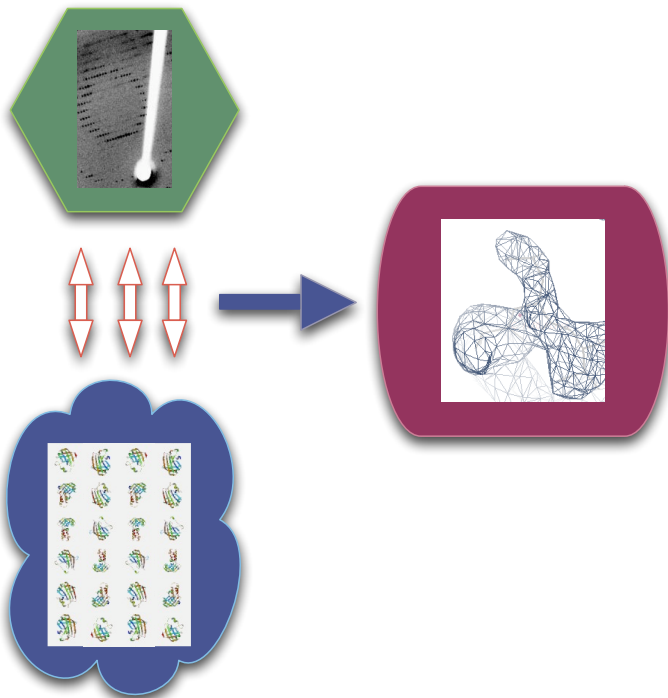


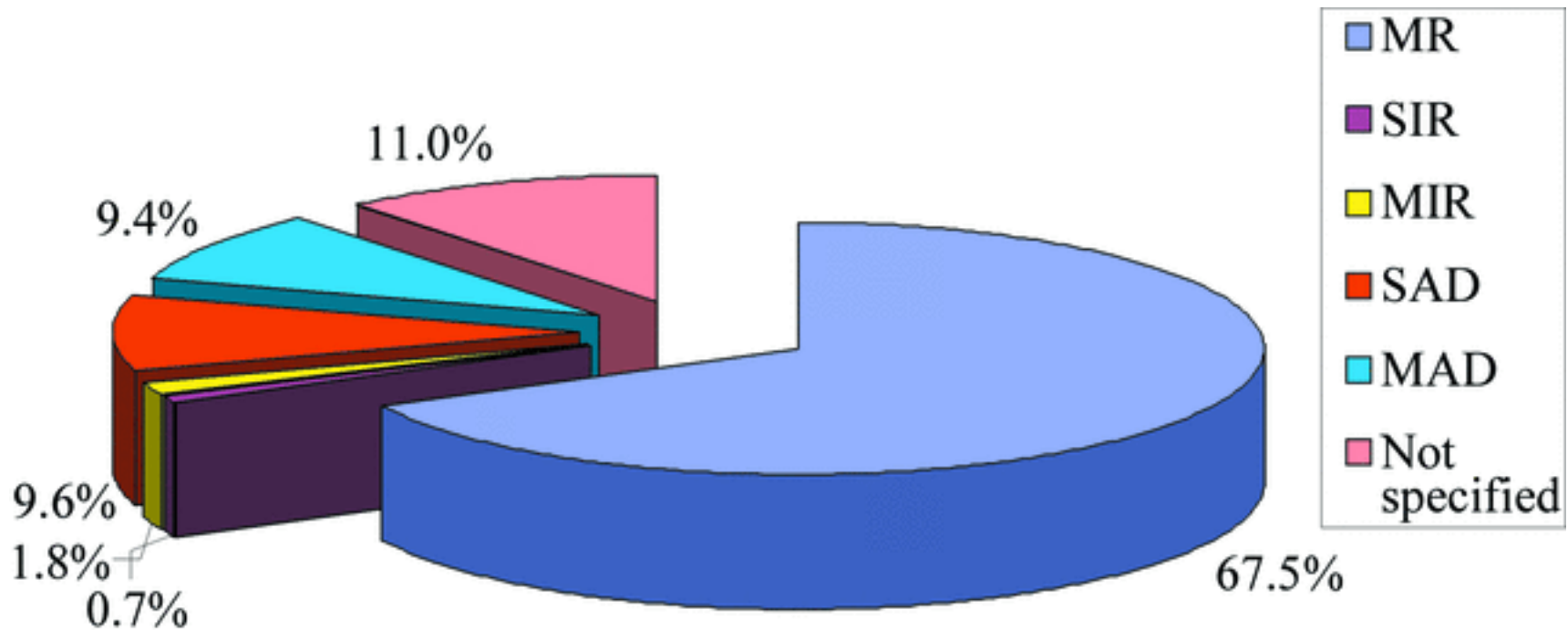
Comparative Protein Structure Modeling for Molecular Replacement in X-ray Crystallography



Ursula Pieper, David Eramian, Ben Webb, Narayanan Eswar, Jeff Bonanno, David Agard, Steve Almo, Stephen Burley, Andrej Sali
PSI2 Bottleneck Workshop
Bethesda, April 2008

Role of Molecular Replacement in X-ray Crystallography

X-ray Structure Solution Methods in 2006:



- Long F, Vagin AA, Young P, Murshudov GN. BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr.* **64**:125-132, 2008.

Why use Models for MR?

- Advantages of MR
 - **No need for Se-Met or heavy atom derivatives**
 - **More rapid structure determination**
- Disadvantages of MR
 - **Need a “good” related PDB structure or model**
 - **Possibly longer refinement effort due to phase bias**
- Impact of more suitable models for MR on PSI
 - **Less experiments**
 - **Less synchrotron time**
 - **More structures**

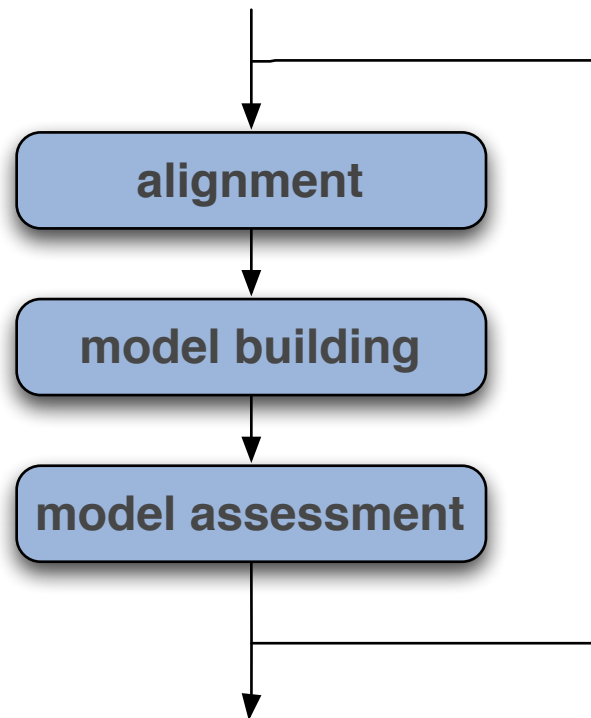
Limits of Molecular Replacement

MR depends on overcoming errors in comparative modeling:

- generally > 30% sequence identity (mostly correct alignment)
- no major shifts in secondary structure elements

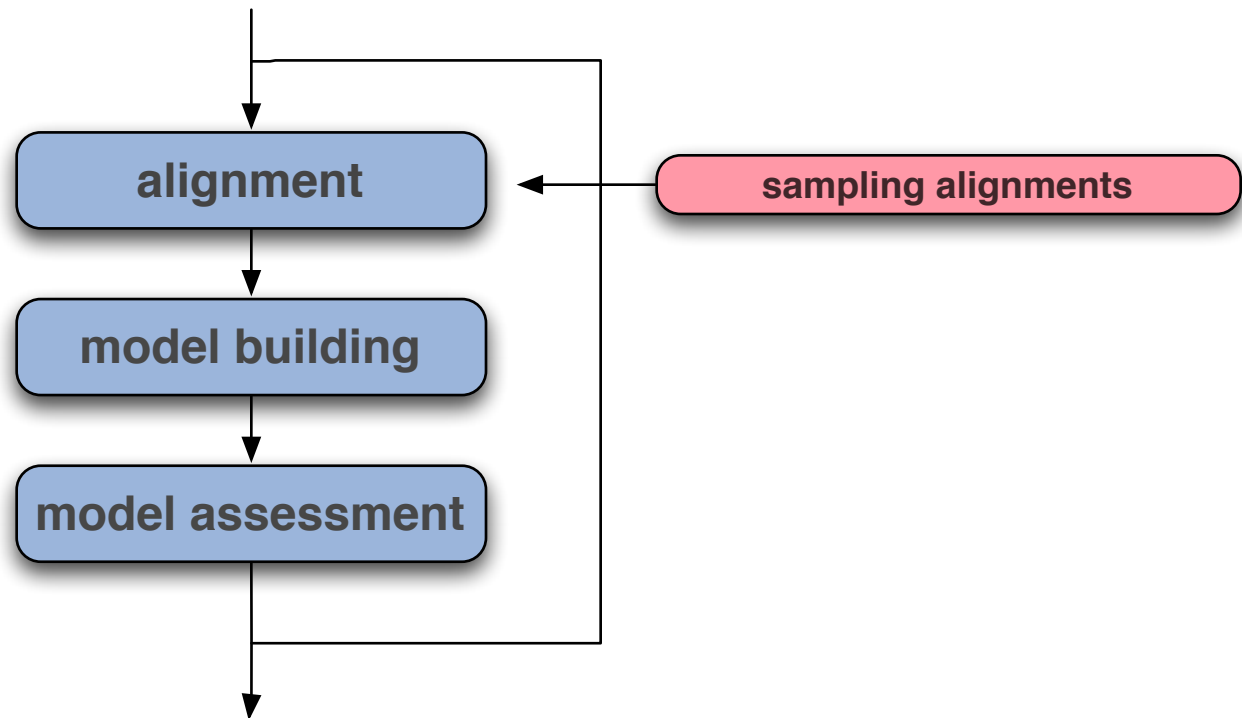
Solution Concept: Sampling the Alignment and Conformation Space

MOULDER: A genetic algorithm protocol for sampling alignments and conformations guided by a model fitness function



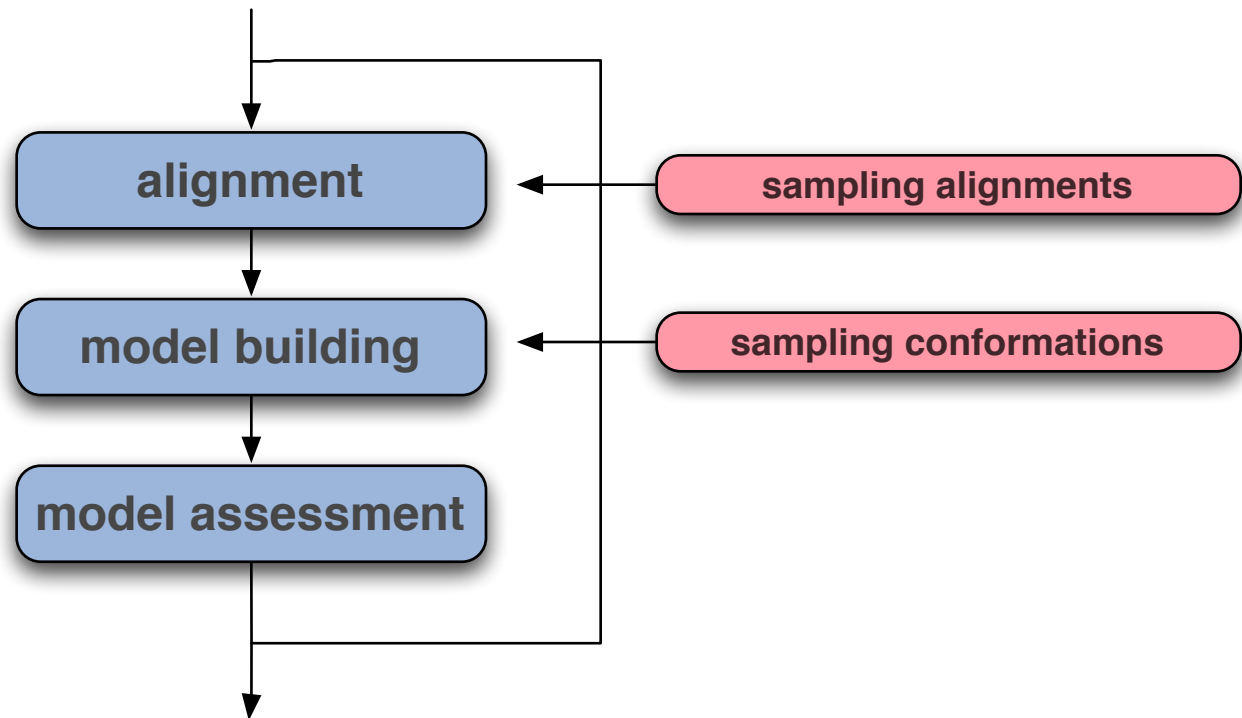
Solution Concept: Sampling the Alignment and Conformation Space

MOULDER: A genetic algorithm protocol for sampling alignments and conformations guided by a model fitness function



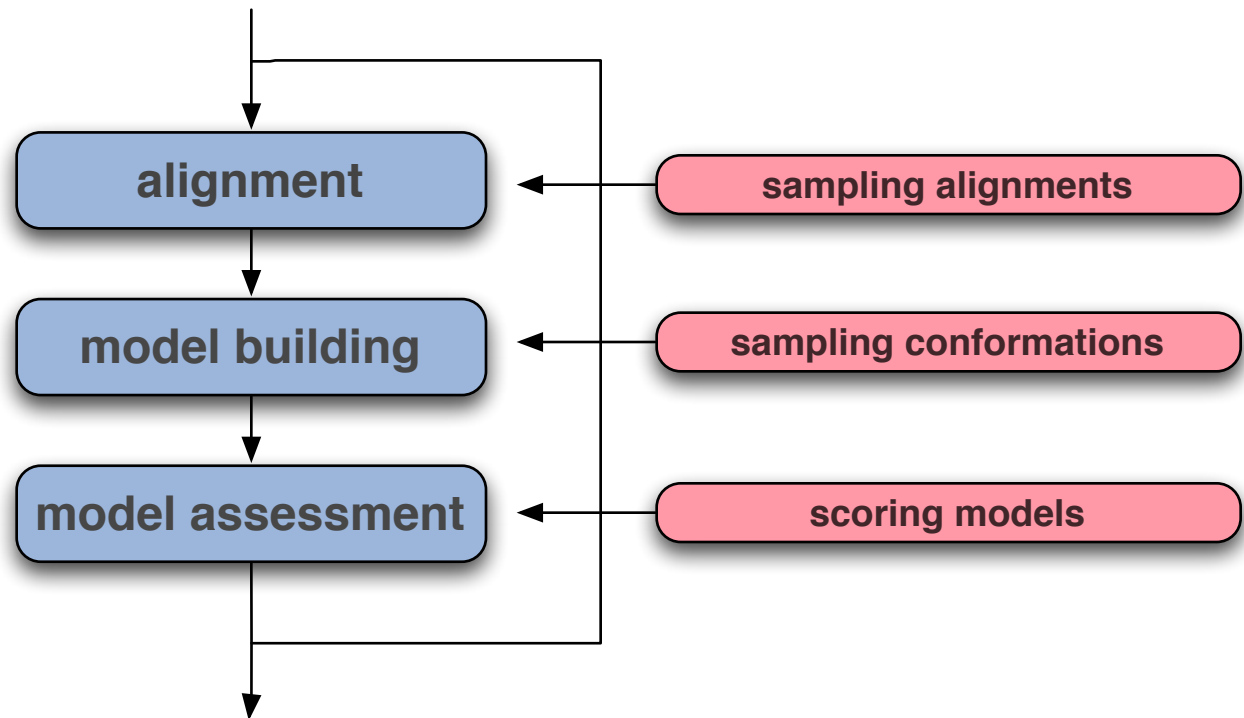
Solution Concept: Sampling the Alignment and Conformation Space

MOULDER: A genetic algorithm protocol for sampling alignments and conformations guided by a model fitness function

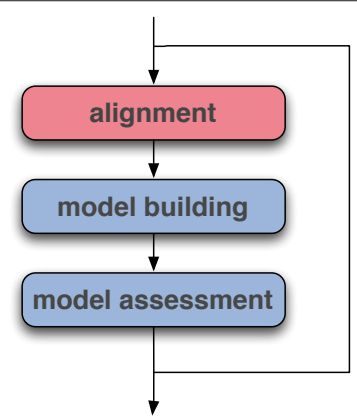


Solution Concept: Sampling the Alignment and Conformation Space

MOULDER: A genetic algorithm protocol for sampling alignments and conformations guided by a model fitness function



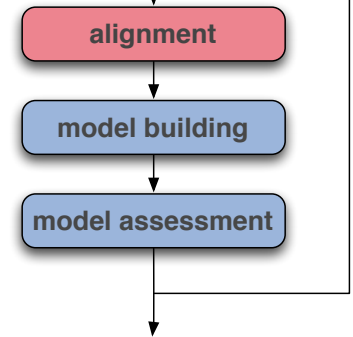
Sampling the Alignment Space



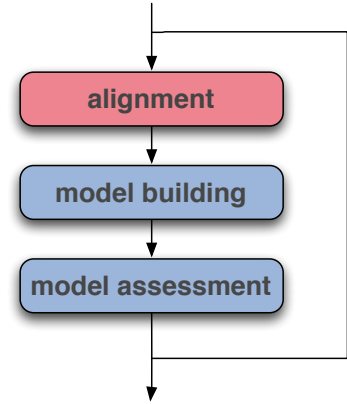
Sampling the Alignment Space

Parents

...TSSQNMKLGVFWGY... ...TSSQN--MKLGVFWGY...
...VSSCNGDLHMKVGV... ...VSSCNGDLHMKVG--V...



Sampling the Alignment Space

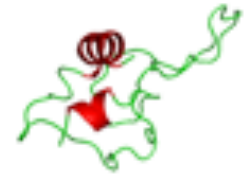


Parents

...TSSQNMKLGVFWGY... ...TSSQ--MKLGVFWGY...
...VSSCNGDLHMKVGV... ...VSSCNGDLHMKV--V...

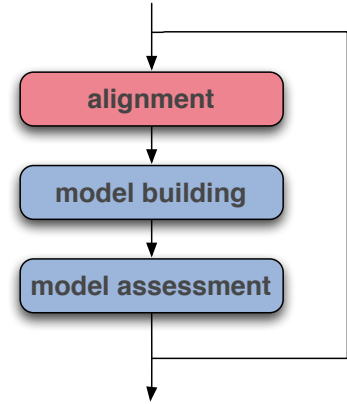
Children

..TSSQ-NMK---LGVFWGY.. ..TSSQ-NMCLGVFWGY---.. ..TSSQNMK---LGVFWGY.. ..TSSQNMKLGVFWGY---..
..V-SSCNGDLHMKV---GV.. ..V-SSCN---GDLHMKVGV.. ..VSSCNGDLHMKV---GV.. ..VSSCN---GDLHMKVGV..



Fitness function: Model evaluation score

Sampling the Alignment Space

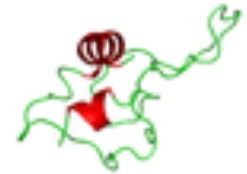


Parents

...TSSQNMKLGVFWGY... ...TSSQN--MKLGVFWGY...
 ...VSSCNGDLHMKVGV... ...VSSCNGDLHMKVGV--V...

Children

..TSSQ-NMK---LGVFWGY.. ..TSSQ-NMCLGVFWGY----.. ..TSSQNMK---LGVFWGY..
 ..V-SSCNGDLHMKV---GV.. ..V-SSCN---GDLHMKVGV.. ..VSSCNGDLHMKV---GV..
 ..TSSQNMKLGVFWGY----.. ..VSSCN---GDLHMKVGV..



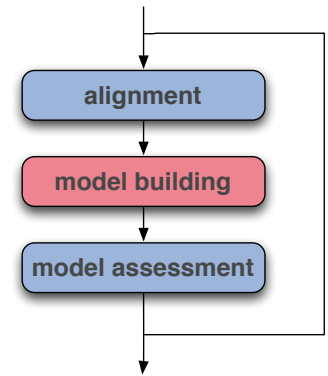
Fitness function: Model evaluation score

Most Fit Children

..TSSQ-NMK---LGVFWGY.. ..TSSQ-NMCLGVFWGY----.. ..TSSQNMK---LGVFWGY..
 ..V-SSCNGDLHMKV---GV.. ..V-SSCN---GDLHMKVGV.. ..VSSCNGDLHMKV---GV..
 ..TSSQNMKLGVFWGY----.. ..VSSCN---GDLHMKVGV..



Sampling the Conformational Space



- **Current model building sampling (Modeller):**
 - **Conjugate gradients, MD, simulated annealing¹**
- **Improved sampling for model building in progress:**
 - **Langevin Dynamics²** (adds “random noise” terms to MD sampling)
 - **Self-Guided Langevin Dynamics³** (adds self-adapting terms based on current fluctuations to MD sampling)
 - **Replica Exchange MD⁴** (parallel MD simulations (replicas) at different temperatures / exchange replicas for better sampling)

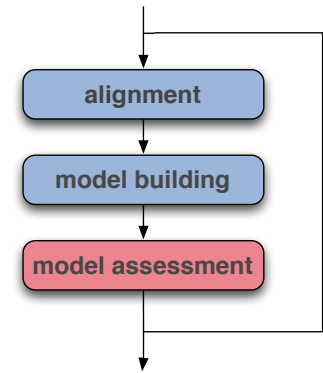
1. Sali, A, Blundell, TL (1993). J. Mol. Biol. 234, 779-815.

2. Loncharich, RJ, Brooks, BR, & Pastor, R. W. (1992). Biopolymers 32, 523-535.

3. Wu, X & Brooks, BR (2003). Chem. Phys. Lett. 381, 512-518.

4. Sugita, Y, Okamoto Y (1999). Chem. Phys. Lett. 314: 141-151.

Model assessment for MR



Model fitness function:

- **DOPE¹ (Modeller): Atomic distance-dependent statistical potential**
- **MPQS² (MODPIPE): Composite Score (DOPE, sequence similarity, etc)**
- **TSVMod³ (IMP): Tailored Support Vector Machine Score**
- **MR scores⁴: from MR programs (Phaser, MolRep, Amore)**

1. Shen, MY, Sali A (2006). Protein Science **15**, 2507 - 2524.

2. Eswar N *et al* (2003). Nucleic Acids Research **31**, 3375-3380.

3. Eramian, D, Eswar, N, Shen, MY, Sali, A, submitted.

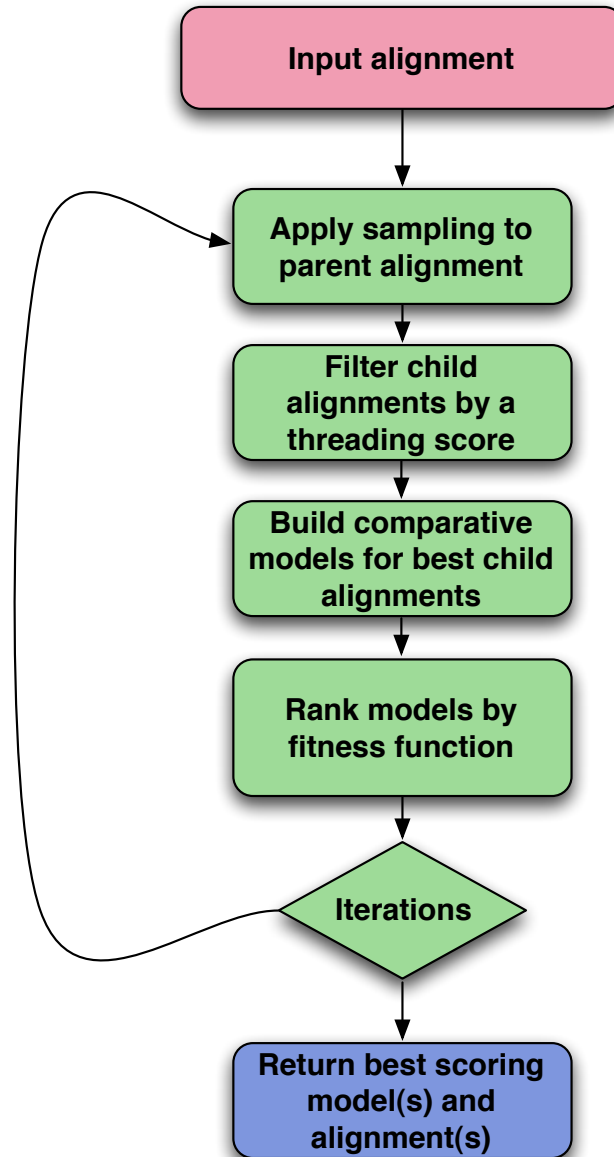
4. Molecular Replacement:

Storoni LC, McCoy AJ, Read RJ. (2004). Acta Cryst. D Biol. Cryst. **60**, 432-438.

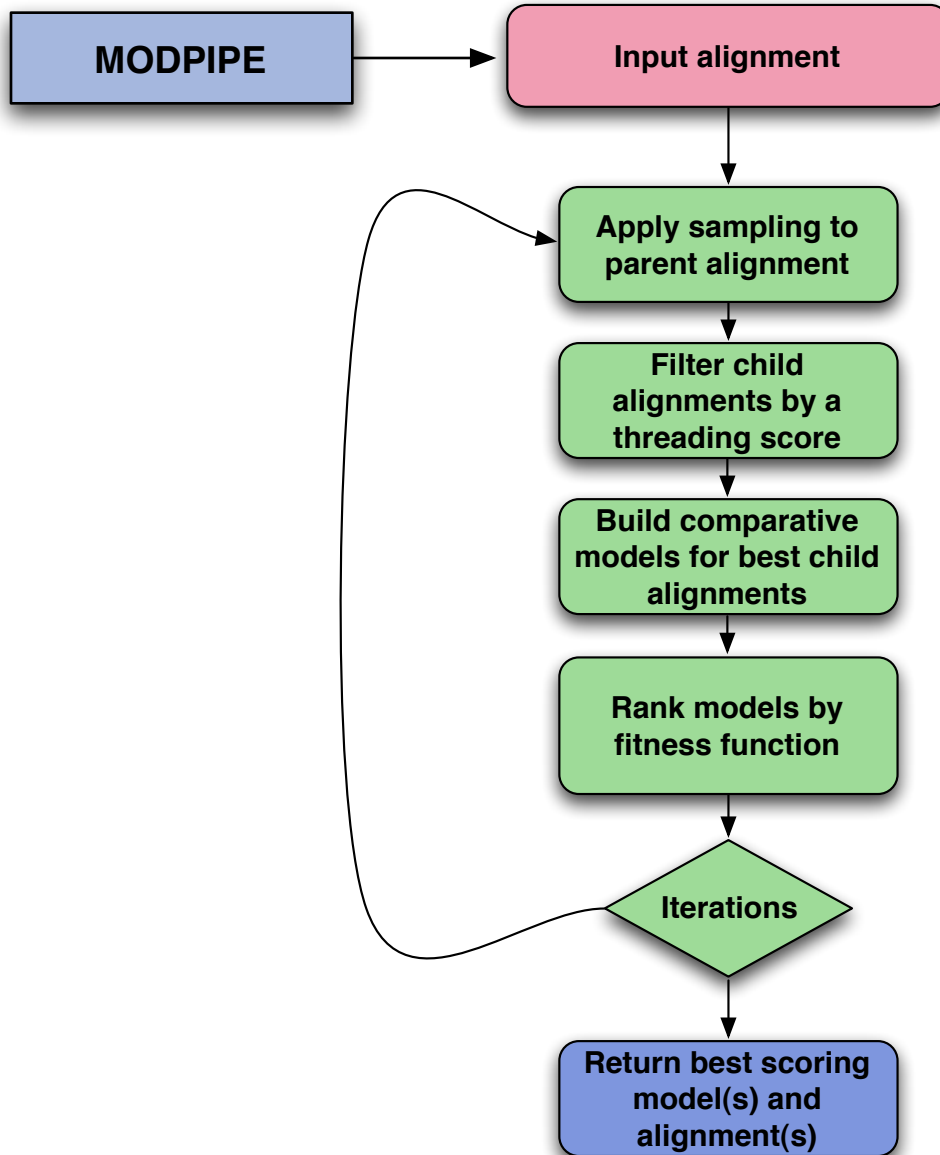
Trapani S, Navaza J. (2008). Acta Cryst. D Biol. Cryst. **64**,11-16.

Vagin A, Teplyakov A. (2000). Acta Cryst. D Biol. Cryst. **56**,1622-1624.

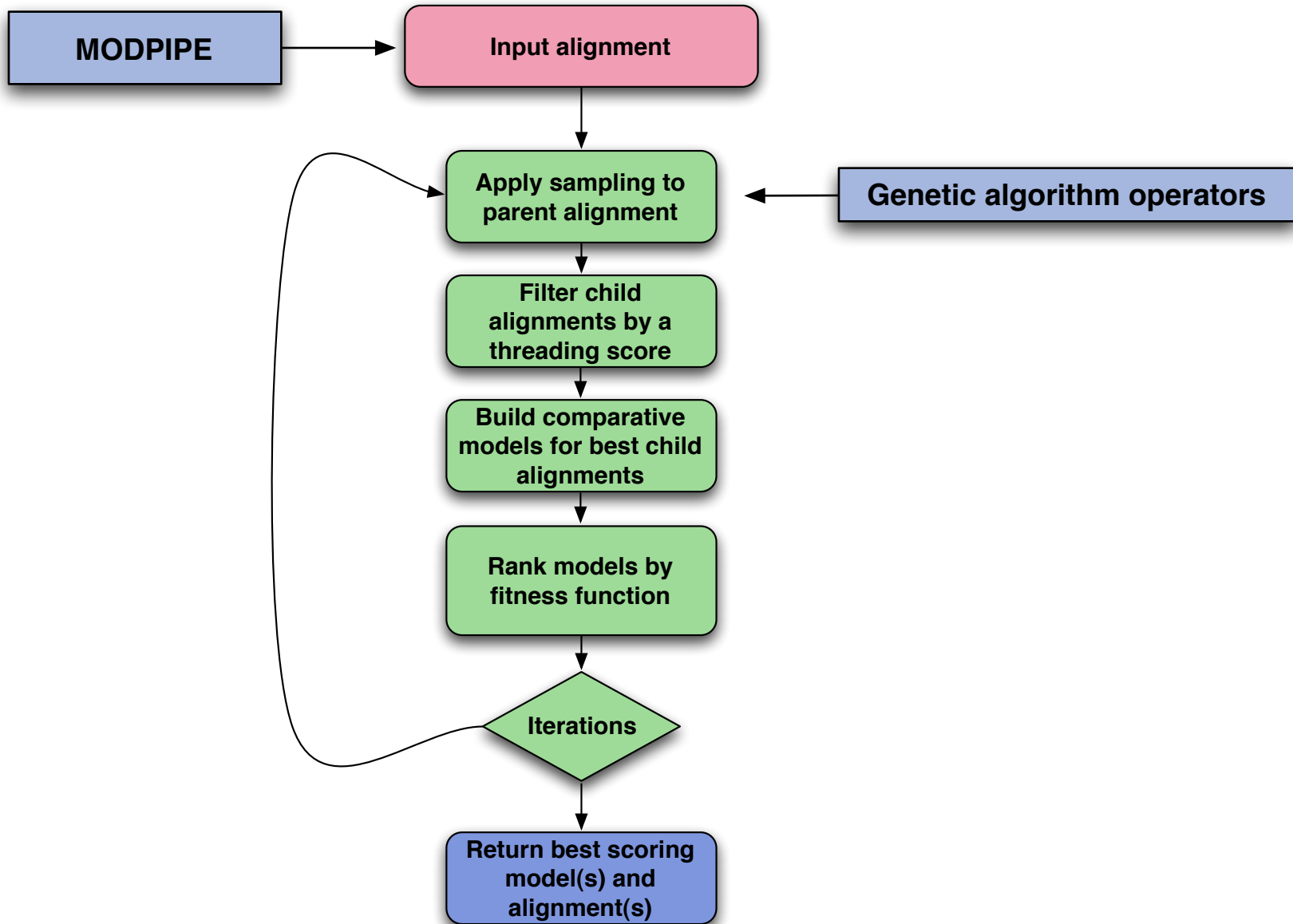
Flowchart of MOULDER for MR



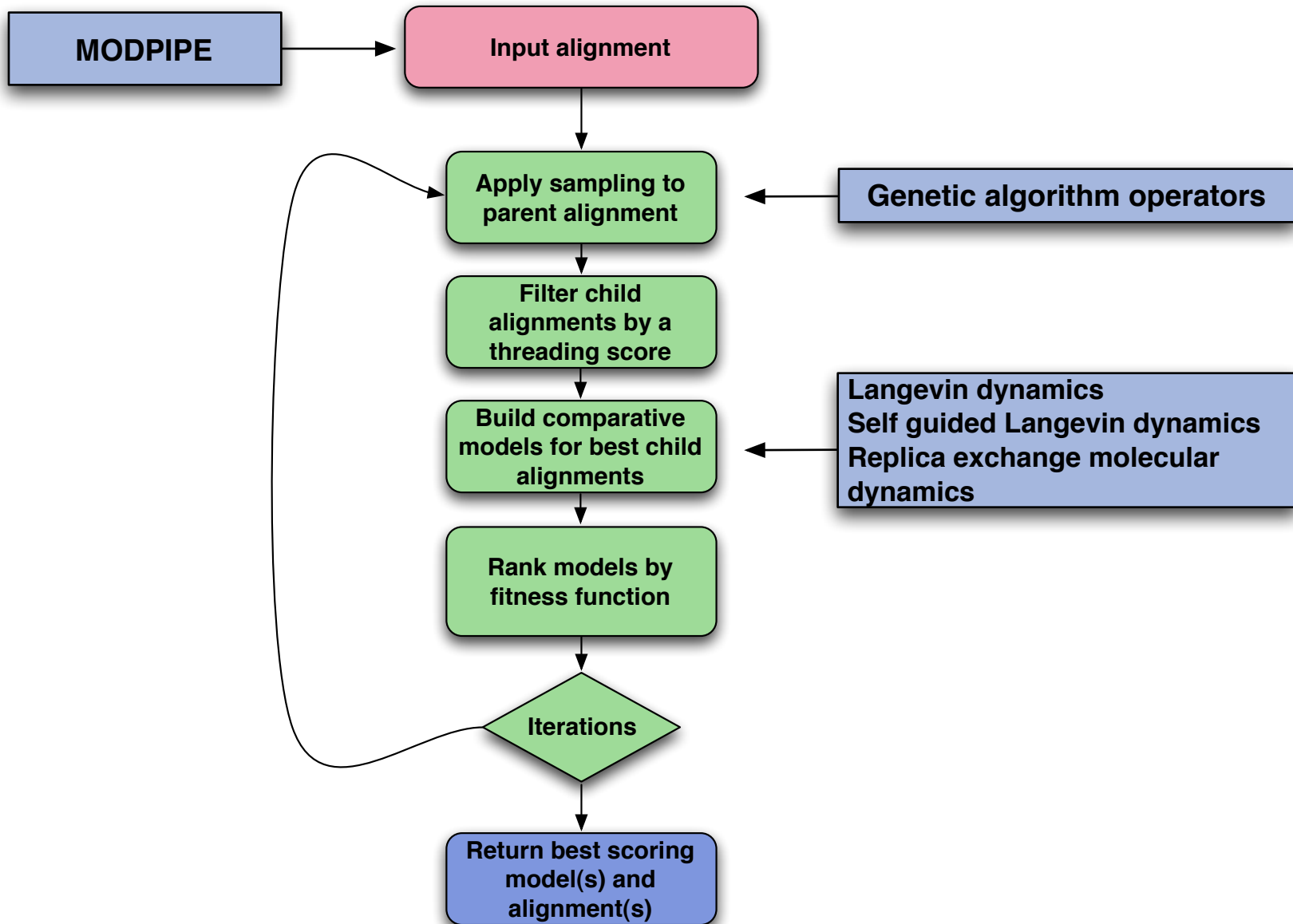
Flowchart of MOULDER for MR



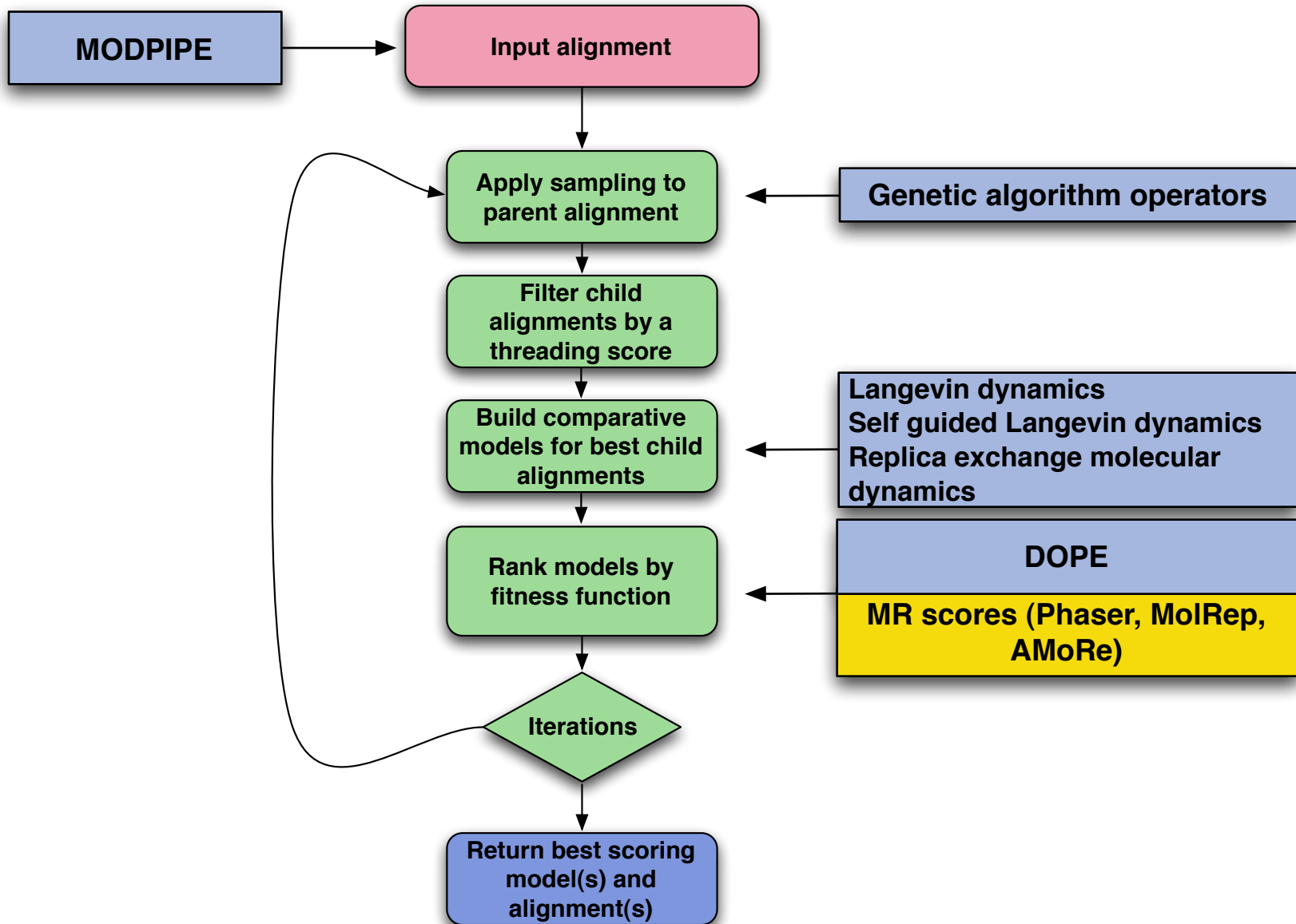
Flowchart of MOULDER for MR



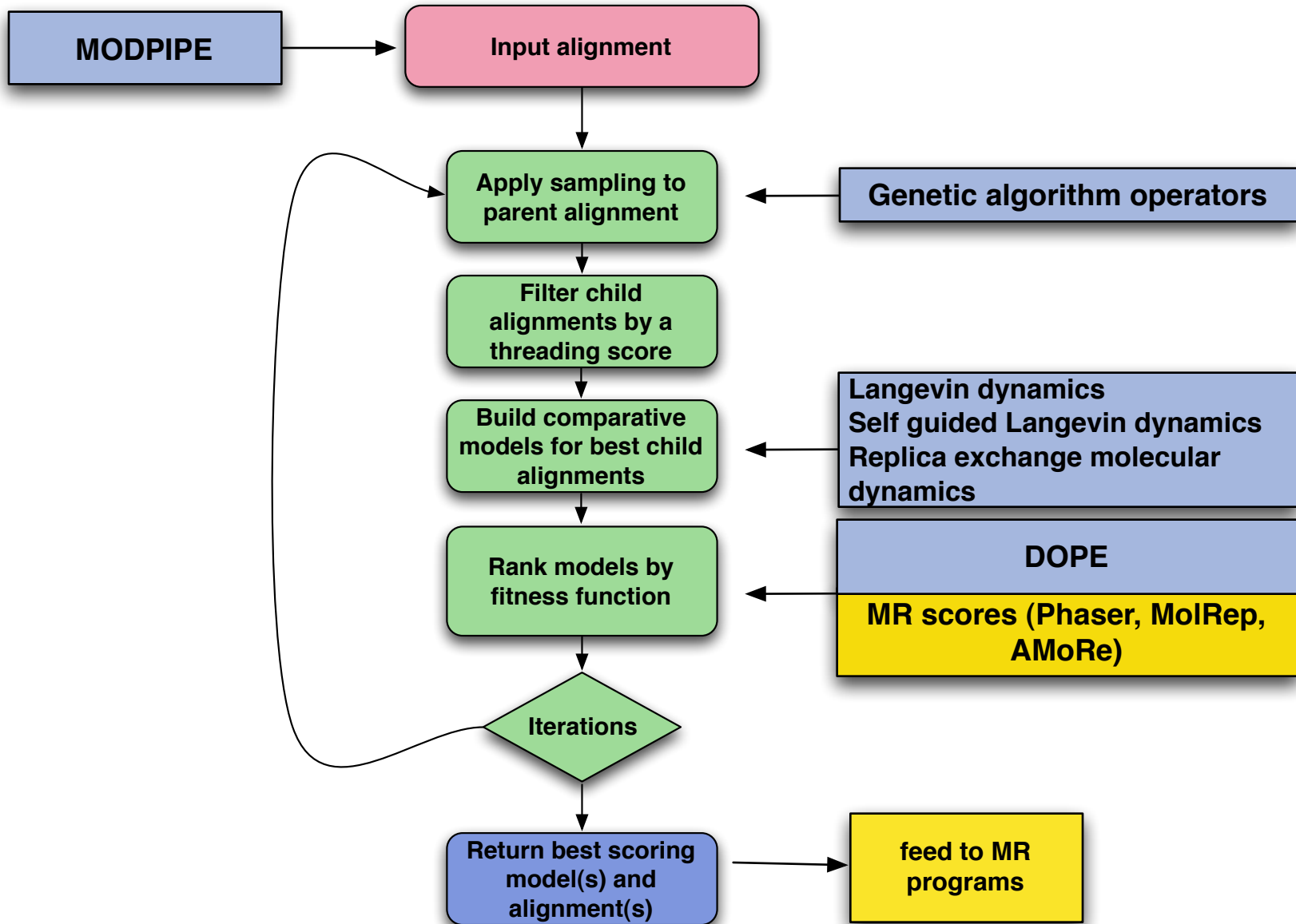
Flowchart of MOULDER for MR



Flowchart of MOULDER for MR

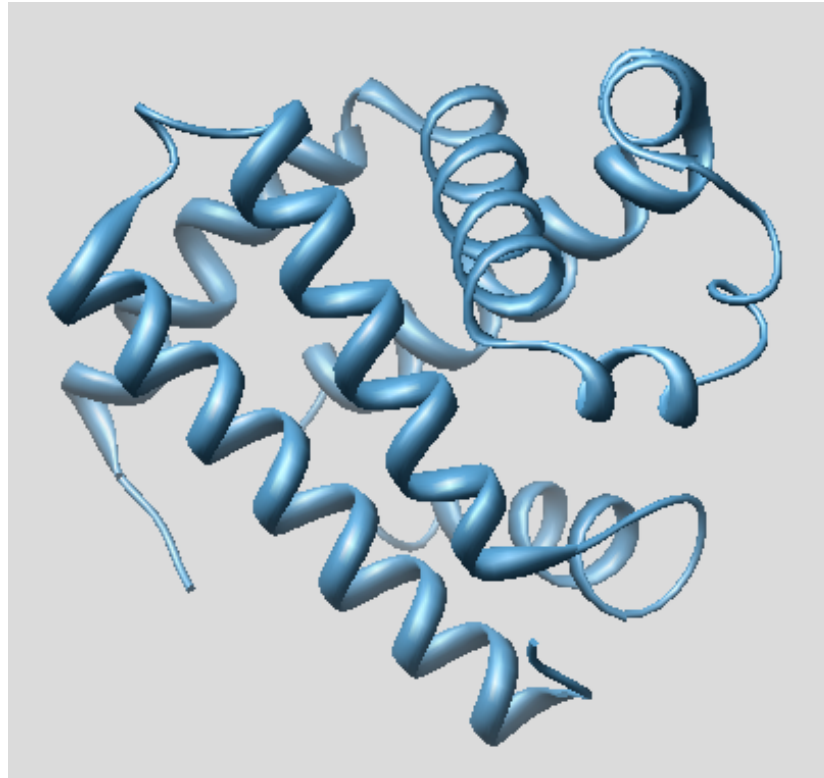


Flowchart of MOULDER for MR

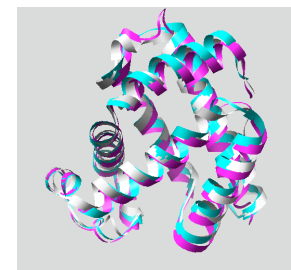


First Test Case: PDB code 106m (Sperm whale myoglobin)

- Space group: P6
- R value: 15.5%
- Resolution: 1.99 Å
- One molecule in the asymmetric unit
- 156 amino acid residues, all helical



MR Results for 106m



PDB File	Model Type	Sequence Identity	RFZ	TFZ	LLG	Solved
1ux9	Initial	30%	3.0	4.0	20/20	no
1c0h	Initial	24%	3.9	3.8	18/19	no
1a0w	Initial	23%	3.0	3.8	18/19	no
1ecn	Initial	20%	4.4	-	17/17	no
1ux9	Moulder	30%	3.1	5.1	19/19	no
1c0h	Moulder	24%	4.1	4.0	20/21	no
1a0w	Moulder	23%	3.4	3.7	17/17	no
1ecn	Moulder	20%	3.8	3.9	16/16	no
Combined	Initial		3.3	4.0	17/17	no
Combined	Moulder		3.3	9.6	51/52	yes
Combined 3	Moulder		3.9	3.5	15/15	no

MR scores using PHASER:

RFZ: Rotation function Z score

TFZ: Translation function Z score

LLG: Log likelihood gain

Initial: MODPIPE model

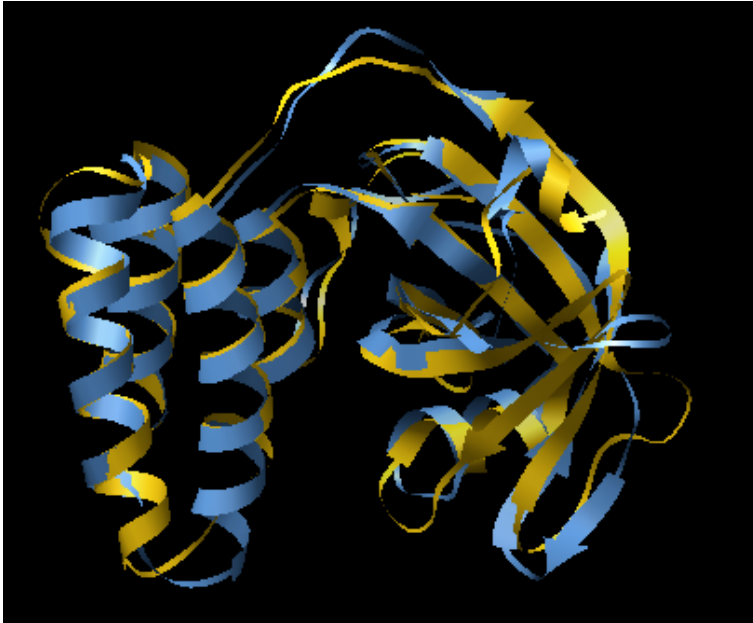
Moulder: MOULDER model

Combined: search ensembles from several models

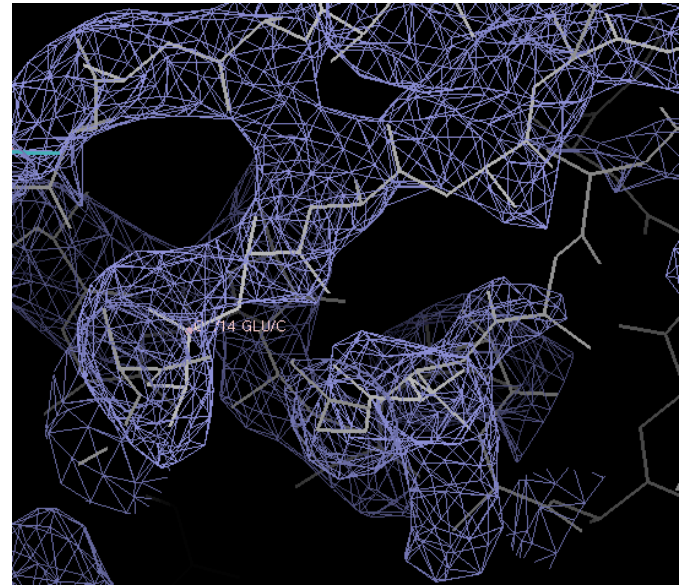
Example: NYSGXRC 10069e1

- I222 or I2₁2₁2₁, 4 molecules in the asymmetric unit, 251 aa residues
- Closest structures:
 - 2nr4 (33% sequence identity)
 - 2iml (32%)
 - 2ptf (23%)
- Best MR Scores: RFZ: 4.5, TFZ: 8.1 (but with clashes)
- Best solution without clashes after MOULDER model refinement (backbone only):
 - I222,
 - RFZ: 3.8-4.5
 - TFZ: 4.3-6.0
 - LLG: 114

Example: 10069e1



Two search models



Electron density snippet for backbone-only search model: some promising features, but a lot of chain breaks. Keep trying ...

Benchmark

(in progress)

- 7 structures in the twilight zone, from PDB
- X-ray data resolution range: 1.4-2.9 Å
- One molecule in the asymmetric unit

approximately 10,000 CPU hours

Benchmark

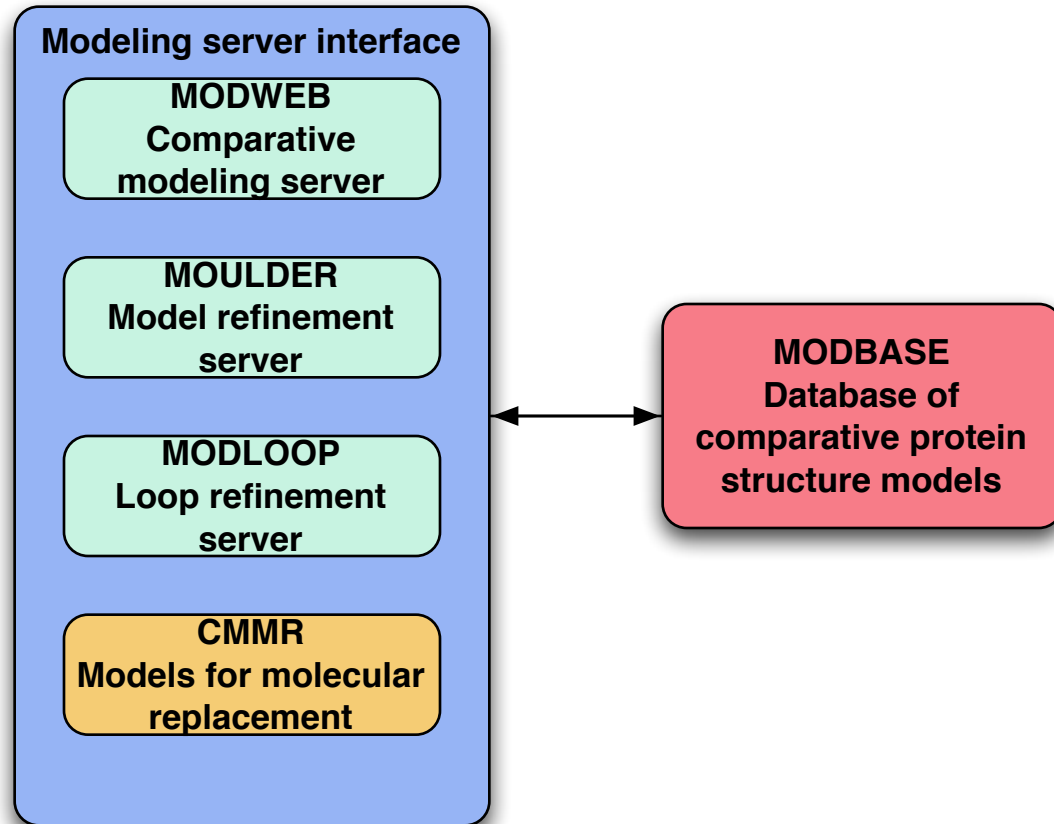
(in progress)

- 7 structures in the twilight zone, from PDB
- X-ray data resolution range: 1.4-2.9 Å
- One molecule in the asymmetric unit

Project	# Models	Max Sequence Identity	Min Sequence Identity	# Moulder Calculations	# MR tests
1c3k	8	30	26	2	4
1kw3	5	27	21	1	4
1mtz	12	28	23	3	4
1xp3	9	32	16	1	5
1yz7	7	29	20	1	6
2a0j	7	27	24	1	5
1yk9	6	29	24	1	5

approximately 10,000 CPU hours

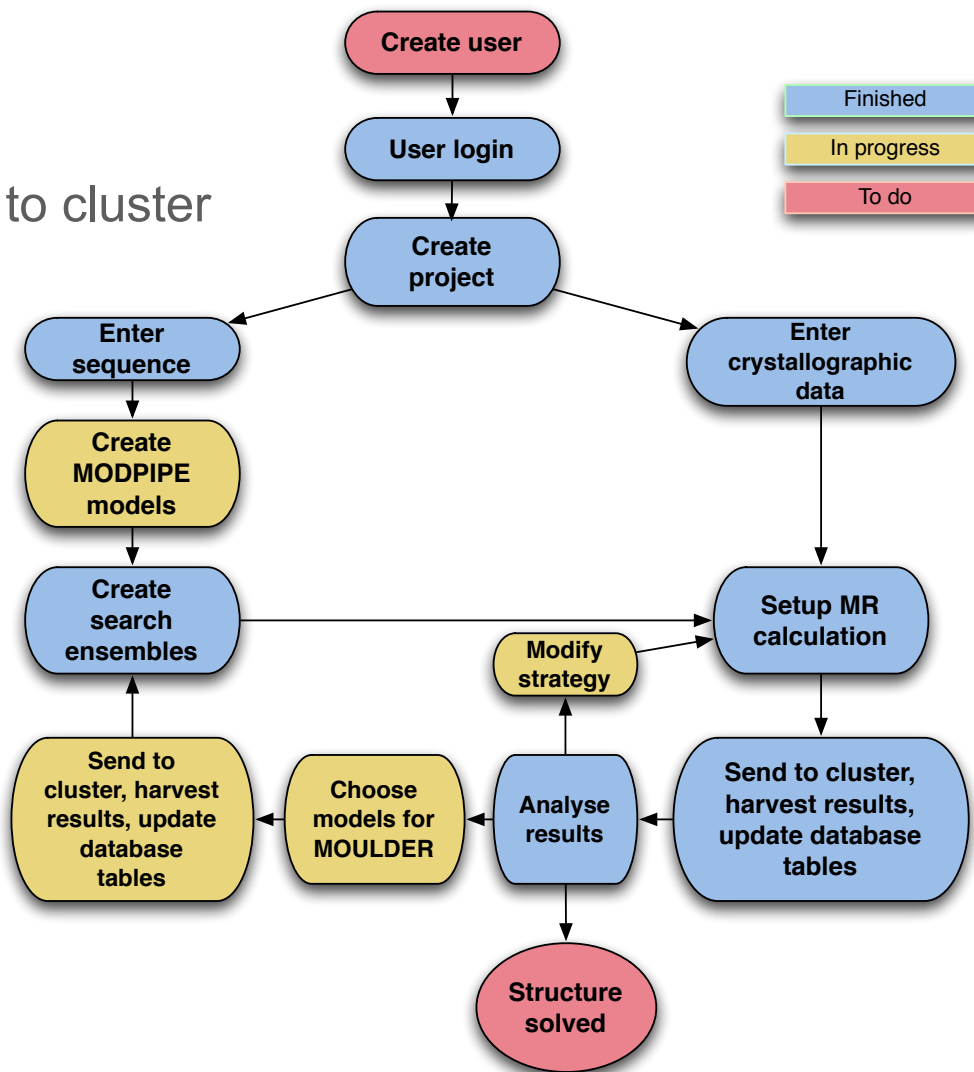
Modeling web server



CMMR Server

Interactive web server:

- Upload data
- Submit calculations to cluster
- Retrieve results



CMMR Server: Screen Shots

Select Models

Model Selection for 1kw3



Currently selected Models:

Database ID Model Count

1kw3 5

Sequence	PDB Code	Begin	End	SeqIdent	GA341	MPQS	zDope	Score	Value	Method	Description
1kw3	1f1xA	2	284	21	1	1.19	0.11			modpipe2.0	four with full coverage
1kw3	2ig9A	2	284	21	1	1.19	0.12			modpipe2.0	four with full coverage
1kw3	1f1uA	2	284	22	1	1.19	0.18			modpipe2.0	four with full coverage
1kw3	2ig9A	70	288	27		0.92	0.88	1.02		modpipe2.0	five with full coverage
1kw3	2ig9A	70	288	27		dope	0.532045		moulder	moulder	model

MODEL	PDB Code	Target Beg	Target End	SeqIdent	MPQS	zDope	Evalue	GA341
-------	----------	------------	------------	----------	------	-------	--------	-------

Model Description:

<input type="checkbox"/>	MODEL	PDB Code	Target Beg	Target End	SeqIdent	MPQS	zDope	Evalue	GA341
<input type="checkbox"/>		1kw3B	1	288	100.00	2.17	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	100.00	2.19	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	100.00	2.22	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	100.00	2.22	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	100.00	2.18	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	98.00			6e-46	1.00
<input type="checkbox"/>		1igtA	1	288	67.00	1.86	-1.18	0	1.00
<input type="checkbox"/>		1igtA	1	288	67.00	1.87	-1.18	0	1.00
<input type="checkbox"/>		1igtA	1	288	67.00	1.81	-1.22	0	1.00
<input type="checkbox"/>		1igtA	1	288	67.00	1.84	-1.18	0	1.00
<input type="checkbox"/>		1igtA	1	288	65.00	1.79	-1.03	0	1.00
<input checked="" type="checkbox"/>		2ig9A	70	288	27.00	0.88	1.02	7.8e-05	0.92
<input type="checkbox"/>		1nkiA	2	118	26.00	0.62	0.42	0	0.97
<input type="checkbox"/>		1nkiA	2	116	25.00	0.58	0.31	4.7e-12	0.95

CMMR Server: Screen Shots

Select Models

Model Selection for 1kw3



Currently selected Models:

Database ID Model Count

Sequence	PDB Code	Begin	End	SeqIdent	GA341	MPQS	zDope	Score	Value	Method	Description
1kw3	5										
1kw3	1f1xA	2	284	21	1	1.19	0.11			modpipe2.0	four with full coverage
1kw3	2ig9A	2	284	21	1	1.19	0.12			modpipe2.0	four with full coverage
1kw3	1f1uA	2	284	22	1	1.19	0.18			modpipe2.0	four with full coverage
1kw3	2ig9A	70	288	27	0.92	0.88	1.02			modpipe2.0	five with full coverage
1kw3	2ig9A	70	288	27				dope	0.532045	moulder	moulder model

MODEL	PDB Code	Target Beg	Target End	SeqIdent	MPQS	zDope	Evalue	GA341
-------	----------	------------	------------	----------	------	-------	--------	-------

Model Description: Enter Model

<input type="checkbox"/>		1kw3B	1	288	100.00	2.17	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	100.00	2.19	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	100.00	2.22	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	100.00	2.22	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	100.00	2.18	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	98.00			6e-46	1.00
<input type="checkbox"/>		1igtA	1	288	67.00	1.86	-1.18	0	1.00
<input type="checkbox"/>		1igtA	1	288	67.00	1.87	-1.18	0	1.00
<input type="checkbox"/>		1igtA	1	288	67.00	1.81	-1.22	0	1.00
<input type="checkbox"/>		1igtA	1	288	67.00	1.84	-1.18	0	1.00
<input type="checkbox"/>		1igtA	1	288	65.00	1.79	-1.03	0	1.00
<input checked="" type="checkbox"/>		2ig9A	70	288	27.00	0.88	1.02	7.8e-05	0.92
<input type="checkbox"/>		1nkiA	2	118	26.00	0.62	0.42	0	0.97
<input type="checkbox"/>		1nkiA	2	116	25.00	0.58	0.31	4.7e-12	0.95

Create Ensemble

MODEL Comment PDB Code Target Beg Target End SeqIdent MPQS zDope Evalue GA341 Weight

Ensemble Description:

Exclude Loop Residues (by loop size)

<input type="checkbox"/>		five with full coverage	2ig9A	70	288	27	0.88	1.02	7.8e-05	0.92	<input type="text" value="10"/>
<input type="checkbox"/>		four with full coverage	1f1uA	2	284	22	1.19	0.18	0	1	<input type="text" value="10"/>
<input type="checkbox"/>		four with full coverage	1f1xA	2	284	21	1.19	0.11	0	1	<input type="text" value="10"/>
<input type="checkbox"/>		four with full coverage	2ig9A	2	284	21	1.19	0.12	0	1	<input type="text" value="10"/>
<input type="checkbox"/>		moulder model	2ig9A	70	288	27					<input type="text" value="10"/>

CMMR Server: Screen Shots

Select Models

Model Selection for 1kw3



Currently selected Models:

Database ID Model Count

Sequence	PDB Code	Begin	End	SeqIdnt	GA341	MPQS	zDope	Score	Value	Method	Description
1kw3	1f1xA	2	284	21	1	1.19	0.11			modpipe2.0	four with full coverage
1kw3	2ig9A	2	284	21	1	1.19	0.12			modpipe2.0	four with full coverage
1kw3	1f1uA	2	284	22	1	1.19	0.18			modpipe2.0	four with full coverage
1kw3	2ig9A	70	288	27	0.92	0.88	1.02			modpipe2.0	five with full coverage
1kw3	2ig9A	70	288	27				dope	0.532045	moulder	moulder model

MODEL	PDB Code	Target Beg	Target End	SeqIdnt	MPQS	zDope	Value	GA341
-------	----------	------------	------------	---------	------	-------	-------	-------

Model Description: Enter Model

<input type="checkbox"/>		1kw3B	1	288	100.00	2.17	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	100.00	2.19	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	100.00	2.22	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	100.00	2.22	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	100.00	2.18	-1.46	0	1.00
<input type="checkbox"/>		1kw3B	1	288	98.00			6e-46	1.00
<input type="checkbox"/>		1qtA	1	288	67.00	1.86	-1.18	0	
<input type="checkbox"/>		1qtA	1	288	67.00	1.87	-1.18	0	
<input type="checkbox"/>		1qtA	1	288	67.00	1.81	-1.22	0	
<input type="checkbox"/>		1qtA	1	288	67.00	1.84	-1.18	0	
<input type="checkbox"/>		1qtA	1	288	65.00	1.79	-1.03	0	
<input checked="" type="checkbox"/>		2ig9A	70	288	27.00	0.88	1.02	7.8e-05	0.92
<input type="checkbox"/>		1nkiA	2	118	26.00	0.62	0.42	0	
<input type="checkbox"/>		1nkiA	2	116	25.00	0.58	0.31	4.7e-12	

Moulder Submit

Submit Moulder Calculation for 1kw3

modbase

MODEL	PDB Code	Target Beg	Target End	SeqIdnt	MPQS	zDope	Value	GA341
-------	----------	------------	------------	---------	------	-------	-------	-------

Moulder Description: Submit Moulder

<input type="checkbox"/>	2ig9A	70	288	27	0.88	1.02	7.8e-05	0.92
<input type="checkbox"/>	1f1uA	2	284	22	1.19	0.18	0	1
<input type="checkbox"/>	1f1xA	2	284	21	1.19	0.11	0	1
<input type="checkbox"/>	2ig9A	2	284	21	1.19	0.12	0	1
<input type="checkbox"/>	2ig9A	70	288	27				

Submit Moulder

Moulder Parameters (optional)

Iterations (20)

Alignments per Iteration (250)

Models per Iteration (250)

Intermediate Models per Alignment (7)

Final Models per Alignment (11)

Scoring Method

Create Ensemble

MODEL	Comment	PDB Code	Target Beg	Target End	SeqIdnt	MPQS	zDope	Value	GA341	Weight
-------	---------	----------	------------	------------	---------	------	-------	-------	-------	--------

Ensemble Description: Create Ensemble

Exclude Loop Residues (by loop size)

<input type="checkbox"/>		five with full coverage	2ig9A	70	288	27	0.88	1.02	7.8e-05	0.92	<input type="text" value="10"/>
<input type="checkbox"/>		four with full coverage	1f1uA	2	284	22	1.19	0.18	0	1	<input type="text" value="10"/>
<input type="checkbox"/>		four with full coverage	1f1xA	2	284	21	1.19	0.11	0	1	<input type="text" value="10"/>
<input type="checkbox"/>		four with full coverage	2ig9A	2	284	21	1.19	0.12	0	1	<input type="text" value="10"/>
<input type="checkbox"/>		moulder model	2ig9A	70	288	27					<input type="text" value="10"/>

Create Ensemble

CMMR Server: Screen Shots

Select Models

Model Selection for 1kw3



Currently selected Models:

Database ID Model Count

Sequence	PDB Code	Begin	End	Seqldent	GA341	MPQS	zDope	Score	Value	Method	Description
1kw3	5										
1kw3	1f1xA	2	284	21	1	1.19	0.11			modpipe2.0	four with full coverage
1kw3	2ig9A	2	284	21	1	1.19	0.12			modpipe2.0	four with full coverage
1kw3	1f1uA	2	284	22	1	1.19	0.18			modpipe2.0	four with full coverage
1kw3	2ig9A	70	288	27	0.92	0.88	1.02			modpipe2.0	five with full coverage
1kw3	2ig9A	70	288	27				dope	0.532045	moulder	moulder model

MODEL	PDB Code	Target Beg	Target End	Seqldent	MPQS	zDope	Value	GA341
-------	----------	------------	------------	----------	------	-------	-------	-------

Model Description:

MODEL	PDB Code	Target Beg	Target End	Seqldent	MPQS	zDope	Value	GA341
<input type="checkbox"/>	1kw3B	1	288	100.00	2.17	-1.46	0	1.00
<input type="checkbox"/>	1kw3B	1	288	100.00	2.19	-1.46	0	1.00
<input type="checkbox"/>	1kw3B	1	288	100.00	2.22	-1.46	0	1.00
<input type="checkbox"/>	1kw3B	1	288	100.00	2.22	-1.46	0	1.00
<input type="checkbox"/>	1kw3B	1	288	100.00	2.18	-1.46	0	1.00
<input type="checkbox"/>	1kw3B	1	288	98.00			6e-46	1.00
<input type="checkbox"/>	1igtA	1	288	67.00	1.86	-1.18	0	
<input type="checkbox"/>	1igtA	1	288	67.00	1.87	-1.18	0	
<input type="checkbox"/>	1igtA	1	288	67.00	1.81	-1.22	0	
<input type="checkbox"/>	1igtA	1	288	67.00	1.84	-1.18	0	
<input type="checkbox"/>	1igtA	1	288	65.00	1.79	-1.03	0	
<input checked="" type="checkbox"/>	2ig9A	70	288	27.00	0.88	1.02	7.8e-05	
<input type="checkbox"/>	1nkiA	2	118	26.00	0.62	0.42	0	
<input type="checkbox"/>	1nkiA	2	116	25.00	0.58	0.31	4.7e-12	

Submit Moulder Calculation for 1kw3

modbase

MODEL	PDB Code	Target Beg	Target End	Seqldent	MPQS	zDope	Value	GA341
Moulder Description:								
<input type="checkbox"/>	2ig9A	70	288	27	0.88	1.02	7.8e-05	0.92
<input type="checkbox"/>	1f1uA	2	284	22	1.19	0.18	0	1
<input type="checkbox"/>	1f1xA	2	284	21	1.19	0.11	0	1
<input type="checkbox"/>	2ig9A	2	284	21	1.19	0.12	0	1
<input type="checkbox"/>	2ig9A	70	288	27				

Moulder Parameters (optional)

Iterations (20)

Alignments per Iteration (250)

Models per Iteration (250)

Intermediate Models per Alignment (7)

Final Models per Alignment (11)

Scoring Method

Create Ensemble

MODEL Comment PDB Code Target Beg Target End Seqldent MPQS zDope Evaluate GA341 Weight

Ensemble Description:

Exclude Loop Residues (by loop size)

<input type="checkbox"/>	five with full coverage	2ig9A	70	288	27	0.88	1.02	7.8e-05	0.92	10
<input type="checkbox"/>	four with full coverage	1f1uA	2	284	22	1.19	0.18	0	1	10
<input type="checkbox"/>	four with full coverage	1f1xA	2	284	21	1.19	0.11	0	1	10
<input type="checkbox"/>	four with full coverage	2ig9A	2	284	21	1.19	0.12	0	1	10
<input type="checkbox"/>	moulder model	2ig9A	70	288	27					10

Moulder Submit

Submit to MR Programs

Choose Dataset

Dataset ID SG Res Z Cell
 | 4 2 2 1.45 1 121.76 121.76 108.91 90 90 90

Choose Ensemble:

Ensemble: best 27%
 PDB Code Chain Begin End Code Seqldent GA341 MPQS zDope Loop Cut
 2ig9 A 70 288 27 0.92 0.88 1.02 0

Ensemble: 22%
 PDB Code Chain Begin End Code Seqldent GA341 MPQS zDope Loop Cut
 1f1u A 2 284 22 1 1.19 0.18 0

Ensemble: all four
 PDB Code Chain Begin End Code Seqldent GA341 MPQS zDope Loop Cut
 1f1x A 2 284 21 1 1.19 0.11 0
 2ig9 A 2 284 21 1 1.19 0.12 0
 2ig9 A 70 288 27 0.92 0.88 1.02 0
 1f1u A 2 284 22 1 1.19 0.18 0

Ensemble: moulder model
 PDB Code Chain Begin End Code Seqldent GA341 MPQS zDope Loop Cut
 2ig9 A 70 288 27 0

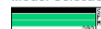
Choose MR program

phaser
 Run Description
 Stage (FRF), FTF, AUTO
 Allowed clashes
 Search for Z molecules

CMMR Server: Screen Shots

Select Models

Model Selection for 1kw3



Currently selected Models:

Database ID Model Count

Sequence	PDB Code	Begin	End	SeqIdnt	GA341	MPQS	zDope	Score	Value	Method	Description
1kw3	1f1xA	2	284	21	1	1.19	0.11			modpipe2.0	four with full coverage
1kw3	2ig9A	2	284	21	1	1.19	0.12			modpipe2.0	four with full coverage
1kw3	1f1uA	2	284	22	1	1.19	0.18			modpipe2.0	four with full coverage
1kw3	2ig9A	70	288	27		0.92	0.88	1.02		modpipe2.0	five with full coverage
1kw3	2ig9A	70	288	27		dope	0.532045	moulder		moulder model	

MODEL	PDB Code	Target Beg	Target End	SeqIdnt	MPQS	zDope	Value	GA341
-------	----------	------------	------------	---------	------	-------	-------	-------

Model Description:

MODEL	PDB Code	Target Beg	Target End	SeqIdnt	MPQS	zDope	Value	GA341
<input type="checkbox"/>	1kw3B	1	288	100.00	2.17	-1.46	0	1.00
<input type="checkbox"/>	1kw3B	1	288	100.00	2.19	-1.46	0	1.00
<input type="checkbox"/>	1kw3B	1	288	100.00	2.22	-1.46	0	1.00
<input type="checkbox"/>	1kw3B	1	288	100.00	2.22	-1.46	0	1.00
<input type="checkbox"/>	1kw3B	1	288	100.00	2.18	-1.46	0	1.00
<input type="checkbox"/>	1kw3B	1	288	98.00		6e-46		1.00
<input type="checkbox"/>	1igtA	1	288	67.00	1.86	-1.18	0	
<input type="checkbox"/>	1igtA	1	288	67.00	1.87	-1.18	0	
<input type="checkbox"/>	1igtA	1	288	67.00	1.81	-1.22	0	
<input type="checkbox"/>	1igtA	1	288	67.00	1.84	-1.18	0	
<input type="checkbox"/>	1igtA	1	288	65.00	1.79	-1.03	0	
<input checked="" type="checkbox"/>	2ig9A	70	288	27.00	0.88	1.02	7.8e-05	
<input type="checkbox"/>	1nkiA	2	118	26.00	0.62	0.42	0	
<input type="checkbox"/>	1nkiA	2	116	25.00	0.58	0.31	4.7e-12	

Submit Moulder Calculation for 1kw3

modbase

MODEL	PDB Code	Target Beg	Target End	SeqIdnt	MPQS	zDope	Value	GA341
Submit Moulder	2ig9A	70	288	27	0.88	1.02	7.8e-05	0.92
Submit Moulder	1f1uA	2	284	22	1.19	0.18	0	1
Submit Moulder	1f1xA	2	284	21	1.19	0.11	0	1
Submit Moulder	2ig9A	2	284	21	1.19	0.12	0	1
Submit Moulder	2ig9A	70	288	27				

Moulder Parameters (optional)
 Iterations (20)
 Alignments per Iteration (250)
 Models per Iteration (250)
 Intermediate Models per Alignment (7)
 Final Models per Alignment (11)
 Scoring Method

Create Ensemble

MODEL Comment PDB Code Target Beg Target End SeqIdnt MPQS zDope Evaluate GA341 Weight

Ensemble Description:

Exclude Loop Residues (by loop size)

<input type="checkbox"/>	five with full coverage	2ig9A	70	288	27	0.88	1.02	7.8e-05	0.92	10
<input type="checkbox"/>	four with full coverage	1f1uA	2	284	22	1.19	0.18	0	1	10
<input type="checkbox"/>	four with full coverage	1f1xA	2	284	21	1.19	0.11	0	1	10
<input type="checkbox"/>	four with full coverage	2ig9A	2	284	21	1.19	0.12	0	1	10
<input type="checkbox"/>	moulder model	2ig9A	70	288	27					10

Moulder Submit

Submit to MR Programs

Choose Dataset

Dataset ID SG Res Z Cell
 | 4 2 2 1.45 1 121.76 121.76 108.91 90 90 80

Choose Ensemble:

PDB Code Chain Begin End Code SeqIdnt GA341 MPQS zDope Loop Cut
 2ig9 A 70 288 27 0.92 0.88 1.02 0

PDB Code Chain Begin End Code SeqIdnt GA341 MPQS zDope Loop Cut
 1f1u A 2 284 22 1 1.19 0.18 0

PDB Code Chain Begin End Code SeqIdnt GA341 MPQS zDope Loop Cut
 1f1x A 2 284 21 1 1.19 0.11 0
 2ig9 A 2 284 21 1 1.19 0.12 0
 2ig9 A 70 288 27 0.92 0.88 1.02 0
 1f1u A 2 284 22 1 1.19 0.18 0

PDB Code Chain Begin End Code SeqIdnt GA341 MPQS zDope Loop Cut
 2ig9 A 70 288 27 0

Choose MR program

Run Description
 Stage (IFR), FTF, AUTO
 Allowed clashes
 Search for Z molecules

Analyze Results

RUN_ID	RUN DESC	Z	MAX RFZ	MAX TFZ	SPGRP	ENSEMBLE	DATE
cd04962c2e11bd2fab33b8dd39c9bcff	27% best	1	3.60	5.10	1 4 2 2	best 27%	2008-04-06 20:25:10
6d1f07e25ebe67c14f9a23782b1db328	all four	1	3.60	5.40	1 4 2 2	all four	2008-04-06 20:26:12
da882513feb4a8c81018db4c6ff36f62	moulder model	1	4.40	6.00	1 4 2 2	moulder model	2008-04-07 13:48:09
6d08049088865c35e17e781bb4710dae	22%	1	4.10	4.10	1 4 2 2	22%	2008-04-06 20:25:17

Future Directions

- Complete the benchmarking
- Explore MR scores in guiding MOULDER
- Explore additional sampling methods
- Finish the web-server interface
- Applying the tool ourselves to NYSGXRC structures
- Distributing the tool to many others

Acknowledgments

(<http://salilab.org>)

Sali Lab

**Dave Eramian
Ben Webb
Narayanan Eswar
Andrej Sali**

NYSGXRC

**Jeff Bonanno
Steve Almo
Swaminathan
Stephen Burley**

UCSF

**David Agard
Franz Gruswitz (CSMP)**