hungary)	Hong Kong University	Pfizer	USC	Univ. Queensland (Australia)

# Looking Ahead: *SnB* vers. 2.0

---

- ◆ **Speed: 5 - 10X**
- ◆ **Includes Inverse FFT**
- ◆ **Additional Density Modification Options**
- ◆ **Improved Fourier Recycling: “Twice Baking”**
- ◆ **Input Improvements:  $|E|$  Calculation**
- ◆ **Output Improvements: Compatibility**
- ◆ **(SIR/SAS/MAD Invariants with est. values)**



# Summary

---

*SnB*

- ◆ *Shake-and-Bake*: Enhanced Direct Methods
- ◆ Targeted at 100-800 atom structures
- ◆ Publicly available
- ◆ LEVY / EVAL (Bob Blessing) recommended for calculating  $|E|$ 's in *SnB* vers. 1.5
- ◆ *SnB* version 2.0
- ◆ Addresses:
  - <http://www.hwi.buffalo.edu/SnB/>
  - [snb-requests@hwi.buffalo.edu](mailto:snb-requests@hwi.buffalo.edu)
  - [snb-comments@hwi.buffalo.edu](mailto:snb-comments@hwi.buffalo.edu)

**SnB**



# Running Times

(Average Seconds/Cycle on Iled)



Machine	Total	StrFac	Refine	FFT	Pk Pk
DEC/Alcor 266	0.672	0.160	0.248	0.196	0.029
SGI Pwr Ind 2	1.247	0.210	0.426	0.485	0.068
DEC/Alpha 150	1.822	0.437	0.682	0.520	0.075
<b>SGI Indigo</b>	<b>2.975</b>	<b>0.881</b>	<b>0.760</b>	<b>1.037</b>	<b>0.163</b>
Sparc 10	3.140	0.821	0.919	1.111	0.162
Pentium 75	3.535	1.181	0.995	1.205	0.154
Sparc 5	3.759	0.765	1.087	1.661	0.210
i486dx2-66	11.396	2.175	2.423	5.467	0.711

# ToxII: Running Times

---

*SnB*

- ◆ **SGI Indigo (R4000):**
  - 50 seconds per cycle
  - 200 minutes per trial
  - *7 CPU months for 1600 trials*
- ◆ **IBM SP2 Projection:**
  - *4 hours for 1600 trials*