X線結晶構造解析における構造バイオインフォマティクス

(1) 分子置換法を使ってタンパク質の立体構造を決定してみよう。

(2) Coot で分子モデルを電子密度に合わせてみよう。

東京大学 大学院農学生命科学研究科 応用生命化学専攻 食品工学研究室 永田 宏次

1. 背景と目的

PDBには 60,000 個以上のタンパク質立体構造が登録されている。この情報を利用して、すでにアミノ酸配列類似タンパク質の立体構造が報告されているタンパク質の X 線結晶構造 解析を分子置換法により行う。分子置換法を用いれば、配列相同性 30%以上の類似タンパ ク質の立体構造情報をモデル(鋳型)として、たいていの場合、目的タンパク質の立体構 造解析が可能である。分子置換法で構造が解けない場合は、単波長・多波長異常分散法、 重原子同型置換法等により構造解析を行う。

今回、X 線結晶構造解析に用いるソフトウェアパッケージ CCP4 はフリーウェアで、Unix, Linux, Mac OSX, Windows で動くので、パソコンでも構造解析が可能である。 http://www.ccp4.ac.uk/

2. 流れ

目的タンパク質の選択-human S100A13 発現系作成 発現・精製・結晶化 X線回折データ取得・処理 X線結晶構造解析(分子置換法、単波長・多波長異常分散法、重原子同型置換法) 構造精密化・確認・PDBへの登録の仕方の説明

3. 実習

以下の実習で最終的に作成した PDB ファイルを aknagata@mail.ecc.u-tokyo.ac.jp に送ってください。

 アグリバイオ講義 HP から、圧縮ファイル 100507.lzh をデスクトップにダウンロードし、 100507.lzh のアイコンをダブルクリックして解凍する。□
 http://www.iu.a.u-tokyo.ac.jp/lectures/AG04/index.html
 デスクトップ上のフォルダ 100507 には、以下の 7 個のファイルが入っている。□
 s100a13.seq
 human S100A13 のアミノ酸配列ファイル (一文字表記)
 s100a13.sca

(Denzo/HKL2000 フォーマット)

s100a13yobi.mtz	human S100A13	のX線回折データファイル(予備)
	(CCP4 フォーマ	·ット)
1XK4_A.pdb	human calgranu	lin A の原子座標ファイル
1XK4_C.pdb	human calgranu	lin B の原子座標ファイル
1XK4_C_molrep1_re	fmac2yobi.pdb	構造精密化途中の原子座標ファイル
s100a13_refmac2yob	i.mtz	構造精密化途中のX線回折データファイル

2. [この作業は時間節約のため、永田が実行するのを見るだけにしてください] Blast を使って、PDB(すなわち立体構造情報が登録されているタンパク質)から human S100A13にアミノ酸配列の類似したタンパク質を検索する。□

http://www	.expasy.org	/tools/blast/
11000 11 11 11 11	.onpubjioig	10001010101000

Choose the appropriate BLAST program	am and 🤊 database:
O blastp - query against the UniProt Knowle	dgebase (Swiss-Prot + TrEMBL)
Taxonomic groups (not available for PDB and translation)	ited EST):
select a database Complete database - To re- the database gives	strict the search to a particular taxon, it is much faster to se op-down list on the left than to specify your own taxonomic _t more accurate statistics.
or specify a Enter taxonomic OX III group	a species name, a TaxID or the latin name of a taxonomic res) to restrict your search to a particular taxon. You may er Example: Fungi, Homo sapiens.
or select a microbial proteome	
Search only Swiss-Prot (curated sequences)	Exclude fragment sequences
⊙ blastp - query against another protein dat	abase PDB 🔻 Please, supply an email addres
• tblastn - query against the six-frame translat Taxonomic groups: All	ion of a nucleotide database All EMBL + GSS (witho
or select a microbial genome	
 Your email address: If an e-mail address is provided, results will be auto tblastn searches). 	omatically mailed back (recommended for
Run BLAST OF Reset Form	

S100A13 にアミノ酸配列相同性が高く、かつ立体構造情報が PDB に登録されている タンパク質のリストが出力される。

この中から、S100A13の立体構造情報は除外する(S100A13の結晶構造は未知と仮定して講義しているため)。□

また、NMR で決定された溶液構造は、結晶構造に比べて正確さと精密さで劣るので、 分子置換法のモデルとして用いるには不向きである。ゆえに除外する。□ 結果として、

10 個目の 1XK4-C (PDB entry: 1XK4 の chain C) が最良のモデルと考えられる。 まずはこの座標をモデル(鋳型)として用いて分子置換を試みる。□ 失敗したら、次の候補 1IRJ-A をモデルとして用いる。 Send selected sequences to Clustal W (multiple alignment) 💿 実行 Select up to...

□ Include query sequence

Db AC Description

Score E-value

	-				
		pdb	2KI4-C FGF1_HUMAN Chain C, Fgf1-S100a13 Complex Structure: Key Componen	166	2e-42
		pdb	2KI4-B FGF1_HUMAN Chain B, Fgf1-S100a13 Complex Structure: Key Componen	166	2e-42
		pdb	2H2K-A S10AD_HUMAN Chain A, Crystal Structure Analysis Of Human S100a13	166	2e-42
		pdb	1YUT-A S10AD_HUMAN Chain A, Solution Structure Of Calcium-S100a13 (Minim	166	2e-42
		pdb	1YUR-A S10AD_HUMAN Chain A, Solution Structure Of Apo-S100a13 (Minimized	166	2e-42
		pdb	1YUR-B S10AD_HUMAN Chain B, Solution Structure Of Apo-S100a13 (Minimized	164	6e-42
		pdb	2CXJ-A S10AD_MOUSE Chain A, 3d Solution Structure Of S100a13 >gi 1105910	144	9e-36
		pdb	2KAX-A S10A5_HUMAN Chain A, Solution Structure And Dynamics Of S100a5 In	59	5e-10
		pdb	1NSH-A S10AB_RABIT Chain A, Solution Structure Of Rabbit Apo-S100a11 (19	57	2e-09
		pdb	1XK4-C S10A8_HUMAN Chain C, Crystal Structure Of Human Calprotectin(S100	57	2e-09
		pdb	1IRJ-A S10A9_HUMAN Chain A, Crystal Structure Of The Mrp14 Complexed Wit	57	2e-09
		pdb	10DB-A S10AC_HUMAN Chain A, The Crystal Structure Of Human S100a12 - Cop	56	4e-09
		pdb	1E8A-A SIOAC_HUMAN Chain A, The Three-Dimensional Structure Of Human SIO	56	4e-09
		pdb	2WCE-A S10AC_HUMAN Chain A, Calcium-Free (Apo) S100a12 >gi 241913116 pdb	55	9e-09
		pdb	3C1V-A S10A4_HUMAN Chain A, The 1.5 A Crystal Structure Of Ca2+-Bound S1	55	9e-09
		pdb	1M31-A S10A4_HUMAN Chain A, Three-Dimensional Solution Structure Of Apo	55	9e-09
		pdb	2JPT-A S10A1_BOVIN Chain A, Structural Changes Induced In Apo-S100a1 Pro	54	1e-08
		pdb	1K2H-A S10A1_RAT Chain A, Three-Dimensional Solution Structure Of Apo	54	2e-08
		pdb	2RGI-A S10A2_HUMAN Chain A, Crystal Structure Of Ca2+-Free S100a2 At 1.6	53	3e-08
		pdb	1QLS-A ANXA1_HUMAN Chain A, S100c (S100a11), Or Calgizzarin, In Complex W	53	3e-08
		pdb	1CFP-A S100B_BOVIN Chain A, S100b (S100beta) Nmr Data Was Collected From	50	2e-07
		pdb	15YN-A S100B_RAT Chain A, 3-D Solution Structure Of Reduced Apo-S100b	50	3e-07
		pdb	1MHO-A S100B BOVIN Chain A, The 2.0 A Structure Of Holo S100b From Bovin	50	3e-07
		pdb	1958-A S100B BOVIN Chain A, Solution Structure Of Calcium Loaded S100b C	50	3e-07
		pdb	1020-A S100P HUMAN Chain A, Three-Dimensional Solution Structure Of Apo	49	4e-07
		pdb	1J55-1 S100P HUMAN Chain A, The Crystal Structure Of Ca+-Bound Human S10	49	4e-07
		pdb	2H61-A S100B HUMAN Chain A, X-Ray Structure Of Human Ca2+-Loaded S100b >	49	4e-07
		pdb	2K70-A S100B RAT Chain A, Ca2+-S100b, Refined With Rdcs >gi 213424020	49	5e-07
		pdb	- 1KSM-A S100G BOVIN Chain A, Average Nmr Solution Structure Of Ca Ln Calb	49	5e-07
		pdb	- 1HT9-A S100G BOVIN Chain A, Domain Swapping Ef-Hands >gi 14277911 pdb 1H	49	5e-07
		pdb	2H61-C S100B HUMAN Chain C, X-Ray Structure Of Human Ca2+-Loaded S100b >	49	6e-07
		pdb	1N65-A S100G BOVIN Chain A, Family Of Nmr Solution Structures Of Ca Ce C	49	6e-07
	□	pdb	1MQ1-A CAZA1 HUMAN Chain A, Ca2+-S100b-Trtk-12 Complex >gi 27573963 pdb	49	6e-07
	-	pdb	1CB1-A S100G PIG Chain A, Three-Dimensional Solution Structure Of Ca2+	48	1e-06
		pdb	1BOC-A S100G BOVIN Chain A, The Solution Structures Of Mutant Calbindin	48	1e-06
		pdb	4ICB-A S100G BOVIN Chain A, Proline Cis-Trans Isomers In Calbindin D9k 0	47	1e-06
		pdb	1IG5-A S100G BOVIN Chain A, Bovine Calbindin D9k Binding Mc2+ >ci 142783	47	1e-06
•					1

複数のペプチド鎖を含む場合は、似ているペプチド鎖だけの情報を抽出して、別名で 保存する。例:1xk4 C.pdb。□

14 / 02 0	P 3		• · I• • • • • 0					
ATOM	1434	Ν	LYS C	4	9. 892	70. 055 167. 750	1. 00 50. 82	Ν
ATOM	1435	CA	LYS C	4	9. 141	68. 965 168. 427	1. 00 49. 76	C
ATOM	1436	С	LYS C	4	9.606	67. 576 167. 983	1. 00 46. 94	C
(途中	省略)							
ATOM	2176	CD	GLU C	92	33. 783	49. 554 166. 930	1. 00 44. 92	C
ATOM	2177	0E1	glu c	92	34. 739	48. 784 167. 199	1.00 45.60	0
ATOM	2178	0E2	GLU C	92	33. 105	50. 143 167. 813	1.00 45.73	0

アミノ酸配列のアラインメントをとると、以下の通り。配列相同性は30%弱。

 1
 10
 20
 30
 40
 50
 60
 70
 80
 90
 100
 110
 115

 S100R13
 1xK4;C
 1xK4;K
 1xK4;C
 1xK4;K
 1xK4;K

3. [お待たせしました。CCP4 を用いて、分子置換を行います。ここから皆さんに実行してもらいます]

まず、デスクトップ上の CCP4i アイコンをダブルクリックして CCP4i (CCP4Interface)を起動する。□

CCP4 Program Suite 6.1.3 CCP4In	nterface 2	0.6 running on iu-pc87 Project: 100507		Change Project Help
Refinement	_	Project Database Job List - currently no jobs	-	Directories&ProjectDir
Model Preparation				View Any File
Restraint Preparation				View Files from Job
Run Refmac5				Search/Sort Database
Run NCS Phased Refinement				Graphical View of Project
Model Completion & Analysis				Delete/Archive Files
				Kill Job
				ReRun Job
				Edit Job Data 💷
				Preferences
				System Administration
			-	Mail CCP4 Exit

4. 作業ファイルを扱うディレクトリを設定する。

右上にある Directories&ProjectDir ボタンを押すと以下のウインドウが開く。□ Add project ボタンをクリックして、追加された空行に以下のように記入する。□ Project: 100507 uses directory: C:/Users/iu/Desktop/100507/ Project for this session of CCP4Interface 2.0.6 として 100507 を選択する。□ その後、Apply&Exit ボタンを押す。□

CCP4Interface 2.0.6 Directories & Project Directory	
	Help
Enter one-word alias and full directory path for your Project directory(s).	<u>^</u>
Deleting these project definitions will not delete the actual directories.	
Project PROJECT uses directory: C:/Users/iu/Desktop Bro	wse
Project 100507 uses directory: C:/Users/iu/Desktop/100507/ Bro	wse
Edit list 🛁 Add pro	ject
Project for this session of CCP4Interface 2.0.6 100507 -	
Enter one-word alias and full directory path for other directories you use regularly.	
Alias: TEMPORARY for directory: C:/Ccp4Temp Bro	wse
Edit list 🛁 Add directory a	lias 🚽
Apply&Exit Quit	

5. X線回折データのフォーマット変換(Denzo/HKL2000 → CCP4)を行う。
 左側の作業メニューの黄色いバーをクリックすると種々のメニューが現れる。□
 Data Reduction → Import Integrated Data → Import Merged Data を選択すると
 ImportScaled のウィンドウが開く。□

CCP4 Program Suite 6.1.3 CCP4Interface :	2.0	0.6 running on iu-pc87 Project: 100507				×
					Change Project	Help
Data Reduction —	4	Project Database Job List - currently no jobs		Direct	ories&ProjectDir	
Data Processing using Mosfim	-			١	/iew Any File	
▼ Import Integrated Data				View Files fro	m Job	
Import Unmerged Data (Pointless)				Search/Sort D)atabase	
Import Unmerged Data (Combat)				Graphical Vie	w of Project	
Import Merged Data				Delete/Archiv	e Files	
Find or Match Laue Group				Kill Job		
Scale and Merge Intensities				ReRun Job.		
Utilities				Edit Job Data		
Automated Data Processing				Luit oob Data		
Check Data Quality				Preferences		
				System Admi	nistration	
						-
	-	4	•	Mail CCI	P4 Ex	it

以下のようにチェックする。□

- $\hfill\square$ Use anomalous data
- Run Ctruncate to convert intensities to structure factors
- Keep the input intensities in the output MTZ file
- Ensure unique data & add FreeR column for 0.05 fraction of data.
- $\hfill\square$ Copy FreeR from another MTZ
- \Box Extend reflections to higher resolution:

(次ペー	ジ	に続	<)
())	~	1 - 196	``	/

ImportScaled - Import Scaled Data from Denzo or d*trek		
		Help
Job title		A
Convert scaled data output from Scalepack (DENZO) — into MTZ format		
Use anomalous data		
Run Ctruncate — to convert intensities to structure factors		
Keep the input intensities in the output MTZ file		
Ensure unique data & add FreeR column for 0.05 fraction of data. 🗖 Copy FreeR from a	another I	NTZ
Extend reflections to higher resolution:		
In 100507 - s100a13.sca	Browse	View
Out 100507 - s100a13.mtz	Browse	View
Use dataset name as identifier to annend to column labels		
MT7 Project Crustal Dataset Names & Data Harvesting		
Create harvest file in project harvesting directory		
Crystal s100a13 belonging to Project 100507		
Dataset name s100a13_01		
Extra Information for MTZ File		
Space group p212121		
Cell dimensions 39.775 59.289 77.628 90.0 90.0 90.0		
Data collected at wavelength 1.0 Angstroms		
Estimated number of residues in the asymmetric unit		
Edit or Transform Input Data		
Log File Output		
Run Save or Restore	Close	

入力ファイルとして、s100a13.sca を選択する。Browse ボタンを使うと楽。□ 出力ファイル名が、勝手に指定される(拡張子が.mtz に変わっただけ)。

In 100507: s100a13.sca

Out 100507: s100a13.mtz

その他、入力が必要な項目は、Extra information for MTZ file の波長の値。有効数字 を考慮して、1.0000 (Angstrom)と入力するが、勝手に 1.0 に変換される。□

Data collected at wavelength: 1.0 Angstroms

Run ボタンを押して、フォーマット変換を実行すると、ファイル s100a13.mtz が作成 される。□

ImportScaled - Import Scaled Data from Denzo or d*trek	
	Help
Job title	A
Convert scaled data output from Scalepack (DENZO) — into MTZ format	
Use anomalous data	
Run Ctruncate — to convert intensities to structure factors	
Keep the input intensities in the output MTZ file	
Ensure unique data & add FreeR column for 0.05 fraction of data Conv FreeR from a	another MT7
Extend reflections to higher resolution:	
	Browse View
Out 100507 - s100a13.mtz	Browse View
Use dataset name — as identifier to append to column labels	
MTZ Project, Crystal, Dataset Names & Data Harvesting	
Create harvest file in project harvesting directory	
Crystal s100a13 belonging to Project 100507	
Dataset name s100a13_01	
Extra Information for MTZ File	
Space group p212121	
Cell dimensions 39.775 59.289 77.628 90.0 90.0 90.0	
Estimated number of residues in the asymptotic unit 09	
Edit or Transform Input Data	
Run Save or Restore	Close

6. 分子置換法の準備として、非対称単位中の S100A13 分子数を見積もる。

左側の作業メニューから、Molecular Replacement → Analysis → Cell Content Analysis を選択すると Matthews のウィンドウが開く。□ MTZ file として、s100a13.mtz を選択する。□ Use molecular weight: estimated from number of residues にして Number of residues: 98 と入力する。□ Run Now ボタンを押すと、下の白い枠に、非対称単位中のタンパク質分子数、 Matthews 係数、溶媒含有率、確率(2 通り)が表示される。□

Matthew	s – Cell Content Ar	nalysis					_ 🗆 ×
							Help
Job title Ca	Iculation for protein	only using 98 re:	sidues				4
Calculate M	latthews coefficient f	or protei	n only –	_			
Read cryst	al parameters from l	ATZ file:					
MTZ file	2008 - 510	0a132.mtz				Browse	View
Space grou	p P 21 21 21						
Cell a 39.77	750 b 59.2890 c	77.6280 alph	a <mark>90.0000</mark> b	eta 90.0000	gamma 90.0000		
High resolu	tion limit 1.888	•			- 1979 - 1979 - 1979 - 1979 - 1979 - 1979 - 1979 - 1979 - 1979 - 1979 - 1979 - 1979 - 1979 - 1979 - 1979 - 197		
Lise molecu	ılar weight estim	ated from numb	er of residues				
Number of	raeiduae 09						
Fationst and	testudes 90						
Cell rol	18306	3 006					-
CEIL VOI	ume. 10000	3.900					
Molecular	weight estimated	from number	of residues:	98			
Nmol/asym	Matthews Coeff	%solvent	P(1.89)	P(tot)			
1	4 15	70.39	0.01	0.05			
2	2.08	40.78	0.99	0.95			
3	1.38	11.16	0.00	0.00			-
	Reset		Run Now		Clos		
	Tieset		TUIT NOW		Cius		

この場合、非対称単位中 S100A13 が 2 分子含まれると確定した。□

7. 非対称単位中の残基数 196 を入力し、Data Reduction → Import Integrated Data → Import Merged Data を再実行する。

左側の作業メニューから Data Reduction → Import Integrated Data → Import Merged Data を選択すると ImportScaled のウィンドウが開く。□ 基本的に5と同じ設定だが、Extra information for MTZ file 中の Estimated number of residues in the asymmetric unit に 196 と入力する。□ その後、Run ボタンを押すと、すでに同じ名称の出力ファイルが存在するという警告 メッセージが出るが、Continue ボタンを押して、上書きする。□

ImportScaled - Import Scaled Data from Denzo or d*trek		×
	He	lp
Job title		A
Convert scaled data output from Scalepack (DENZO) — into MTZ format		
Use anomalous data		
Run Ctruncate — to convert intensities to structure factors		
Keep the input intensities in the output MTZ file		
Ensure unique data & add FreeR column for 0.05 fraction of data. 🗌 Copy FreeR from a	another MTZ	
Extend reflections to higher resolution:		
In 100507 - s100a13.sca	Browse View	
Out 100507 - s100a13.mtz	Browse View	
Use dataset name as identifier to append to column labels		
MTZ Project, Crystal, Dataset Names & Data Harvesting		
Create harvest file in project harvesting directory —		
Crystal s100a13 belonging to Project 100507		
Dataset name s100a13_01		
Extra Information for MTZ File		
Space group p212121		
Cell dimensions 39.775 59.289 77.628 90.0 90.0 90.0		
Data collected at wavelength 1.0 Angstroms		
Estimated number of residues in the asymmetric unit 196		
Edit or Transform Input Data		
Log File Output		-
Run Save or Restore	Close	

8. Molrepを用いて分子置換を実行する。

作業メニューから Molecular Replacement → Model Generation → Run Molrep auto MR を選択すると、Molrep のウィンドウが開く。□

以下のように設定する。□

Do: molecular replacement performing: rotation and translation function

Get structure factors from MTZ file

- $\hfill\square$ Input fixed model
- $\hfill\square$ Multi-copy search
- Use sequence

入力ファイルは以下の3つ。□

MTZ in: 100507: s100a13.mtz

Model in: 100507: 1XK4_C.pdb

Seq in: 100507: s100a13.seq

出力ファイル名は自動で設定される。

Coords out: 100507: 1XK4_C_molrep1.pdb

Run ボタンを押すと計算が始まる。□

Molrep - Molecular Replacement	- • ×
	Help
This interface is for version 9.2 of Molrep	^
Job title	
Do molecular replacement — performing rotation and translation function	
Get input structure factors from MTZ file —	
Input fixed model	
Multi-copy search	
✓ Use sequence	
MTZ in 100507 - s100a13.mtz	Browse View
Use 🗖 Intensities	
FP F_s100a13_01 - SIGFP SIGF_s100a13_01	
Model in 100507 - 1XK4_C.pdb	Browse View
Coords out 100507 - 1XK4_C_molrep1.pdb	Browse View
Experimental Data (Resolution,ANISO,DIFF,BADD,INVER,DSCALE,)	
The Model (SIM,COMPL,SURF,NMR,NCSM,DSCALEM)	
Search Parameters (NMON,NP,NPT,PST,STICK,LOCK,)	
Parameter for SEQ	
Seq in 100507 - s100a13.seq	Browse View
Infrequently Used Parameters (MODE,SAPTF,RAD,PACK,SCORE,LMIN,NOSG)	
Run 🖵 Save or Restore 🖵	Close

CCP4Interface の中央の作業ログで、molrep の行を選択した後、右側の View Files from Job ボタンをクリックし、プルダウンメニューの View Log File をクリックする と計算の過程を追うことができる。□

非対称単位中に S100A13 分子を 1 個置いたときの解。ログファイル最終行近くにある 解 Sol_の Rfac と Scor の値に注目すると、上位 2 個の値が良い。□

CCP4I fileview	ver 3_molre	ep.log									×
											Help
Summary											
S_ RF TF	theta	phi	chi	tx	ty	tz	TFcnt v	Rfac	Scor		
5_2_1 1	164.08 -	133.11	162.46	0.105	0.238	0.145	4.15	0.636	0.402		
S_1_1 2	92.63 -:	123.21	23.92	0.287	0.129	0.080	6.26	0.636	0.401		
S_11_1 3	140.96	-78.66	122.72	0.239	0.279	0.223	2.12	0.661	0.352		
S_10_3 4	43.16 -3	159.67	95.67	0.121	0.174	0.389	1.94	0.659	0.351		
S_8_5 5	35.08	65.74	114.00	0.260	0.120	0.003	1.89	0.662	0.345		
S3_3_6	14.28	15.79	171.90	0.070	0.455	0.040	2.04	0.659	0.345		
S_5_7 7	137.94 -:	173.54	171.33	0.156	0.117	0.467	2.21	0.671	0.338		
S_7_1 8	13.24	31.51	160.90	0.059	0.431	0.047	2.40	0.666	0.337		
S_9_11 9	142.26	134.89	53.80	0.171	0.325	0.361	2.21	0.658	0.337		
S_4_1 10	44.08 -	167.85	139.47	0.055	0.117	0.261	2.26	0.663	0.336		
S6_13 11	142.23	121.47	51.05	0.090	0.479	0.274	2.45	0.667	0.333		
Contrast =	4.64										
After stick c	orrection	:									
Move closer	to origin										1000
I sym operat	or :	1									_
new position	(frac):	0.105	0.238	0.145							
Sol_											
S_ Nmon RF TF	theta	phi	chi	tx	ty	tz	TFcnt	wRfac	Scor		-
S_1 2 1	164.08	-133.11	162.46	0.105	0.238	0.145	4.15	0.636	0.402		➡
Find		Show	Log Graph	ns		Show	/ Summa	ry		Quit	

最上位の解を採用し(S100A13分子を非対称単位中に1個置き)、2個目の分子を置いたときの解。RfacとScorの値に注目すると、最上位の解(さきほどの2位の解)が飛びぬけて良いので、これを採用する。□

2 個目の分子を置くと、1 個だけの時よりも、wRfac 値が下がり、Scor 値が上がる。□ このように Molrep を用いる分子置換法により、非対称単位中に S100A13 分子を 2 個 置くことができた。□

CCP4I fileviewer 3_molrep.log	
	Help
Summary	
S_RFTF theta phi chi tx ty f	tz IFont wRiac Scor
S_1_1 1 92.63 -123.21 23.92 0.789 0.129 0.	.580 18.74 0.579 0.507
S_8_1 2 35.08 65.74 114.00 0.970 0.587 0.	.950 2.04 0.635 0.403
S_13_14 3 171.12 -162.05 141.11 0.843 0.611 0.	.021 1.90 0.631 0.401
S3_2 4 14.28 15.79 171.90 0.025 0.095 0.	.637 3.32 0.632 0.401
S_12_4 5 68.87 88.34 38.24 0.218 0.877 0.	.033 1.75 0.634 0.400
S_11_10 6 140.96 -78.66 122.72 0.503 0.929 0.	.292 2.14 0.631 0.398
S_5_1 7 137.94 -173.54 171.33 0.918 0.633 0.	.974 2.68 0.638 0.398
S_10_9 8 43.16 -159.67 95.67 0.574 0.090 0.	.380 0.68 0.638 0.394
S7_14 9 13.24 31.51 160.90 0.827 0.869 0.	.940 0.90 0.637 0.387
S9_11 10 142.26 134.89 53.80 0.421 0.663 0.	.056 0.37 0.637 0.387
S6_5 11 142.23 121.47 51.05 0.030 0.599 0.	.026 0.62 0.638 0.386
S_4_2 12 44.08 -167.85 139.47 0.614 0.216 0.	.399 -0.01 0.632 0.386
S_2_1 13 164.08 -133.11 162.46 0.105 0.238 0.	.145 4.15 0.636 -3.000
Contrast = 8.62	
After stick correction (more information in molrep.doc Move closer to molecule N : 1 Dimer Rotation to molecule N (polar angles) : 74.15 -77. Distance between centres (A) : 23.9 Ortogonal axis distance and parallel (A) : 23.9 0. I_sym_operator : 3	c): .18 179.86 .0
new position(frac): 0.289 0.371 0.420	
Sol_	ta TEast (Bfag Seen
S 2 1 1 100 04 0 56 166 08 0 280 0 281 4	12 1FORT WEIGC SCOP
convert "molrep.crd" to "molrep.pdb"	J.420 10.74 0.379 0.307
Find Show Log Graphs	Show Summary Quit

9. Refmac を用いて、まず rigid body 構造精密化を行う。

作業メニューから Refinement → Run Refmac5 を選択すると、Run Refmac5 のウィ ンドウが開く。□

以下のように設定する。□

Do: rigid body refinement using: no prior phase information

 $\hfill\square$ Input fixed TLS parameters

no twin refinement

入力ファイルは以下の2つ。□

MTZ in: 100507: s100a13.mtz

PDB in: 100507: 1XK4_C_molrep1.pdb

出力ファイル名は自動で設定される。

MTZ out: 100507: s100a13_refmac1.mtz

PDB out: 100507: 1XK4_C_molrep1_refmac1.pdb

Refiment Parameters で refinement のサイクル数を 20 から 5 に減らしても良い。

Run ボタンを押すと計算が始まる。□

Run Refmac5	
	нер
Job title	
Do restrained refinement — using no prior phase information — input	
Input fixed TLS parameters	
no — twin refinement	
Run Coot:findwaters to automatically add/remove waters to refined structure	
MTZ in 100507 - s100a13.mtz	Browse View
FP F_s100a13_01 - Sigma SIGF_s100a13_01	
MTZ out 100507 - s100a13_refmac1.mtz	Browse View
PDB in 100507 - 1XK4_C_molrep1.pdb	Browse View
PDB out 100507 = 1XK4 C molrep1 refmac1.pdb	Browse View
LIB in 100507 Merge LIBINS	Browse View
	Desuga Misur
	Browse View
Include keyword file 100507	Browse View
Data Harvesting	
Refinement Parameters	
Setup Geometric Restraints	
Setup Non-Crystallographic Symmetry (NCS) Restraints	
No NCS restraints are currently defined	
Edit list 🔟 Add M	ICS restraint
Monitoring and Output Options	
Scaling	
Geometric parameters	
Run - Save or Restore -	Close

CCP4Interface の中央の作業ログで、refmac5 の行を選択した後、右側の View Files from Job ボタンをクリックし、プルダウンメニューの View Log File をクリックする と計算の過程を追うことができる。□

Job が FINISHED になった後、View Log Graph で構造精密化の過程を視覚的に追え て分かりやすい。□

10. Refmac を用いて、次に restrained 構造精密化を行う。

作業メニューから Refinement → Run Refmac5 を選択すると、Run Refmac5 のウィ ンドウが開く。□

以下のように設定する。□

Do: restrained refinement using: no prior phase information

 \Box Input fixed TLS parameters

no twin refinement

入力ファイルは以下の2つ。

MTZ in: 100507: s100a13_refmac1.mtz

PDB in: 100507: 1XK4_C_molrep1_refmac1.pdb

出力ファイル名は自動で設定される。

MTZ out: 100507: s100a13_refmac2.mtz

PDB out: 100507: 1XK4_C_molrep1_refmac2.pdb

Refiment Parameters で refinement のサイクル数を 20 から 50 に増やした方がよい。 Run ボタンを押すと計算が始まる。□

Run Refmac5							
		Help					
Job title							
Do restrained refinement — using no prior phase information — input	t						
Input fixed TLS parameters							
no 🛁 twin refinement							
Run Coot:findwaters to automatically add/remove waters to refined structure							
MTZ in 100507 - s100a13_refmac1.mtz	Browse	View					
FP F_s100a13_01 Sigma SIGF_s100a13_01							
MTZ out 100507 - s100a13_refmac2.mtz	Browse	View					
PDB in 100507 - IXK4_C_molrep1_refmac1.pdb	Browse	View					
PDB out 100507 - IXK4_C_molrep1_refmac2.pdb	Browse	View					
LIB in 100507 - Merge LIBINs	Browse	View					
Output lib 100507 - 1XK4_C_molrep1_refmac1.cif	Browse	View					
Include keyword file 100507 -	Browse	View					
Data Harvesting							
Refinement Parameters							
Do 50 cycles of maximum likelihood restrained refinement							
Use hydrogen atoms: use if present in file — and 🗆 output to coordinate file							
Resolution range from minimum 47.118 to 1.888							
▼ Use automatic weighting ▼ Use experimental sigmas to weight Xray terms							
Refine isotropic isotropic							
✓ Exclude data with freeR label FreeR_flag → with value of 0							
Use the free — set of reflections for fitting the SigmaA estimate							
Setup Geometric Restraints							
Sotun Non Cruetallographic Summatry (MCS) Destraints	-						
Run	Close						

CCP4Interfaceの中央の作業ログで、2回目の refmac5 の行を選択した後、右側の View Files from Job ボタンをクリックし、プルダウンメニューの View Log File をクリック すると計算の過程を追うことができる。□

Job が FINISHED になった後、View Log Graph で構造精密化の過程を視覚的に追え て分かりやすい。□

View Files from Job → View Log Graphs でロググラフを開き、Tables in File → Rfactor analysis, stats vs cycle を選択する。Graphs in Selected Table → <Rfactor> vs cycle でサイクル毎に R factor が低下していく様子を確認できる。Graphs in Selected Table → <Rfactor> vs cycle でサイクル毎に FOM vs cycle でサイクル毎に FOM (位相の確からしさ) が向上していく様子を確認できる。□







Refmac5 を用いた restrained refinement の結果、R factor は 32%、free R factor は 37%まで下がった。□

 11. さらに構造精密化を進めるために、Coot を用いて、視覚的に、分子モデルを電子 密度に合わせていく。□
 Coot Tutorial で Coot の使い方を一通り説明した後、Run Refmac5 の View from Job
 → Output files ..の PDB ファイルと MTZ ファイルを使って、立体構造モデルを電子 密度に合わせて行きます。

Coot (クロガモ=鳥) アイコンをダブルクリックして、Coot を起動。Close。No。 WinCoot: File → Open Coordinates...。 Select Coordinates File: AgriBio_2_molrep_dimer_refmac2.pdb \rightarrow OK_o

Desktop		Name	¥	Modified	-
		AgriBio_2_molrep_dimer.pdb		Today	1
Volume 1 (D:)		AgriBio_2_molrep_dimer_refmac1.pdb		Today	
TOSHIRAZSEM (E-1)		AgriBio_2_molrep_dimer_refmac2.pdb		Today	
GA		AgriBio_3_align.pdb		Today	
	1	AgriBio_3_molrep_dimer.pdb		Today	П
		R AgriBio_4_align.pdb		Today	
Fin 7.1		AgriBio_4_molrep_dimer.pdb		Today	ш
DesBCDesCN		AgriBio_5_align.pdb		Today	
	×.	AariBio 5 molree dimer.odb		Todav	×
A				rdinato filos	-



WinCoot: File \rightarrow Auto Open MTZ....

Select Dataset File: s100a13_refmac2.mtz $\,\rightarrow\,$ OK_{\circ}

Desktop	Name	 Modified
	AgriBio_10_nagata_s.refmac.cif	Today
Volume 1 (D:\)	AgriBio 11 nagata s.refmac.cif	Today
TOSHIBA256M (E:V)	AgriBio_12_nagata_s.refmac.cif	Today
GA GA	AgriBio_6_nagata_s.refmac.cif	Today
H:\	AgriBio_7_nagata_s.refmac.cif	Today
E LA	AgriBio_8_nagata_s.refmac.cif	Today
B Z \	AgriBio_9_nagata_s.refmac.cif	Today
DnaBCDnaCN	s100a13.mtz	2007/01/27
S100A13	S100a13_refmac1.mtz	Today
TT2238	🚽 🗐 s100a13_refmac2.mtz	Today



WinCoot: Edit \rightarrow Map Parameters....

Global map properties window: Map Radius: 20.0 Angstroem \rightarrow OK_o

Map Radius: 20.0 Angstroems Apply Increment Size 0.0500 e/A*3 Diff Map Increment 0.0050 e/A*3 Sampling Rate: 1.5000		Density	/ Settings	
Increment Size 0.0500 e/A*3 Diff Map Increment 0.0050 e/A*3 Sampling Rate: 1.5000 Dynamic Map Sampling Dynamic Map Display Size	Map Radius:	20.0	Angstroems	Apply
Diff Map Increment 0.0050 e/A*3 Sampling Rate: 1.5000 Dynamic Map Sampling Dynamic Map Display Size	Increment Siz	e 0.0500 e/.	A^3	
Sampling Rate: 1.5000 Dynamic Map Sampling Dynamic Map Display Size	Diff Map Increi	ment 0.0050	e/A^3	
☐ Dynamic Map Sampling ☐ Dynamic Map Display Size	Sampling Rate	e: 1.5000		
	□ Dynamic N □ Dynamic N	1ap Sampling 1ap Display Siz	e	
			Concol	



WinCoot: Draw \rightarrow Cell & Symmetry.... Show Symmetry?: 20.0 Angstroem \rightarrow OK. Symmetry/Master Switch: Show Symmetry Atoms? \rightarrow Yes. Symmetry Atom Display Radius: 30 A \rightarrow OK.





WinCoot: Measures \rightarrow Environment Distances...

Environment Distances:
Show Residue Environment?

■ Label Atom? \rightarrow OK_o





WinCoot: Draw \rightarrow Go To Atom....

Go To Atom...: +・Chain A \rightarrow A 5 PRO \rightarrow Apply \rightarrow Close。 右ドラッグ (左から右へ) で指定したアミノ酸残基を中心に拡大する。 右ドラッグ (右から左へ) で指定したアミノ酸残基を中心に縮小する。

🕅 Go To	Atom			- 🗆 ×
	Defin	e an Atom fo	or Centering:	
0Bio	_2_molre	ep_dimer_refi	mac2.pdb 🖨	Molecule
A	Chain			
5	Resid	lue Number		
CA	Atom	Name		
		Next Res	idue	
		Previous Re	esidue	
- Cha - A1	in A 5 PRO 6 LEU 7 THR 8 GLU 9 LEU 10 GLU 11 GLU 12 SER			
•				
	Apply		Close	



左ドラッグ(上←→下、左←→右)で指定したアミノ酸残基を中心に回転する。 スペースバーを押すと次のアミノ酸残基に移動する。 Shift + スペースバーを押すと前のアミノ酸残基に移動する。 スペースバーを何回も押して、21 PHE/A まで移動してください。 21 PHE/A は、分子モデルの側鎖と電子密度とが合っていません。これを合わせます。



WinCoot: Calculate \rightarrow Model/Fit/Refine...。 Model/Fit/Refine: Mutate & Auto Fit...。 WinCoot: CA/21 PHE/A 原子をクリック。 Resi...: PHE (F)。



分子モデルの側鎖と電子密度とが合いましたか? 合ったことを確認してください。 今の方法は簡単過ぎるので、別の方法で合わせてみましょう。 Model/Fit/Refine: Undo を 2 回クリックして、分子モデルの側鎖を元に戻します。 Model/Fit/Refine: Simple Mutate...。 WinCoot: CA/21 PHE/A 原子をクリック。 Resi...: PHE (F)。



さきほどと違って、分子モデルと電子密度とが微妙にずれています。

Model/Fit/Refine: Real Space Refine $Zone_{\circ}$

WinCoot: 21 PHE/A の任意の原子をダブルクリック。

Accept Refinement?: $Accept_{\circ}$



分子モデルと電子密度とが完全に合いました。

課題: 22 THR/A の分子モデルの側鎖と電子密度とが合っていません。これを合わせてください。



このようにして、N 末端から C 末端まで、すべてのアミノ酸残基の分子モデルと電子密度 とを合わせていきます。

その後、WinCoot: Validateの種々のメニューを使って、立体構造を修正します。

WinCoot: Validate \rightarrow Ramachandran Plot \rightarrow *****.pdb

Dynarama: Ramachandran Plot (Phi-Psi Plot)で Disallowed Region にあるアミノ酸残基
■にカーソルを合わせると、そのアミノ酸残基を表示する。87 ILE A, 88 ARG A, 8 GLU B, 88 ARG B の 4 残基。いずれもペプチド鎖末端付近のアミノ酸残基なので、修正が難しい。



WinCoot: Validate \rightarrow Geometry analysis \rightarrow *****.pdb

Geometry Graphs: 各アミノ酸残基の理想の geometry からのずれが表示されている。赤い アミノ酸残基があれば、そのバーをクリックし、その残基の分子モデルを修正する。



同様に、Peptide omega analysis、Temp. fact. variance analysis、Rotamer analysis を行 い、すべての項目について validate された分子モデルが得られたら、ファイルに保存。 WinCoot: File → Save Coordinates...

Save Coordinates Molecule Selector: Save Molecule Number to Save: 0 *****.pdb \rightarrow Select Filename...

Save Filename for Saved Coordinates: Name: デフォルトのまま(*****-coot-0.pdb) \rightarrow Save in folder: CCP4 で指定したフォルダ \rightarrow OK

修正された分子モデルを使って、Refmac5により構造精密化すると、R factor および free R の値が以前より小さくなっている(改善されている)はず。この後、小さな電子密度にリガンドや水分子を当てはめ、Refmac5で精密化し、最終構造を求める。