

# 蛋白質の物理化学的性質と アミノ酸配列解析

2010年5月18日(火)

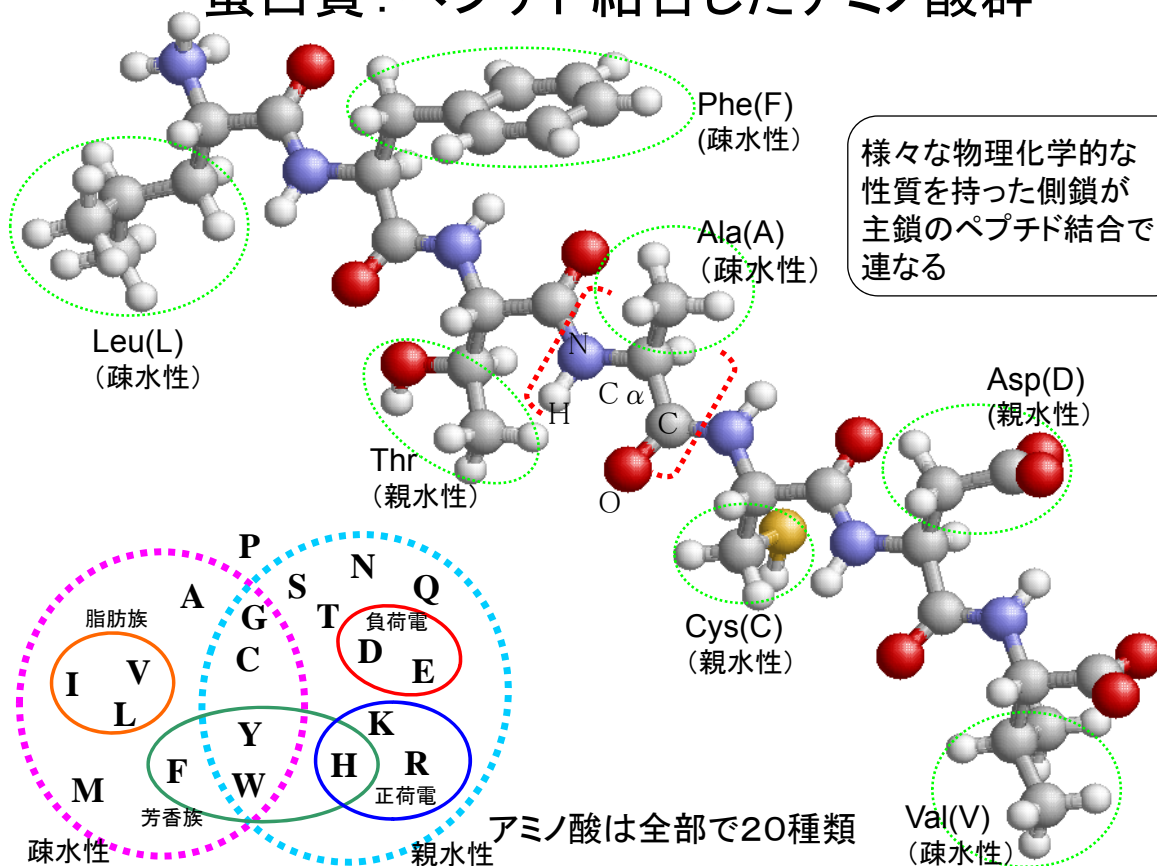
奈良先端大・情報・蛋白質機能予測学講座

川端 猛

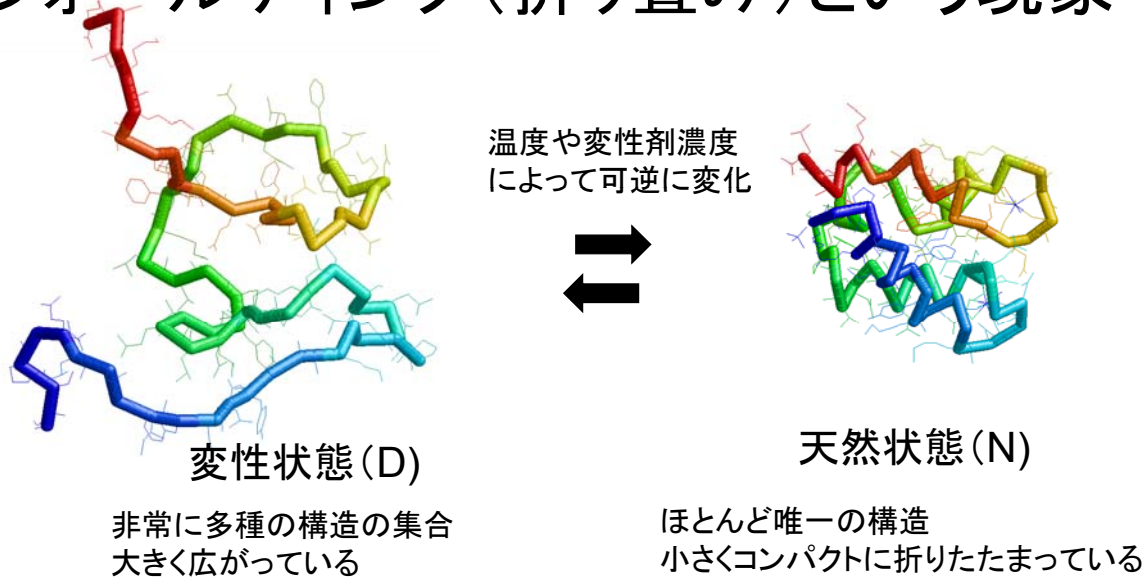
takawaba@is.naist.jp

<http://isw3.naist.jp/IS/Kawabata-lab/lec-ja.html>

## 蛋白質：ペプチド結合したアミノ酸群



# フォールディング(折り畳み)という現象

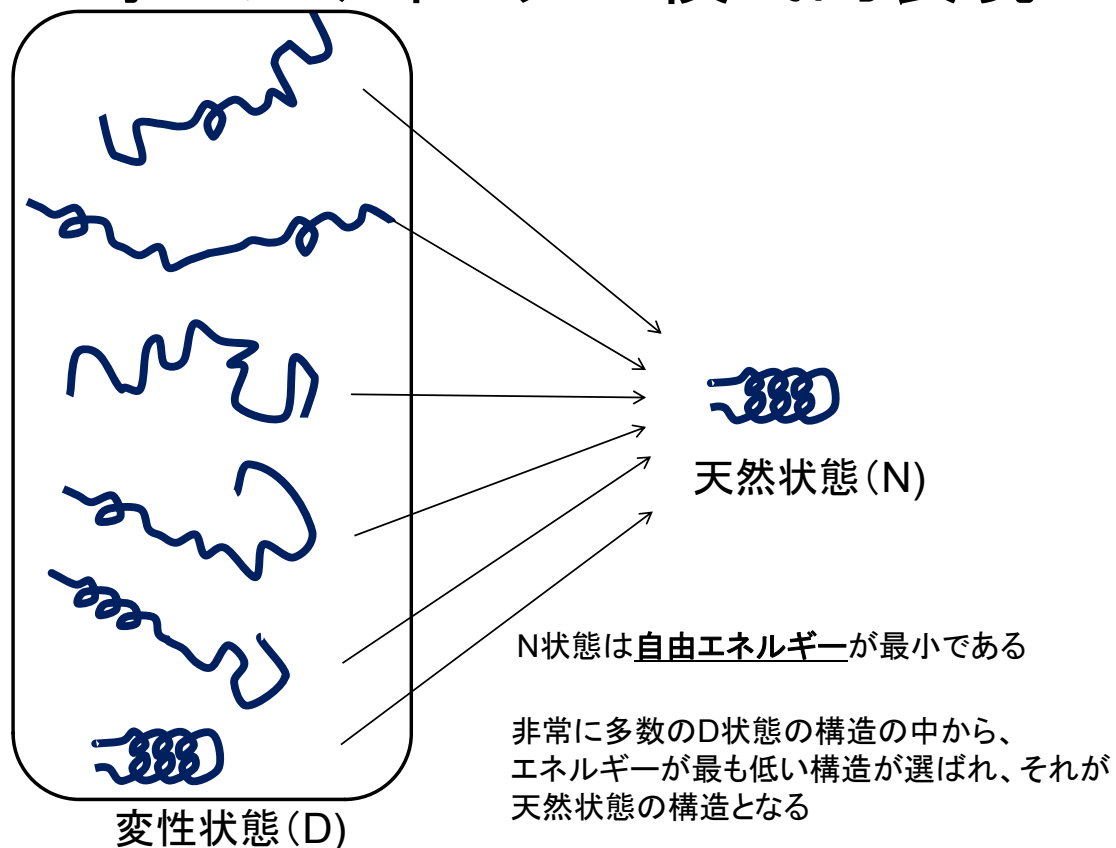


・折り畳みは、原則としてその蛋白質以外の分子の介助を必要としない

## アミノ酸配列の情報だけで、天然状態の立体構造が決定される

- ・どうやってコンパクトになるか? → (1)主鎖の水素結合(2)側鎖間の疎水性相互作用
- ・あるアミノ酸配列がどうやって一つの構造を決めるのか???

# フォールディングの模式的表現



## フォールディングをめぐる疑問

- アミノ酸配列だけの情報から立体構造が決まる  
→ 立体構造はどのようにアミノ酸配列にコードされているのだろうか？
- 折り畳みに要する時間は1ミリ秒(1/1000秒)  
→ どうやって、そんなに速く折り畳めるのか？
- 自然にあるタンパク質のN構造はほぼ一つに決まるが、アミノ酸を人間が並べた人工タンパク質では、そうならない(いくつかの構造の混合になる)ことが多い。  
→ N構造が一つに決まる配列の条件は何か？

## アミノ酸の疎水性の分布による 蛋白質の特徴づけ

# タンパク質の構造・性質による大きな分類

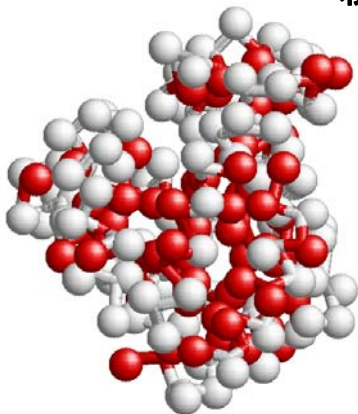
		タンパク質の例	配列上の特徴
水溶性 タンパク質 soluble protein	球状タンパク質 globular protein	多くの酵素、抗体、 転写因子など	疎水性、親水性のアミノ酸がバランスよく含まれている。
	不定形タンパク質 Intrinsically disordered protein	真核生物の転写因子の非ドメイン領域など	親水性アミノ酸が多い。特にEPQSRKが多い。
膜タンパク質 membrane protein		レセプター、トランスポーターなど	疎水性アミノ酸が多く、連続する疎水性アミノ酸領域(膜貫通ヘリックス)が観察されることが多い。

## 球状タンパク質の疎水性アミノ酸

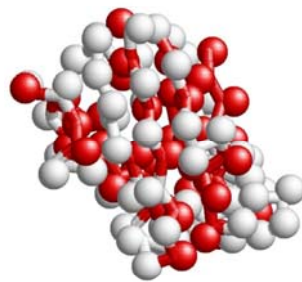
球状タンパク質では、分子内部に疎水性アミノ酸、分子表面に親水性アミノ酸が分布する傾向にある。

**疎水基** : ACILMFWV

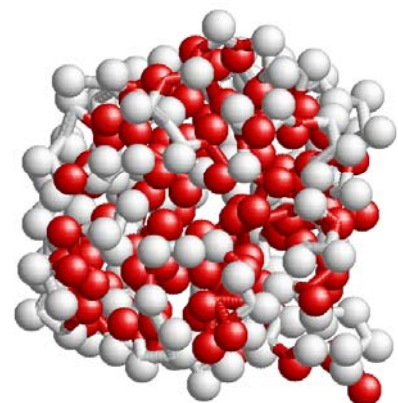
**親水基** : RNDEQGHKPSTY



1mbdA:Myoglobin



7pcyA:plastocyanin



1timA:Triose phosphate isomerase

### 疎水性相互作用(hydrophobic interaction) :

水分子と親和性の少ない非極性(non-polar、疎水性)基が水溶液中で互いに集まろうとする相互作用。電気や磁気と異なり符号はなく、疎水性どうしのものは相手を問わず集合する。

# 疎水性指標による内外予測

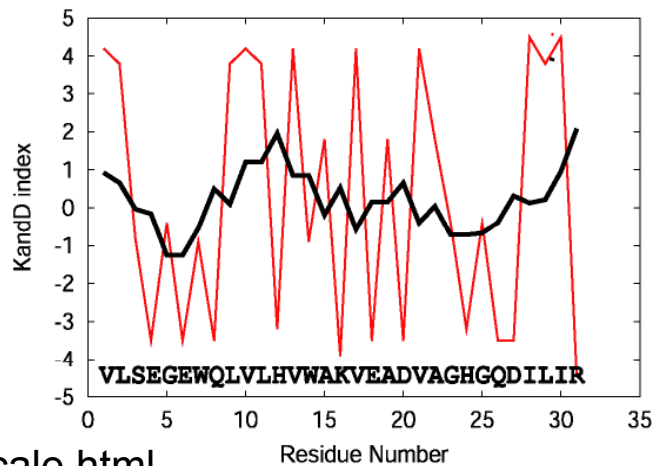
Kyte and Doolittle(1982)の疎水性指標

I	4.5	V	4.2	L	3.8	F	2.8	C	2.5
M	1.9	A	1.8	G	-0.4	T	-0.7	S	-0.8
W	-0.9	Y	-1.3	P	-1.6	H	-3.2	Q	-3.5
N	-3.5	E	-3.5	D	-3.5	K	-3.9	R	-4.5

## 方法

- ・ 配列を横軸にとって、縦軸に対応する疎水性指標をプロット。
- ・ 前後数残基でスムージング。

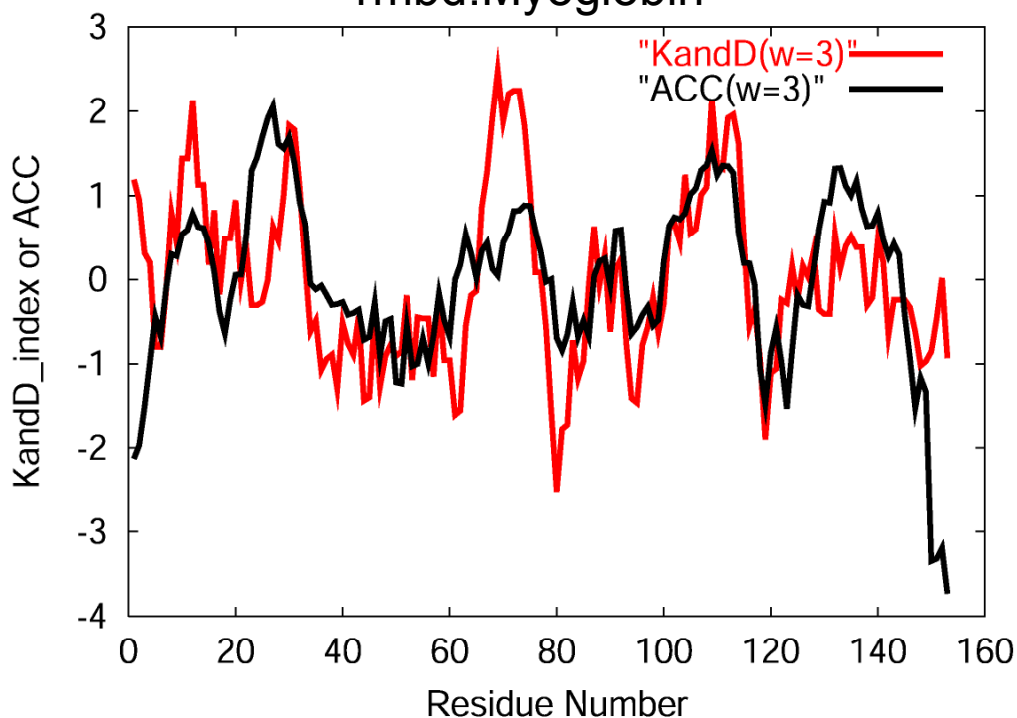
$$V[i] = \frac{1}{2w+1} \sum_{k=-w}^w KD\{Seq[i+k]\}$$



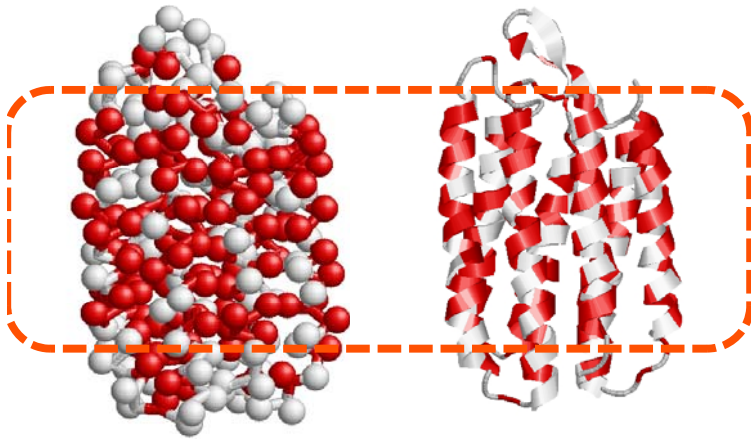
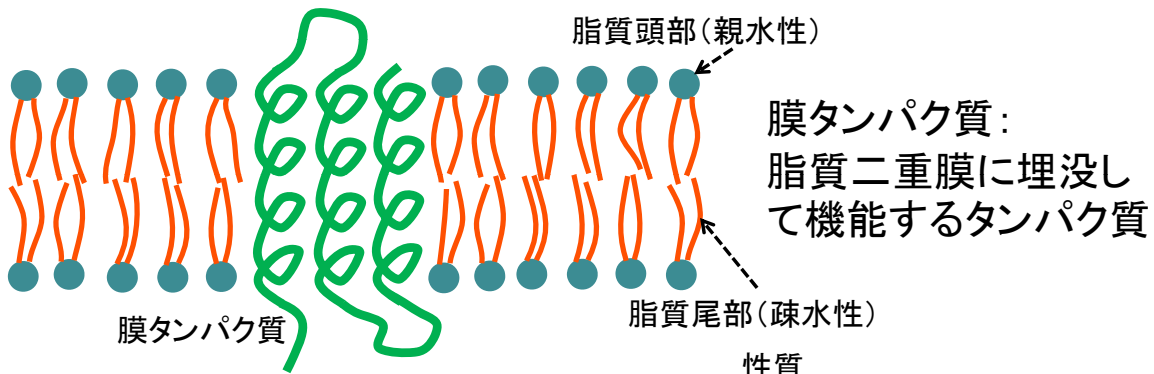
<http://kr.expasy.org/tools/protscale.html>

## 予測結果と実際の埋もれ度との比較

1mbd:Myoglobin



# 膜タンパク質



性質

(1) 疎水的な脂質尾部と相互作用するため、タンパク質表面には疎水性アミノ酸が多い。

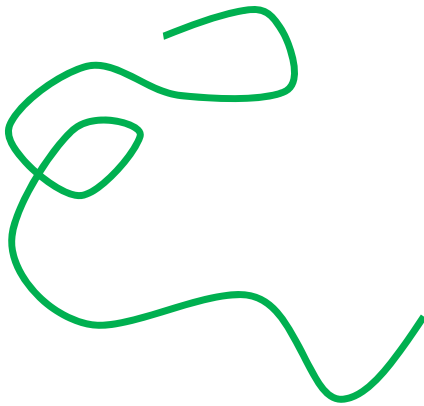
(2) 水に溶けにくく、凝集しやすい。

(3) 多くは膜貫通ヘリックスを持つ

(4) 膜貫通ヘリックスの長さは15-30アミノ酸ぐらい。アミノ酸配列からも連続した疎水性アミノ酸領域としてある程度、予測可能

# 不定形タンパク質

## intrinsically disordered protein



親水性アミノ酸が多く、疎水性アミノ酸が少ないため、天然状態でもコンパクトに折たたまらない。

親水性アミノ酸の中でも、EPQSRK

球状タンパク質ドメインをつなぐ領域としてよく見られる。

真核生物に多く、原核生物には少ない。

転写因子に特に多くみられる。

生物学的な機能はよくわかっていないが、タンパク質間相互作用に重要な働きをするといわれている。

# タンパク質の構造・性質による大きな分類

		タンパク質の例	配列上の特徴
水溶性 タンパク質 soluble protein	球状タンパク質 globular protein	多くの酵素、抗体、 転写因子など	疎水性、親水性のアミノ酸がバランスよく含まれている。
	不定形タンパク質 Intrinsically disordered protein	真核生物の転写因子の非ドメイン領域など	親水性アミノ酸が多い。特にEPQSRKが多い。
膜タンパク質 membrane protein		レセプター、トランスポーターなど	疎水性アミノ酸が多く、連続する疎水性アミノ酸領域(膜貫通ヘリックス)が観察されることが多い。

アミノ酸疎水性による配列特徴解析: 学籍番号: \_\_\_\_\_ 氏名: \_\_\_\_\_

(1)以下の配列の疎水性アミノ酸 (**ACFILMVW**) を○で囲み、そのタンパク質のタイプを判定せよ。

(2)ヘリックスの周期的疎水パターン「X〇〇X〇〇X」, 「×〇〇〇×〇〇」, 「×〇〇××〇〇」を□で囲め  
>1mbdA (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

V L S E G E W Q L V L H V W A K V E A D V A G H G Q D I L I R L F K S H P  
E T L E K F D R F K H L K T E A E M K A S E D L K K H G V T V L T A L G A  
I L K K K G H H E A E L K P L A Q S H A T K H K I P I K Y L E F I S E A I  
I H V L H S R H P G D F G A D A Q G A M N K A L E L F R K D I A A K Y K E  
L G Y Q G

>7pcyA (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

A A I V K L G G D D G S L A F V P N N I T V G A G E S I E F I N N A G F P  
H N I V F D E D A V P A G V D A D A I S A E D Y L N S K G Q T V V R K L T  
T P G T Y G V Y C D P H S G A G M K M T I T V Q

>1bm1A (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

R P E W I W L A L G T A L M G L G T L Y F L V K G M G V S D P D A K K F Y  
A I T T L V P A I A F T M Y L S M L L G Y G L T M V P F G G E Q N P I Y W  
A R Y A D W L F T T P L L L L D L A L L V D A D Q G T I L A L V G A D G I  
M I G T G L V G A L T K V Y S Y R F V W W A I S T A A M L Y I L Y V L F F  
G F T S K A E S M R P E V A S T F K V L R N V T V V L W S A Y P V V W L I  
G S E G A G I V P L N I E T L L F M V L D V S A K V G F G L I L L R S R

>ZN428\_HUMAN (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

M T E T R E P A E T G G Y A S L E E D D E D L S P G P E H S S D S E Y T L  
S E P D S E E E E D E E E E E E E T T D D P E Y D P G Y K V K Q R L G G G  
R G G P S R R A P R A A Q P P A Q P C Q L C G R S P L G E A P P G T P P C  
R L C C P A T A P Q E A P A P E G R A L G E E E E E P P R A G E G R P A G  
R E E E E E E E E G T Y H C T E C E D S F D N L G E L H G H F M L H A R  
G E V



アミノ酸疎水性による配列特徴解析: 学籍番号: \_\_\_\_\_ 氏名: \_\_\_\_\_

(1)以下の配列の疎水性アミノ酸 (ACFILMVW) を○で囲み、そのタンパク質のタイプを判定せよ。

(2)ヘリックスの周期的疎水パターン「X〇〇X〇〇X」、「×〇〇〇×〇〇」、「×〇〇××〇〇」を□で囲め  
>1mbdA (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

V L S E G E W Q L V L H V W A K V E A D V A G H G Q D I L I R L F K S H P  
E T L E K F D R F K H L K T E A E M K A S E D L K K H G V T V L T A L G A  
I L K K K G H H E A E L K P L A Q S H A T K H K I P I K Y L E F I S E A I  
I H V L H S R H P G D F G A D A Q G A M N K A L E L F R K D I A A K Y K E  
L G Y Q G

>7pcyA (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

A A I V K L G G D D G S L A F V P N N I T V G A G E S I E F I N N A G F P  
H N I V F D E D A V P A G V D A D A I S A E D Y L N S K G Q T V V R K L T  
T P G T Y G V Y C D P H S G A G M K M T I T V Q

>1bm1A (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

R P E W I W L A L G T A L M G L G T L Y F L V K G M G V S D P D A K K F Y  
A I T T L V P A I A F T M Y L S M L I G Y G L T M V P F G G E Q N P I Y W  
A R Y A D W I F T T P L L L L D L A L L V D A D Q G T I L A L V G A D G I  
M I G T G L V G A L T K V Y S Y R E V W W A I S T A A M L Y I L Y V L F F  
G F T S K A E S M R P E V A S T F K V L R N V T V V L W S A Y P V V W L I  
G S E G A G I V P L N I E T L L F M V L D V S A K V G F G L I L L R S R

>ZN428\_HUMAN (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

M T E T R E P A E T G G Y A S L E E D D E D L S P G P E H S S D S E Y T L  
S E P D S E E E E D E E E E E E E T T D D P E Y D P G Y K V K Q R L G G G  
R G G P S R R A P R A A Q P P A Q P C Q L C G R S P L G E A P P G T P P C  
R L C C P A T A P Q E A P A P E G R A L G E E E E E P P R A G E G R P A G  
R E E E E E E E E G T Y H C T E C E D S F D N L G E L H G H F M L H A R  
G E V

アミノ酸疎水性による配列特徴解析: 学籍番号: \_\_\_\_\_ 氏名: \_\_\_\_\_

(1)以下の配列の疎水性アミノ酸 (ACFILMVW) を○で囲み、そのタンパク質のタイプを判定せよ。

(2)ヘリックスの周期的疎水パターン「X〇〇X〇〇X」、「×〇〇〇×〇〇」、「×〇〇××〇〇」を□で囲め  
>1mbdA (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

V L S E G E W Q L V L H V W A K V E A D V A G H G Q D I L I R L F K S H P  
E T L E K F D R F K H L K T E A E M K A S E D L K K H G V T V L T A L G A  
I L K K K G H H E A E L K P L A Q S H A T K H K I P I K Y L E F I S E A I  
I H V L H S R H P G D F G A D A Q G A M N K A L E L F R K D I A A K Y K E  
L G Y Q G

>7pcyA (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

A A I V K L G G D D G S L A F V P N N I T V G A G E S I E F I N N A G F P  
H N I V F D E D A V P A G V D A D A I S A E D Y L N S K G Q T V V R K L T  
T P G T Y G V Y C D P H S G A G M K M T I T V Q

>1bm1A (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

R P E W I W L A L G T A L M G L G T L Y F L V K G M G V S D P D A K K F Y  
A I T T L V P A I A F T M Y L S M L I G Y G L T M V P F G G E Q N P I Y W  
A R Y A D W I F T T P L L L L D L A L L V D A D Q G T I L A L V G A D G I  
M I G T G L V G A L T K V Y S Y R E V W W A I S T A A M L Y I L Y V L F F  
G F T S K A E S M R P E V A S T F K V L R N V T V V L W S A Y P V V W L I  
G S E G A G I V P L N I E T L L F M V L D V S A K V G F G L I L L R S R

>ZN428\_HUMAN (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

M T E T R E P A E T G G Y A S L E E D D E D L S P G P E H S S D S E Y T L  
S E P D S E E E E D E E E E E E E T T D D P E Y D P G Y K V K Q R L G G G  
R G G P S R R A P R A A Q P P A Q P C Q L C G R S P L G E A P P G T P P C  
R L C C P A T A P Q E A P A P E G R A L G E E E E E P P R A G E G R P A G  
R E E E E E E E E G T Y H C T E C E D S F D N L G E L H G H F M L H A R  
G E V



アミノ酸疎水性による配列特徴解析: 学籍番号: \_\_\_\_\_ 氏名: \_\_\_\_\_

(1)以下の配列の疎水性アミノ酸 (ACFILMVW) を○で囲み、そのタンパク質のタイプを判定せよ。  
 (2)ヘリックスの周期的疎水パターン「X○○X○○X」、「×○○○×○○」、「×○○××○○」を□で囲め  
 >1mbdA (球状タンパク質・膜タンパク質・不定形タンパク質) (どれか一つを選択)

V L S E G E W Q L V L H V W A K V E A D V A G H G Q D I L I R L F K S H P  
 E T L E K F D R F K H L K T E A E M K A S E D L K K H G V T V L T A L G A  
 I L K K K G H H E A E L K P L A Q S H A T K H K I P I K Y L E F I S E A I  
 I H V L H S R H P G D F G A D A Q G A M N K A L E L F R K D I A A K Y K E  
 L G Y Q G

>7pcyA (球状タンパク質・膜タンパク質・不定形タンパク質) (どれか一つを選択)

A A I V K L G G D D G S L A F V P N N I T V G A G E S I E F I N N A G F P  
 H N I V F D E D A V P A G V D A D A I S A E D Y L N S K G Q T V V R K L T  
 T P G T Y G V Y C D P H S G A G M K M T I T V Q

>1bm1A (球状タンパク質・膜タンパク質・不定形タンパク質) (どれか一つを選択)

R P E W I W L A L G T A L M G L G T L Y F L V K G M G V S D P D A K K F Y  
 A I T T L V P A I A F T M Y L S M I L G Y G L T M V P F G G E Q N P I Y W  
 A R Y A D W I F E T T P L L L L D L A L L V D A D Q G T I L A L V G A D G I  
 M I G T G L V G A L T K V Y S Y R E V W W A I S T A A M L Y I L Y V L F E  
 G F T S K A E S M R P E V A S T F K V L R N V T V V L W S A Y P V V W L I  
 G S E G A G I V P L N I E T L L F M V L D V S A K V G F G L I L L R S R

>ZN428\_HUMAN (球状タンパク質・膜タンパク質・不定形タンパク質) (どれか一つを選択)

M T E T R E P A E T G G Y A S L E E D D E D L S P G P E H S S D S E Y T L  
 S E P D S E E E E D E E E E E E T T D D P E Y D P G Y K V K Q R L G G G  
 R G G P S R R A P R A A Q P P A Q P C Q L C G R S P L G E A P P G T P P C  
 R L C C P A T A P Q E A P A P E G R A L G E E E E E P P R A G E G R P A G  
 R E E E E E E E E G T Y H C T E C E D S F D N L G E L H G H F M L H A R  
 G E V

# アミノ酸配列から膜貫通ヘリックスを予測するサーバ

SOSUIサーバ

<http://bp.nuap.nagoya-u.ac.jp/sosui/>

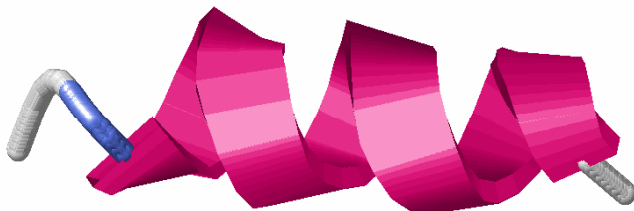
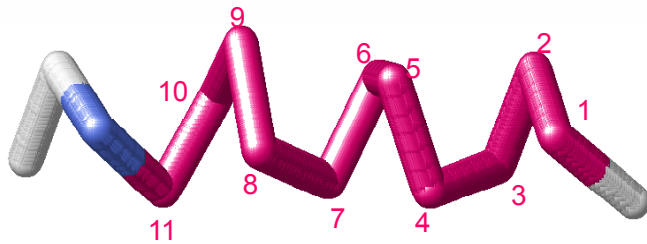
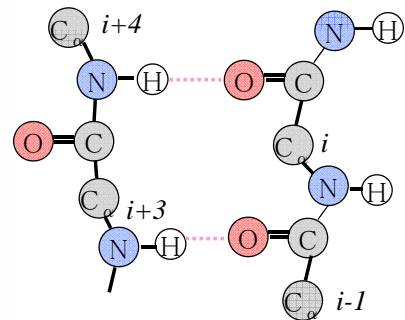
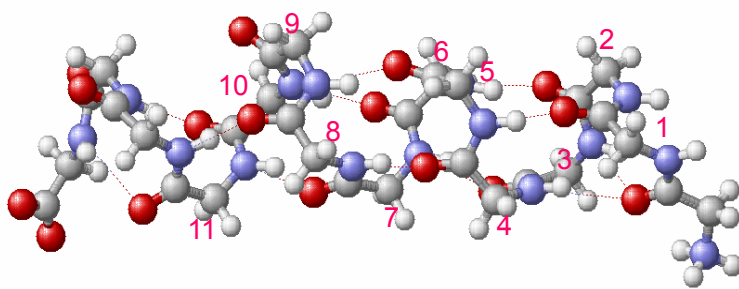
The screenshot shows the SOSUI web server interface. The main window displays the title "Classification and Secondary Structure Prediction of Membrane Proteins". Below the title, there is a "Contents" section with links to "SOSUI system", "SOSUI (SOSUI eng)", "SOSUI (Batch) (SOS)", and "SOSUI (signal)".

The "SOSUI: Submit a protein sequence" form is visible, with fields for "Enter a title or comment for the sequence" and "Enter your sequence with one-letter symbol (by copy & paste)". The sequence entered is: RREVTILALGTALMGLTYFLVWKGVSDDAKVFAITLLVPAIAFTMYLMLLQGLTNPFGSEOPTYWARVAADLFITPILLLOLALLVDADQGITLALVGDGIMIGLVLGALTKVYSYRFFVWAISTAAMLYILVLFPGFTSKAESMRPEVASIFKVLNRVTVLWSAYPVVLLIGSEGAGIVPLNIEITLFWLWVSAKVGFLILLRSR.

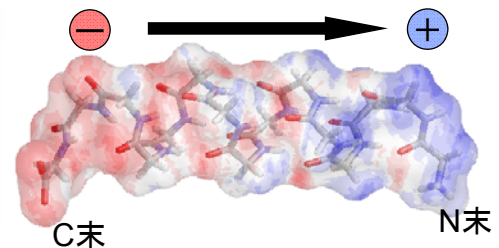
The "SOSUI: Result" window shows the "Hydropathy profile" graph, which is a line plot with peaks indicating transmembrane helices. The x-axis is labeled with residue numbers 60, 100, and 150. Below the graph, a secondary structure diagram shows the predicted primary and secondary helices, with the N-terminus (NH2) and C-terminus (COOH) labeled.

# 蛋白質の二次構造 ( $\alpha$ ヘリックス、 $\beta$ シート)

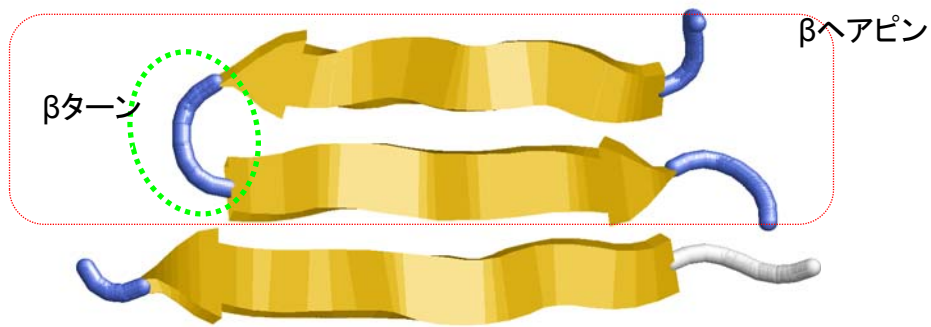
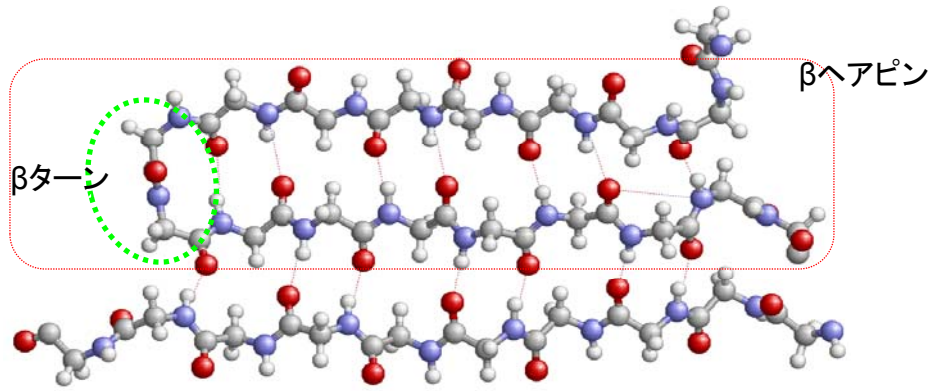
## 主鎖の水素結合： $\alpha$ ヘリックス



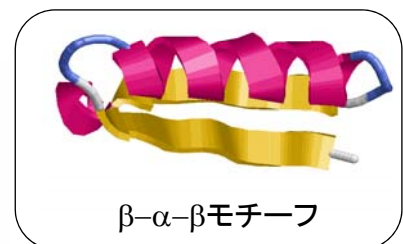
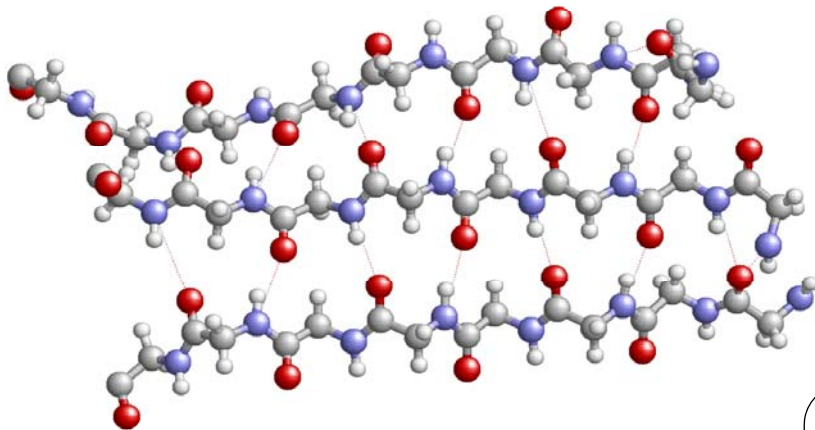
- $i$ 番目のCOが $i+4$ 番目のNHと水素結合を形成
- 3.6残基が1周期の右巻きらせん
- $\text{NH}\cdots\text{O}=\text{C}$ が同じ向きに並ぶため、ヘリックス全体に電気双極子が発生



# 主鎖の水素結合：逆平行βシート

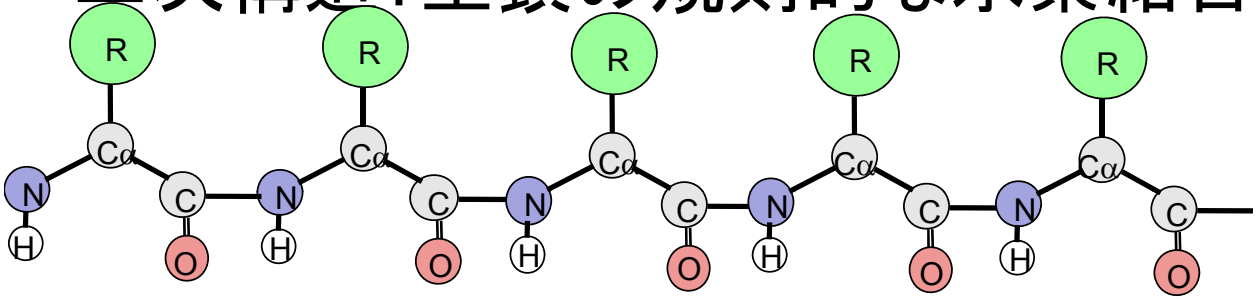


# 主鎖の水素結合：平行βシート



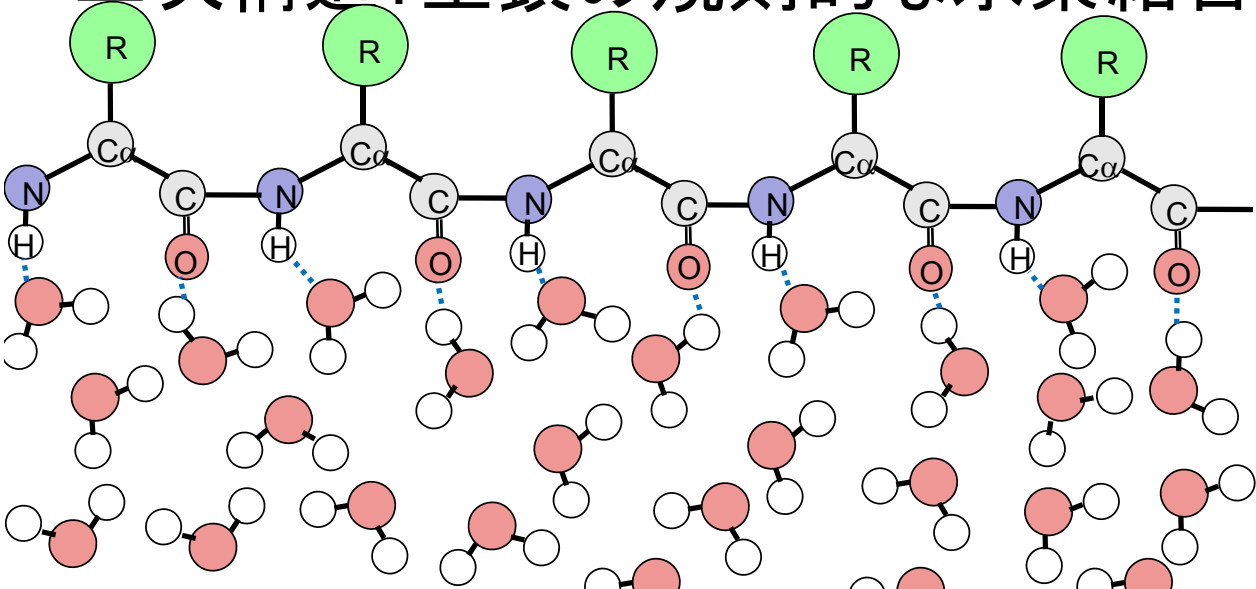
平行βシートがαヘリックスで  
接続された構造モチーフ

## 二次構造：主鎖の規則的な水素結合



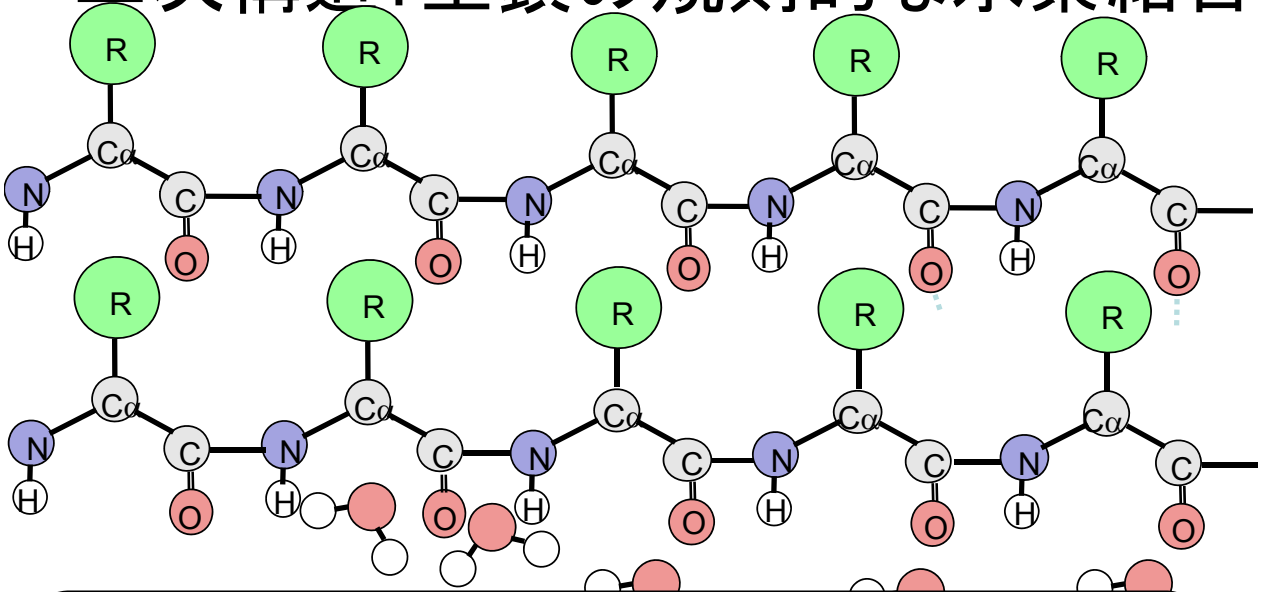
・主鎖はNH基とCO基の繰り返しであるため、極めて親水的

## 二次構造：主鎖の規則的な水素結合



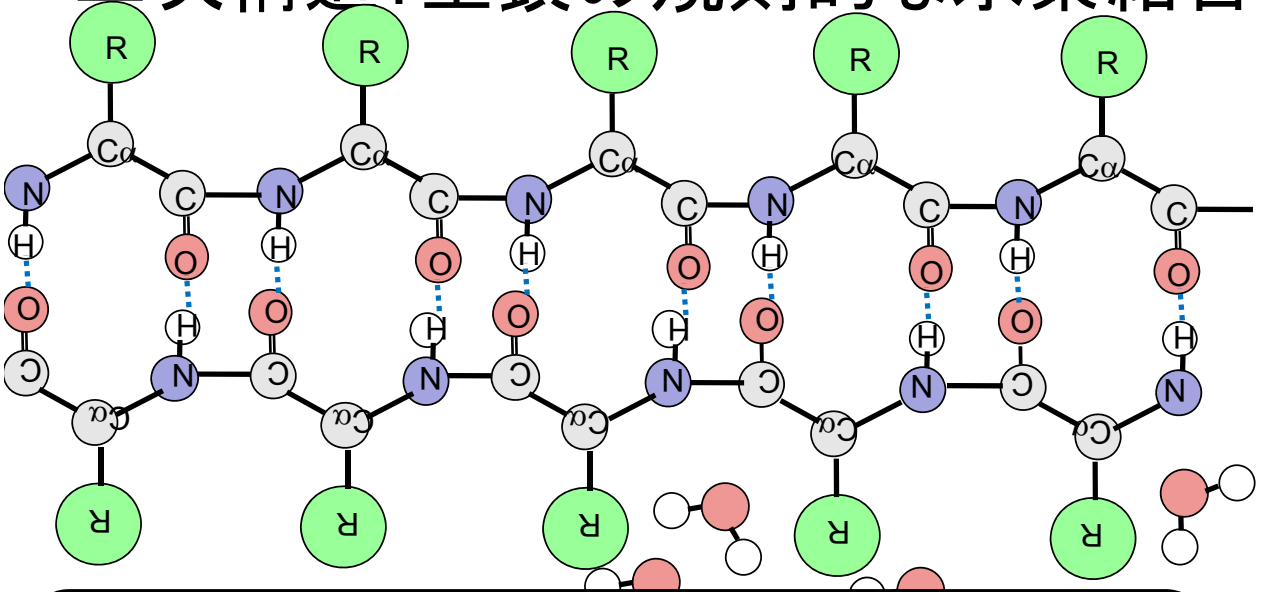
・主鎖はNH基とCO基の繰り返しであるため、極めて親水的  
→ 水分子が水素結合しやすい

## 二次構造：主鎖の規則的な水素結合



- ・主鎖はNH基とCO基の繰り返しであるため、極めて親水的  
→ 水分子が水素結合しやすい
- ・折りたたまれ、水が排除されると水素結合は切断される  
→ エネルギー的に不利

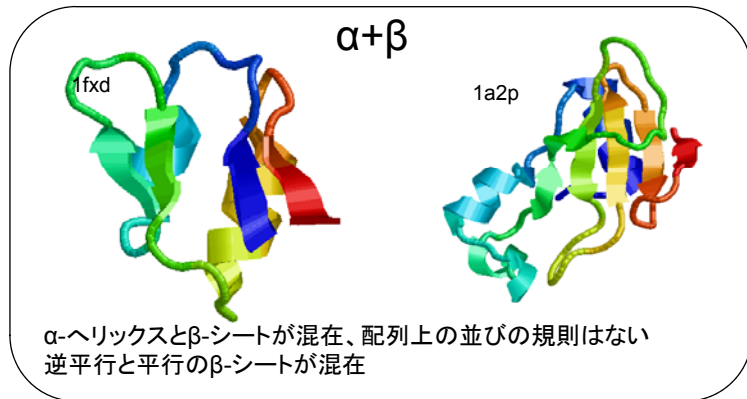
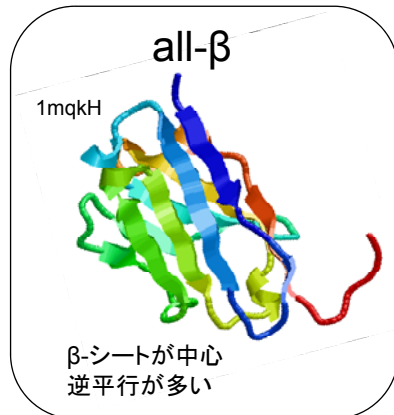
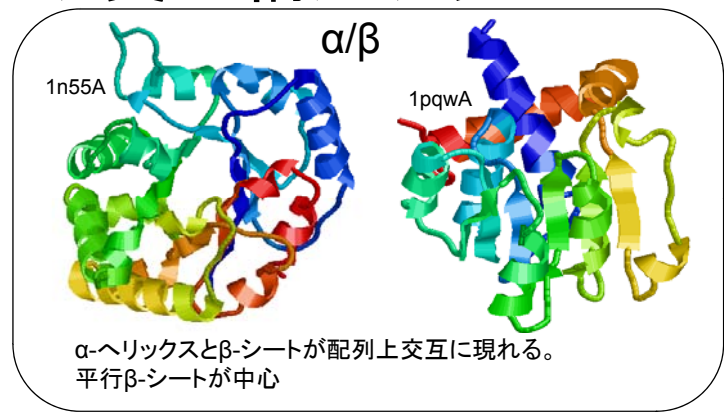
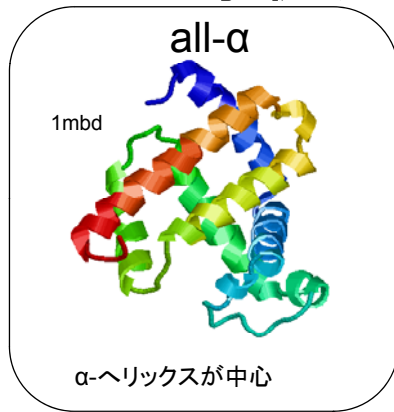
## 二次構造：主鎖の規則的な水素結合



- ・主鎖はNH基とCO基の繰り返しであるため、極めて親水的  
→ 水分子が水素結合しやすい
- ・折りたたまれ、水が排除されると水素結合は切断される  
→ エネルギー的に不利



# 球状タンパク質の構造クラス



## アミノ酸ごとに2次構造のなりやすさに差がある

Chou-Fasmanのパラメータ  $P_{Glu}(\alpha) = \frac{f(Glu \cap \alpha)}{f(\alpha)}$

(scop1.59 30%list:3077chainsを用いて再計算した値)

$\alpha$ ヘリックス	A	1.42	E	1.37	L	1.33	Q	1.30	M	1.27
	R	1.21	K	1.16	W	1.07	I	1.06	F	0.97
	Y	0.96	V	0.92	H	0.87	D	0.82	C	0.80
	T	0.77	S	0.77	N	0.74	G	0.43	P	0.42

$\beta$ シート	V	1.95	I	1.77	F	1.46	Y	1.45	C	1.30
	W	1.29	T	1.20	L	1.11	M	1.01	H	0.99
	R	0.90	S	0.83	K	0.79	Q	0.77	A	0.76
	E	0.70	G	0.64	N	0.61	D	0.53	P	0.42

コイル	P	1.66	G	1.55	N	1.35	D	1.33	S	1.23
	H	1.09	T	1.07	C	1.01	K	0.98	R	0.90
	Q	0.90	E	0.88	Y	0.84	W	0.83	F	0.82
	M	0.81	A	0.81	L	0.72	V	0.65	I	0.62



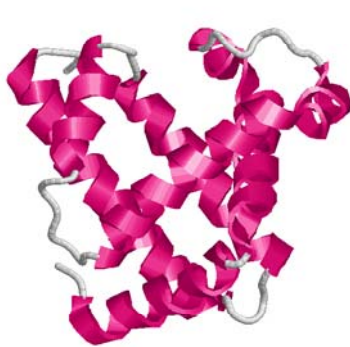
(3)  $\alpha$ ヘリックスを好むアミノ酸AEQLMを○で囲みなさい。この二つがall- $\alpha$ かall- $\beta$ のどちらかだとしたら、どちらに属すると考えられるか、○をつけて答えよ。

>1mbdA (all- $\alpha$ クラス、 all- $\beta$ クラス)(どちらか一つを選択)

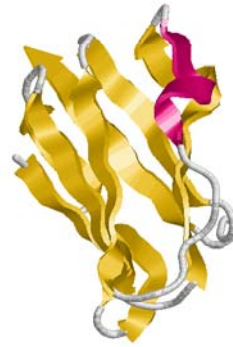
V L S E G E W Q L V L H V W A K V E A D V A G H G Q D I L I R L F K S H P  
 E T L E K F D R F K H L K T E A E M K A S E D L K K H G V T V L T A L G A  
 I L K K K G H H E A E L K P L A Q S H A T K H K I P I K Y L E F I S E A I  
 I H V L H S R H P G D F G A D A Q G A M N K A L E L F R K D I A A K Y K E  
 L G Y Q G

>7pcyA (all- $\alpha$ クラス、 all- $\beta$ クラス)(どちらか一つを選択)

A A I V K L G G D D G S L A F V P N N I T V G A G E S I E F I N N A G F P  
 H N I V F D E D A V P A G V D A D A I S A E D Y L N S K G Q T V V R K L T  
 T P G T Y G V Y C D P H S G A G M K M T I T V Q



all- $\alpha$



all- $\beta$

生命情報学 (2009.4.28)

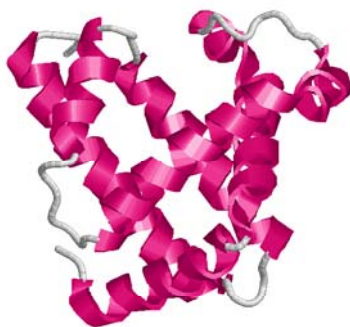
(3)  $\alpha$ ヘリックスを好むアミノ酸AEQLMを○で囲みなさい。この二つがall- $\alpha$ かall- $\beta$ のどちらかだとしたら、どちらに属すると考えられるか、○をつけて答えよ。

>1mbdA (all- $\alpha$ クラス、 all- $\beta$ クラス)(どちらか一つを選択)

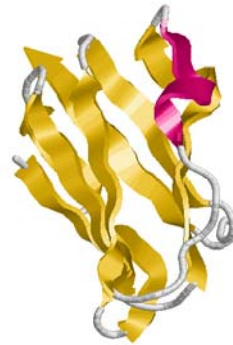
V L S (E) G (E) W (Q) L V (L) H V W (A) K V (E) A D V (A) G H G (Q) D I (L) I R (L) F K S H P  
 (E) T (L) (E) K F D R F K H (L) K T (E) A (E) M K (A) S (E) D (L) K K H G V T V (L) T (A) (L) G (A)  
 I (L) K K K G H H (E) A (E) L K P (L) A (Q) S H (A) T K H K I P I K Y (L) E F I S (E) A I  
 I H V (L) H S R H P G D F G (A) D (A) Q G (A) M N K (A) L (E) L F R K D I (A) (A) K Y K (E)  
 (L) G Y (Q) G

>7pcyA (all- $\alpha$ クラス、 all- $\beta$ クラス)(どちらか一つを選択)

(A) (A) I V K (L) G G D D G S (L) A F V P N N I T V G (A) G (E) S I (E) F I N N (A) G F P  
 H N I V F D (E) D (A) V P (A) G V D (A) D (A) I S (A) E D Y (L) N S K G (Q) T V V R K (L) T  
 T P G T Y G V Y C D P H S G (A) G (M) K (M) T I T V (Q)



all- $\alpha$

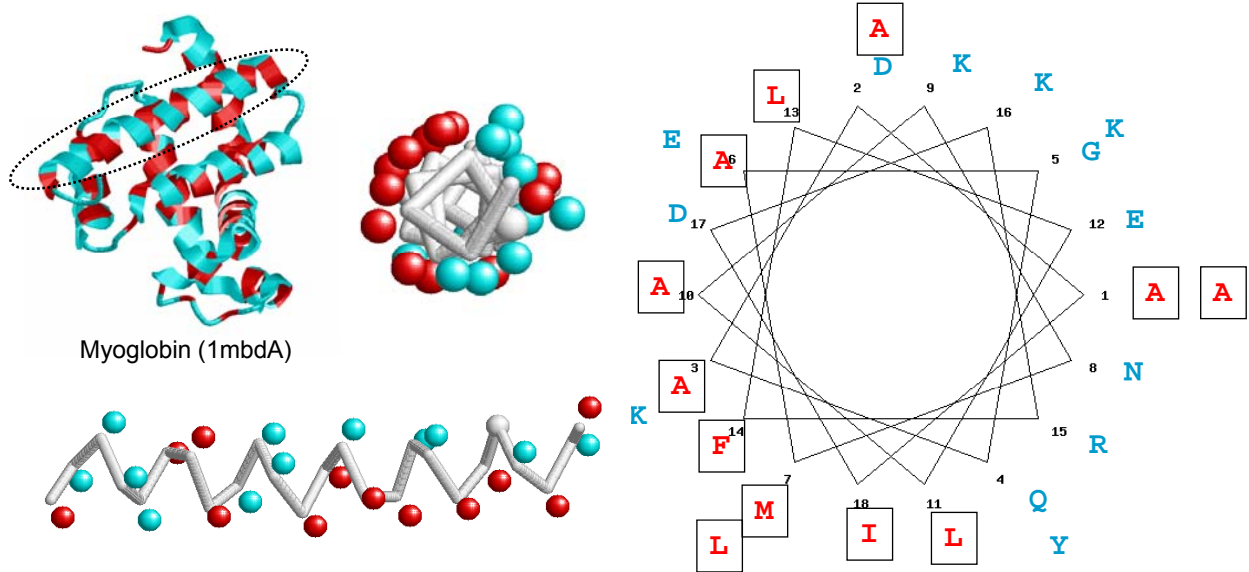


all- $\beta$

生命情報学 (2009.4.28)

# 疎水性の車輪図(Helical Wheel)

1mbdAの最後のヘリックスの配列: **ADAQGAMNKALELFRKDI**AAKY**KEL**



両親媒性ヘリックス : 片側が疎水的、反対の側が親水的になっているヘリックスのこと  
 ヘリックスの1回転の周期が3. 6アミノ酸なので、  
 3から4アミノ酸周期の疎水性 → ヘリックス構造を示唆

## 2次構造ごとに疎水性パターンに特徴がある

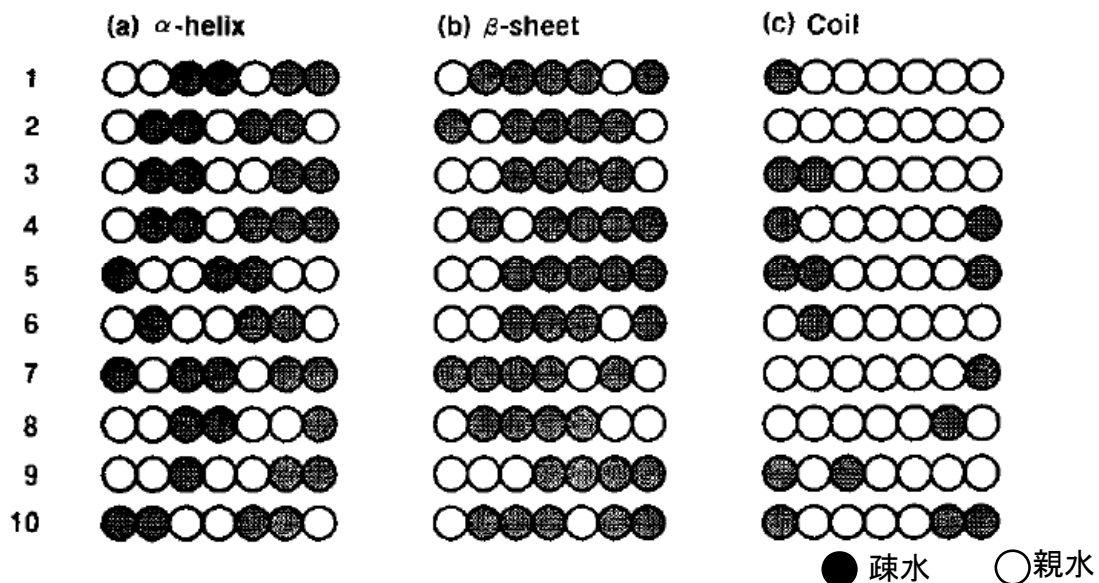


Fig. 9. Binary words that prefer the  $\alpha$  helix (a), the  $\beta$  sheet (b), and the coil (c). They are ranked by the value of the binary word information. The top 10 are shown.

アミノ酸疎水性による配列特徴解析: 学籍番号: \_\_\_\_\_ 氏名: \_\_\_\_\_

(1)以下の配列の疎水性アミノ酸 (ACFILMVW) を○で囲み、そのタンパク質のタイプを判定せよ。

(2)ヘリックスの周期的疎水パターン「XOOXOOX」、「XOOOXOO」、「XOOX×OO」を□で囲め  
>1mbdA (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

V L S E G E W Q L V L H V W A K V E A D V A G H G Q D I L I R L F K S H P  
E T L E K F D R F K H L K T E A E M K A S E D L K K H G V T V L T A L G A  
I L K K K G H H E A E L K P L A Q S H A T K H K I P I K Y L E F I S E A I  
I H V L H S R H P G D F G A D A Q G A M N K A L E L F R K D I A A K Y K E  
L G Y Q G

>7pcyA (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

A A I V K L G G D D G S L A F V P N N I T V G A G E S I E F I N N A G F P  
H N I V F D E D A V P A G V D A D A I S A E D Y L N S K G Q T V V R K L T  
T P G T Y G V Y C D P H S G A G M K M T I T V Q

>1bm1A (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

R P E W I W L A L G T A L M G L G T L Y F L V K G M G V S D P D A K K F Y  
A I T T L V P A I A F T M Y L S M I L G Y G L T M V P F G G E Q N P I Y W  
A R Y A D W I F T T P L L L L D L A L L V D A D Q G T I L A L V G A D G I  
M I G T G L V G A L T K V Y S Y R E V W W A I S T A A M L Y I L Y V L F F  
G F T S K A E S M R P E V A S T F K V L R N V T V V L W S A Y P V V W L I  
G S E G A G I V P L N I E T L L F M V L D V S A K V G F G L I L L R S R

>ZN428\_HUMAN (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

M T E T R E P A E T G G Y A S L E E D D E D L S P G P E H S S D S E Y T L  
S E P D S E E E E D E E E E E E E T T D D P E Y D P G Y K V K Q R L G G G  
R G G P S R R A P R A A Q P P A Q P C Q L C G R S P L G E A P P G T P P C  
R L C C P A T A P Q E A P A P E G R A L G E E E E E P P R A G E G R P A G  
R E E E E E E E E G T Y H C T E C E D S F D N L G E L H G H F M L H A R  
G E V

アミノ酸疎水性による配列特徴解析: 学籍番号: \_\_\_\_\_ 氏名: \_\_\_\_\_

(1)以下の配列の疎水性アミノ酸 (ACFILMVW) を○で囲み、そのタンパク質のタイプを判定せよ。

(2)ヘリックスの周期的疎水パターン「XOOXOOX」、「XOOOXOO」、「XOOX×OO」を□で囲め  
>1mbdA (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

V L S E G E W Q L V L H V W A K V E A D V A G H G Q D I L I R L F K S H P  
E T L E K F D R F K H L K T E A E M K A S E D L K K H G V T V L T A L G A  
I L K K K G H H E A E L K P L A Q S H A T K H K I P I K Y L E F I S E A I  
I H V L H S R H P G D F G A D A Q G A M N K A L E L F R K D I A A K Y K E  
L G Y Q G

>7pcyA (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

A A I V K L G G D D G S L A F V P N N I T V G A G E S I E F I N N A G F P  
H N I V F D E D A V P A G V D A D A I S A E D Y L N S K G Q T V V R K L T  
T P G T Y G V Y C D P H S G A G M K M T I T V Q

>1bm1A (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

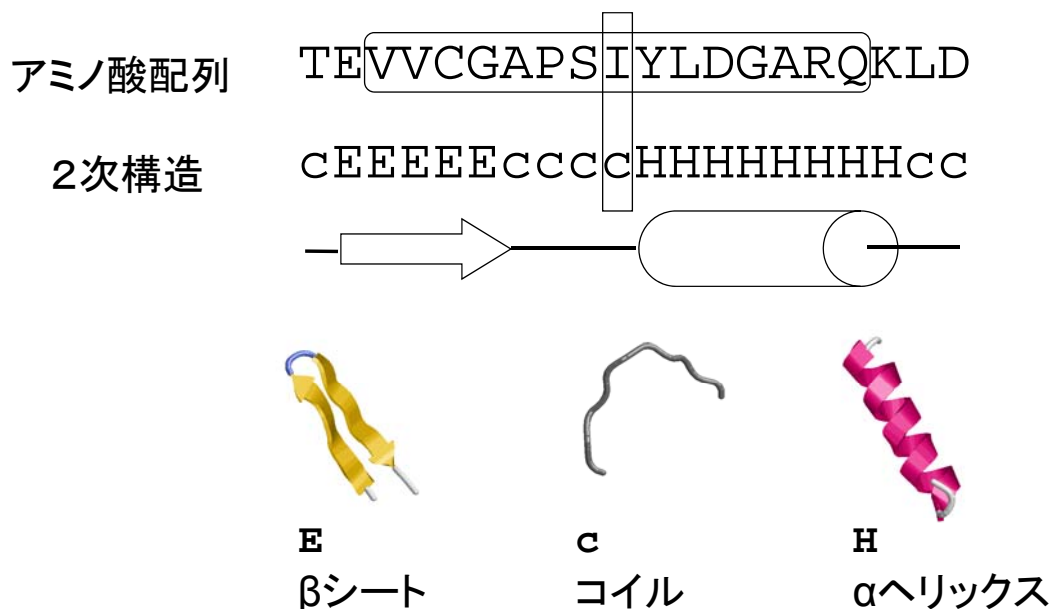
R P E W I W L A L G T A L M G L G T L Y F L V K G M G V S D P D A K K F Y  
A I T T L V P A I A F T M Y L S M I L G Y G L T M V P F G G E Q N P I Y W  
A R Y A D W I F T T P L L L L D L A L L V D A D Q G T I L A L V G A D G I  
M I G T G L V G A L T K V Y S Y R E V W W A I S T A A M L Y I L Y V L F F  
G F T S K A E S M R P E V A S T F K V L R N V T V V L W S A Y P V V W L I  
G S E G A G I V P L N I E T L L F M V L D V S A K V G F G L I L L R S R

>ZN428\_HUMAN (球状タンパク質・膜タンパク質・不定形タンパク質)(どれか一つを選択)

M T E T R E P A E T G G Y A S L E E D D E D L S P G P E H S S D S E Y T L  
S E P D S E E E E D E E E E E E E T T D D P E Y D P G Y K V K Q R L G G G  
R G G P S R R A P R A A Q P P A Q P C Q L C G R S P L G E A P P G T P P C  
R L C C P A T A P Q E A P A P E G R A L G E E E E E P P R A G E G R P A G  
R E E E E E E E E G T Y H C T E C E D S F D N L G E L H G H F M L H A R  
G E V

# 2次構造予測

- アミノ酸配列から、2次構造を予測すること。



## 2次構造予測の現状のまとめ

- Chou-Fasman等70年代に開発された方法の多くは正答率は**60%以下**
- ニューラルネットワークや改良GOR法を用いると**63~68%**ぐらいの正答率
- マルチプルアライメントを入力とする方法は正答率は**70%**を超える。現状では**76%**ぐらいまで到達。
- 一般にαヘリックスに比べ、βシートの予測は困難

現状のベストの方法

マルチプルアライメントを入力とする ニューラルネットワーク

Blast, ClustalW,  
PSI-BLAST, HMMer

アンサンブル学習  
再帰ネットワーク  
K-NNとの組み合わせ

## 2次構造予測の例

Single : 1つの配列を入力とするニューラルネットワーク [Q3=66.6%]

Profile : 複数の配列を入力とするニューラルネットワーク [Q3=83.5%]

```
>1n55A [c.1.1.1] ISOMERASE A: - - - TRIOSEPHOSPHATE ISOMERASE
AAseq :AKPQPIAAANWKCNGTTASIEKLVQVFNEHTISHDVQCQVAPTFFVHIPLVQAKLRNPKYV
Single :ccccccccccccccccHHHHHHHHHHHHccccccEEEEccccccccHHHHccccccH
Profile:ccccEEEEcHccccHHHHHHHHHHHHccccccccEEEEccHHHHHHHHHHccccccE
Observe:ccccEEEEccccccHHHHHHHHHHHHccccccccEEEEccccccHHHHHHccccccEE

AAseq :ISAQNIAIAKSGAFTGEVSMPIKLDIGVHWVILGHSERRTYGETDEIVAQKVSEACKQGF
Single :ccccccccccccccccHHHHHHHHHHHHccccccEEEEccccccccHHHHccccccH
Profile:EEccccccccccccccccHHHHHHccccEEEEccccccccccccHHHHHHHHHHHHccc
Observe:EEEccccccccccccccccHHHHHHccccEEEEccHHHHHHccccHHHHHHHHHHHHccc

AAseq :MVIACIGETLQQREANQTAKVVLSTSAIAAKLTKDAWNQVVLAYEPVWAIGTGKVATPE
Single :EEEEHHHHHHHHHHccccHEHHEHccHHHHHHHHHHHHHHHHccccccEEccccccHH
Profile:cEEEEccccHHccccHHHHHHHHHHHHHHccccccccEEEEccccccccccccccHH
Observe:EEEEccHHHHccccHHHHHHHHHHHHccccccccEEEEccccccccccccccHH

AAseq :QAQEVHLLLRKWVSENIGTDVAAKLRILYGGSVNAANAATLYAKPDINGFLVGGASLKPE
Single :HHHHHHHHHHHHHHHHccccHHHHHHHHHHccccHHHHHHccccccccccccccccccH
Profile:HHHHHHHHHHHHHHHHccccccccEEEEccccHHHHHHHHHHccccccccccccccHH
Observe:HHHHHHHHHHHHHHHHccHHHHHHcEEEEccccccccHHHHccccccccEEEEcccccccc

AAseq :FRDIIDATR
-----
```

## 参考図書

### 構造生物学一般について

- Bluce Alberis他著、(中村桂子、松原謙一監訳)「Essential 細胞生物学 原書第2版」(2005)、第2章、第4章、南江堂
- 松澤洋(編集)「タンパク質工学の基礎」(2004)、東京化学同人
- C.Branden & J.Tooze (勝部幸輝ら訳)「タンパク質の構造入門」(2000)、ニュートンプレス
- G.A.Petsko & D.Ringe (横山茂之監訳)「タンパク質の構造と機能 ゲノム時代のアプローチ」(2005)、メディカル・サイエンス・インターナショナル
- Arthur M.Lesk (高木淳一訳)「ポストゲノム時代のタンパク質科学 –構造・機能・ゲノミクスー」(2007)化学同人
- 樋口芳樹、中川敦史「構造生物学 –原子構造からみた生命現象の営み-」(2010)、化学同人

### 構造バイオインフォマティクス

- 藤博幸 編「はじめてのバイオインフォマティクス」(2006) 講談社サイエンティフィク
- 郷通子、高橋健一 編「基礎と実習 バイオインフォマティクス」(2004)、共立出版
- Arthur M.Lesk(岡崎康司、坊農秀雄 監訳)「バイオインフォマティクス基礎講義 一歩進んだ発想をみがくために」(2003)、メディカル・サイエンス・インターナショナル
- 美宅成樹・榊佳之「バイオインフォマティクス」(2003)、第6、7章、東京化学同人
- 日本バイオインフォマティクス学会編「バイオインフォマティクス事典」第10章、共立出版(2006)