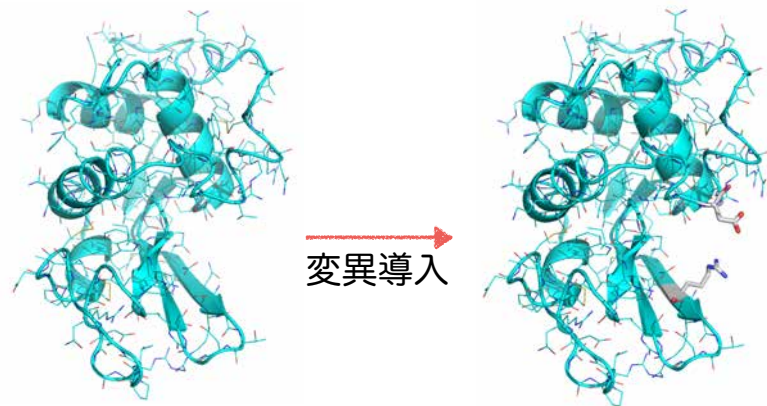# 構造生物学

13. 蛋白質工学

蛋白質工学

Protein engineering

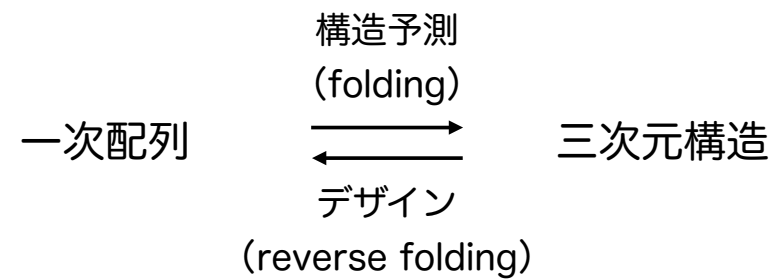

野生型　　　　変異導入　　　　新しい機能の蛋白質
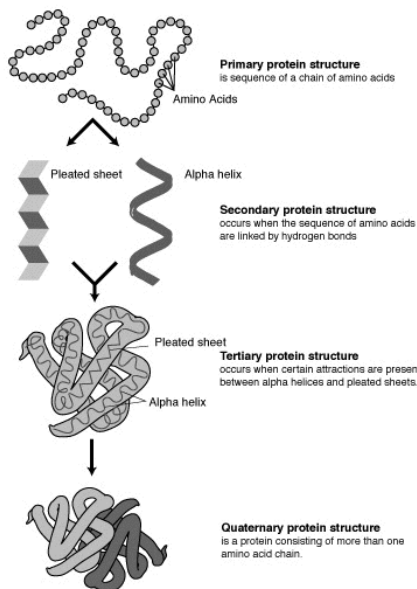

構造予測

第二の遺伝暗号解読問題

構造予測
（folding）

一次配列 　⇄　 三次元構造

デザイン
（reverse folding）

スライド5：構造予測

構造予測

二次構造予測

三次構造予測

四次構造予測

Primary protein structure is sequence of a chain of amino acids
Amino Acids

Pleated sheet / Alpha helix
Secondary protein structure occurs when the sequence of amino acids are linked by hydrogen bonds

Pleated sheet
Tertiary protein structure occurs when certain attractions are present between alpha helices and pleated sheets.
Alpha helix

Quaternary protein structure is a protein consisting of more than one amino acid chain.

http://en.wikipedia.org/wiki/Protein_structure_prediction

スライド6：ホモロジーモデリング

ホモロジーモデリング
（homology modeling）

相同な蛋白質

一次構造が似ている　── 既知の三次元構造　データベース

三次元構造（と機能）も似ている

同じコア構造　──　ループ部の変化

スライド7：相同な蛋白質の構造の類似性

相同な蛋白質の構造の類似性

$0.4 \exp[0.0187(100 - I)]$

x Calmoldulins
x Immunoglobulins
x Other

RMSD, R (Å)

Sequence identity, I (%)

53,383蛋白質
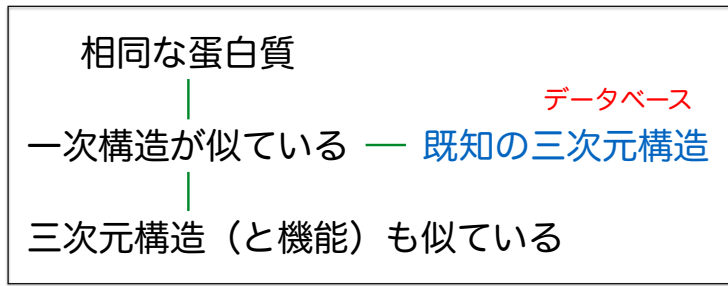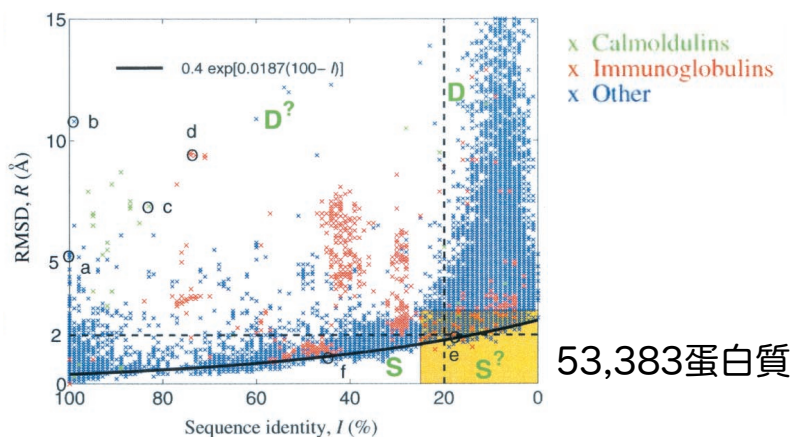
H.H. Gan et al., Biophys. J. 83, 2781-2791 (2002)　Fig. 17-1

スライド8：相同性検索をやってみよう

相同性検索をやってみよう

BLAST

http://www.genome.jp/tools/blast/

hin:HI1392

```
MIDCIYNSDSIXEIKKLDSNSIHAIISDIPYGIDYDDWDILHSNTNSALGGTSSAQHKTS
LFKRRGKPLNGWSEADKKRPQEYQEWVESWSNEWFRVLKSGSSVFVFAGRQFAHRVVVAF
ENSGFTFKDMLSWEKDKAPHRAQRISCVFERRGDIANTNKWVGWRVANLRPLFEPILWFQ
KPYKTGSTLADNLIKHEVGAWNENSLTHWNIQQGALNHSNILKVRITSEDKGYHVAQKPL
NLMKLLIDLVTKEEQIVLDPFAGSGTTLLAAKELNRHFIGYEKNNGIYNIAVNRLGIEKN
NCFYNKEKK
```

BLAST Search

Hin:HI1392

| BLAST | FASTA | KEGG2 |

Compute Clear

Enter query sequence: (in one of the three forms)

Sequence ID [hin:HI1392] (Example) mja:MJ_1041

Local file name [ファイルを選択] ファイル未選択

Sequence data [または，ここに配列をペースト]

Select program and database:

- BLASTP (prot query vs prot db)
- BLASTX (nucl query vs prot db)
- KEGG GENES
  - Favorite organism code or category
- KEGG MGENES
- KEGG VGENES
- nr-aa (GenBank, UniProt, RefSeq, PRF and PDBSTR)
  - Swiss-Prot ○ UniProt ○ RefSeq ○ PRF
  - PDBSTR

9

BLASTP Search Result   CLUSTALW

Database: nr-aa

Protein sequence database entries related to hin:HI1392 - 500 hits

Show alignment

Entry	bits	E-val
Top 10 Clear CLUSTALW Exec		
rs:WP_010869200 [WP_010869200] modification methylase [Haemophil...	639	0.0
sp:MTH3_HAEIN [P43871] RecName: Full=Modification methylase Hind...	638	0.0
rs:WP_005692451 [WP_005692451] modification methylase [Haemophil...	632	e-179
rs:WP_015940015 [WP_015940015] modification methylase [Haemophil...	531	e-148
rs:WP_016528561 [WP_016528561] Modification methylase HindIII [H...	531	e-148
rs:WP_021113702 [WP_021113702] modification methylase HindIII [H...	525	e-147
rs:WP_000861683 [WP_000861683] modification methylase [Streptoco...	455	e-125
rs:WP_003000263 [WP_003000263] modification methylase [Streptoco...	435	e-119
rs:WP_021958730 [WP_021958730] dNA methylase [Brachyspira sp. CA...	383	e-104
rs:WP_014573488 [WP_014573488] class II DNA-methyltransferase M...	382	e-104
rs:WP_020003654 [WP_020003654] hypothetical protein [Brachyspira...	381	e-103
rs:WP_022069711 [WP_022069711] dNA (Cytosine-5-)-methyltransfera...	379	e-102
rs:WP_022088592 [WP_022088592] dNA (Cytosine-5-)-methyltransfera...	363	4e-98
rs:WP_002687479 [WP_002687479] hypothetical protein [Treponema d...	361	2e-97
rs:WP_009271308 [WP_009271308] modification methylase [Erysipelo...	353	4e-95
rs:WP_022304644 [WP_022304644] dNA (Cytosine-5-)-methyltransfera...	350	3e-94
rs:WP_016243916 [WP_016243916] hypothetical protein [Escherichia...	350	5e-94
rs:WP_015059043 [WP_015059043] type II DNA- methyltransferase Ec...	342	9e-92
rs:WP_023410579 [WP_023410579] hypothetical protein [uncultured ...	334	2e-89

10

CLUSTALW Result

[clustalw.aln][clustalw.dnd][readme]

[Rooted phylogenetic tree with branch length (UPGMA)] Exec

Rooted phylogenetic tree (UPGMA)

```
CLUSTAL 2.1 Multiple Sequence Alignments

Sequence type explicitly set to Protein
Sequence format is Pearson
Sequence 1: hin_HI1392      309 aa
Sequence 2: rs_WP_010869200 309 aa
Sequence 3: sp_MTH3_HAEIN   309 aa
Sequence 4: rs_WP_005692451 310 aa
Sequence 5: rs_WP_015940015 331 aa
Sequence 6: rs_WP_016528561 332 aa
Sequence 7: rs_WP_021113702 311 aa
Sequence 8: rs_WP_000861683 303 aa
Sequence 9: rs_WP_003000263 296 aa
Sequence 10: rs_WP_021958730  308 aa
Sequence 11: rs_WP_014573488  296 aa
Start of Pairwise alignments
Aligning...


Sequences (1:2) Aligned. Score: 100
Sequences (1:3) Aligned. Score: 99.6764
Sequences (1:4) Aligned. Score: 98.7055
Sequences (1:5) Aligned. Score: 79.6117
Sequences (1:6) Aligned. Score: 79.6117
Sequences (1:7) Aligned. Score: 79.288
Sequences (1:8) Aligned. Score: 68.9769
Sequences (1:9) Aligned. Score: 66.2162
Sequences (1:10) Aligned. Score: 58.1169
Sequences (1:11) Aligned. Score: 59.7973
```
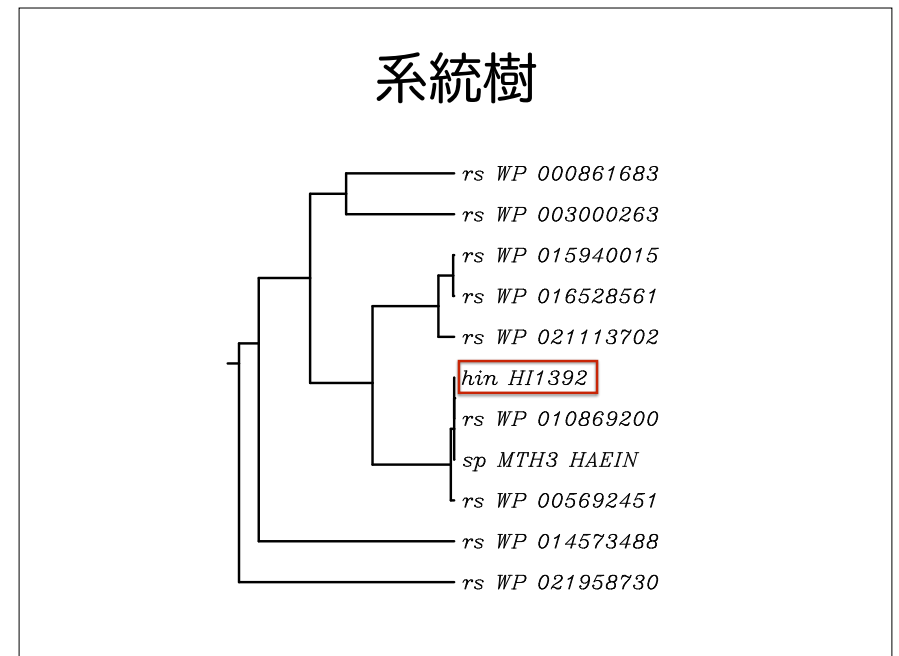
11

系統樹

rs WP 000861683
rs WP 003000263
rs WP 015940015
rs WP 016528561
rs WP 021113702
hin HI1392
rs WP 010869200
sp MTH3 HAEIN
rs WP 005692451
rs WP 014573488
rs WP 021958730

12

Slide 13

PDBデータベースでの検索

BLAST Search

| BLAST | FASTA | KEGG2 |

Enter query sequence: (in one of the three forms)

Sequence ID: hin:HI1392 (Example) mja:MJ_1041

Local file name: ファイルを選択 ファイル未選択

Sequence data:

Select program and database:

- ● BLASTP (prot query vs prot db)
- ○ BLASTX (nucl query vs prot db)
- ○ KEGG GENES
- ○ Favorite *organism code or category*
- ○ KEGG MGENES
- ○ KEGG VGENES
- ○ nr-aa (GenBank, UniProt, RefSeq, PRF and PDBSTR)
  - ○ Swiss-Prot ○ UniProt ○ RefSeq ○ PRF
  - ● PDBSTR

PDBSTR

13

Slide 14

BLASTP Search Result

Database: pdbstr

Protein sequence database entries related to hin:HI1392 - 49 hits

Show alignment

```
          Entry                                                      bits  E-val
------------------------------------------------------------------  ----- -------
Top 10 ▾  Clear   Select operation            ▾  Exec
☑ pdbstr:1G60B CRYSTAL STRUCTURE OF METHYLTRANSFERASE MBOIIA (MORA...  49  1e-05
☑ pdbstr:1G60A CRYSTAL STRUCTURE OF METHYLTRANSFERASE MBOIIA (MORA...  49  1e-05
☑ pdbstr:1EG2A CRYSTAL STRUCTURE OF RHODOBACTER SPHEROIDES (N6 ADE...  48  3e-05
☑ pdbstr:1NW7A STRUCTURE OF THE BETA CLASS N6-ADENINE DNA METHYLTR...  48  3e-05
☑ pdbstr:1NW6A STRUCTURE OF THE BETA CLASS N6-ADENINE DNA METHYLTR...  48  3e-05
☑ pdbstr:1NW5A STRUCTURE OF THE BETA CLASS N6-ADENINE DNA METHYLTR...  48  3e-05
☑ pdbstr:1NW8A STRUCTURE OF L72P MUTANT BETA CLASS N6-ADENINE DNA ...  47  5e-05
☑ pdbstr:2ZIFB CRYSTAL STRUCTURE OF TTHA0409, PUTATIVE DNA MODIFIC...  47  5e-05
☑ pdbstr:2ZIFA CRYSTAL STRUCTURE OF TTHA0409, PUTATIVE DNA MODIFIC...  47  5e-05
☑ pdbstr:2ZIGB CRYSTAL STRUCTURE OF TTHA0409, PUTATIVE DNA MODIFIC...  47  5e-05
☐ pdbstr:2ZIGA CRYSTAL STRUCTURE OF TTHA0409, PUTATIVE DNA MODIFIC...  47  5e-05
☐ pdbstr:2ZIEB CRYSTAL STRUCTURE OF TTHA0409, PUTATATIVE DNA MODIF...  46  1e-04
☐ pdbstr:2ZIEA CRYSTAL STRUCTURE OF TTHA0409, PUTATATIVE DNA MODIF...  46  1e-04
☐ pdbstr:1BOOA PVUII DNA METHYLTRANSFERASE (CYTOSINE-N4-SPECIFIC) ...  38  0.022
☐ pdbstr:1SEJE CRYSTAL STRUCTURE OF DIHYDROFOLATE REDUCTASE-THYMID...  30  6.3
☐ pdbstr:1SEJD CRYSTAL STRUCTURE OF DIHYDROFOLATE REDUCTASE-THYMID...  30  6.3
☐ pdbstr:1SEJC CRYSTAL STRUCTURE OF DIHYDROFOLATE REDUCTASE-THYMID...  30  6.3
```

14

Slide 15

```
>pdbstr:1G60B CRYSTAL STRUCTURE OF METHYLTRANSFERASE MBOIIA (MORAXELLA BOVIS) ↑ Top
            MOL_ID: 1; MOLECULE: ADENINE-SPECIFIC METHYLTRANSFERASE
            MBOIIA; CHAIN: A, B; SYNONYM: MODIFICATION METHYLASE
            MBOIIA; EC: 2.1.1.72; ENGINEERED: YES (2 of  2  ( B of
            A,B )) /RESOLUTION 1.74 ANGSTROMS.
            Length = 260

 Score = 48.9 bits (115), Expect = 1e-05,   Method: Compositional matrix adjust.
 Identities = 27/67 (40%), Positives = 40/67 (59%), Gaps = 3/67 (4%)

Query: 234 HVAQKPLNLMKLLIDLVTKEEQIVLDPFAGSGTTLLAAKELNRHFIGYEKNNGIYNIA-- 291
           H+  KP +L++ +I    +   +VLD F GSGTT + AK+L R+FIG + N    N A
Sbjct: 193 HITPKPRDLIERIIRASSNPNDLVLDCFMGSGTTAIVAKKLGRNFIGCDMNAEYVNQANF 252

Query: 292 -VNRLGI 297
            +N+L I
Sbjct: 253 VLNQLEI 259


>pdbstr:1G60A CRYSTAL STRUCTURE OF METHYLTRANSFERASE MBOIIA (MORAXELLA BOVIS) ↑ Top
            MOL_ID: 1; MOLECULE: ADENINE-SPECIFIC METHYLTRANSFERASE
            MBOIIA; CHAIN: A, B; SYNONYM: MODIFICATION METHYLASE
            MBOIIA; EC: 2.1.1.72; ENGINEERED: YES (1 of  2  ( A of
            A,B )) /RESOLUTION 1.74 ANGSTROMS.
            Length = 260

 Score = 48.9 bits (115), Expect = 1e-05,   Method: Compositional matrix adjust.
 Identities = 27/67 (40%), Positives = 40/67 (59%), Gaps = 3/67 (4%)

Query: 234 HVAQKPLNLMKLLIDLVTKEEQIVLDPFAGSGTTLLAAKELNRHFIGYEKNNGIYNIA-- 291
           H+  KP +L++ +I    +   +VLD F GSGTT + AK+L R+FIG + N    N A
Sbjct: 193 HITPKPRDLIERIIRASSNPNDLVLDCFMGSGTTAIVAKKLGRNFIGCDMNAEYVNQANF 252

Query: 292 -VNRLGI 297
            +N+L I
Sbjct: 253 VLNQLEI 259
```

ホモロジーモデリングに進めないので，ちょっと例としては良くないかも…

15

Slide 16

GenomeNet

Database: PDBSTR
Entry: 1G60B
LinkDB: 1G60B

```
MEMBER      1G60B        260    PROTEIN    1G60 00/11/02  02/05/01
DEFINITION  CRYSTAL STRUCTURE OF METHYLTRANSFERASE MBOIIA (MORAXELLA
            BOVIS)
            MOL_ID: 1;
            MOLECULE: ADENINE-SPECIFIC METHYLTRANSFERASE MBOIIA;
            CHAIN: A, B;
            SYNONYM: MODIFICATION METHYLASE MBOIIA;
            EC: 2.1.1.72;
            ENGINEERED: YES
SOURCE      MOL_ID: 1;
            ORGANISM_SCIENTIFIC: MORAXELLA BOVIS;
            ORGANISM_TAXID: 476;
            EXPRESSION_SYSTEM: ESCHERICHIA COLI;
            EXPRESSION_SYSTEM_TAXID: 562;
            EXPRESSION_SYSTEM_STRAIN: BL21(DE3)PLYSS;
            EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
            EXPRESSION_SYSTEM_PLASMID: PET24A
SEGMENT     2 of  2  ( B of A,B )
DEPOSITOR   J.OSIPIUK,M.A.WALSH,A.JOACHIMIAK,MIDWEST CENTER FOR
            STRUCTURAL GENOMICS (MCSG)
RESOLUTION  1.74 ANGSTROMS.
FEATURES         FROM      TO       DESCRIPTION
  HELIX    1    11     19    1
  HELIX    5    38     42    2
  HELIX    1    45     63    3
  HELIX    1    75     89    4
  HELIX    5   131    135    5
  HELIX    1   140    152    6
  HELIX    1   198    210    7
  HELIX    1   224    233    8
  HELIX    1   243    256    9
  HELIX    1    11     19   10
  HELIX    5    38     42   11
  HELIX    1    45     63   12
  HELIX    1    75     89   13
```
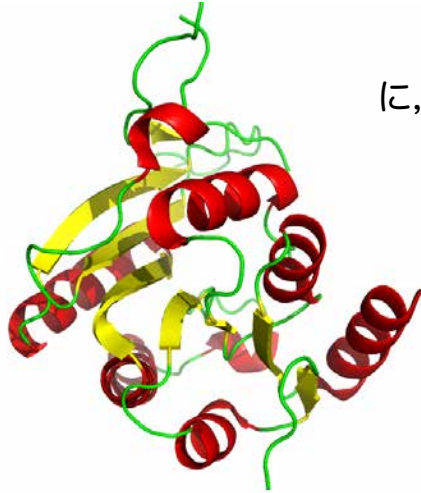
All links

Chemical reaction (1)
KEGG ENZYME (1)
3D Structure (1)
PDB (1)
All databases (2)

Download RDF

16

スライド 17

Methyltransferase MbolIa (Moraxella bovis)の構造

に，似ているだろう…

@1g60.txt

PDBID: 1G60

スライド 18

第二の遺伝暗号解読問題

立体構造予測

二次構造予測法

フォールド認識法

Ab initio 構造予測

スライド 19

三次元構造予測のための二次構造予測法

Chou-Fassman 法 (1974) 　　　　〜50%

GOR方法 (1978) 　　Garinier-Osguthorpe-Robson 　〜60%

経験的方法:
構造の分かっている蛋白質の構造情報からそれぞれのアミノ酸が
どの二次構造を好むかの指標を定義して使用.

Lim の方法（1974） 　　　　〜55%

立体化学的方法：
パターンマッチング

スライド 20

フォールド認識法

3D-1D法　Eisenberg，（1991）

アミノ酸配列(1D)と，既知の立体構造
のライブラリー(3D)を直接比較

三次元構造中の残基（の位置）を
18種類の状態に分類

3D−1Dスコア

蛋白質の改変

（蛋白質工学）

例：熱安定性の向上

熱安定化に寄与する変異

- SS結合の導入
- Gly と Pro

   Gly → Ala，Proを増やす

- ヘリックスの双極子

   N末側　⊖，C末側　⊕

- 疎水性コア（分子内空洞）の改変は困難…

T4 リゾチームの熱安定化

折畳まれた時とほどけた時のエントロピー差を削減

S-S結合の導入

ほどけた時に取れる構造の自由度を減らす
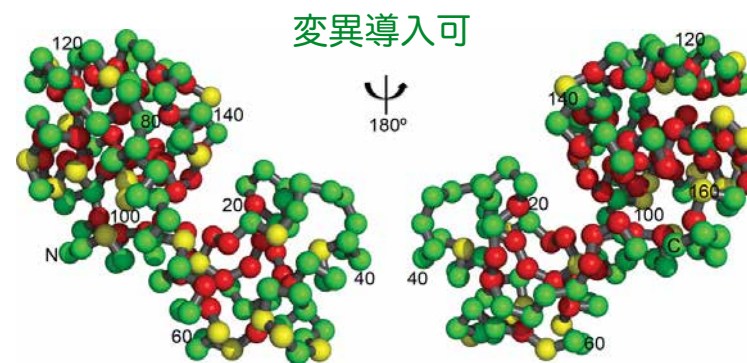
変異導入 T4 リゾチーム

変異導入可

Lessons from the lysozyme of phage T4
Baase WA et al.，Protein Sci. 2010, 19: 631–641.
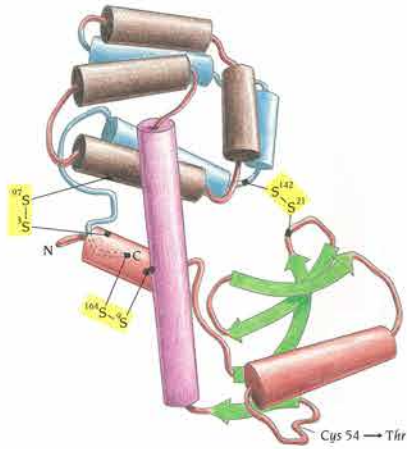
SS 変異導入 T4 リゾチーム

Fig. 17-3
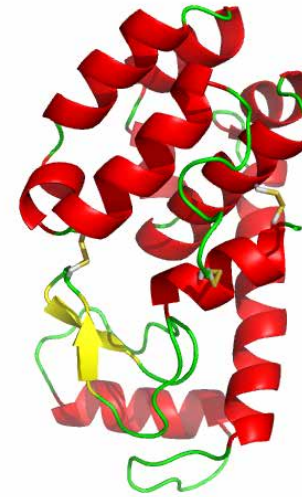
SS 変異導入 T4 リゾチーム

I3C
I9C
T21C
C54T
T142C
L164C

@152l-1.txt
PDBID: 152L

Fig. 17-3
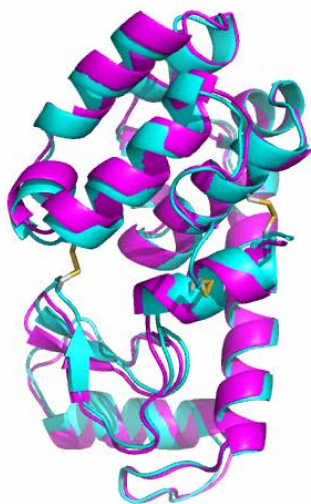
変異導入の影響

「構造」への影響が
無いことが重要

野生型
変異導入型

@152l-2.txt
PDBID: 152L

変異導入の影響

野生型
変異導入型

I3C

@152l-3.txt
PDBID: 152L
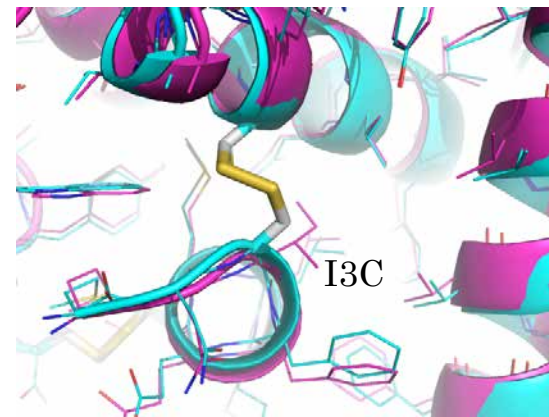
SS 変異導入 T4 リゾチームの熱安定性
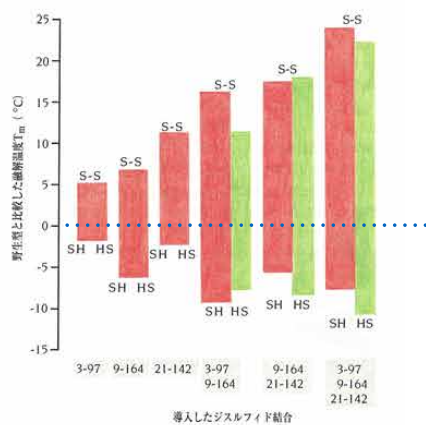
Fig. 17-4

αヘリックスの双極子安定化変異導入

Fig. 17-5

他の変異導入

ΔT=8.3℃
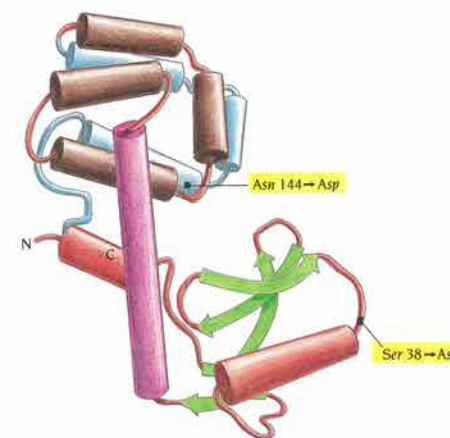
αヘリックス双極子安定化

プロリンの導入
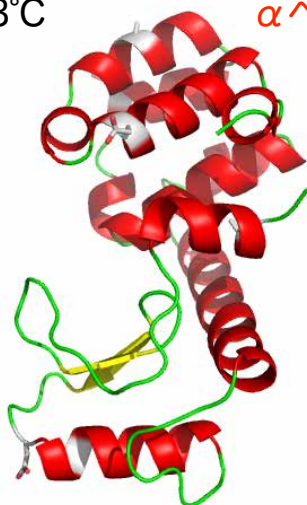
I3L
S38D
A41V
A82P
N116D
V131A
N144D

@189l-1.txt

PDBID: 189L

変異導入 T4 リゾチーム

Table I. *Thermodynamic, Crystallographic, and Other Key Information for Representative T4 Lysozyme Mutants and Complexes*

Protein	PDB ID	Activity (%)	ΔT_m (°C)	$\Delta\Delta G$ (kcal/mol)	pH	Space group	Resolution (Å)	Comment	Reference
WT (100K)	3FA0					P3$_2$21	0.98	High resolution WT	103,111
I3C/I9C/T21C/C54T/ T142C/L164C ox.	152L	0	23.4		2.0	P2$_1$2$_1$2$_1$	2.0	Triple S-S bridge	9,63
I3L/S38D/A41V/ A82P/N116D/ V131A/N144D	189L	2	8.32	3.57	5.42	P2$_1$2$_1$2$_1$	2.6	Cumulative stabilization	64
T21C/S38D/L99A/ M102E/E108V/ S117V/T142C/N114D	3GUI		−8.8	−3.2	5.3	P4$_1$2$_1$2	1.45	Buried charge in cavity	112
S38D	1L19	80	1.6	0.6	6.7	P3$_2$21	1.7	Helix dipole interaction	15,16,64, 92
S44F/WT*	137L	~WT	0.18	0.06	3.0	P2$_1$	1.85	Helix propensity analysis	49,50
S44W/WT*	216L	~WT	0.15	0.05	3.0	C2	2.1	Helix propensity analysis	28,49,50
K48-[HP]/WT*	201L	~WT	−7.0	−2.4	5.45	P2$_1$	2.0	Insertion "recovery"	52,56
A73-[AAA]/WT*	209L		−15.9	−5.0	5.4	P6$_2$22	2.7	Insertion	72
R96H (100K)	3F8V		−8.3	−3.1	5.35	P3$_2$21	1.08	Site 96 survey	103,104
L99A/WT* (200 MPa)	2B6T					P3$_2$21	2.1	Lysozyme cavity under pressure	108,109
L99A/WT* + C$_6$F$_5$I	3DN3					P3$_2$21	1.8	Halogenated benzene binding	101
L99F/M102L/V111I/ F153L/WT*	1L82	87	−1.82	−0.54	3.01	P3$_2$21	2.1	Designed core repacking	34
M102K/WT*	1L54	35	−20.3	−6.9	5.3	P3$_2$21	1.9	Buried lysine	35,77
S117A/WT*	165L		3.64	1.27	3.01	P3$_2$21	1.75	Polyalanine helix 115–123	61

Lessons from the lysozyme of phage T4

Baase WA et al., Protein Sci. 2010, 19: 631–641.

課題

下記の一次構造の蛋白質は，どういうものか．
相同検索を行って推察せよ．

```
TSNLQGAVPTNSWESSILWNQYSLPIYAHPLTFKFKAEGIEVGKPALGGS
GIAYFGAHKNDFTVGHSSVYTFPDARADKISDFAVDAVMASGSGSIKATL
MKGSPYAYFVFTGGNPRIDFSGTPTVFYGDSGSQCLGVTINGVNYGLFAP
SGSKWQGIGTGTITCILPAGKNYFSIAVLPDNTVSTLTYYKDYAYCFVTD
TKVEWSYNETESTLTTTFTAEVSVKEGTNKGTILALYPHQWRNNPHILPL
PYTYSTLRGIMKTIQGTSFKTVYRYHGILPNLPDKGTYDREALNRYINEL
ALQADAPVAVDTYWFGKHLGKLSCALPIAEQLGNISAKDRFISFMKSSLE
DWFTAKEGETAKLFYYDSNWGTLIGYPSSYGSDEELNDHHFHYGYFLHAA
AQIALRDPQWASRDNWGAMVELLIKDIANWDRNDTRFPFLRNFDPYEGHS
WASGHAGFADGNNQESSSEAINAWQAIILWGEATGNKTIRDLGIYLYTTE
VEAVCNYWFDLYKDIFSPSYGHNYASMVWGGKYCHEIWWNGTNSEKHGIN
FLPITAASLYLGKDPNYIKQNYEEMLRECGTSQPPNWKDIQYMYYALYDP
AAAKNMWNESIVPEDGESKAHTYHWICNLDSLG
```