

2006年度修士論文

# 生体分子の分子動力学に対する時系列解析

奈良女子大学人間文化研究科物理科学専攻  
複雑系の物理学研究室

修士2回生 大林 加奈子

# 目次

<b>1</b>	<b>はじめに</b>	<b>3</b>
1.1	生命への探求	3
1.2	タンパク質	4
1.2.1	タンパク質の基礎	4
1.2.2	タンパク質の理論的研究	6
1.3	非平衡性の必要性	7
1.4	問題意識	9
<b>2</b>	<b>解析方法</b>	<b>10</b>
2.1	解析対象	10
2.1.1	酵素とは	10
2.1.2	AKとは	10
2.1.3	AKのデータについて	11
2.2	解析手段	14
2.2.1	窓フーリエ変換	14
2.2.2	ウェーブレット変換	17
2.2.3	連続ウェーブレット変換	17
2.2.4	離散ウェーブレット変換	18
2.2.5	Haar ウェーブレット	19
2.2.6	トレンドとは	20
2.3	数値計算	23
2.3.1	wavelet とは	23
2.3.2	レベルとは	25
2.3.3	データの扱い	27
<b>3</b>	<b>オープン型構造のAK</b>	<b>28</b>
3.1	慣性テンソル	28
3.2	エネルギー	28
3.2.1	wavelet を通した際のエネルギーのゆらぎ	28
3.2.2	レベルごとの平均エネルギー	32
3.2.3	エネルギー等分配則	32
3.2.4	エネルギーの累積分布	34
3.3	集団運動性	36
3.3.1	指標 1 ( $C_{\text{ang}}$ )	36

3.3.2	指標 2 ( $C_{vel}$ )	41
3.4	実際の集団運動	46
3.5	まとめ	53
<b>4</b>	<b>AK の構造転移</b>	<b>54</b>
4.1	慣性テンソル	54
4.2	多項式フィッティング	55
4.3	フィッティングからのずれ	57
4.4	フィッティングの慣性テンソル	60
4.5	開いていく過程	61
4.6	集団運動性	68
4.7	実際の集団運動	73
4.7.1	場所相関 (1)	74
4.7.2	場所相関 (2)	76
4.8	まとめ	78
<b>5</b>	<b>総括と展望</b>	<b>79</b>
<b>6</b>	<b>謝辞</b>	<b>80</b>
<b>7</b>	<b>Appendix A</b>	<b>81</b>
7.1	慣性テンソル	81
<b>8</b>	<b>Appendix B</b>	<b>82</b>
8.1	エネルギー (lake)	82
8.1.1	エネルギー (1)	83
8.1.2	エネルギー (2)	86

# 1 はじめに

## 1.1 生命への探求

誰もが一度は考えたことがあるのではないだろうか。「自分は何故生きているのだろう。」と。勿論こんなことを考えなくとも、我々は食事をしたり睡眠をとったりと日々の生命活動を行うことができる。この世に生を与えられ命が果てるまで、意識的でなくともその生命は続いていくのである。無論、このことは我々人間だけに限ったことではなく、我々の周りにある多種多様な生物についても同様のことなのである。

では、そもそも生きているとは、生命とは何なのか。この問いは単純にして究極の問題なのではないだろうか。我々は生命を以ってして生きているのだから、これは一見簡単な謎解きのように思える。しかしながら、この問いかけは古代ギリシアの時代から投げかけられてきたにもかかわらず、未だ我々人類は明確な答えを得ていない。ここで、その探求の歴史を少し振り返ってみることとしよう。

古代では、病気は神や悪魔がもたらすものであると考えられていた。そんな中、古代ギリシアにおいて最初の合理的かつ科学的な医学が芽生えた。医学の祖である Hippokrates は初めて病気を自然現象として捉え、四液説や動物解剖によって得た解剖学的知見など様々な業績を残した。そのすぐ後、生物学の開祖と呼ばれる Aristoteles は比較解剖学と発生学を打ち立て、生物学的知識をそれまでになく豊かなものにし、広きに渡る研究方法を開拓した。西洋紀元のはじめ頃になると、Galenos が解剖学や生理学の基礎を築いた。彼の著作は 100 を超え、その影響力から中世を通じ医学の権威と仰がれた。このころから解剖への関心は高まり、中世には世界初の解剖学の専門家が誕生した。ボロニア大学の教授であった Mondino である。彼が 1316 年に著した解剖書『アナトミア』は中世における解剖学の最高水準を示すものである。Leonardo da Vinci は、この Mondino の解剖書を用いて人体解剖を学び、自らの観察により人体をあるがままの姿で認識し描いた数々の人体解剖図を残した。彼の解剖図はそれまでの歴史の中で最も正確で最高のものであった。そのようななかで解剖学は発展し、Vesalius によって近代解剖学が確立された。そうして、この時代から人は人体の構造と機能の真実を求めるようになったのである。そんな中、Jansen らによって顕微鏡が発明され、17 世紀半ば過ぎになると対象物を拡大して見る顕微鏡が登場し、その研究は急速に発展した。様々な「目に見えないもの」が見えるようになり、よりミクロなものへと人々の関心は高まるようになった [12]。

1869 年、Miescher によってついに核酸が発見された。その後、核酸の化学的、物理学的性質について、様々な研究がなされた。そうして、20 世紀の半ば、Watson と Crick によって DNA の分子構造が明らかになり、20 世紀の最後の年にはヒトゲノムが解読され、ヒトが持つ塩基配列の全てが決定された。しかしながら、生体内で実際に機能を果たしているのは塩基配列ではなく、タンパク質などのように立体構造をもつものなのである。核酸、とりわけ DNA にはタンパク質の設計図は書かれているが、実際にどのような構造で機能を果たしているのかという

ことまでは書かれていない。つまり、履歴書を見ただけで人物を判断出来ないのと同様、遺伝子の情報だけでは実際のタンパク質を知ることは出来ないのだ。生命現象を本当に知るためには、遺伝子だけを見ては不十分であり、タンパク質自体を研究する必要があるということである。タンパク質の理解は生命への探求にもつながると言っても過言ではないのではないだろうか。次節では、タンパク質についての概説を行うこととする。

## 1.2 タンパク質

### 1.2.1 タンパク質の基礎

タンパク質とは、アミノ酸が重縮合した鎖状の生体高分子のことである。重縮合した結果、アミノ酸どうしはペプチド結合と呼ばれる結合でつながっている。まず、アミノ酸とはどのような物質なのかを概説したいと思う。

アミノ酸とは、分子内に塩基としての性質をもつアミノ基 ( $-\text{NH}_2$ ) と、酸としての性質をもつカルボキシル基 ( $-\text{COOH}$ ) をもっている物質の総称である。アミノ酸はカルボキシル基に対するアミノ基の結合位置により、 $\alpha$ -アミノ酸、 $\beta$ -アミノ酸、 $\gamma$ -アミノ酸などと呼ばれる。生体内には、 $\alpha$ -アミノ酸、 $\beta$ -アミノ酸も存在し、重要な役割を担っているが、タンパク質に関係するアミノ酸は  $\alpha$ -アミノ酸である。タンパク質はすべて 炭素原子にアミノ基が結合した  $\alpha$ -アミノ酸から構成されており、通常、アミノ酸の構造式はアミノ酸を左側に、カルボキシル基を右側を書く (図 1)。タンパク質を構成するアミノ酸は 20 種類あり、それぞれ異なる側鎖を持っている。地球上のすべての生物のタンパク質は、この同じ 20 種類のアミノ酸からできている [13]。

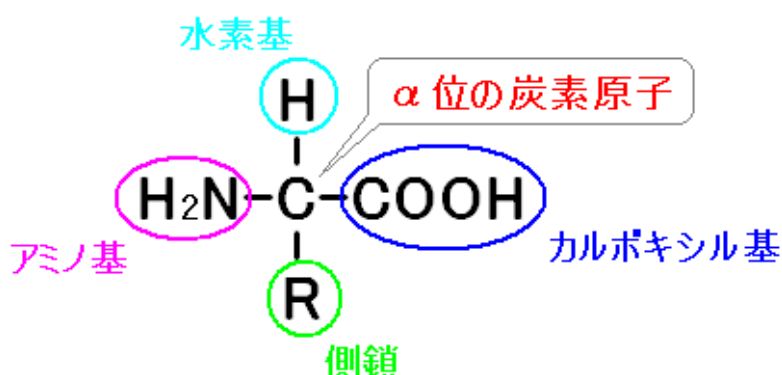


図 1:  $\alpha$ -アミノ酸の構造式

アミノ酸のカルボキシル基と、隣のアミノ酸のアミノ基との間で縮合が起こり、水分子が遊離し、ペプチド結合(-CONH-)が形成される(図2)。

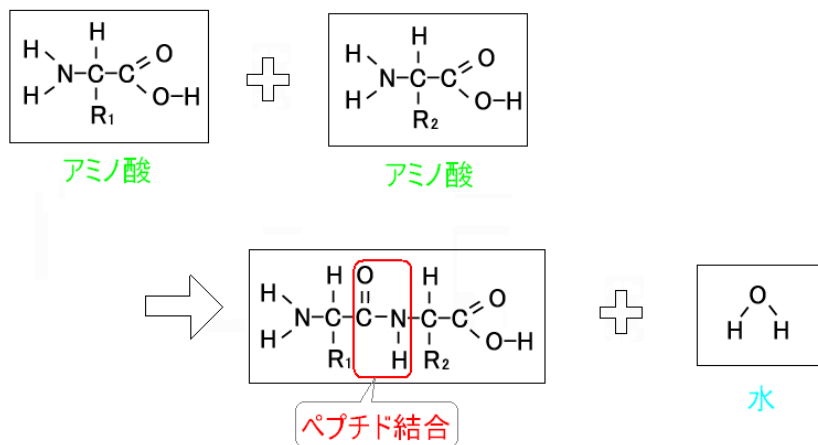


図 2: ペプチド結合の形成

出来上がった分子をペプチドと呼ぶ。ポリペプチドは、アミノ酸がペプチド結合により重合した分子であり、ポリペプチドが即ちタンパク質である。タンパク質分子の上で、もとのおのおののアミノ酸に当たる部分をアミノ酸残基と呼ぶ。

タンパク質は、常に、遊離のアミノ基の存在する側、すなわちアミノ末端を左側に、遊離のカルボキシル基の存在する側、すなわちカルボキシル末端を右側にして書く。そして、アミノ末端より数えて、1番目のアミノ酸残基、2番目のアミノ酸残基、.....、 $n$ 番目のアミノ酸残基とする。つまり、アミノ酸残基の配列は、アミノ末端の残基を1番目とし、カルボキシル末端へ向かって、右の方へ示すことになっている(図3)。それぞれのアミノ酸残基に、20種類のアミノ酸のうちの1個がくるので、アミノ酸配列は様々になるのである。具体的には、短いものでは数十、長いものでは数千個ものアミノ酸が連なり、千~数十万個の原子でできる巨大で複雑な分子を構成し、タンパク質となる。

ところで、図3に挙げたような構成は何によって決められているのだろうか。その答えは「遺伝子」である。アミノ酸の並び方は、タンパク質が細胞の中で作られる際に、遺伝子に書き込まれた遺伝情報に基づいて定められている。つまり、遺伝子がタンパク質のアミノ酸の並び方を決めているのだ。遺伝子によって与えられたアミノ酸の並びの順序をタンパク質の「一次構造」という。しかしながら、実際に生体内で機能を果たしている構造は「三次構造(立体構造)」なのである。つまり、タンパク質を考える際には一次構造だけでは不十分であり、立体構造を

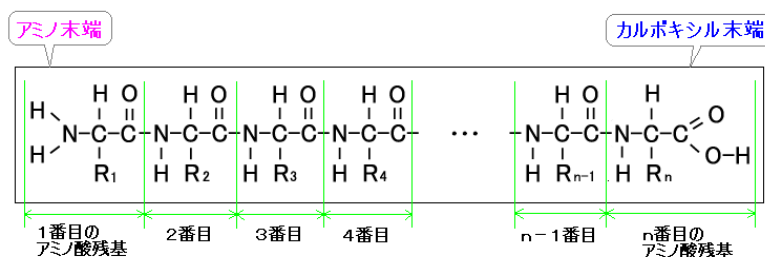


図 3: タンパク質の形成

考える必要があるのだ。ところで、アミノ酸の配列としてコードされた一次元的な遺伝情報がタンパク質としての機能をもった天然の三次元的な立体構造に変わっていく過程をタンパク質のフォールディング過程という。次節では、タンパク質のフォールディングについての研究などタンパク質についての理論的な研究について紹介することにする。

### 1.2.2 タンパク質の理論的研究

さまざまな遺伝病の原因は、タンパク質の間違ったフォールディングに由来するなど、我々の生命に関わる要因になりうるということもあり、フォールディングの研究は様々な形で行われている。そのなかでも、フォールディングの理論的な研究には、主として2つの流れがある。1つは、出来るだけ簡単なモデルを用いて現象の本質を探ろうという試みであり、もう1つは、計算機を用いてタンパク質の運動を出来るだけ正確に再現しようという試みである。前者の代表的な研究としては Sasai の研究が挙げられるであろう。Sasai は、炭素だけを考慮し、そのほかの原子の自由度を無視したモデルによるカルシウム結合タンパク質の構造計算を経験的ポテンシャルを用いて行った [6]。これは、簡単なモデルを用いて、天然構造へと向かっていくフォールディング過程を探ろうというものであった。この実験により、天然構造が際立ったエネルギー谷であるということだけではフォールディングするための十分条件ではないことや、エネルギー面が広い範囲にわたってうまく設計されている必要があることなどが示唆された。また、後者の代表的な研究としては Kitao らの研究が挙げられるであろう。タンパク質の動きは、1つの構造状態近傍でのゆらぎだけではなく、構造状態間の繊維によっても生じてくる。このような複雑さは、タンパク質の分子のみに起因するのではなく、溶媒環境によっても大きく影響されている。Kitao らは、タンパク質だけでなく周辺の溶媒分子にも着目し、分子シミュレーションを使って、タンパク質の動きの複雑性と構造状態の階層性に関する研究を行った [7][8]。その結果、タンパク質と水の運動のカップリングは複雑で、時間スケールの階層的な影響を及ぼしあっていることが分かった。しかしながら、Sasai や Kitao らの研究はどちらも自由エネ



ルギーを前提としたものである。これは、局所平衡を仮定したものであるが、この仮定の正当性は必ずしもない。また、Kideraらによって、X線結晶構造解析はタンパク質の形を精密に決めるだけでなく、その動きの情報も温度因子を通して与えていると考えられた [9][10][11]。この温度因子を用いると、動的なタンパク質の描像を得ることが出来ることも提示されている。

他にも、タンパク質の機能にかかわるダイナミクスの研究は盛んに行われている。しかしながら、未だその全容は明らかにされていない。本研究の究極の目的は、まさに、タンパク質の機能と dynamics の関係を知るところにあるのである。

### 1.3 非平衡性の必要性

One of the best established facts in thermodynamics is that it is impossible in a system enclosed in an envelope which permits neither change of volume nor passage of heat, and in which both the temperature and the pressure are everywhere the same, to produce any inequality of temperature or of pressure without the expenditure of work. This is the second law of thermodynamics, and it is undoubtedly true as long as we can deal with bodies only in mass, and have no power of perceiving or handling the separate molecules of which they are made up. But if we conceive a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are still as essentially finite as our own, would be able to do what is at present impossible to us. For we have seen that the molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower ones to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics.

This is only one of the instances in which conclusions which we have drawn from our experience of bodies consisting of an immense number of molecules may be found not to be applicable to the more delicate observations and experiments which we may suppose made by one who can perceive and handle the individual molecules which we deal with only in large masses.

In dealing with masses of matter, while we do not perceive the individual molecules, we are compelled to adopt what I have described as the statistical method of calculation, and to abandon the strict dynamical method, in which we follow every motion by the calculus. ([17] より引用)

「熱力学において最も確立された事象の一つは、温度が一様であり外からの力は一様な圧力だけであり体積変化も熱の出入りもない閉じた系において、何らかの仕事を加えずして温度や圧



力の不均衡を生み出すことは不可能であるということだ。これが熱力学の第二法則であり、我々が系を集団としてしか扱えず、またその集団を構成する一つ一つの分子を知覚し手を触れる能力を持たない限り、これは疑いもない真実である。しかし、もし仮にすべての分子の軌道を追跡することができるくらい鋭い能力をもった“存在”がいるとしたら、この“存在”の特性が我々同様本質的に限定されているとしても、その“存在”は現時点では我々には不可能なことができるだろう。というのも、たとえそのなかの任意に選ばれた多数の分子の平均速度がほとんど正確に一樣だったとしても、一樣な温度の空気が満たされた容器の中の分子が、すべて同じ速度で運動しているわけではないからである。そこでこの容器を小さな穴のある仕切りで二つの部分 A,B に分け、件の個々の分子を見分けることのできる“存在”がこの小さな穴を開閉し、速度の速い分子は A から B へ、速度の遅い分子は B から A へと通すとしよう。こうしてこの“存在”は全く仕事を伴わずに熱力学の第二法則に反した行為、即ち B の温度を下げ、A の温度を上げる行為を行うこととなる。

これは多くの分子からなる系についての我々の経験から導かれた結論が、我々が大きな集団としてしか扱うことのできない分子を知覚し、手を触れることのできる“存在”によって行われたと思われる、繊細な観察と実験に対してはあてはまらないという一例にすぎない。

我々は物質の集団を扱い、一方で個々の分子を知覚しないことで、計算の統計的方法として述べることを採用し、計算ですべての運動を追うような、厳密な力学的手法を捨てることを強要されている。」 [19]

この一節で、Maxwell は、熱力学の第二法則に反するような存在を“a being”と表現し、その存在を示唆している。“a being”とは、後に“Maxwell の悪魔”と呼ばれるようになるのだが、“Maxwell の悪魔”に関しては、今日まで様々な議論が行われている [18]。

ここで、Maxwell の議論についてまとめる。いま、閉じた系であるマクロな大きさの箱があるとす。この箱の中は、温度一定の熱平衡状態であるとす、箱の中に仕切りをつけて小さな穴を開けておく。次に、“a being”は、「分子を識別できて、その速度を測れる」という能力を持っているとする。そのため、“a being”は、

- $|v| > v_0$  のときには、小さな穴に分子を通す
- $|v| < v_0$  のときには、小さな穴に分子を通さない

という作業ができるものとする。すると、一樣な温度の系から非一樣な温度をもつ系が作られることになる。つまり、“a being”という存在は、「2つの部分が違っていれば、ある時間の後に熱平衡状態に緩和する」という熱力学の第二法則に矛盾した現象を引き起こしていることになるのだ。

しかしながら、生体内では、しばしばこのような現象が確認できる。例えば、細胞膜は、特定の分子のみを通し細胞の内外に濃度差を作り、様々な生命現象を引き起こしている。このこ

とは、生体分子の分子識別能力の存在はマクロな現象を変えるということを意味している。ただし、この分子識別能力とは、“Maxwellの悪魔”的なものであろうか。生体分子が“Maxwellの悪魔”のような仕組みを持っているだけでその機能を果たすことが出来るのだろうか。

近年、タンパク質などの生体分子が機能を果たすためには、非平衡性が必要だという議論がなされている。それは、“Maxwellの悪魔”を含めた全系ではエントロピーが増大するが、部分系では非平衡系が維持され、その非平衡性が生体分子の方向性を持った動きや生体内での様々な動きに必要なのではないかと、いうものである [19]。我々は、生体分子が機能を果たすためには、非平衡性が必要であるだろうという立場にたって考えている。

## 1.4 問題意識

前節でも述べたように、熱平衡状態では生体分子は分子識別能力を発揮できないため、生体分子が機能を発現するためには、非平衡性が必要であると考えられる。また、分子レベルでのミクロなものと人間など個体レベルでのマクロなものとを分離することは出来ないため、分子レベルでの分子識別について考える必要がある。

「どのようなメカニズムで分子識別は起こっているのだろうか？」

この問いに答えるべく、本研究では、生体分子の集団運動に注目し、その運動の中から重要と考えられる少数の集団運動の特徴を抽出する理論的な方法を得ることを目的とする。そのため、アデニル酸キナーゼという生命現象に欠かせない酵素に着目し、この運動に対して解析を行う。また、この酵素は機能を発現する際に大きな変形が見られるため、その集団運動に注目することにより、生体分子の機能に関する運動と機能に関係しない運動のメカニズムを知ることが出来るのではないかと考える。このメカニズムの解明は分子識別のメカニズムの解明にも繋がるであろう。そして、さらには、生化学反応のコントロール可能性についても模索したい。

## 2 解析方法

### 2.1 解析対象

#### 2.1.1 酵素とは

生物の体の中では様々な化学反応が起こっている。その中でも特に重要な役割を果たしているのは酵素であろう。酵素は特定の物質（基質）に作用し、特定の化学反応を進行させる触媒作用をもつ物質である。では、どのようなメカニズムで酵素は触媒機能を発揮しているのだろうか。

酵素反応のメカニズムを説明すべく、いろいろなモデルや考え方が生み出されてきた。歴史的に最も古いのが「鍵と鍵穴」説であろう。「基質と酵素の活性部位は、ちょうど鍵と鍵穴の関係にあって、ぴったりと適合する。適合しない物質は、基質にはなりえない。」という説である [15]。しかし、この説はタンパク質を剛体と捉え、タンパク質の「やわらかさ」を考慮に入れていないことに問題点があった。研究が進むにつれ、酵素自身の鍵穴ははじめから不動のものとして出来ているものではなく、基質を迎え入れる際に酵素自身の構造も特定の変化をし、しっかりと包み込んで触媒反応を起こすことが分かってきた。今日では、X線結晶構造解析やNMRにより酵素ごとの構造を知ることは出来る。しかしながら、アミノ酸配列からどのような立体構造ができ、そこからどのような酵素活性が生ずるのかというような一般的な理論は、未だ打ち立てられていない。

#### 2.1.2 AKとは

AKとは、アデニル酸キナーゼという酵素のことである。AKは、生命体内のエネルギー反応に関係しているという点において、酵素のなかでも特に重要な酵素なのである。

生命を営むほとんどすべての細胞は、化学的エネルギーをATP（アデノシン三リン酸）に一時的に蓄え、エネルギーを使用する際にはリン酸が3つ結合しているATPから、リン酸を1つはずしてADP（アデノシン二リン酸）、2つはずしてAMP（アデノシン一リン酸）とし、その時に出てきたエネルギーを利用する [16]。また、細胞内では、解糖系、クエン酸回路、酸化リン酸化反応によるATP生産系と、筋肉を動かしたり、化学反応を起こしたりするATP利用系があり、複雑な調節機構により、お互いに均衡を保っている。この均衡を保つのに一役買っているのが、AKなのである。アデニル酸キナーゼ（AK）という酵素は、3つのリン酸をもつATPと1つのリン酸をもつAMPから、2つのリン酸をもつADPを2分子生成するリン酸転移反応を可逆的に触媒する酵素なのである（図4）。

約200個のアミノ酸残基により構成されるAKは、基質の結合に伴って立体構造が大きく変化することが知られている。AKには、基質が結合していない場合に安定であるオープン型構造

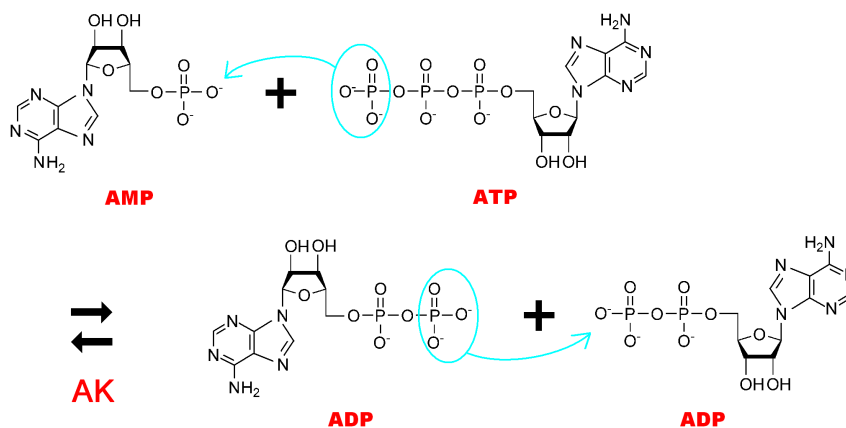


図 4: AK の触媒反応

と基質が結合したときに安定であるクローズ型構造がある。PDB(Protein Data Bank)のIDで、オープン型構造を 4ake と呼び、クローズ型構造を 1ake と呼ぶ。

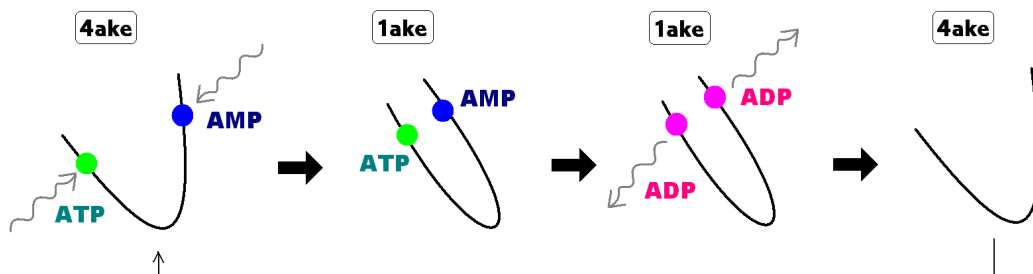


図 5: AK の模式図:黒い線は AK を表す。上図は、AK が図 4 の右方向へ触媒する際の模式図である。

本研究では、オープン型構造 (4ake) の AK とクローズ型構造 (1ake) からオープン型構造 (4ake) へ構造変化する際の AK の dynamics について考える。解析する際に用いている AK のデータの詳細については、次節に記すこととする。

### 2.1.3 AK のデータについて

本研究で用いている AK の時系列データは、横浜市立大学大学院の木寺詔紀氏と瀧上壮太郎氏によって作成されたものである。以下にデータの詳細を記す。

前節で、AK には 2 つの構造、つまり、オープン型構造 (4ake) とクローズ型構造 (1ake) があると述べた。オープン型構造は基質がない状態で、クローズ型構造は基質がある状態であるとも述べたが、今回用いているデータはどちらも基質が無いものである。つまり、4ake という

のは、オープン型構造の AK のシミュレーションであり、1ake というのは、クローズ型構造を初期構造として用いた基質無しの AK のシミュレーションということである。基質がないということは、オープン型構造へと構造緩和することが当然期待されるが、本研究で用いているデータはまさにその期待通りに得られた 1ake のトラジェクトリの結果を用いている。

次に、系の詳細について紹介する。まずはじめに、AK (214 残基、3341 原子) を水分子の箱に入れた系を作る。この際、静電相互作用の計算の都合上、系を中性にする必要があるため、適当なイオンを追加する。今回の場合は、AK の電荷が 4 であるため、2 つの系の構成は以下のようになっている。

4ake: タンパク質 (214 残基、3341 原子)

水分子 (23044 分子、69132 原子)、ナトリウムイオン (4 原子)

計 72477 原子

box size: 90.6 [ ] の立方体

1ake: タンパク質 (214 残基、3341 原子)

水分子 (15972 分子、47916 原子)、ナトリウムイオン (4 原子)

計 51261 原子

box size: 80.52 [ ] の立方体

シミュレーションの実行には、横浜市立大学大学院の池口氏によって開発された分子動力学シミュレーションソフトウェア MARBLE を使い、力場には CHARMM22 を用いている。また、MARBLE で用いているインテグレータは、エネルギー一定の計算に限らず 2 次のシンプレクティック法である。シミュレーションは多数の CPU を用いた並列計算で実行しているが、空間分割法および動的負荷分散アルゴリズムの導入により、高い並列化効率を実現している。また、MARBLE では、水素原子を含む原子団 ( $-\text{CH}_3$ ,  $-\text{CH}_2-$ ,  $-\text{CH}-$ ,  $\text{NH}_3$ ,  $\text{NH}_2$ ,  $-\text{NH}-$ ,  $-\text{OH}$ ,  $-\text{SH}$  など) を剛体として取り扱うことが出来る (部分剛体分子動力学法)。このことにより、SHAKE 法や RATTLE 法によって原子間距離に拘束をかけるよりも効率的な計算を実現している。つまり、メチル基 ( $-\text{CH}_3$ ) などでは結合距離だけでなく、 $\text{H}-\text{C}-\text{H}$  などの結合角も固定されていることになる。系の境界条件としては、周期境界条件を導入している。これは、タンパク質周辺の水分子の運動が境界の存在によって影響を受け、タンパク質の挙動にさらに人為的要因が混入するのを防ぐためである。また、静電相互作用の計算には、Particle Mesh Ewald (PME) 法を用いている。静電相互作用の計算は全計算時間の大部分を占め、その計算精度がシミュレーション結果の信頼性に直結するが、PME 法では高精度でかつ高速な計算が可能であり、現在の最もスタンダードな方法である。

続いて、計算の手順について紹介する。まずはじめに、エネルギーの最小化を行う。これは、



結晶構造をもとに水素原子の配置を決めたり、水分子の中にタンパク質を埋めたときにできる可能性のある原子間の衝突を避けるためである。エネルギーの最小化を行う際には、単純な最急降下法を用いている。次に、系の平衡化を行う。まず、タンパク質の重原子を拘束し、平衡化を 400[psec] 行う。これは、平衡への緩和過程において、タンパク質が大きく変形してしまうのを避けるためである。続いて、タンパク質に課した拘束を 100[psec] かけて徐々にはずしていく。最後に、拘束がまったくない状態で、500[psec] の計算を行う。500[psec] で十分に平衡状態に達しているかどうかは、別途確認する必要がある。オープン型構造からはじめた場合は、これで十分だが、クローズ型構造からはじめた計算の場合は、平衡状態がオープン型構造を平均構造とするはずなので 500[psec] では不十分であり、引き続き行う本計算において、緩和過程における構造変化を見ている。また、平衡化に用いたアンサンブルは NPT (等温等圧、 $P=1.0[\text{atm}]$ 、 $T=298.15[\text{K}]$ ) で、Nose-Andersen の方法を使っている。上記の平衡化後、アンサンブルを NVE (ミクロカノニカル) にスイッチし、10[nsec] の本計算を行っている。

すべての MD シミュレーションにおいて、時間刻みは 2[fsec] としている。これは、水素原子を固定する場合の標準的な刻み幅である。また、トラジェクトリのデータは、時間刻みごとに保存すると膨大なデータ量になるため、1[psec] ごとに保存している。

今回の解析においては、上記のようにして求められた AK のトラジェクトリのデータのうち炭素のみのダイナミクスに着目している。また、今回用いている AK のデータは、214 個のアミノ酸残基で構成される大腸菌の AK のものである。図 6 は、molscript を用いて AK のトラジェクトリのデータの炭素のみを描いたものである。赤い原子は、疎水性のアミノ酸の炭素であることを表し、青い原子は、親水性のアミノ酸の炭素であることを表す。

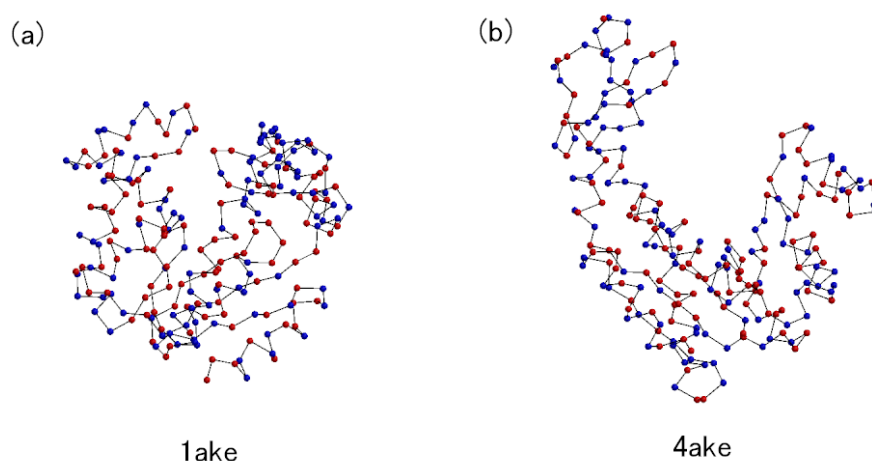


図 6: AK の立体構造

## 2.2 解析手段

本研究では、ウェーブレット変換という手法を用いている。ウェーブレット変換とは、1983年にJ.Morletにより地震波データの解析を行うために提案された手法であり、データや関数、作用素を異なった周波数成分に分け、それぞれの成分をそのスケールに合った解像度で調べる道具である [1]。信号解析がその主な例である。この場合、時間的に徐々に変化する信号のウェーブレット変換は、周波数と時間の二つの変数に依存する。即ち、ウェーブレット変換は時間 - 周波数の局在化の道具を与えているのである。単に、時間 - 周波数の局在化を行いたいだけなら、窓フーリエ変換でも可能である。しかしながら、窓フーリエ変換にはある種の限界があり、ウェーブレット変換であればそれは可能なのである。そこに、本研究でウェーブレット変換を用いた大きな理由が潜んでいる。次節からは、その理由を明らかにすべく、窓フーリエ変換、ウェーブレット変換を概説する。

### 2.2.1 窓フーリエ変換

時間により変化する信号  $f(t)$  をフーリエ変換すると、そのスペクトル  $F(\omega)$  が得られる。

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int dt e^{-i\omega t} f(t) \quad (2.1)$$

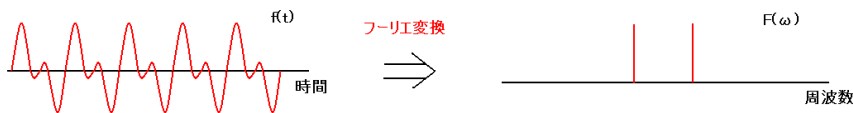


図 7: フーリエ変換

しかし、このスペクトルは、周波数のみの関数となり、もはや時間の情報は失われてしまうのである。したがって、スペクトルの時間的変化を知るためには、信号の一部を窓関数を用いて切り出し、この窓をずらしながら、その区間の信号のスペクトルを次々に解析する必要がある。信号の周波数の時間的変化を解析する手法として、まず窓フーリエ変換が挙げられる。窓フーリエ変換では、一定の大きさの窓関数を用いて信号を切り出し、その結果をフーリエ変換してスペクトルを計算する。

$$(T^{\text{win}} f)(\omega, t) = \int ds f(s) g(s-t) e^{-i\omega s} \quad (2.2)$$

これは窓フーリエ変換であり、時間 - 周波数の局在化の標準的な手法である。しかし、窓フーリエ変換は、周波数によらず一定の時間枠を用いるため、必ずしも効率的な方法ではない。そ



うしたフーリエ変換の限界を補うものが、ウェーブレット変換である。窓フーリエ変換が非効率なのは、周波数に関係なく同じ大きさの窓を使ったことが原因であった。したがって、周波数に応じて窓の大きさを変えてやれば、効率的な分解が可能になる。周期の長い低周波に対しては、窓の枠を大きくとってもよいが、周期の短い高周波に対しては、窓の枠を小さくする必要がる。ウェーブレット変換は、まさにそうした要望に応える手法なのである。次項に、窓フーリエ変換とウェーブレット変換の違いをあらわす模式図を示し、次節からは、ウェーブレット変換について概説する。

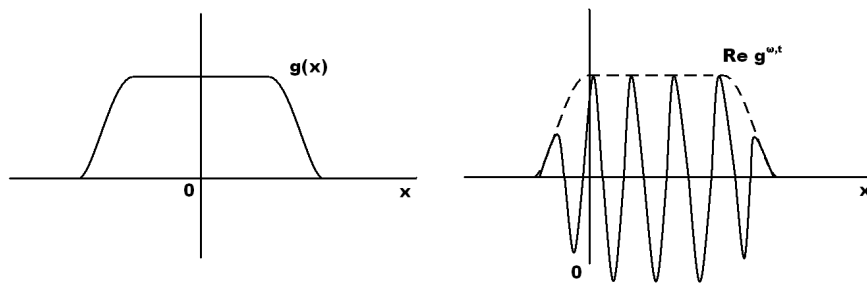


図 8: 窓フーリエ変換  $g^{\omega,t}(x) = e^{-i\omega x}g(x-t)$  ([1] より引用)

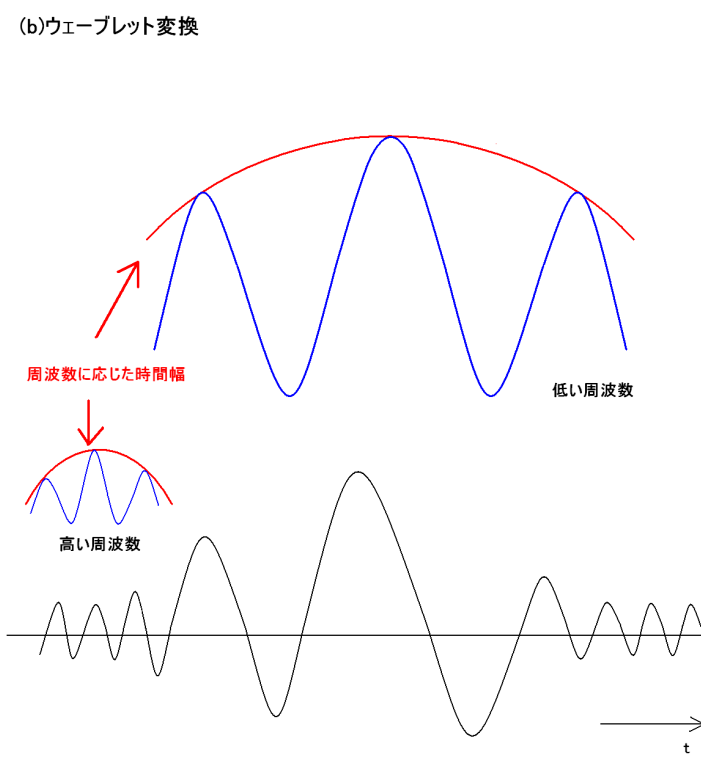
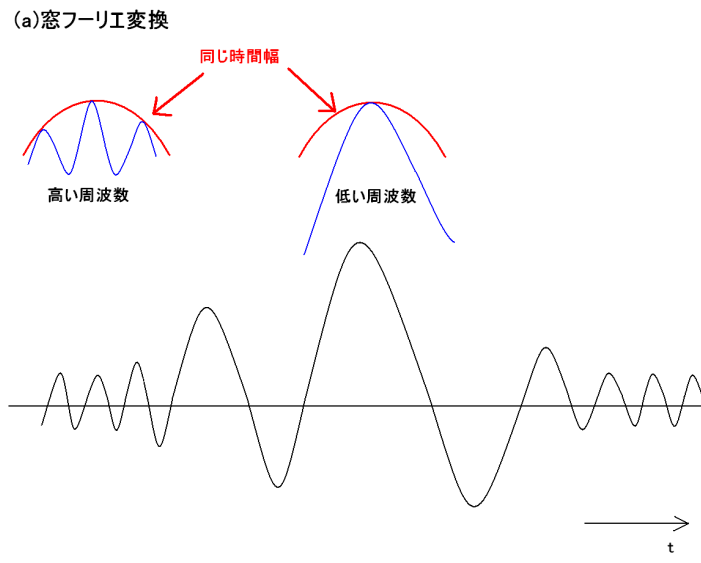


図 9: 窓フーリエ変換とウェーブレット変換の違い：窓フーリエ変換では窓の時間幅は周波数によらず一定だが、ウェーブレット変換では窓の時間幅は周波数により異なる。

## 2.2.2 ウェーブレット変換

ウェーブレット (wavelet) は、「さざなみ」と訳される。ウェーブレットは、時間的な流れの中で、生まれては消えてゆく一時的な波であり、ウェーブレット変換は、任意の時系列データをウェーブレットの和として表現する手法である。一方、フーリエ変換は、任意の時系列データを、無限に続く恒久的な波の和として表現するものである。

フーリエ変換とウェーブレット変換は、どちらかが勝っているというものでもない。例えば、時系列特性が時間を通じて不変で、規則的な変動を繰り返すデータに対しては、時間的な情報を無視して、詳細な周波数解析を行うフーリエ変換は効率的な方法である。しかし、不規則変動を示すデータに対しては、フーリエ変換ではなくウェーブレット変換がその本領を発揮する。本研究では、タンパク質の運動というまさに不規則な運動のデータ解析を行うため、ウェーブレット変換が適していると考える。ウェーブレット変換には、連続ウェーブレット変換と離散ウェーブレット変換がある。以下に、それぞれについての概説を行う。

## 2.2.3 連続ウェーブレット変換

ウェーブレット変換は以下のように定義される。

$$(T^{\text{wav}} f)(a, b) = |a|^{-1/2} \int dt f(t) \psi\left(\frac{t-b}{a}\right) \quad (2.3)$$

ここで、 $a$  は scale、 $b$  は translation を表し、 $\psi$  はマザーウェーブレットといい、 $\psi$  は  $\int dt \psi(t) = 0$  を満たすとする。また、マザーウェーブレット  $\psi$  を拡大縮小し、平行移動したさまざまな参照波のことをウェーブレット  $\psi^{a,b}$  と呼ぶ。元の関数は、以下の "単位の分解" の公式により  $f$  のウェーブレット変換から再構成することができる。

$$f = C_{\psi}^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{da db}{a^2} \langle f, \psi^{a,b} \rangle \psi^{a,b} \quad (2.4)$$

ただし、 $\psi^{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right)$  であり、 $\langle \cdot, \cdot \rangle$  は  $L^2$ -内積を表す。定数  $C_{\psi}$  は  $\psi$  だけに依存し、

$$C_{\psi} = 2\pi \int_{-\infty}^{\infty} d\xi |\hat{\psi}(\xi)|^2 |\xi|^{-1} \quad (2.5)$$

で与えられる。ここで、 $C_{\psi} < \infty$  と仮定する。例えば、ウェーブレットにはメキシカンハット関数などがある (図 10)。

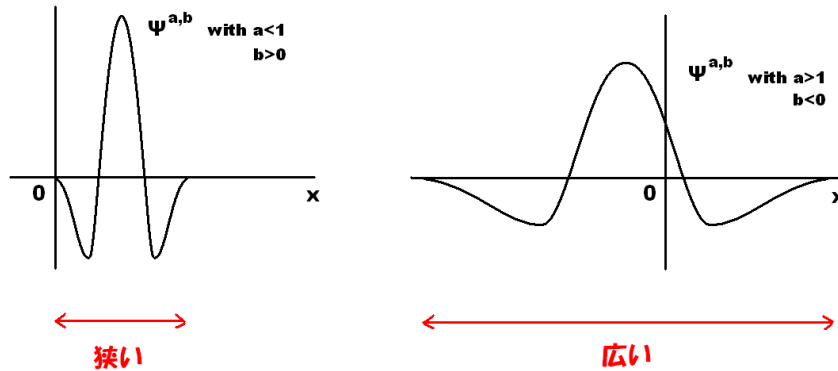


図 10: ウェーブレット  $\psi^{a,b}$  ([1] より引用)

### 2.2.4 離散ウェーブレット変換

信号解析では前述の連続ウェーブレット変換を離散化したものが一般的に用いられている。 $t$  や  $\omega$  を  $t = nt_0$ 、 $\omega = m\omega_0$  と離散化して、(2.3) 式の変換を離散化すると、

$$T_{m,n}^{\text{wav}}(f) = a_0^{-m/2} \int dt f(t) \psi(a_0^{-m}t - nb_0) \quad (2.6)$$

となる。ただし、 $m \in \mathbb{Z}$ 、 $n \in \mathbb{Z}$ 、 $a_0 > 1$ 、 $b_0 > 0$  とする。

先程紹介した連続ウェーブレット変換は、過剰完全系であるため独立の成分に分解することが出来ない。しかし、多重解像度解析などに用いられる離散ウェーブレット変換では、基底が直交完全系であるため独立の成分に分解することが出来る。その代表例として、Haar ウェーブレットが挙げられるであろう。次節では、Haar ウェーブレットを用いての離散ウェーブレット変換を概説する。

### 2.2.5 Haar ウェーブレット

Haar 関数  $\psi(x)$  は

$$\psi(x) = \begin{cases} 1, & 0 \leq x < \frac{1}{2} \\ -1, & \frac{1}{2} \leq x < 1 \\ 0, & \text{その他} \end{cases} \quad (2.7)$$

で定義される関数である。ここで、

$$\psi_{m,n}(x) = 2^{-m/2} \psi(2^{-m}x - n) \quad (2.8)$$

とする (図 10)。

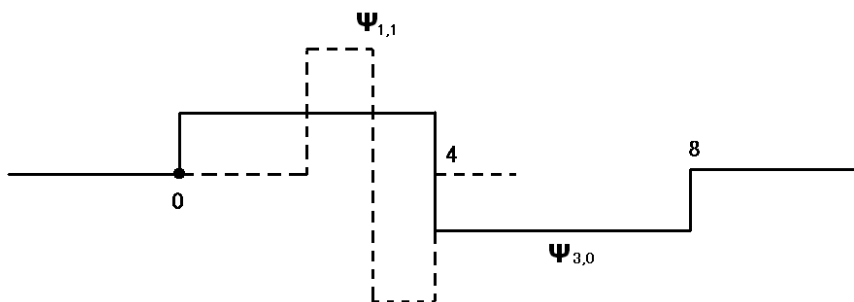


図 11: Haar 関数 ([1] より引用)

いま、レベル  $n$  の近似関数を  $f_n(x)$  とする。レベルについては後ほど詳述するが、今はレベルの値が小さい程、より正確な近似になっていると考えてもらいたい。 $f_n(x)$  を各点の平均値  $f_{n+1}(x)$  と差  $(f_{n+1}(x) - f_n(x))$  に分解する。続いて、 $n+2$ 、 $n+3$  とレベルを上げていくと、より解像度の低いデータが得られる。一つレベルを上げるごとに、解像度が半分になるのである。ここで重要なことは、解像度を落としたとしても、平均値  $f_k(x)$  と差  $(f_{k+1}(x) - f_k(x))$  の値が分かれば、 $f_{k-1}(x)$ 、 $f_{k-2}(x)$ 、... と次々にデータを再構成できるということなのである。以上より、Haar ウェーブレットを用いると、データの分解・再構成が出来ることが分かった (図 12)。本節では、簡単のために Haar ウェーブレットについて紹介したが、本研究では、Daubechies のウェーブレット変換を用いている。その理由としては、Daubechies のウェーブレットは Haar ウェーブレットのように時間で局在しているだけでなく、周波数空間でも局在しているからである。次節では、Daubechies のウェーブレットの性質を用いて、trend という概念について紹介したいと思う。

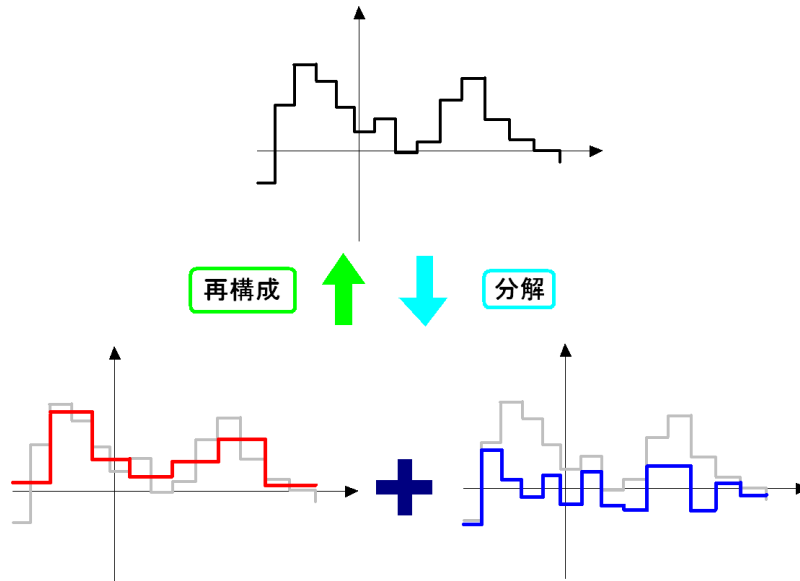


図 12: Haar ウェーブレット

### 2.2.6 トレンドとは

ある物理量  $A(t)$  に対して、多項式のトレンド  $T(t)$  があるとする。すると、

$$A(t) = T(t) + X(t) \quad (2.9)$$

$$( A(t) = a_1t + a_2t^2 + \dots + a_k t^k )$$

と表すことが出来る。ここで、 $k$  は正の整数であり、 $X(t)$  は多項式では表せないトレンドを含む時系列データであるとする。また、 $a_k$  は係数とする。 $A(t)$  に対して、ウェーブレット  $\psi_n(t)$  でウェーブレット変換を行うと、

$$\int_{-\infty}^{\infty} \psi_n(t)A(t)dt = \int_{-\infty}^{\infty} \psi_n(t)(a_1t + a_2t^2 + \dots + a_k t^k)dt + \int_{-\infty}^{\infty} \psi_n(t)X(t)dt \quad (2.10)$$

となる。

ここで、Daubechies のマザーウェーブレット  $\psi$  は、周波数空間で局在しているためモーメントが 0 になるという条件

$$\int_{-\infty}^{\infty} x^l \psi_n(x)dx = 0 \quad (l = 0, 1, \dots, n-1) \quad (2.11)$$

を満たさなければいけないので、

$$\int_{-\infty}^{\infty} \psi_n(t)(a_1t + a_2t^2 + \dots + a_kt^k)dt = 0 \quad (2.12)$$

となる。これを上式に代入すると、

$$\int_{-\infty}^{\infty} \psi_n(t)A(t)dt = \int_{-\infty}^{\infty} \psi_n(t)X(t)dt \quad (2.13)$$

となる。これより、Daubechies のウェーブレットを用いると、多項式のトレンドを取り出せることが分かる。さらに、Daubechies のウェーブレットは  $X(t)$  が持っているトレンドも取り出すことができるということが、Craigmile らによって提示されている [5]。その内容について紹介しよう。



下図は、Craigmireらによって[5]に挙げられているものである。この図について説明をする。ある2次関数に対して、(d)のようなmemoryをもつ雑音を与える。これが(a)である。(a)に対して、Daubechiesのウェーブレット変換を行うと、(b)のようなゆっくりした動きと(c)のような速い動きに分けることができる。また、(c)を周波数領域で見た(e)については、低周波領域での値が著しく小さくなっているのに対し、(d)を周波数領域で見た(f)については、低周波数領域での値が大きくなっていることが読み取れる。これは、始めに与えた雑音 $X(t)$ はtrendを含んでいたが、wavelet変換したあとの雑音はtrendを含まない定常的なものになっていることを示唆している。以上より、Daubechiesのwaveletを用いると、定常的な部分と非定常的な部分に分けることができるということが分かる。

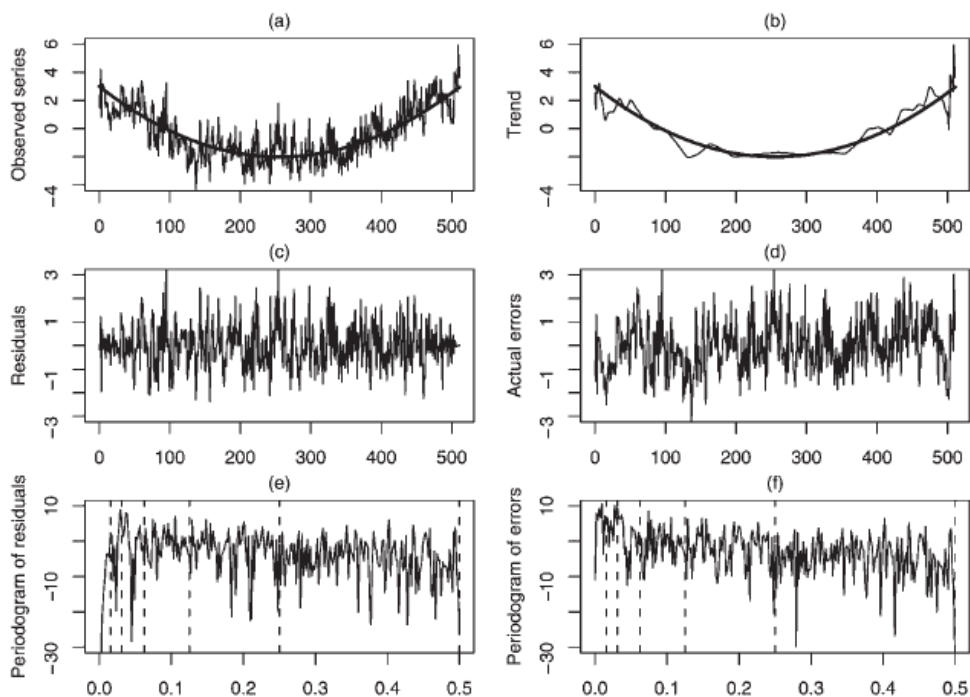


Figure 1. (a) Simulation of a quadratic trend plus an FD(0.4, 1) process ( $N = 512$ ) (jagged line) and the true trend (solid line); (b) the estimated (jagged line) and true trend (smooth line); (c) the estimated residuals; (d) the actual errors; (e) periodogram of the estimated residuals. The vertical dotted lines indicate the approximate bandpass frequencies,  $[2^{-(j+1)}, 2^j]$ , for each wavelet level  $j = 1, \dots, 5$ ; (f) periodogram of the actual realization of the error process

図 13: [2] より引用

## 2.3 数値計算

### 2.3.1 wavelet とは

今、分子動力学により求めた AK を構成する原子の座標の時系列データが 10000 ステップ (10[nsec]) 与えられている。その中から 炭素の座標データのみを抽出し、そのデータに対して wavelet 解析を行っている。座標データの wavelet 変換を行うと、速度ベクトルのデータが得られる。その得られたデータは、以下の模式図にもあるように、周波数順に並んでいて、レベルに分けることができる。レベルについての定義は次節に述べる。

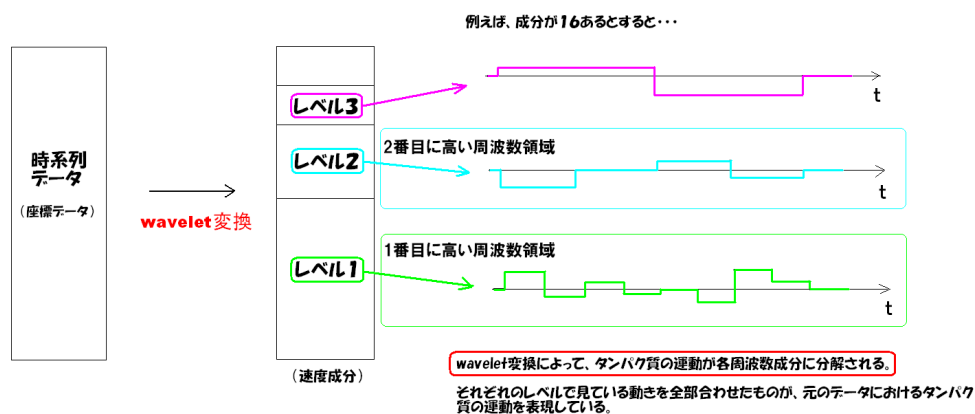


図 14: 計算方法

計算方法としては、low-pass filter と middle-pass filter を用いた。以下が、それらの模式図である。low-pass filter は主に trend 解析に用い、middle-pass filter は 4ake の集団運動性を調べる際になどに用いた。

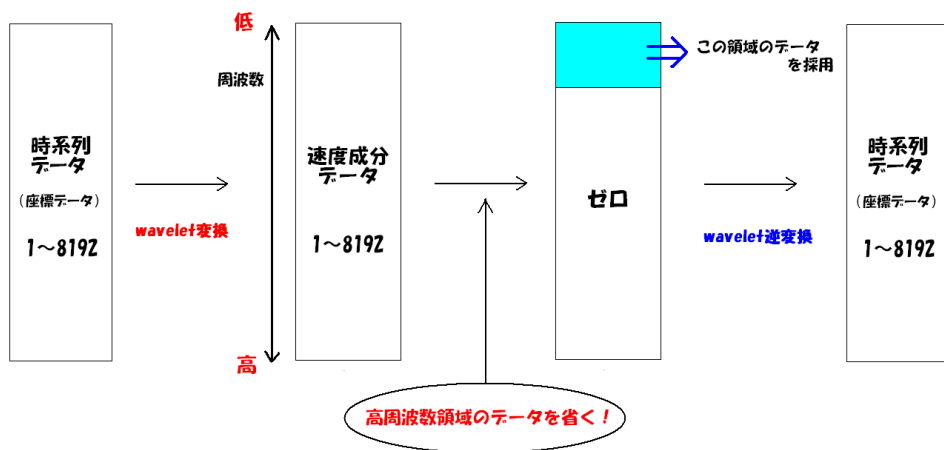


図 15: low-pass filter

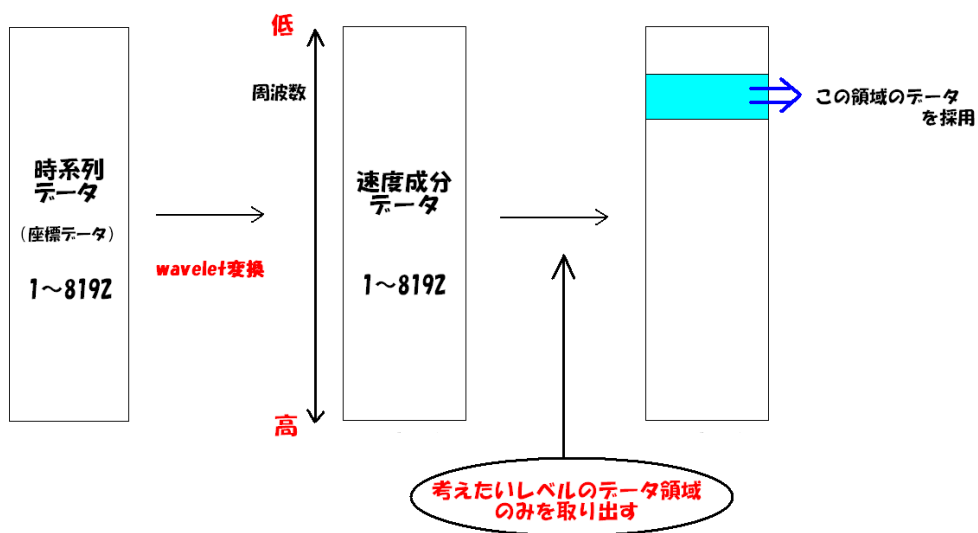


図 16: middle-pass filter

### 2.3.2 レベルとは

wavelet 変換によって、タンパク質の運動を各周波数成分に分解した際に、最も高い周波数領域をレベル1とし、2番目に高い周波数領域をレベル2というふうにして、レベルを定める。また、本研究では、『Numerical Recipes in C』で紹介されている wavelet のサブルーチンを用いているため、2の累乗のデータ数でなければならない。よって、10000ステップのデータのうち  $2^{12} = 8192$  ステップのデータを用いて、wavelet 変換を行うこととする。つまり、レベルとしては12個に分けることが出来るのだが、レベル7になるとデータの個数が単純計算で64個となり、周期境界条件も考え合わせると、統計量が求められないデータ数となってしまう。そのため、本研究では、レベル1~6まで考えることとする。次項に、レベル  $n(n = 1 \sim 6)$  ではどのデータ領域を用いるのかを図式化したものを載せる。ただし、これは middle-pass filter のものである。low-pass filter の場合は、middle-pass filter の上の部分、例えば、レベル1なら1~4096のデータを用い、レベル2なら1~2048までのデータを用いるものとする。

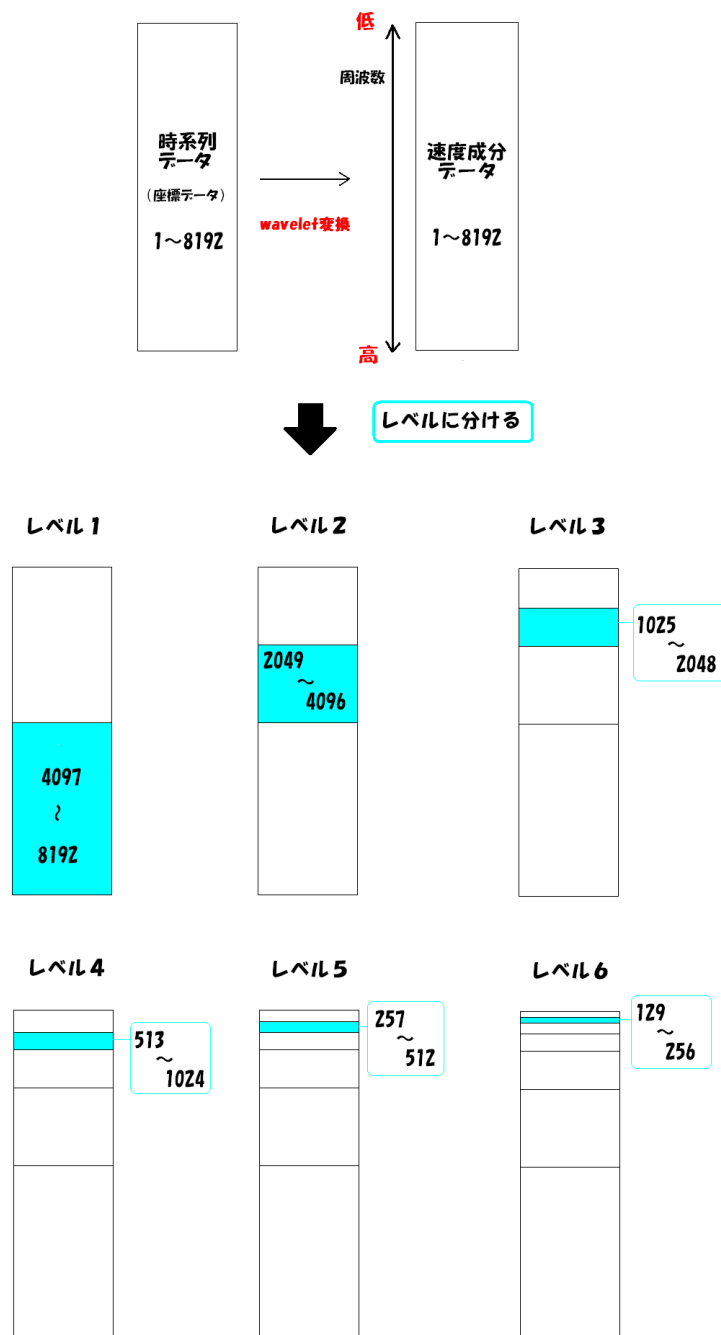


図 17: レベルの模式図：水色の部分がレベルごとに用いるデータの領域を示す。

### 2.3.3 データの扱い

先にも述べた通り、wavelet 解析を行うためには  $2^n$  の長さのデータでなければならないため、10000 ステップのデータを一度に wavelet 変換することが出来ない。そのため、以下のようにしてデータを扱うこととする。ただし、これは3章での方法論である。4章では、構造転移に注目したため、1~8192 ステップまでのデータのみを用いて解析を行っている。

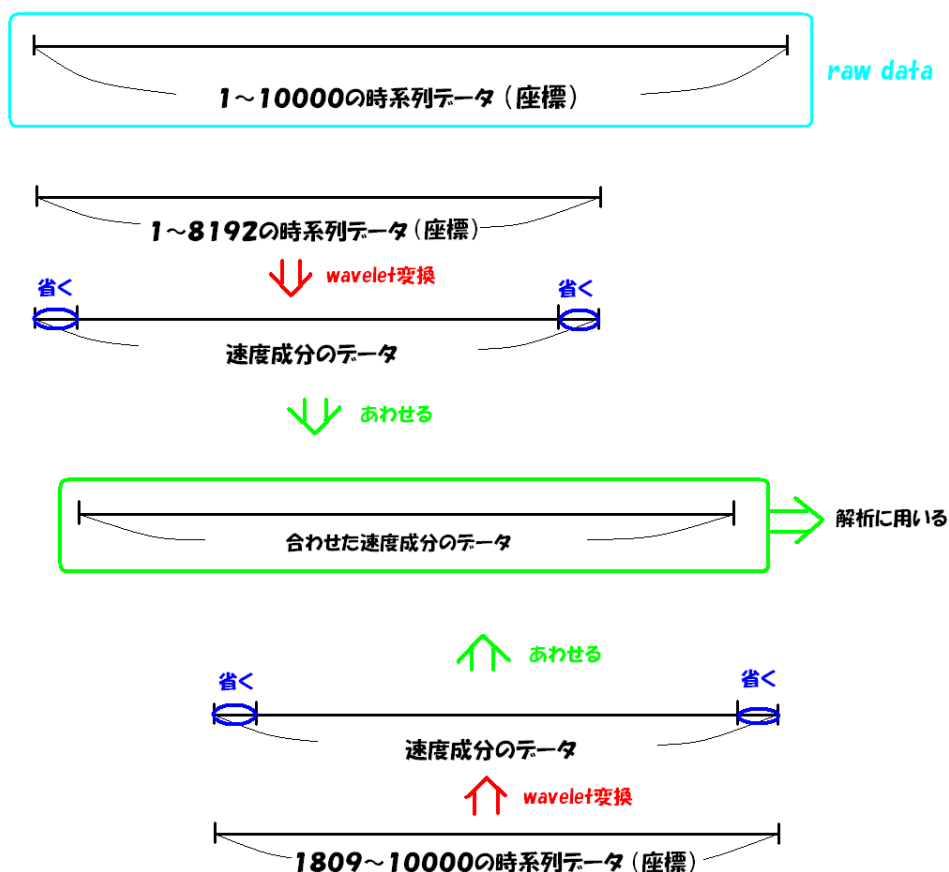


図 18: データの扱い

本計算では、周期境界条件により両端のデータを省かなければならない。省くデータの長さは、Daubechies のウェーブレットの波の長さに依存する。例えば、Daubechies の 20 番 [2] を用いているなら、両端において 20 個ずつのデータを省かなければいけないということである。3章では 20 番を用い、4章では 12 番を用いていることをここに記しておくこととする。

### 3 オープン型構造のAK

#### 3.1 慣性テンソル

以下の結果は、オープン型構造のAK (4ake) の慣性テンソルの固有値を求めたものである。これを見ると、2つの固有値の値が大きく、1つの固有値の値は小さくなっていることが読み取れる。これは、分子の形が細長くなっていることを意味する。また、1[nsec]程度の特徴的な時間スケールでは、非定常であることが確認でき、100[psec]程度の特徴的な時間スケールでは、定常的であることが確認できる。本章では、オープン型構造のAK (4ake) の100[psec]程度の時間スケールでのダイナミクスについて考えることとする。

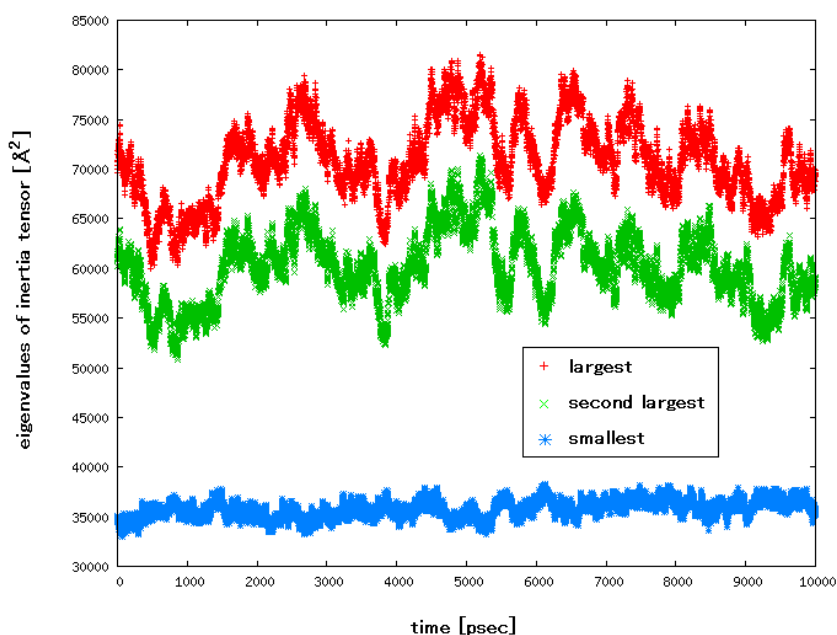


図 19: 4ake

#### 3.2 エネルギー

##### 3.2.1 wavelet を通した際のエネルギーのゆらぎ

レベルごとのエネルギーの時間発展

$$E(l, t) = \sum_{i=1}^{214} |\mathbf{v}_i(l, t)|^2 \tag{3.1}$$



を求めた。ここで、 $l$ はレベルの番号を表し、 $i$ はアミノ酸残基の炭素の番号を表す。また、ここでの wavelet 変換は DAUB20[2] を用いた。

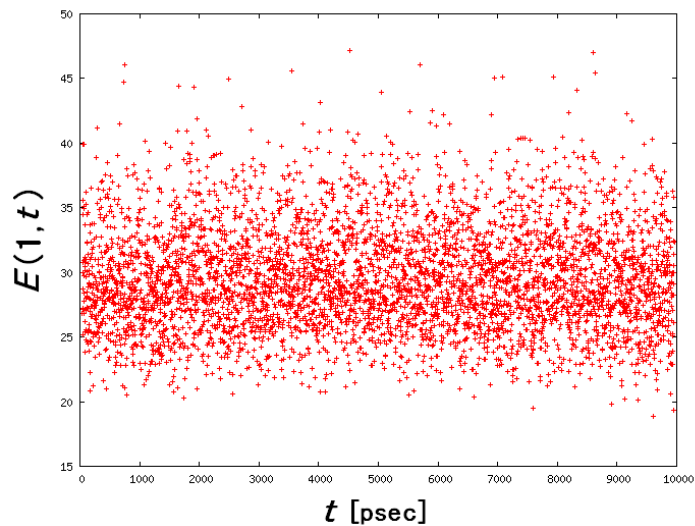


図 20: 4ake:レベル 1

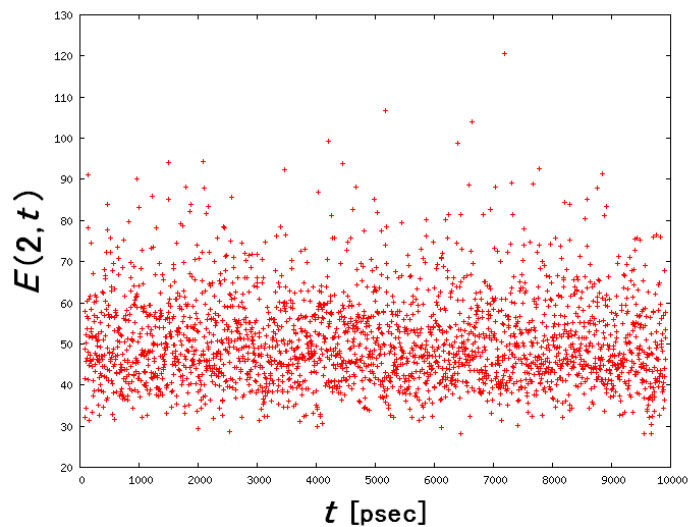


図 21: 4ake:レベル 2

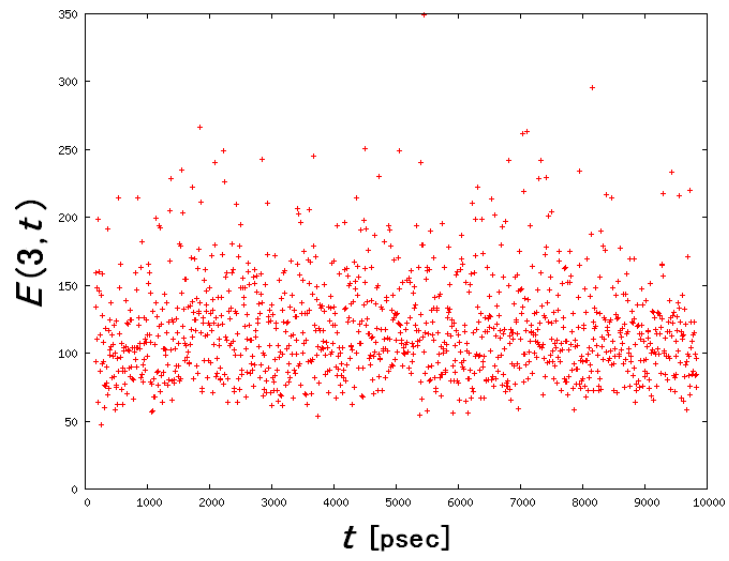


図 22: 4ake:レベル 3

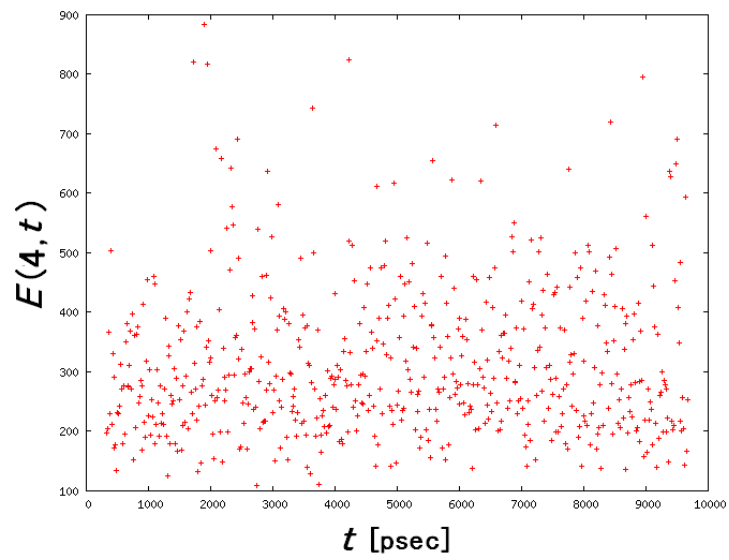


図 23: 4ake:レベル 4

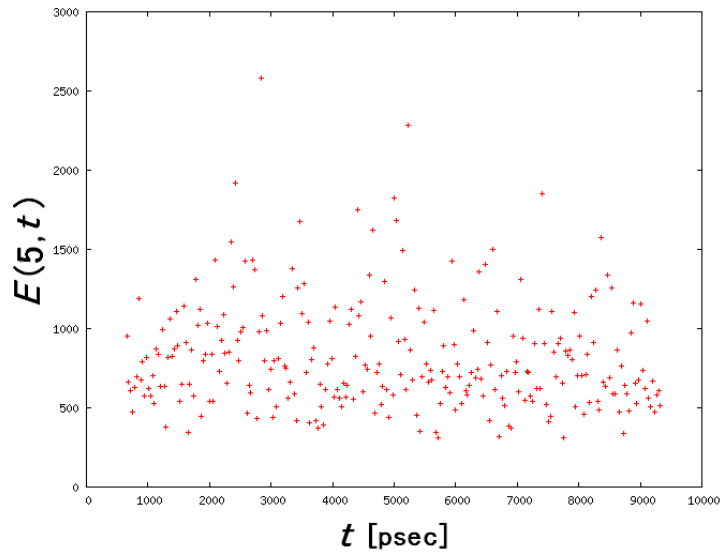


図 24: 4ake:レベル 5

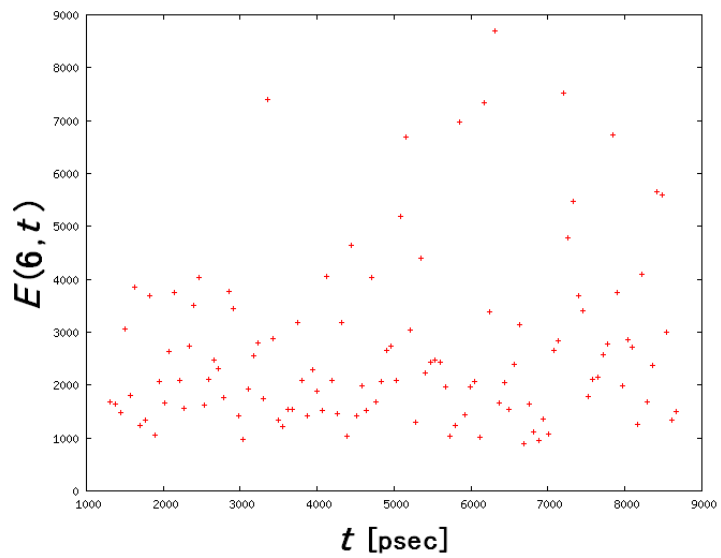


図 25: 4ake:レベル 6

これらの結果からレベルが上がるごとにエネルギーの値が上がっていることが分かる。このことについては次節で考えることとする。

### 3.2.2 レベルごとの平均エネルギー

前節で求めたエネルギーのレベルごとの時間平均

$$\langle E \rangle_t^{(l)} = \frac{1}{T} \sum_{t=1}^T E(l, t) \quad (3.2)$$

を求めた。ここで、 $T$  は考えている時間領域での時間の全ステップ数とする。

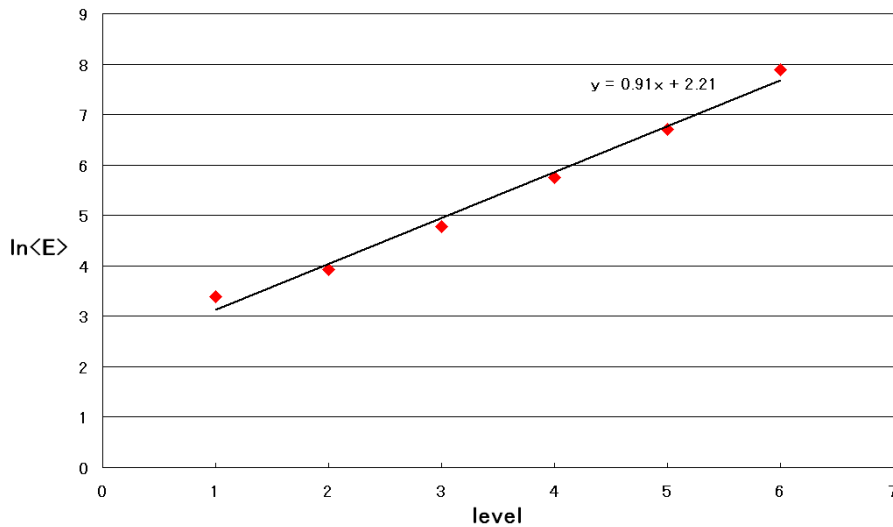


図 26: 平均エネルギー

前節の結果からレベルごとにエネルギーの値が上がっていることが分かったため、レベルごとにエネルギーの平均値の対数をとったものをプロットした。このグラフより、傾き (=  $\frac{\ln\langle E \rangle \text{の变化量}}{\text{レベルの变化量}}$ ) が約 0.9 であることが読み取れる。つまり、レベルが 1 上がるごとに、エネルギーが  $e^{0.9} \approx 2$  倍になっていることが分かる。これはエネルギー等分配則が成り立っていることを意味しているのではないかと考える。その理由は次節で述べることにしよう。

### 3.2.3 エネルギー等分配則

いま、1 成分ごとに  $m\omega^2 x^2(\omega) \propto kT$  のエネルギーが与えられていると考える。つまり、

$$x^2(\omega) \sim \frac{kT}{m\omega^2} \quad (3.3)$$

となる。また、ここで、レベルが 1 つ上がるごとに周波数が 0.5 倍になり、以下の模式図に示すようにレベルが 1 つ上がるごとに成分の自由度が半分になる。

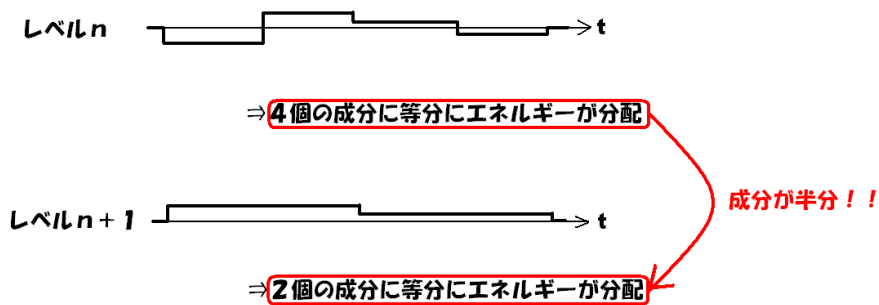


図 27: レベルが 1 つ上がると成分の自由度が半分になる理由

レベルが 1 つ上がると成分の個数が半分になるため、成分の自由度が半分になるのである。これらのことからレベル  $n$  のエネルギーを  $E_n$ 、レベル  $n+1$  のエネルギーを  $E_{n+1}$  とするとそれぞれ以下のように求められる。

・レベル  $n$  のとき (成分の自由度  $N$ )  $E_n = Nx_n^2(\omega) \sim \frac{NkT}{m\omega^2}$

・レベル  $n+1$  のとき (成分の自由度  $\frac{N}{2}$ )  $E_{n+1} = \frac{1}{2}Nx_{n+1}^2(\omega) \sim \frac{NkT}{2m(\frac{\omega}{2})^2} = \frac{2NkT}{m\omega^2}$

このことにより、 $E_{n+1}$  は  $E_n$  の 2 倍になっていることが分かる。つまり、エネルギー等分配則が成り立っているといえる。

### 3.2.4 エネルギーの累積分布

先ほど求めたエネルギー

$$E(l, t) = \sum_{i=1}^{214} |\mathbf{v}_i(l, t)|^2 \quad (3.4)$$

の結果を用いてエネルギーの累積分布を求めた。ここで、

$$\langle E \rangle_t^{(l)} = \frac{1}{T} \sum_{t=1}^T E(l, t) \quad (3.5)$$

$$\langle E^2 \rangle_t^{(l)} = \frac{1}{T} \sum_{t=1}^T E^2(l, t) \quad (3.6)$$

$$\sigma^2 = \langle E^2 \rangle_t^{(l)} - (\langle E \rangle_t^{(l)})^2 \quad (3.7)$$

とする。ここで、 $T$  は考えている時間領域での時間の全ステップ数とする。

また、 $P(X)$  を  $X$  が起こる確率とし、

$$\text{Prob}(X > E) = \int_E^{\infty} P(X) dX \quad (3.8)$$

を求める。

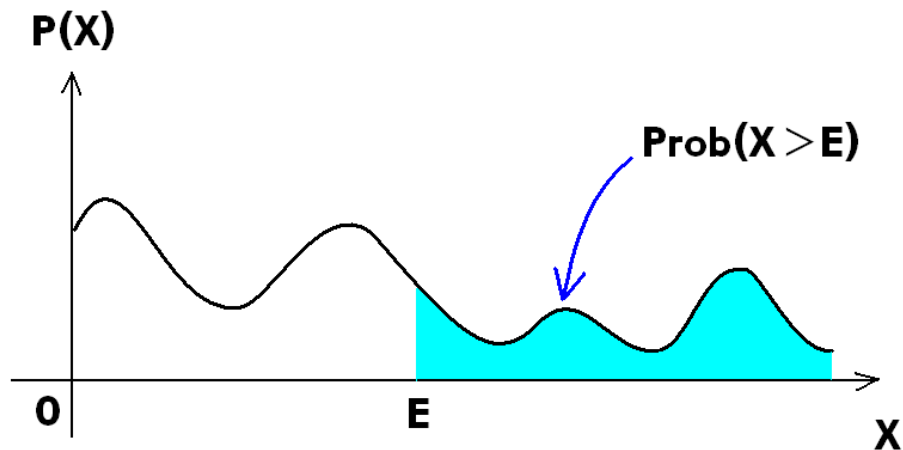


図 28: 累積分布の模式図

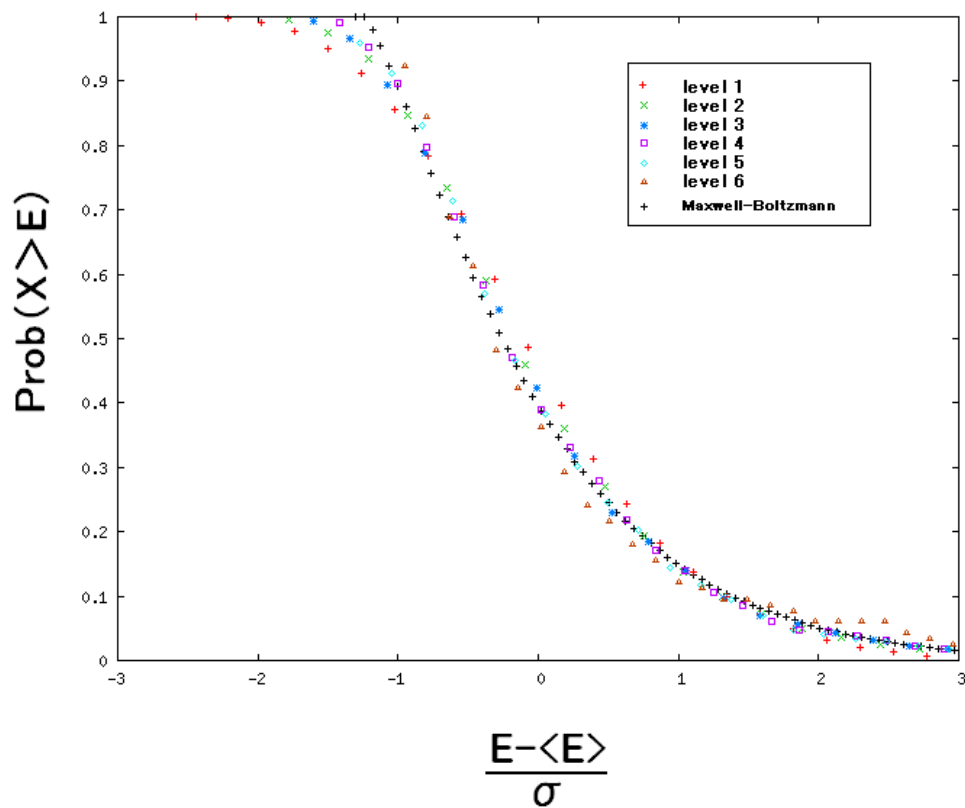


図 29: 4ake のエネルギーの累積分布

レベル1~6について求めたが、どのレベルにおいても Maxwell-Boltzmann 分布に類似する分布が得られた。このことは、レベル1~6の場合、wavelet 変換は AK の定常的な動きを取り出しているということを意味する。

### 3.3 集団運動性

wavelet 変換を行うと AK の集団運動を取り出すことが出来る。本研究では、2 つの指標を用いてこのことを確認した。以下にその結果を記す。

#### 3.3.1 指標 1 ( $C_{\text{ang}}$ )

AK の集団運動性について、次の指標を用いて考える。

$$C_{\text{ang}}(l, t) = \sum_{i < j, |\mathbf{r}_i - \mathbf{r}_j| < d} \frac{\mathbf{v}_i(l, t) \cdot \mathbf{v}_j(l, t)}{|\mathbf{v}_i(l, t)| |\mathbf{v}_j(l, t)|} = \sum_{i < j, |x_i - x_j| < d} \cos \theta_{i, j}(l, t) \quad (3.9)$$

この内積は、 $i$  番目の炭素原子とその原子を中心とした半径  $d$  [ ] の円内にある全ての炭素原子との速度ベクトルの内積の和を求めたものである。ここで、炭素間の平均距離が約 3.8 [ ] であることから、 $d = 10.0$  [ ] とした。以下の模式図は、黄緑色の丸が炭素を表し、黒い線は炭素の結合の手を表す。次項より  $C_{\text{ang}}(l, t)$  の結果を載せる。

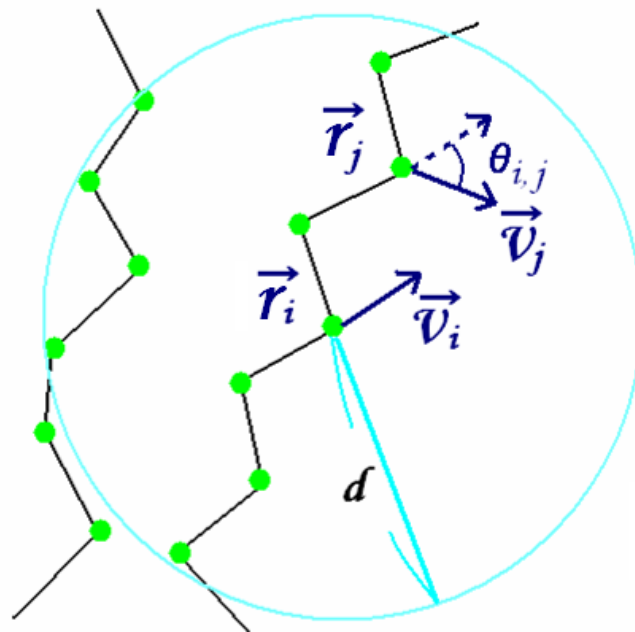
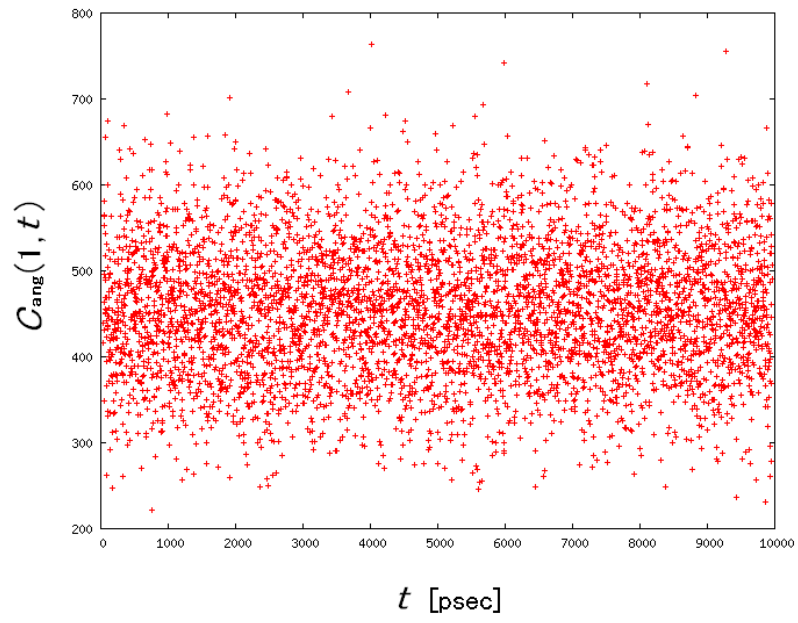
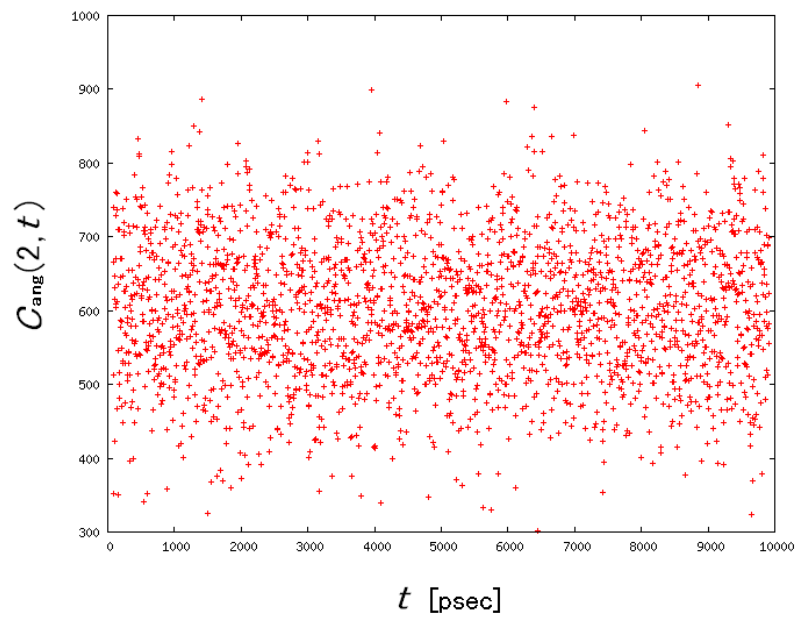


図 30: 内積の求め方の模式図

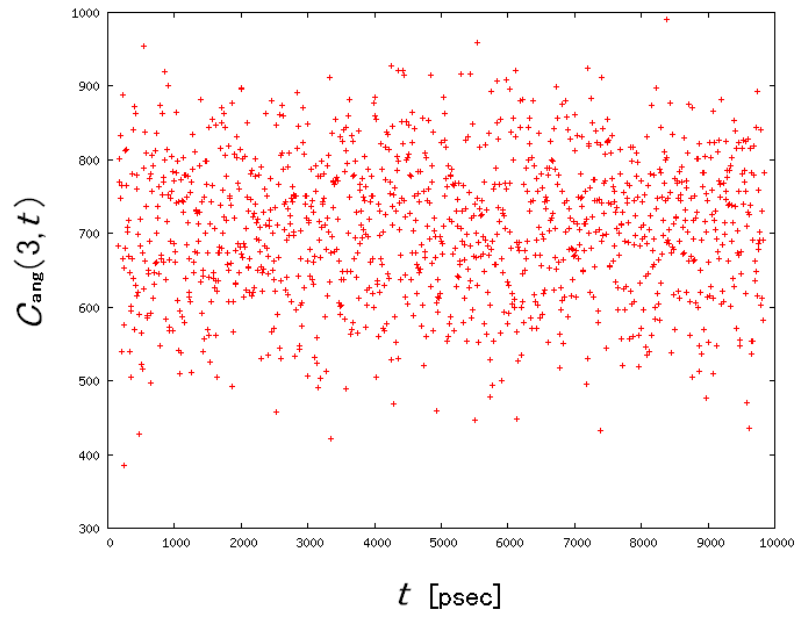




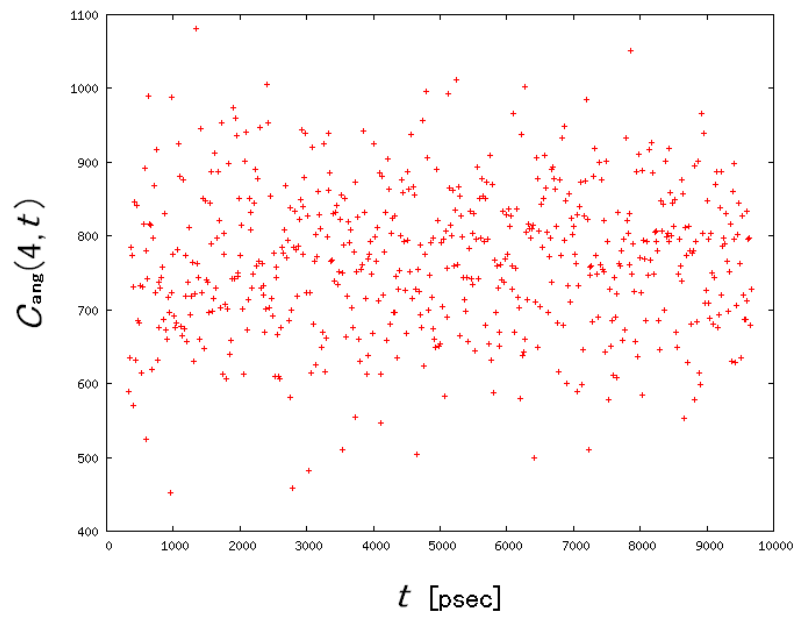
☒ 31: 4ake: $C_{ang}$  ( level 1 )



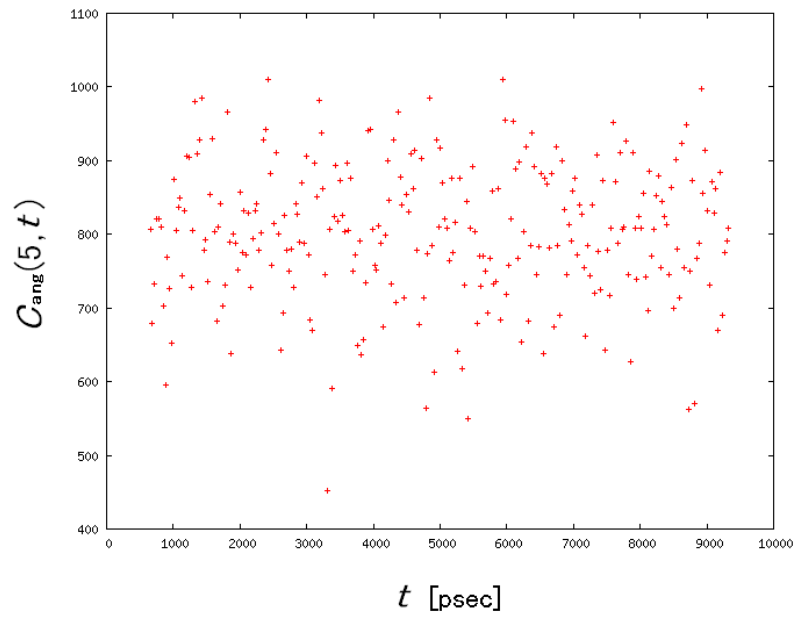
☒ 32: 4ake: $C_{ang}$  ( level 2 )



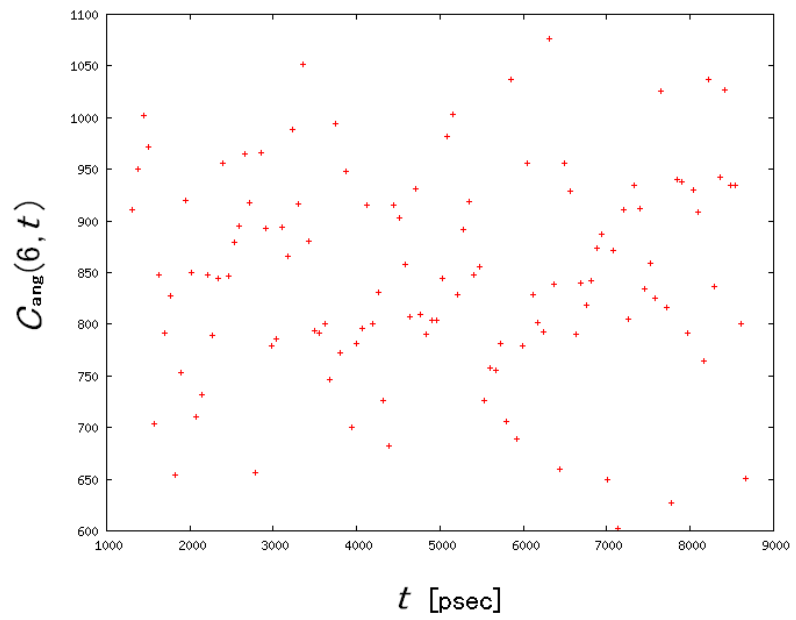
☒ 33: 4ake: $C_{\text{ang}}$  ( level 3 )



☒ 34: 4ake: $C_{\text{ang}}$  ( level 4 )



☒ 35: 4ake: $C_{ang}$  ( level 5 )



☒ 36: 4ake: $C_{ang}$  ( level 6 )

以上の結果より、レベルが上がると共に  $C_{\text{ang}}(l, t)$  の値が上がっていることが分かる。  
次に、レベルごとの  $C_{\text{ang}}(l, t)$  の時間平均

$$\langle C_{\text{ang}} \rangle_t^{(l)} = \frac{1}{T} \sum_{t=1}^T C_{\text{ang}}(l, t) \tag{3.10}$$

を求めたものを示す。ここで、 $T$  は考えている時間領域の時間の全ステップ数とする。また、 $\langle C_{\text{ang}} \rangle_t^{(l)}$  は  $C_{\text{ang}}(l, t)$  の時間平均とする。

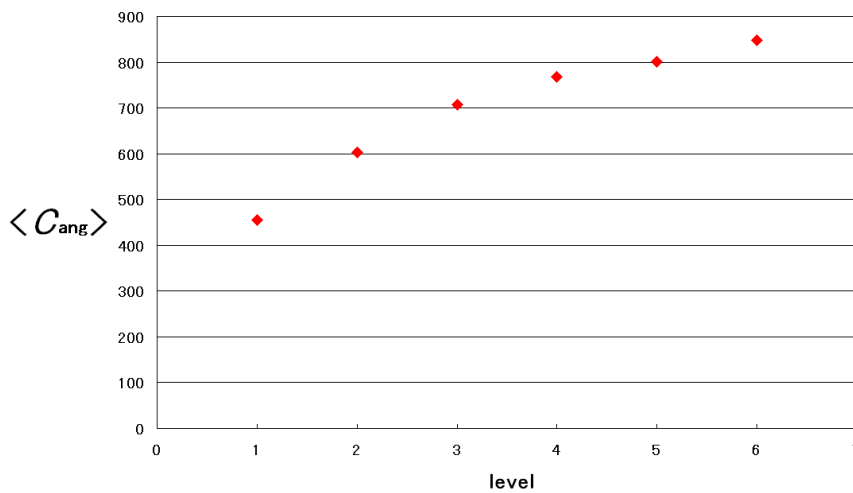


図 37: 4ake:レベルごとの  $C_{\text{ang}}$  の時間平均

この結果より、先程も述べたが、レベルがあがるごとに  $C_{\text{ang}}$  が増えていることが分かる。 $C_{\text{ang}}$  の値が上がっているということは、低周波数領域では原子同士がより同じ方向に動いているということである。つまり、低周波数領域になると、集団運動性が強いということがいえる。このことより、wavelet 変換により AK 内の集団運動が取り出せるということがいえるのではないだろうか。ただし、この  $C_{\text{ang}}$  という指標は、ベクトルの長さを考慮に入れていないため、図 38 の (b) のような運動が頻繁に起こっている場合には不都合が生じるのではないかと考える。何故ならば、図 38 の (a) のように同じ大きさのベクトルがお互いに動きを打ち消しあっている場合は集団運動を妨げる効果があるように考えられるが、図 38 の (b) の  $j$  番目のベクトルのように大きさが小さいベクトルに対しては集団運動を妨げる効果が少ないように考えられるからである。

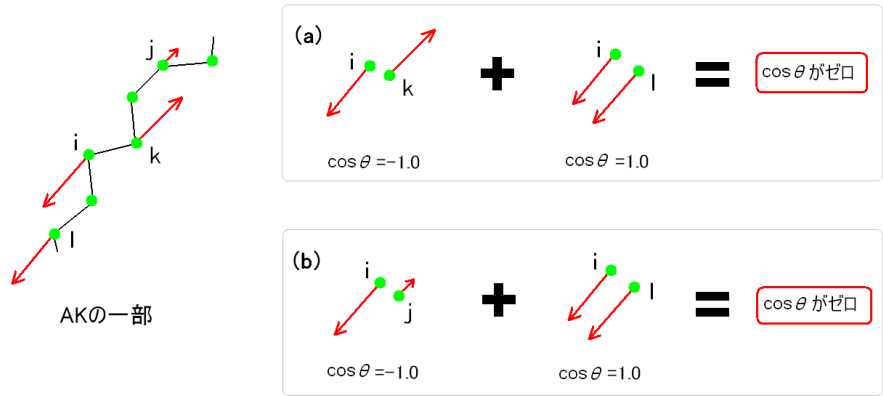


図 38: 模式図

したがって、次節にベクトルの大きさを考慮に入れた指標を用いての結果を示すこととする。

### 3.3.2 指標 2 ( $C_{vel}$ )

集団運動性を見る一つの指標として、

$$C_{vel}(l, t) = \frac{\sum_{i < j, |\mathbf{r}_i - \mathbf{r}_j| < d} \mathbf{v}_i(l, t) \cdot \mathbf{v}_j(l, t)}{\sum_i |\mathbf{v}_i(l, t)|^2} \quad (i, j = 1, 2, \dots, 214) \quad (3.11)$$

を用いた。ここで、 $d = 10.0[ ]$ である。この指標を用いた結果を次項より載せる。

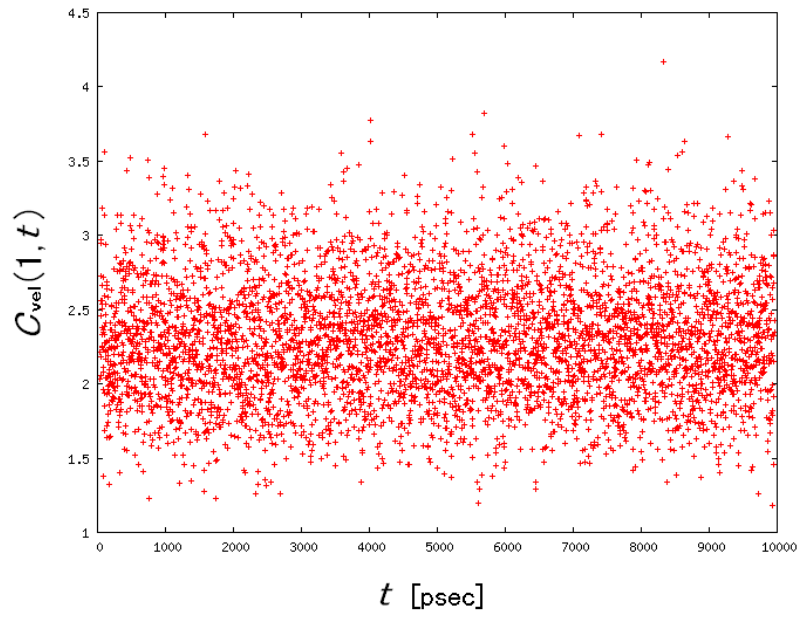


図 39: 4ake: $C_{\text{vel}}$  (レベル 1)

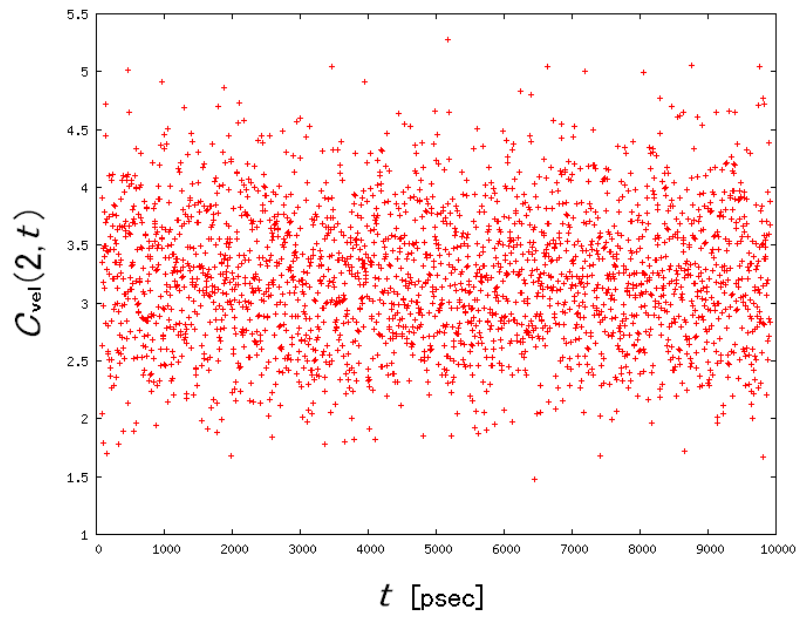


図 40: 4ake: $C_{\text{vel}}$  (レベル 2)

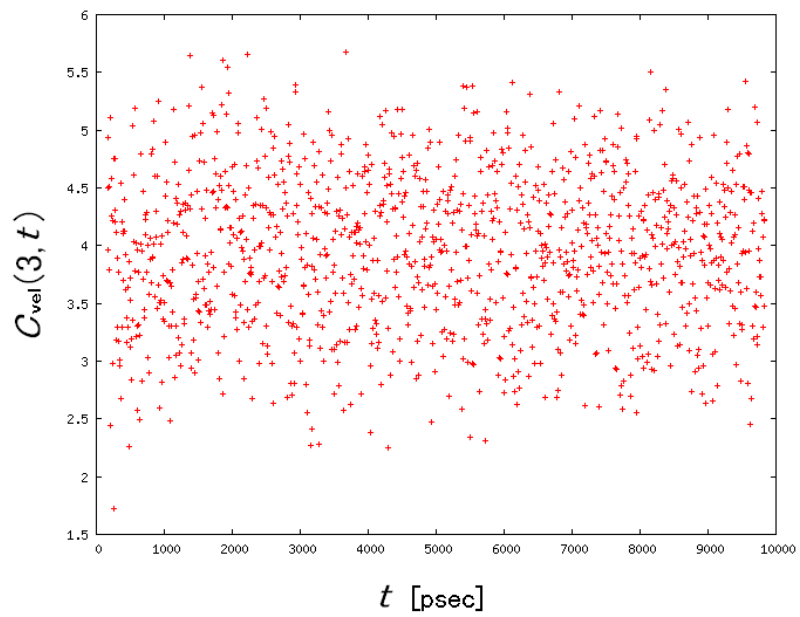


図 41: 4ake: $C_{\text{vel}}$  ( レベル 3 )

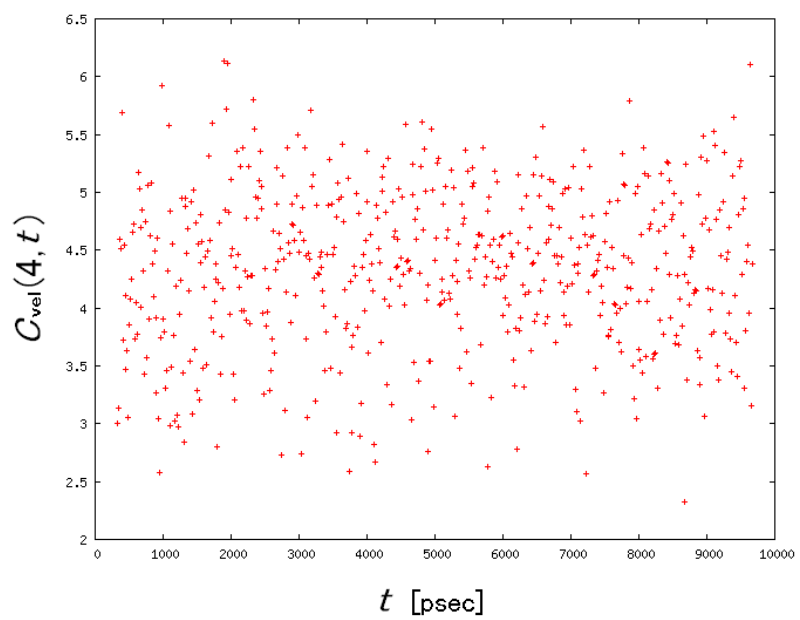


図 42: 4ake: $C_{\text{vel}}$  ( レベル 4 )

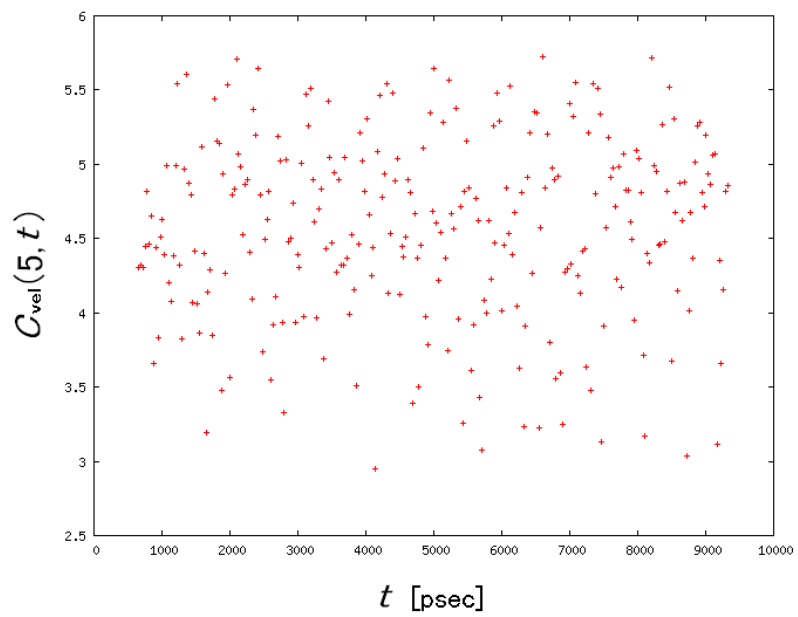


図 43: 4ake: $C_{\text{vel}}$  ( レベル 5 )

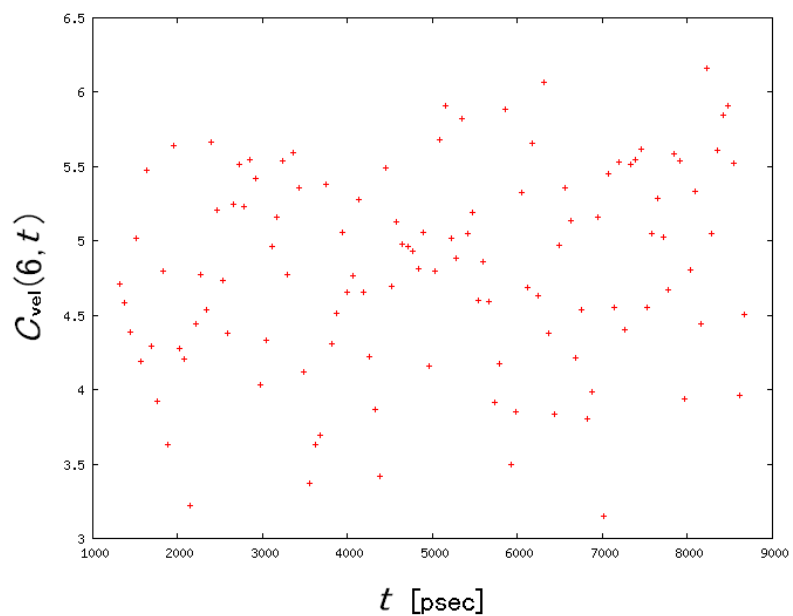


図 44: 4ake: $C_{\text{vel}}$  ( レベル 6 )

以上の結果より、レベルが上がるごとに  $C_{\text{vel}}(l, t)$  の値が上がっていることが分かる。



次に、レベルごとの  $C_{\text{vel}}(l, t)$  の時間平均

$$\langle C_{\text{vel}} \rangle_t^{(l)} = \frac{1}{T} \sum_{t=1}^T C_{\text{vel}}(l, t) \quad (3.12)$$

を求めたものを示す。ここで、 $T$  は考えている時間領域での全ステップ数とする。

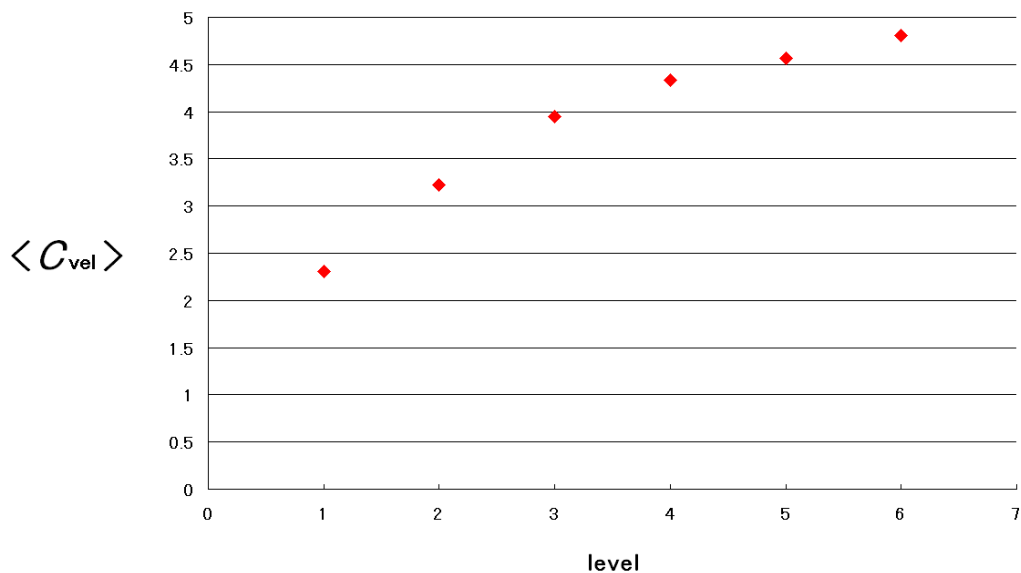


図 45: 4ake: レベルごとの  $C_{\text{vel}}$  の時間平均

この結果より、先ほども述べたが、レベルがあがるごとに  $C_{\text{vel}}$  の値が増えていることがわかる。 $C_{\text{ang}}$  のときと同様に、 $C_{\text{vel}}$  の値が上がっているということは、低周波数領域では原子同士がより同じ方向に動いているということである。つまり、低周波数領域になると、集団運動性が強いということがいえる。前節の結果と考え合わせても、wavelet 変換により AK 内の集団運動が取り出せるということがいえるだろう。

### 3.4 実際の集団運動

前節では、AKの集団運動性を2つの指標を用いてレベル1から6の場合について見た。その結果、それぞれのレベルにおける集団運動性の指標の時間発展は揺らいでおり、指標の時間平均を考えるとレベルが上がるごとに集団運動性が強くなるということが確認できた。では、実際にAKはどのような集団運動をしているのだろうか。ここでは、2つの場所相関の指標  $C_{i,j}^{\text{ang}}(l,t)$ 、 $C_{i,j}^{\text{vel}}(l,t)$  を用いて実際のAKの集団運動について調べた。レベル6について調べた結果、主に4種類の集団運動を見ることが出来た。次項よりその結果を載せる。

ここで、場所相関の指標  $C_{i,j}^{\text{ang}}(l,t)$  を

$$C_{i,j}^{\text{ang}}(l,t) = \frac{\mathbf{v}_i(l,t) \cdot \mathbf{v}_j(l,t)}{|\mathbf{v}_i(l,t)| |\mathbf{v}_j(l,t)|} \quad (i, j = 1, 2, \dots, 214) \quad (3.13)$$

と定義し、 $C_{i,j}^{\text{vel}}(l,t)$  を

$$C_{i,j}^{\text{vel}}(l,t) = \frac{\mathbf{v}_i(l,t) \cdot \mathbf{v}_j(l,t)}{\frac{1}{T} \sum_{t=1}^T \frac{1}{N} \sum_{i=1}^N |\mathbf{v}_i(l,t)|^2} \quad (i, j = 1, 2, \dots, 214 \quad , N = 214) \quad (3.14)$$

と定義する。ここで、 $T$  は考えている時間領域の全ステップ数とする。 $C_{i,j}^{\text{ang}}(l,t)$  が、

$$C_{i,j}^{\text{ang}}(l,t) > 0.7$$

の条件を満たすときに、点  $(i, j)$  を赤色でプロットし、

$$C_{i,j}^{\text{ang}}(l,t) < -0.7$$

の条件を満たすときに、点  $(i, j)$  を青色でプロットする。

また、 $C_{i,j}^{\text{vel}}(l,t)$  が、

$$C_{i,j}^{\text{vel}}(l,t) > 1.0$$

の条件を満たすときに、点  $(i, j)$  を赤色でプロットし、

$$C_{i,j}^{\text{vel}}(l,t) < -1.0$$

の条件を満たすときに、点  $(i, j)$  を青色でプロットする。

ここで、赤色の部分は正の相関があることを意味し、青色の部分は負の相関があることを意味する。また、その求めた場所相関の時間でのAKの3次元空間内での動きも併せて載せる。

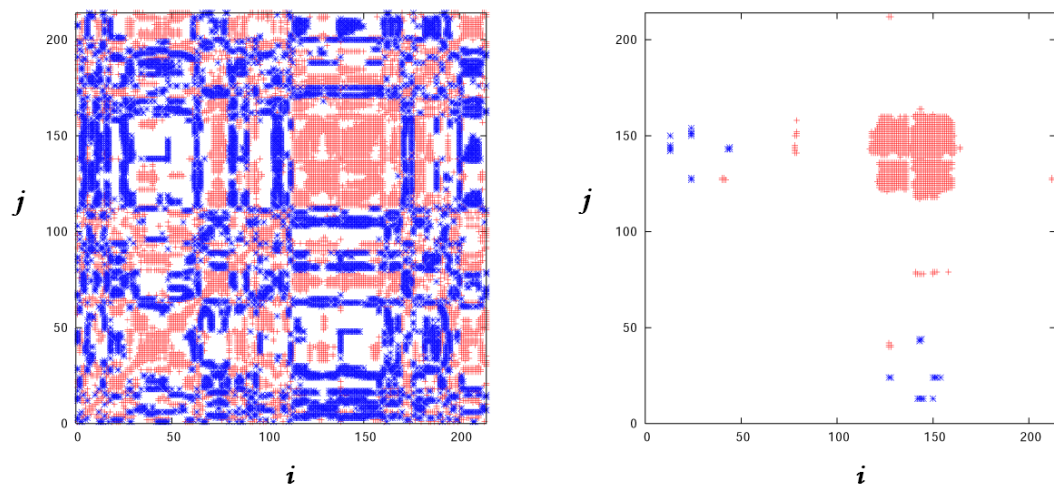


図 46: 4ake: (左図)  $C_{i,j}^{ang}$ : (右図)  $C_{i,j}^{vel}$  (time=2816[psec])

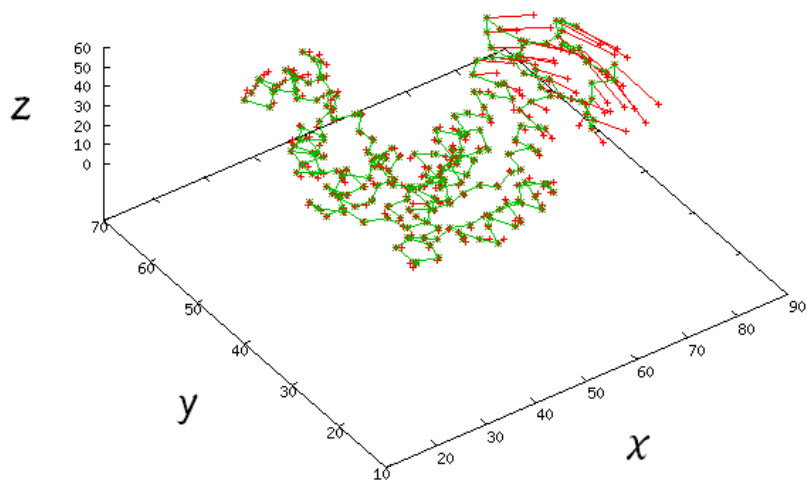


図 47: 4ake:AK の動き (time=2816[psec])

・ 140 番目付近の原子が同一方向を向いて運動している場合。

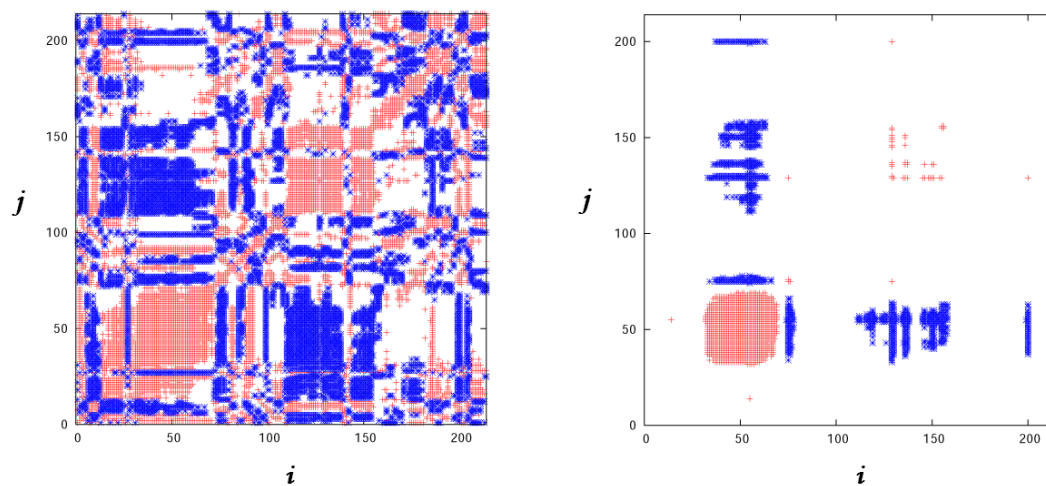


図 48: 4ake: (左図)  $C_{i,j}^{\text{ang}}$ : (右図)  $C_{i,j}^{\text{vel}}$  (time=4800[psec])

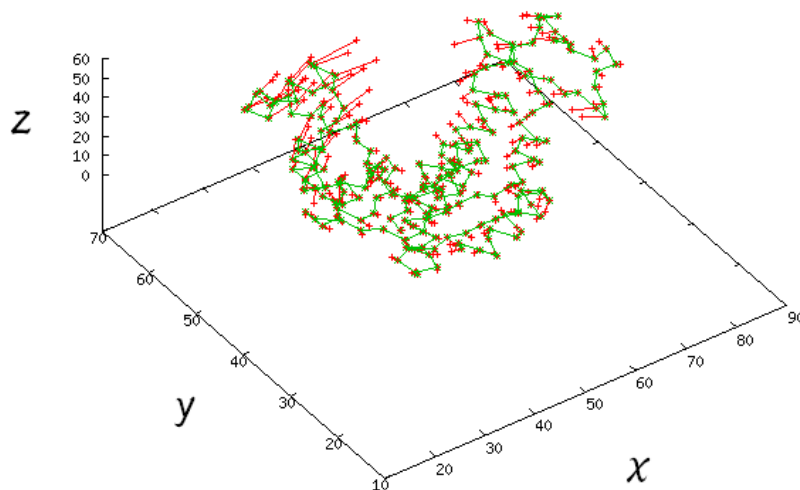


図 49: 4ake:AK の動き (time=4800[psec])

・ 50 番目付近の原子が同一方向を向いて運動している場合。

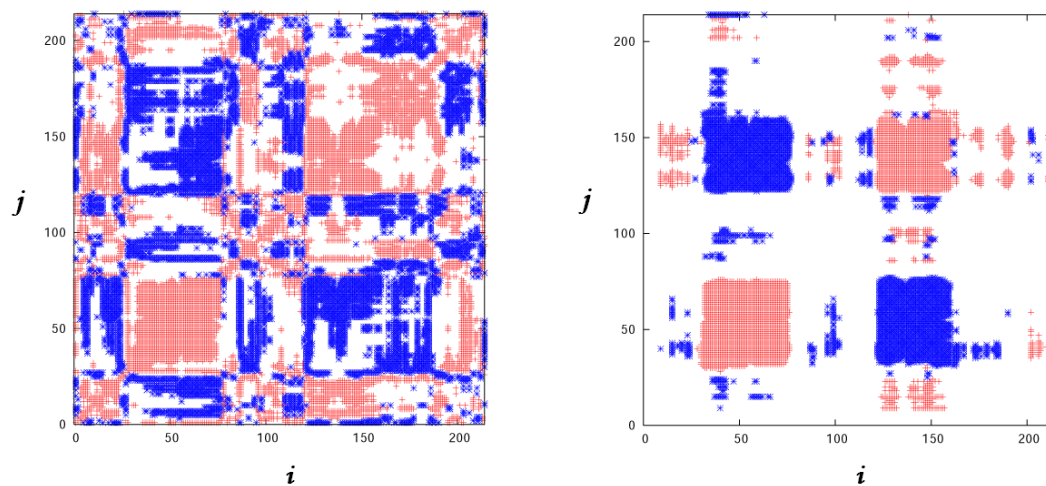


図 50: 4ake: (左図)  $C_{i,j}^{ang}$ : (右図)  $C_{i,j}^{vel}$  (time=7360[psec])

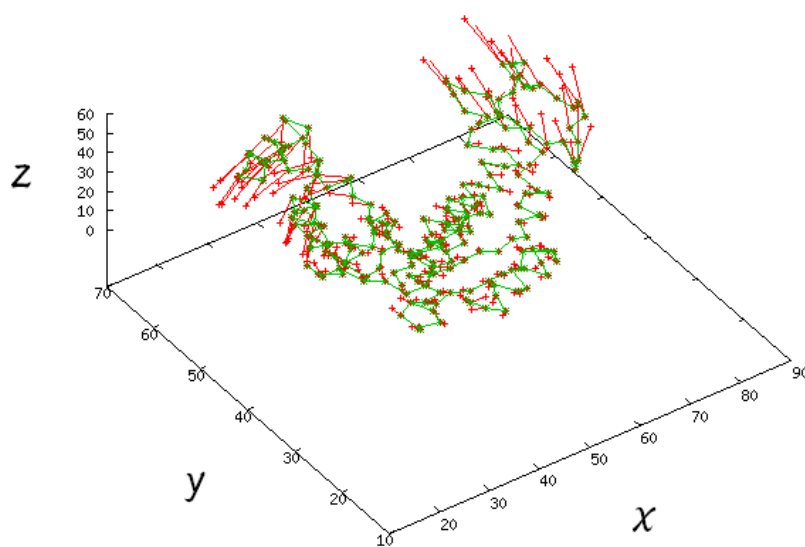


図 51: 4ake:AK の動き (time=7360[psec])

- 140 番目付近の原子と 50 番目付近の原子が反対方向を向いて運動している場合。

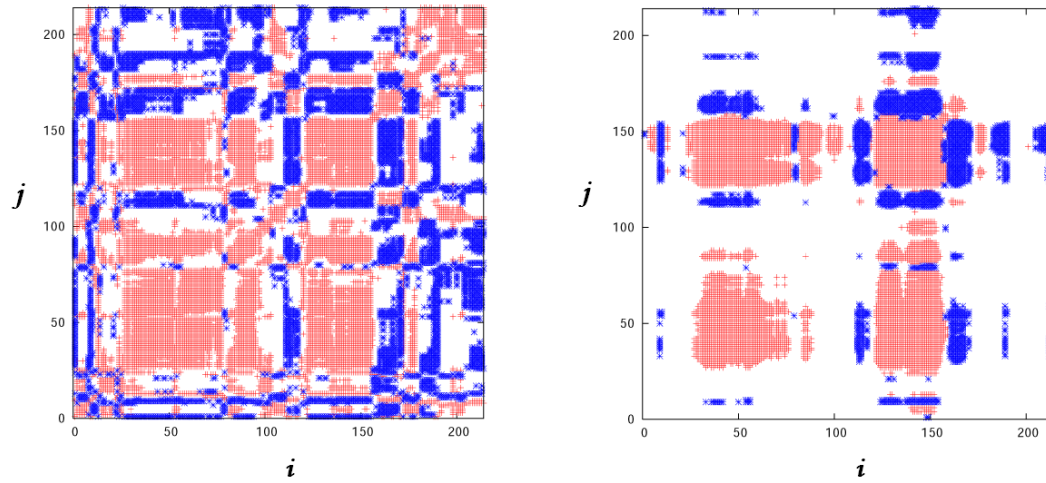


図 52: 4ake: (左図)  $C_{i,j}^{\text{ang}}$ : (右図)  $C_{i,j}^{\text{vel}}$  (time=3392[psec])

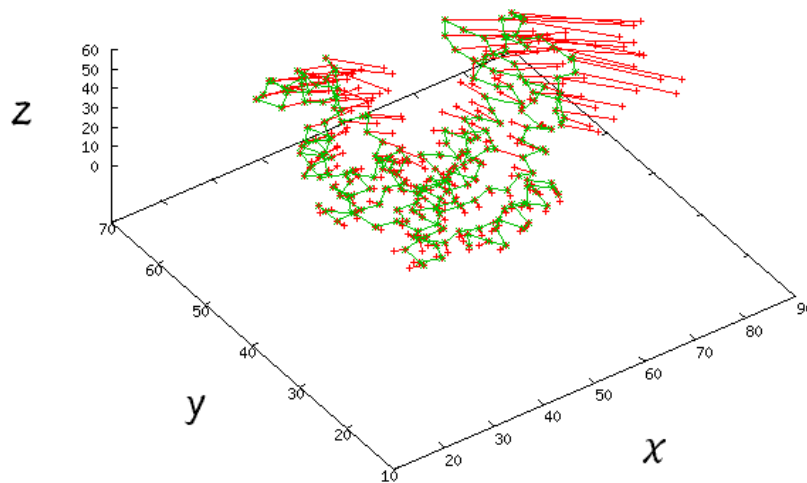


図 53: 4ake:AK の動き (time=3392[psec])

- ・ 140 番目付近の原子と 50 番目付近の原子が同一方向を向いて運動している場合。

以上の結果により、主に 50 番目付近の原子と 140 番目付近の原子が集団運動を形成していることがわかった。また、大きく分けて次の 4 つの場合があることもわかった。

- 140 番目付近の原子が同一方向を向いて運動している場合。
- 50 番目付近の原子が同一方向を向いて運動している場合。
- 140 番目付近の原子と 50 番目付近の原子が反対方向を向いて運動している場合。
- 140 番目付近の原子と 50 番目付近の原子が同一方向を向いて運動している場合。

以上の結果は、レベル 6 について得られたものである。レベル 6 で集団運動が起こっている際に、他のレベルではどのような運動が起こっているのだろうか。ここでは、前節で見た時刻におけるレベル 1 での場所相関を見ることにする。

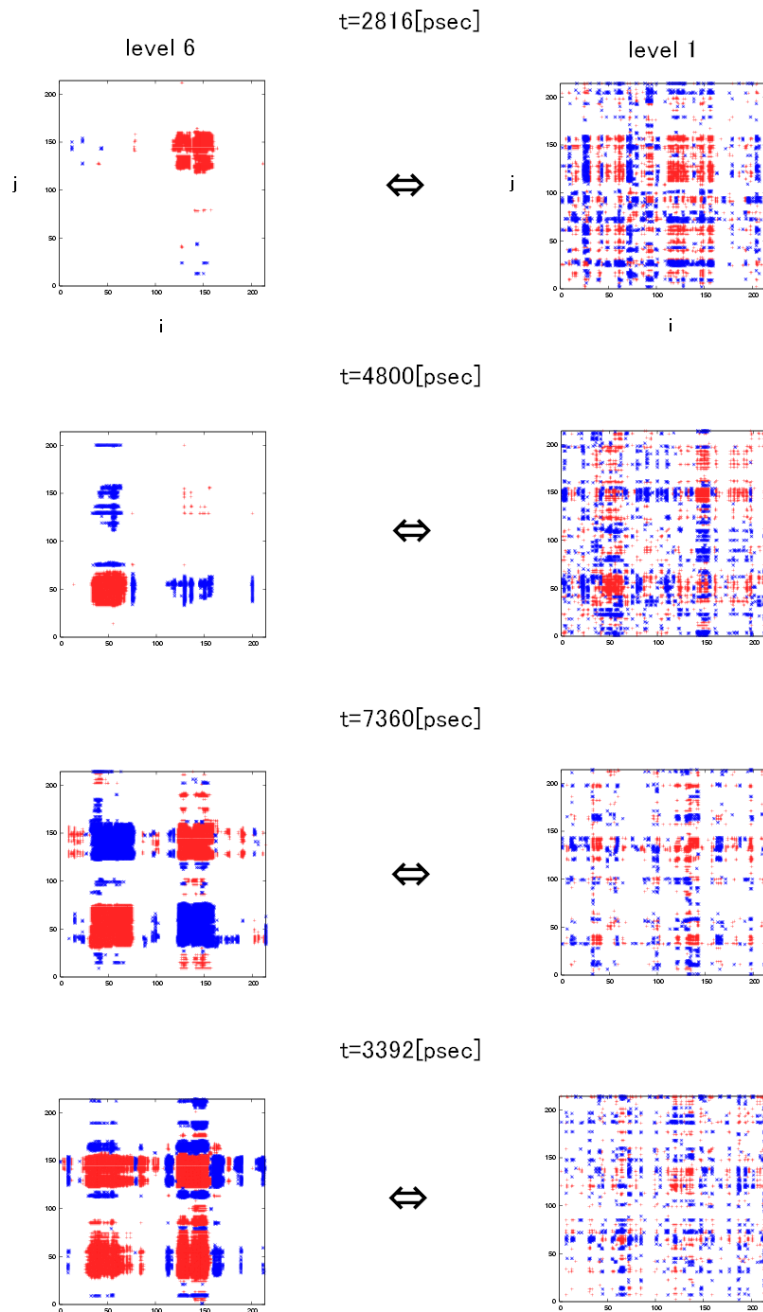


図 54: レベル 1 とレベル 6 の場所相関の違い

レベル 6 では集団運動が見られるのに対し、レベル 1 では集団運動を確認することが出来ない。



### 3.5 まとめ

本章では、オープン型構造の AK (4ake) の dynamics について調べた。以上の結果についてまとめる。

#### 1. エネルギーについて

- (a) エネルギー等分配則により、低周波領域になるほどエネルギーが高くなっていることが確認できた。
- (b) wavelet 変換を行うことによって、熱平衡的な動きが取り出せていることが確認できた。

#### 2. 集団運動について

- (a) 2つの指標  $C_{ang}$  と  $C_{vel}$  について調べた結果、wavelet 変換を用いると、集団運動が取り出せることが分かった。
- (b) 低周波領域になると、集団運動性が高くなることが確認できた。

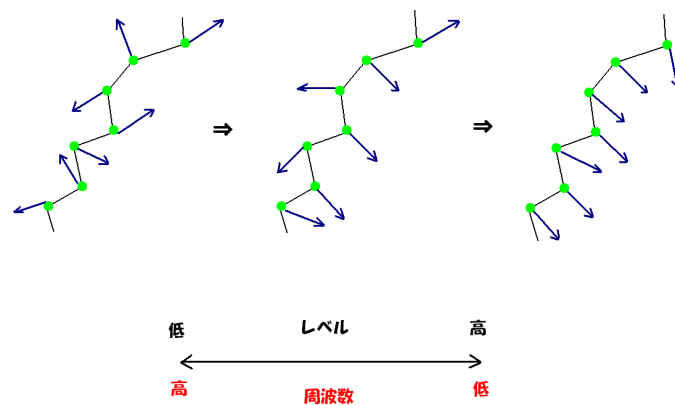


図 55: 低周波領域になると集団運動性が強くなることの模式図

- (c) 実際の集団運動を調べてみた結果、4種類の集団運動が存在することが確認できた。
- (d) レベル1とレベル6の場所相関を比較した結果、レベルが低くなると熱平衡的であるということが確認できた。

## 4 AKの構造転移

本章では、AK がクローズ型構造から (1ake) からオープン型構造 (4ake) に構造転移するプロセスについて調べた結果を載せる。解析には、2つの trajectory の時系列データ、trj01 と trj03 を用いた。trj01 と trj03 は、1ake の 10000 ステップ (= 10[nsec]) の時系列データである。trj01 は、クローズ型構造からオープン型構造へのプロセスに時間がかかる trajectory である。一方、trj03 はクローズ型構造からオープン型構造へのプロセスがスムーズに起こる trajectory である。

まずはじめに、trj01 と trj03 のそれぞれについて慣性テンソルの固有値を求めて、以上のことを確認した。次節にその結果を載せる。

### 4.1 慣性テンソル

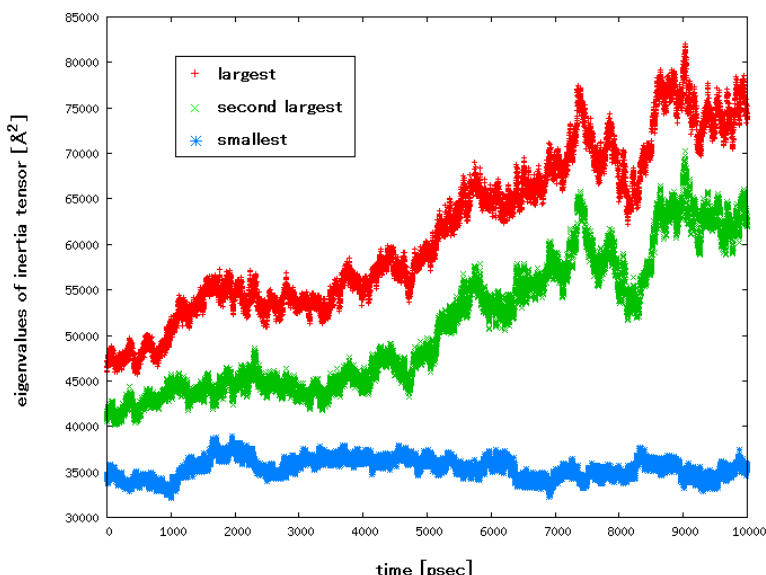


図 56: 1ake trj01

4ake の場合と同様に、2つの慣性テンソルの固有値の値が大きく、1つの固有値の値が小さいので、分子の形が細長くなっていることが分かる。また、2000[psec] から 4000[psec] までの間に2つの固有値の値が停滞し、4000[psec] から 6000[psec] の間に2つの固有値の値が大幅に増加していることが読み取れる。このことは、この間に AK がクローズ型構造からオープン型構造へ転移していることを示唆している。

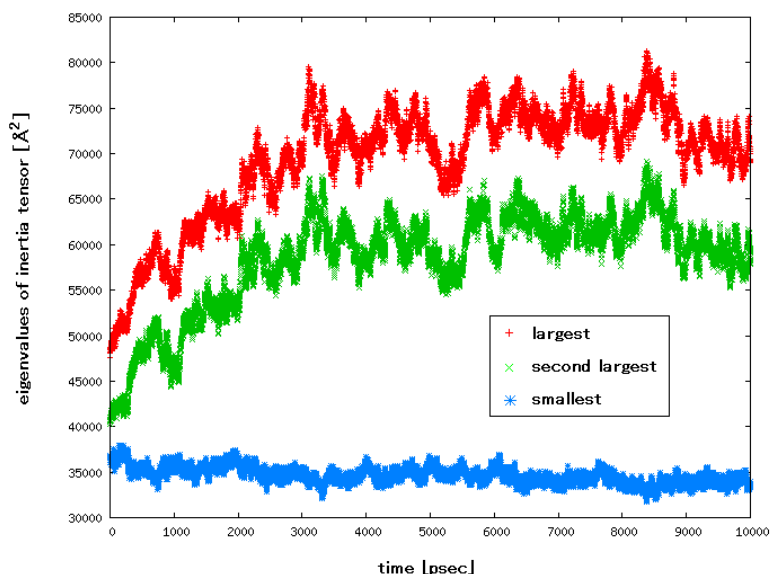


図 57: lake trj03

0[psec] から 2000[psec] までの間に 2 つの固有値の値が大幅に増加している。これは、この間に AK がクローズ構造からオープン構造へ転移していることを示唆している。

以上のことから、慣性テンソルを求めることにより、trj01 は遅い段階でクローズ型構造からオープン型構造へ転移し、trj03 は初期の段階でクローズ型構造からオープン型構造へ転移していることが確認できた。

この章では、クローズ型構造からオープン型構造への転移のプロセスを trend を用いて考えることとする。2 章でも述べたとおり、Daubechies のウェーブレットを用いると、wavelet 変換によって定常的な動きを取り除くことが出来るということが分かっている。AK の熱雑音的な動きを wavelet 変換により取り除きそのデータに対して多項式フィッティングを行うことにより、AK がクローズ型構造からオープン型構造へと転移していくプロセスについて考えていきたい。

## 4.2 多項式フィッティング

wavelet 変換を行って求めた trend が、wavelet 変換を行う前のデータの傾向をどれくらい捉えることが出来ているかを調べた。また、wavelet 変換を行って求めた trend を 3 次関数でフィッティングしたものが、trend の傾向をどれくらい捉えられているかということについても調べた。また、これ以下の計算は特筆しない限り全てレベル 6 について行っており、用いている wavelet

は DAUB12[2] である。これは、trj03 において構造転移が早い段階で起こるため、DAUB20 を用いると周期境界条件によりより多くのデータが削られてしまい構造転移のプロセスが見えにくくなってしまうと考えたからである。そのために、本章では DAUB12 を用いている。

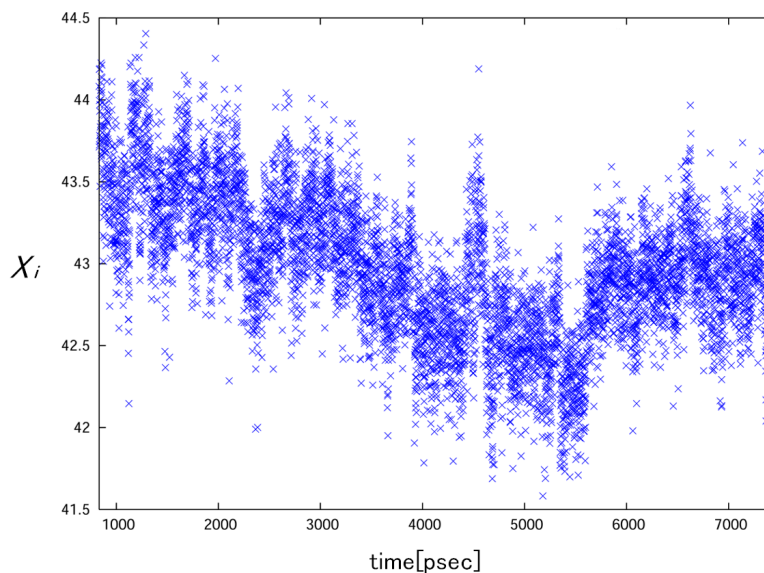


図 58: wavelet 変換を行う前のデータ (16 番目の 炭素の x 座標)

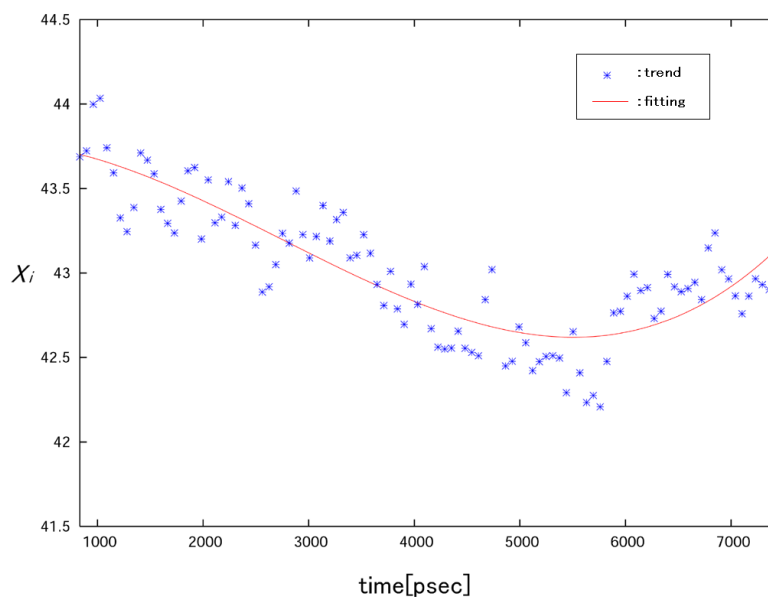


図 59: wavelet 変換を行った後のデータ (16 番目の 炭素の x 座標)

これらの結果より、wavelet 変換を行う前の時系列データの傾向と wavelet 変換を行った後の trend のデータの傾向はよく一致していることが分かる。また、3 次関数でのフィッティングも trend の傾向をよく捉えていることが分かる。しかしながら、3 次関数でのフィッティングは trend からある程度のずれがあることが確認できる。どの程度のずれであるのかを次節で検討する。

### 4.3 フィッティングからのずれ

trend の多項式フィッティングからのずれ

$$D_{\text{abs}}(i) = \frac{1}{T} \sum_{t=T_{\text{min}}}^{T_{\text{max}}} |d_i(t)| \quad (4.1)$$

を求めた。ここで、wavelet 変換での周期境界条件の影響を考え合わせて、 $T_{\text{min}} = 13[\text{step}]$ 、 $T_{\text{max}} = 116[\text{step}]$  とする。また、 $T = T_{\text{max}} - T_{\text{min}} + 1$  とする。本節では、レベル 6 について考えているため、 $T = 128 - 12 \times 2 = 104$  となっている。また、本計算では、各自由度でのずれを見ている。1 個あたりの炭素につき 3 個の自由度があるため、642 個の自由度について計算している。2 次関数と 3 次関数でのフィッティングからのずれ  $D_{\text{abs}}(i)$  を計算した結果を次項に載せる。

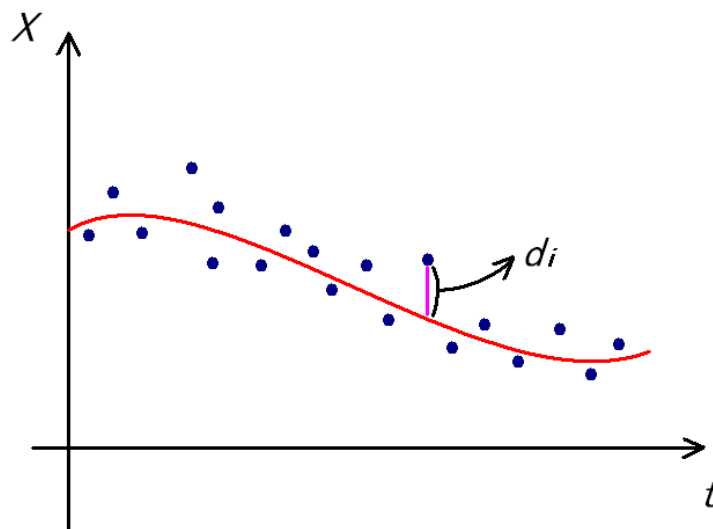


図 60: フィッティングの模式図

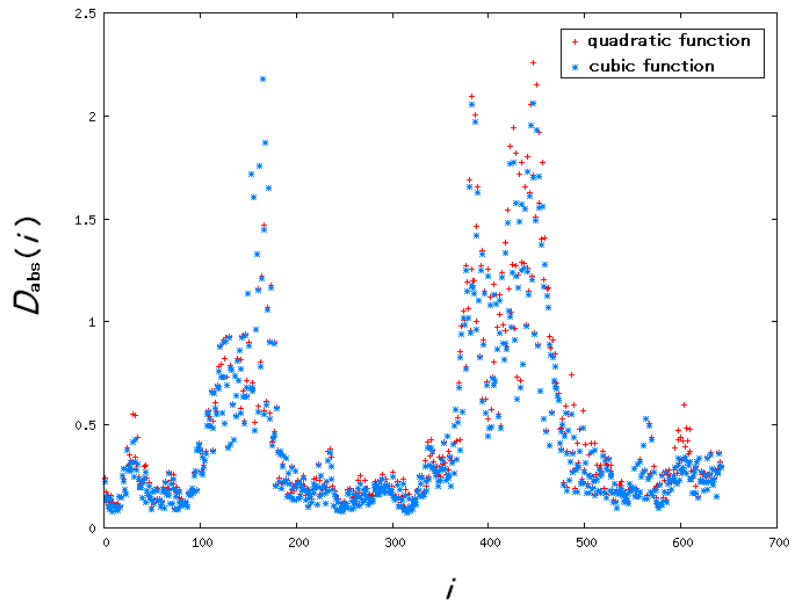


図 61: 絶対的なずれ

この結果より、フィッティングからのずれは高次になるほど著しく小さくなっているというわけではなく、次数にかかわらず同程度のずれをもっていることが分かる。また、アミノ酸残基の炭素の番号で言うと、40番付近と150番付近で大きなずれが確認できる。3章で確認したとおり、40番付近と150番付近は、AKの集団運動が形成されている部分であった。このことにより、大きく動く部分はずれの値も大きくなっていることが分かる。したがって、トレンドの変化量も考慮した相対的な指標

$$D_{\text{rel}}(i) = \frac{1}{T} \sum_{t=T_{\text{min}}}^{T_{\text{max}}} \frac{|d_i(t)|}{d_{\text{max}} - d_{\text{min}}} \quad (4.2)$$

での計算を行うこととする。ここで、ずれの最大値を  $d_{\text{max}}$ 、ずれの最小値を  $d_{\text{min}}$  とする。また、 $T$  などは先程と同様である。

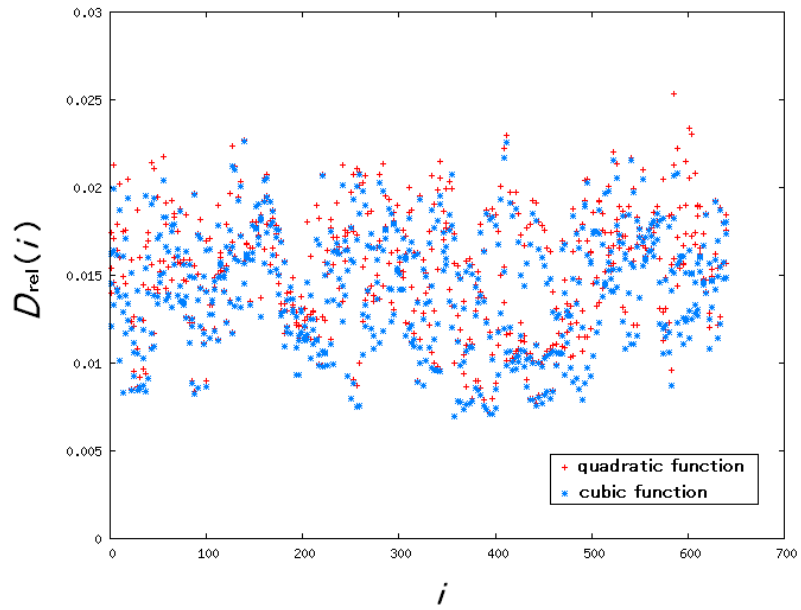


図 62: 相対的なずれ

この結果より、AKの大きく動く部分はゆらぎも大きいことが確認できた。また、次数にかかわらず、相対的なずれは各自由度で小さい値を取ることが確認できた。

## 4.4 フィッティングの慣性テンソル

trend から求めた慣性テンソルと3次関数でフィッティングを行ったときの慣性テンソルの固有値を trj01 と trj03 に対して求めた。ここで求めている慣性テンソルの固有値は3つあるうちで最も大きい値のものである。両結果とも傾向が一致していることが確認できた。

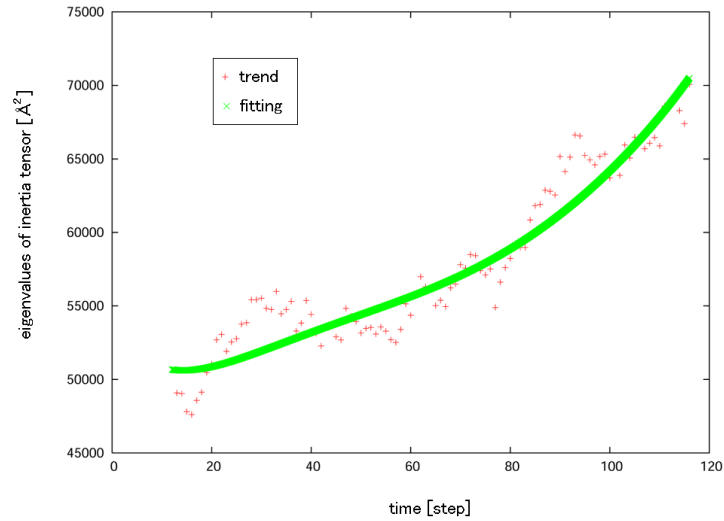


図 63: 慣性テンソルの固有値 ( trj01 )

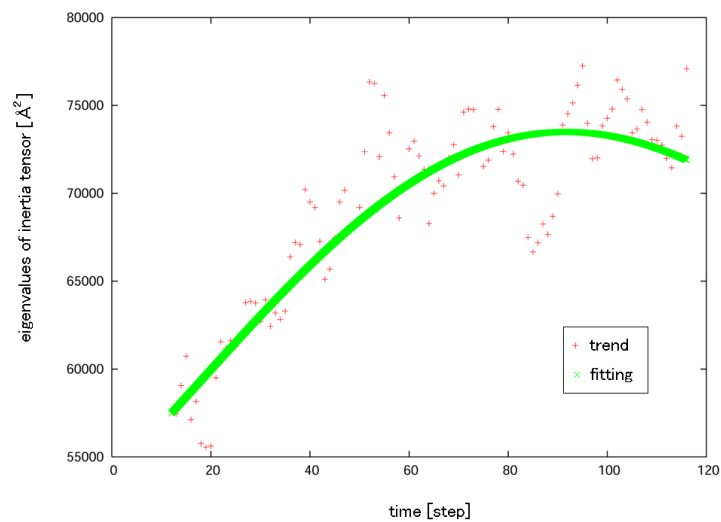


図 64: 慣性テンソルの固有値 ( trj03 )



## 4.5 開いていく過程

前節では、3次関数でフィッティングを行ったあとのデータに対しても、AKがクローズ型構造からオープン型構造へ開いていく際に、慣性テンソルの固有値の値が著しく大きくなっていることを確認した。クローズ型構造からオープン型構造へ構造転移していく際に、AK内の分子の集団運動に変化は見られるのだろうか。AK内の分子の動きが変わる際に、その境目は存在するのだろうか。例えば、小さい分子などは、急に反応座標が変わる際に、分子自身が大きく振動励起しているということが気相反応において確認されている。AKにおいても、そのようなことが確認できるのだろうか。このようなことを確認すべく、以下のような計算を行った。trendで得られた座標ベクトル $\mathbf{r}_{\text{tre}}(t)$ から速度ベクトル $\mathbf{v}_{\text{tre}}(t)$ を求める。

$$\mathbf{v}_{\text{tre}}(t) = \mathbf{r}_{\text{tre}}(t + dt) - \mathbf{r}_{\text{tre}}(t) \quad (4.3)$$

ここで、 $\mathbf{r}_{\text{tre}}(t)$ は炭素原子の位置を表すベクトルである。次に、 $\mathbf{v}_{\text{tre}}(t)$ を多項式フィッティングの空間成分とそれに直交する成分に分解する。

$$\mathbf{v}_{\text{tre}}(t) = \mathbf{v}_1(t) + \mathbf{v}_2(t) \quad (4.4)$$

ここで、

$$\mathbf{v}_1(t) = (t)\mathbf{a} + (t)\mathbf{b} + (t)\mathbf{c}$$

であり、 $\mathbf{a}$ 、 $\mathbf{b}$ 、 $\mathbf{c}$ は3次関数の多項式フィッティングで得られた正規直交系の基底ベクトルである。また、 $\mathbf{v}_2(t)$ は $\mathbf{a}$ 、 $\mathbf{b}$ 、 $\mathbf{c}$ に直交するベクトルである。これより、運動エネルギー $|\mathbf{v}_1(t)|^2$ を求め時間の関数としてグラフにする。

続いて、多項式フィッティングにより得られた位置ベクトル $\mathbf{r}_{\text{fit}}(t)$ を

$$\mathbf{r}_{\text{fit}}(t) = \mathbf{r}_1(t) + \mathbf{r}_2(t) \quad (4.5)$$

のように分解する。ここで、 $\mathbf{r}_1(t)$ は $\mathbf{a}$ 、 $\mathbf{b}$ 、 $\mathbf{c}$ が張る空間のベクトルで、 $\mathbf{r}_2(t)$ はそれに直交する空間のベクトルである。 $\mathbf{r}_{\text{fit}}(t)$ を $\mathbf{a}$ 、 $\mathbf{b}$ 、 $\mathbf{c}$ が張る3次元空間に図示する。また、これと同様に $\mathbf{r}_{\text{tre}}(t)$ も分解し、 $\mathbf{a}$ 、 $\mathbf{b}$ 、 $\mathbf{c}$ が張る3次元空間に図示する。以下にこの結果を載せる。なお、運動エネルギーがピークをとる点と3次元空間での動きも比較した。また、3次元プロットに関しては4つの方向から見たものを載せる。

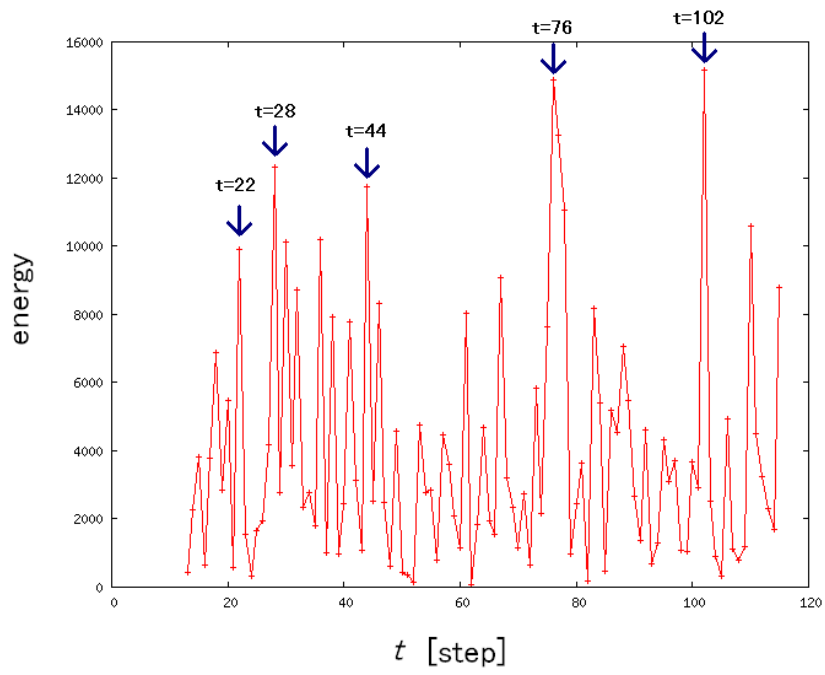


図 65: 運動エネルギー (trj01)

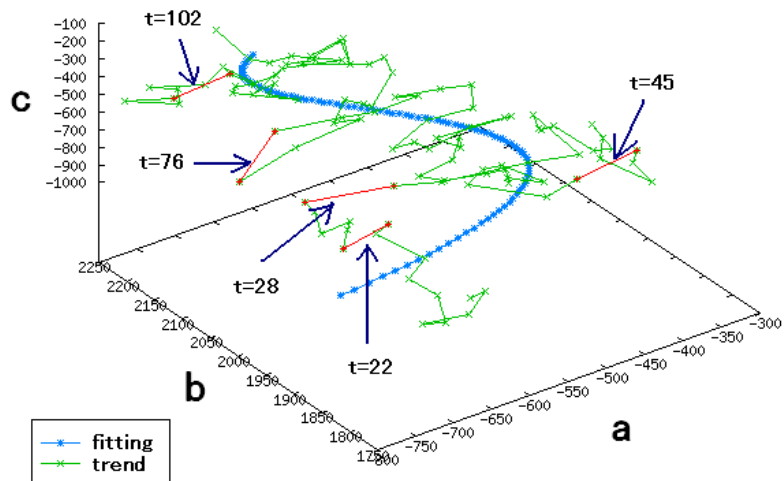


図 66: trj01 の軌跡 (1)

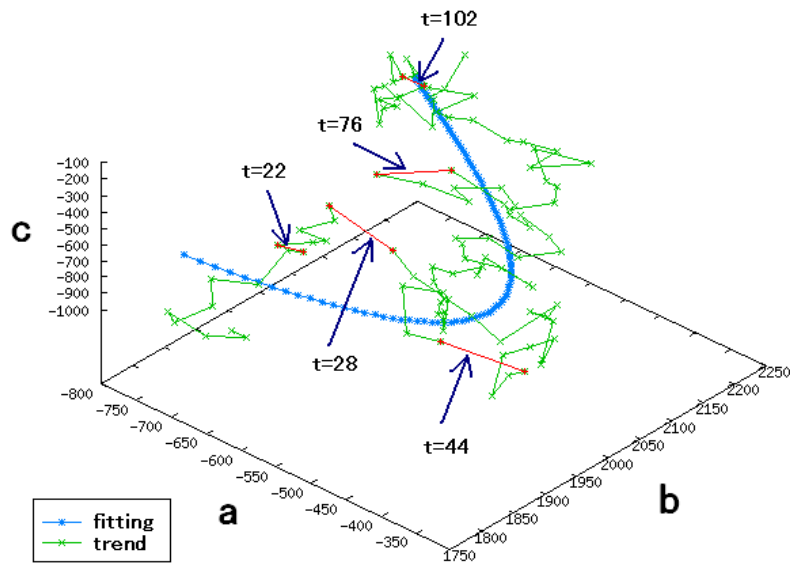


図 67: trj01 の軌跡 (2)

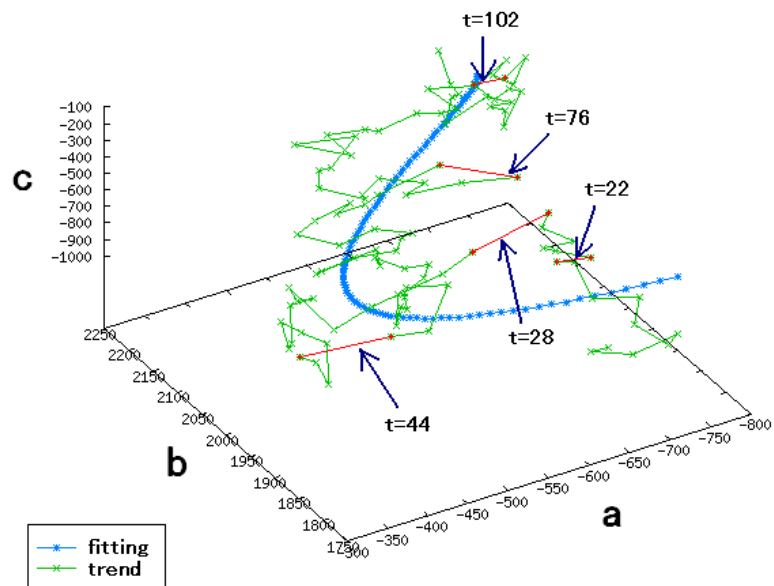


図 68: trj01 の軌跡 (3)

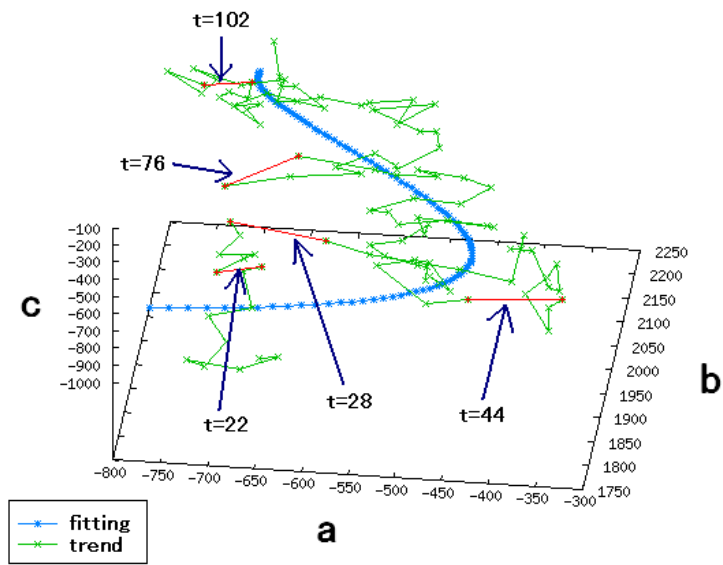


図 69: trj01 の軌跡 (4)

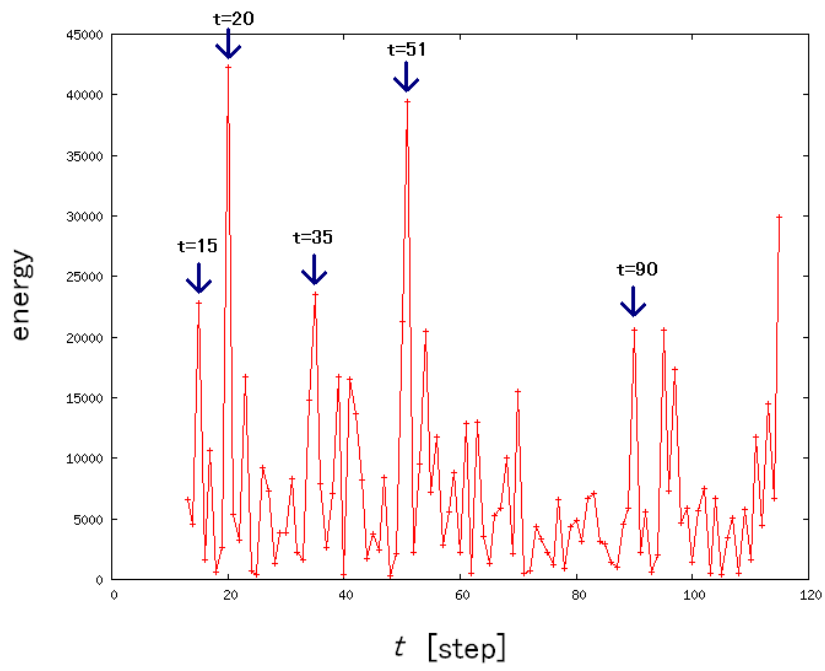


図 70: 運動エネルギー (trj03)

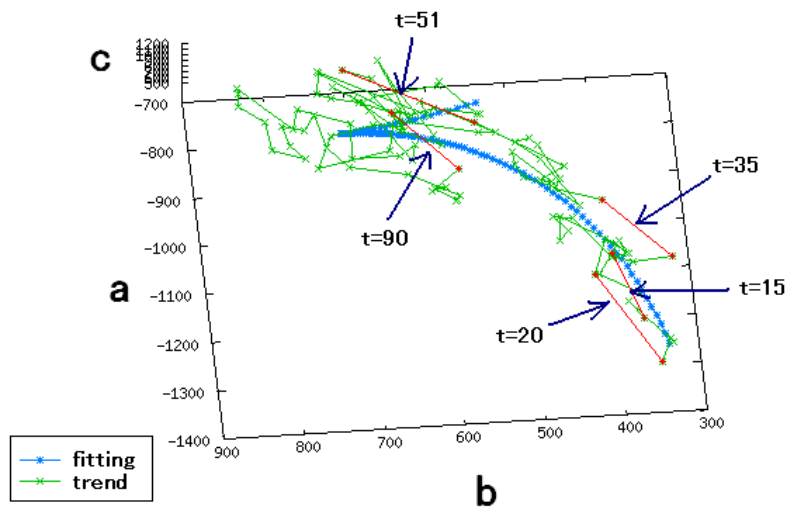


図 71: trj03 の軌跡 (1)

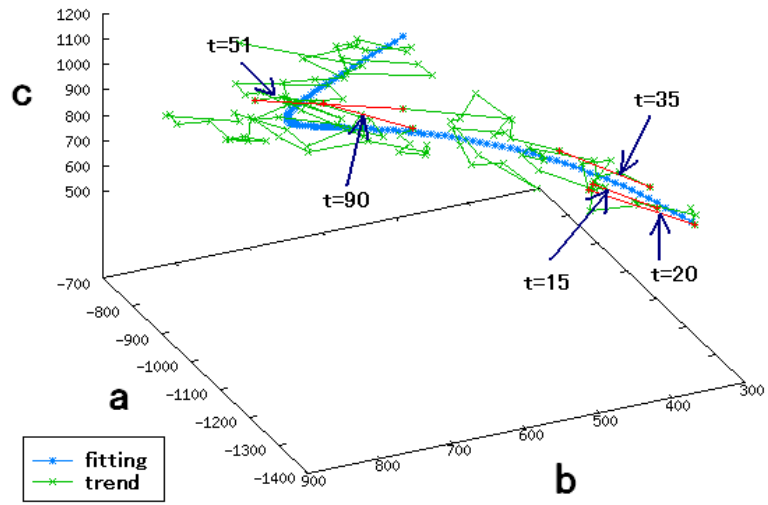


図 72: trj03 の軌跡 (2)

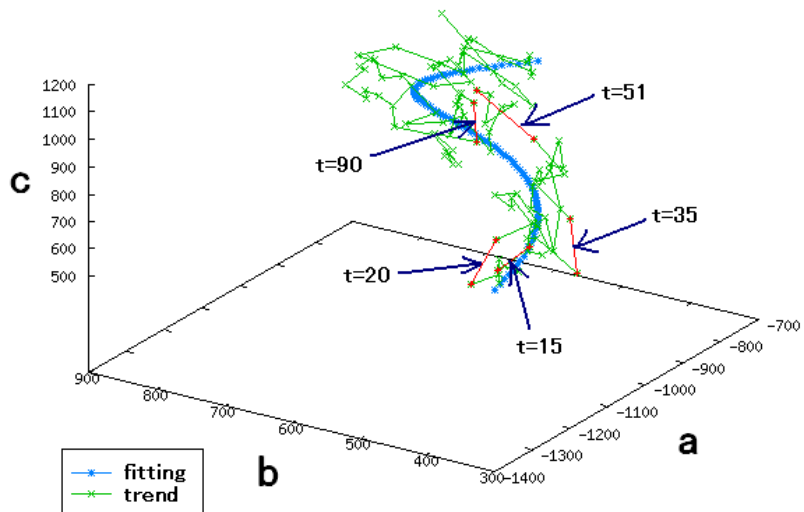


図 73: trj03 の軌跡 (3)

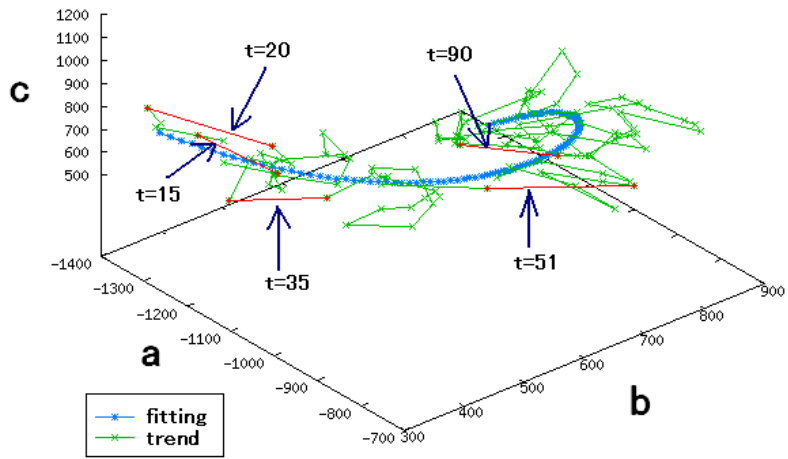


図 74: trj03 の軌跡 (4)

以上の結果より、運動エネルギーがピークをとるところでは軌跡の時間変化も大きいことが確認できた。また、trj03 では trend で求めた軌跡がフィッティングで求めた軌跡によく沿っているのに対し、trj01 では trend で求めた軌跡はフィッティングで求めた軌跡から大きなずれが見られるところがあった。運動エネルギーがピークをとるところでの trend の軌跡の変化においても、trj03 ではフィッティングでの軌跡によく沿っているのに対し、trj01 ではずれていることが確認できた。これらの要因として、trj01 における停滞期の存在が影響していることが挙げられるが、現時点では解析結果が2つの初期値に対してのものしかないため更に解析結果を増やす必要があるといえる。

## 4.6 集団運動性

3章でも調べたように、本章でも、多項式でフィッティングしたあとのデータを用いて、集団運動性について調べることにする。trj01 と trj03 について、それぞれ2次関数、3次関数でフィッティングを行ったデータに対して集団運動性を調べた結果を載せる。以下は、

$$C_{\text{ang}}(l, t) = \sum_{i < j, |\mathbf{r}_i - \mathbf{r}_j| < d} \frac{\mathbf{v}_i(l, t) \cdot \mathbf{v}_j(l, t)}{|\mathbf{v}_i(l, t)| |\mathbf{v}_j(l, t)|} = \sum_{i < j, |\mathbf{r}_i - \mathbf{r}_j| < d} \cos \theta_{i,j}(l, t) \quad (4.6)$$

の結果である。ここで、 $t$  は実際の時間ではなく、時間のステップ数とする。また、 $i, j = 1, 2, \dots, 214$  であり、 $d$  は第3章で用いた値に準ずるものとする。指標そのものは3章で求めたものと同じものを用いている。



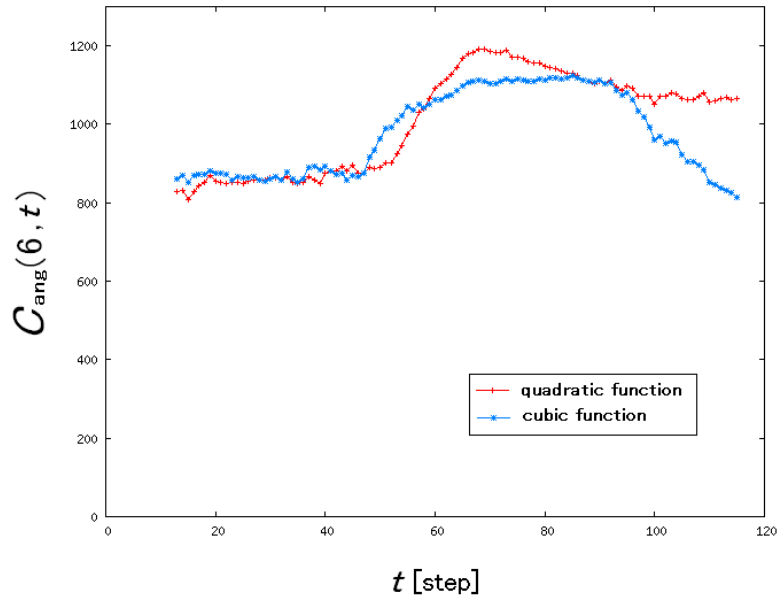


図 75:  $C_{\text{ang}}(\text{trj01})$

上図において、2次関数と3次関数のフィッティングで得られたデータでの  $C_{\text{ang}}$  は初期の段階、つまり開いていく過程において、よく一致している。後半は多少のずれが見られるが、全体的に傾向はよく一致していることが分かる。

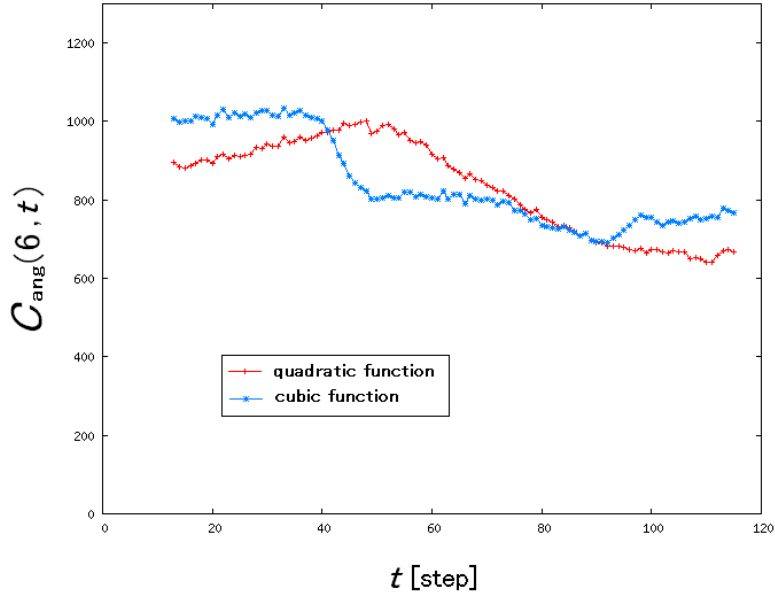


図 76:  $C_{\text{ang}}(\text{trj03})$

上図において、wavelet の周期境界条件により初期過程が見えていない。これは trj01 においても同様のことであるが、trj03 においては慣性テンソルでも確認したとおり初期の段階で構造転移が起こるため、DAUB12 を用いても構造転移の全過程を確認できないことが分かる。

2 次関数と 3 次関数でのフィッティングにおいて傾向がある程度一致しているため、信頼できるフィッティングではあるようだ。3 章でも述べたが、この指標  $C_{\text{ang}}$  は、方向がどのくらい揃っているかということを表す指標であり、速度成分の大きさについては考えられていなかった。次に、速度成分の大きさについても考えた指標  $C_{\text{vel}}$  について調べた結果を載せる。ここでの指標も 3 章で用いた指標と同じものである。

$$C_{\text{vel}}(l, t) = \frac{\sum_{i < j, |\mathbf{r}_i - \mathbf{r}_j| < d} \mathbf{v}_i(l, t) \cdot \mathbf{v}_j(l, t)}{\sum_i |\mathbf{v}_i(l, t)|^2} \quad (i, j = 1, 2, \dots, 214) \quad (4.7)$$

ここで、 $i, j = 1, 2, \dots, 214$  であり、 $d$  は第 3 章で用いた値に準ずるものとする。

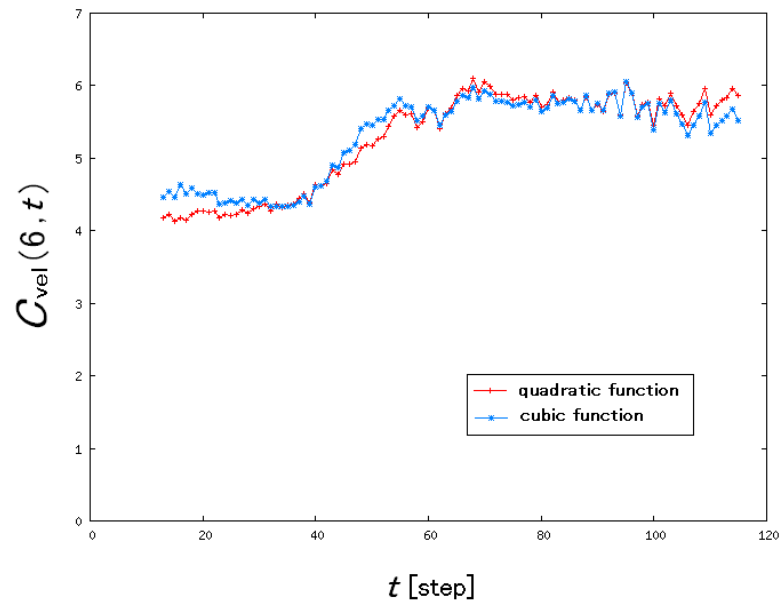


図 77:  $C_{\text{vel}}(\text{trj01})$

上図より、2次関数と3次関数のフィッティングで、全過程においてよく一致していることが分かる。

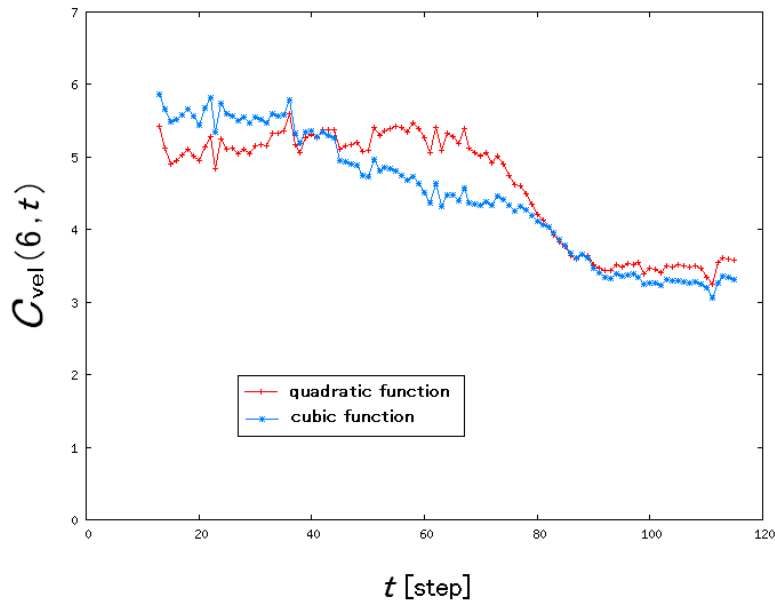


図 78:  $C_{\text{vel}}(\text{trj03})$

多少のずれは見られるが、2次関数と3次関数のフィッティングで全体的に傾向が一致していることが分かる。

trj01では、 $C_{\text{ang}}$ と $C_{\text{vel}}$ の両方において2次関数と3次関数で大体の一致が確認出来た。しかしながら、trj03では、 $C_{\text{ang}}$ と $C_{\text{vel}}$ の両方において大体の傾向は一致していたが、多少のずれが見られた。これは、trj03では初期段階で構造転移が起こることに起因しているのではないだろうか。trj03の解析においては、局所的なデータの解析の必要性も考え得る。

以上の結果より、フィッティング関数として3次関数程度が適当であると考え、以下の結果より3次関数でのフィッティングを用いることとする。

## 4.7 実際の集団運動

多項式フィッティングにより求めた速度成分によって、3章と同様に場所相関の指標を用いて AK の実際の集団運動を見た。場所相関の指標  $C_{i,j}^{\text{ang}}(l, t)$  を3章で用いたものと同様に

$$C_{i,j}^{\text{ang}}(l, t) = \frac{\mathbf{v}_i(l, t) \cdot \mathbf{v}_j(l, t)}{|\mathbf{v}_i(l, t)| |\mathbf{v}_j(l, t)|} \quad (i, j = 1, 2, \dots, 214) \quad (4.8)$$

と定義し、 $C_{i,j}^{\text{vel}}(l, t)$  を

$$C_{i,j}^{\text{vel}}(l, t) = \frac{\mathbf{v}_i(l, t) \cdot \mathbf{v}_j(l, t)}{\frac{1}{T} \sum_{t=T_{\min}}^{T_{\max}} \frac{1}{N} \sum_{i=1}^N |\mathbf{v}_i(l, t)|^2} \quad (i, j = 1, 2, \dots, 214, N = 214) \quad (4.9)$$

と定義する。  $T$  などは先述した通りである。 $C_{i,j}^{\text{ang}}(l, t)$  が、

$$C_{i,j}^{\text{ang}}(l, t) > 0.9$$

の条件を満たすときに、点  $(i, j)$  を赤色でプロットし、

$$C_{i,j}^{\text{ang}}(l, t) < -0.9$$

の条件を満たすときに、点  $(i, j)$  を青色でプロットする。

また、 $C_{i,j}^{\text{vel}}(l, t)$  が、

$$C_{i,j}^{\text{vel}}(l, t) > 0.3$$

の条件を満たすときに、点  $(i, j)$  を赤色でプロットし、

$$C_{i,j}^{\text{vel}}(l, t) < -0.3$$

の条件を満たすときに、点  $(i, j)$  を青色でプロットする。

ここで、赤色の部分は正の相関があることを意味し、青色の部分は負の相関があることを意味する。trj01 と trj03 について代表的な場所相関の図とその求めた場所相関の時間での AK の 3 次元空間内での動きを以下に載せる。また、全時間ステップでの結果は Appendix B に掲載する。

### 4.7.1 場所相関 (1)

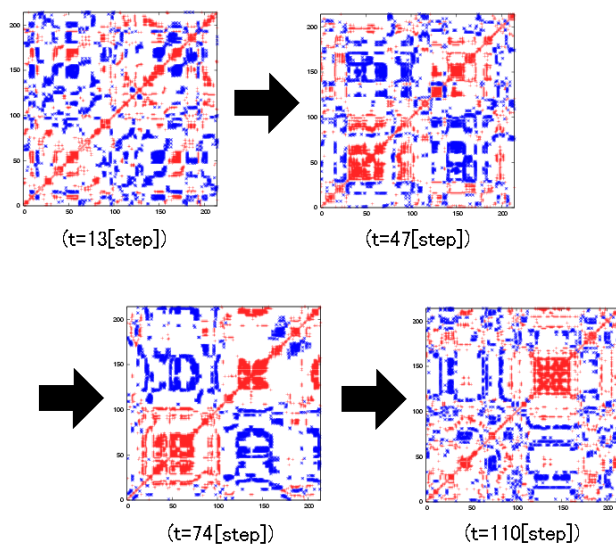


図 79:  $C_{i,j}^{ang}(l,t)$  ( trj01 )

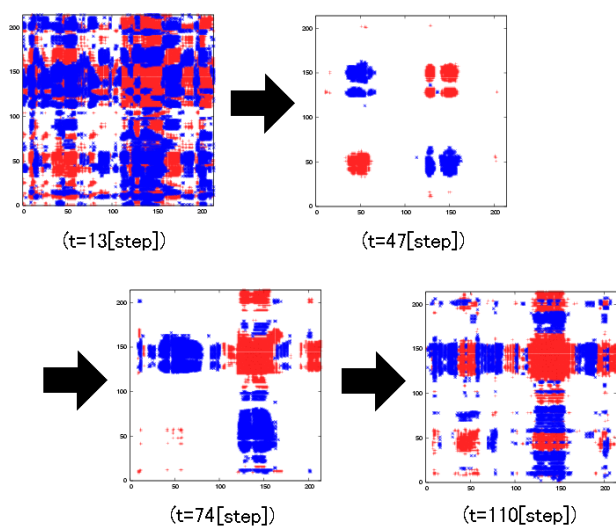


図 80:  $C_{i,j}^{vel}(l,t)$  ( trj01 )

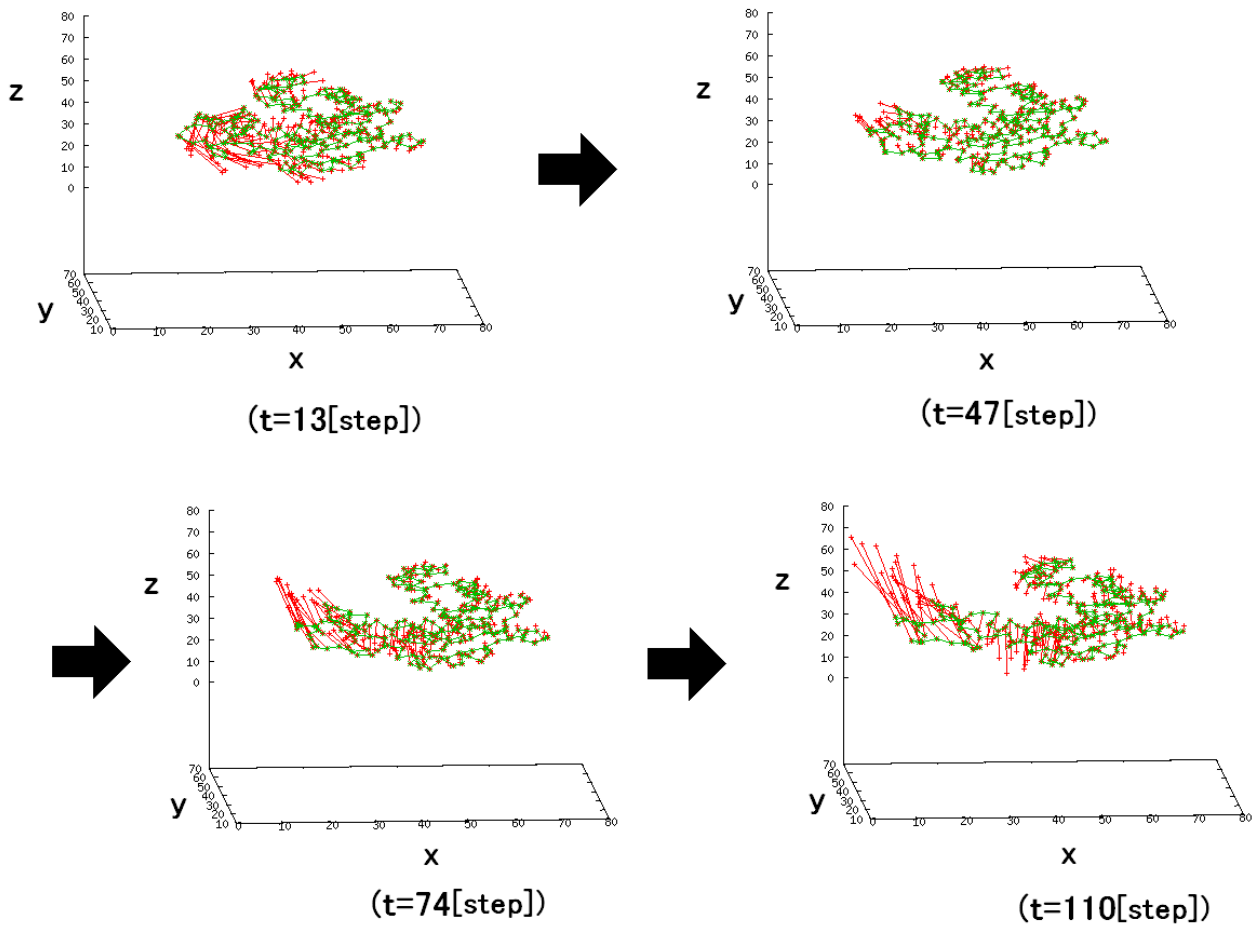


図 81: AK の動き (trj01)

trj01 における代表的な場所相関の 4 例を挙げた。  $t = 13[\text{step}]$  では AK が開き始める前の動きをみており、  $t = 110[\text{step}]$  では AK が開いた後の動きを見ている。  $t = 47[\text{step}]$  から  $t = 74[\text{step}]$  にかけて、集団運動性が強くなっていることが分かる。これは、慣性テンソルを見たときにも確認できたことである。

## 4.7.2 場所相関 (2)

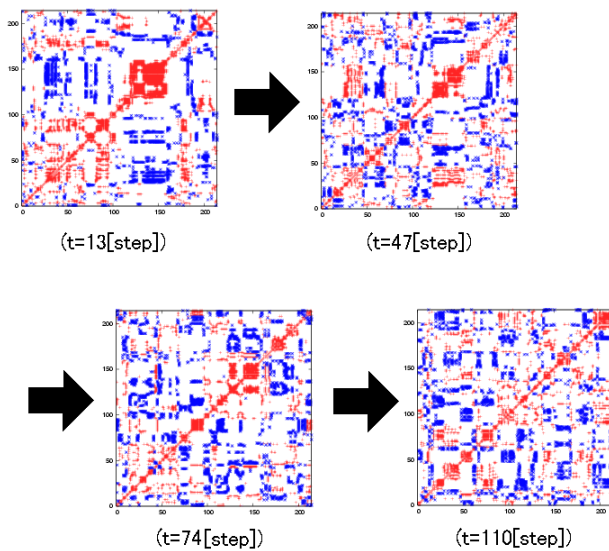


図 82:  $C_{i,j}^{ang}(l, t)$  ( trj03 )

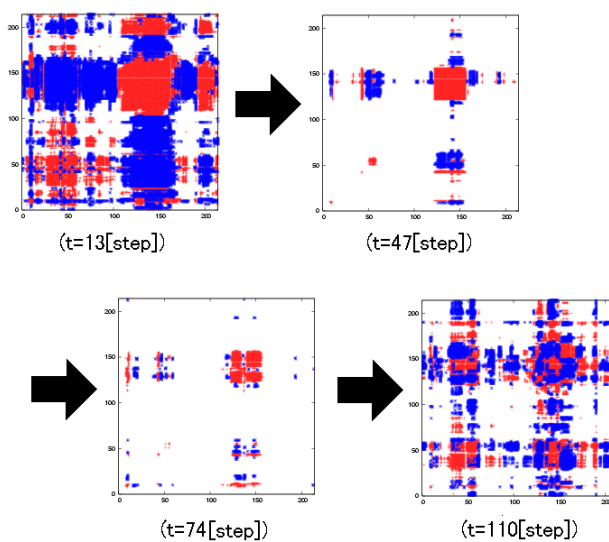


図 83:  $C_{i,j}^{vel}(l, t)$  ( trj03 )



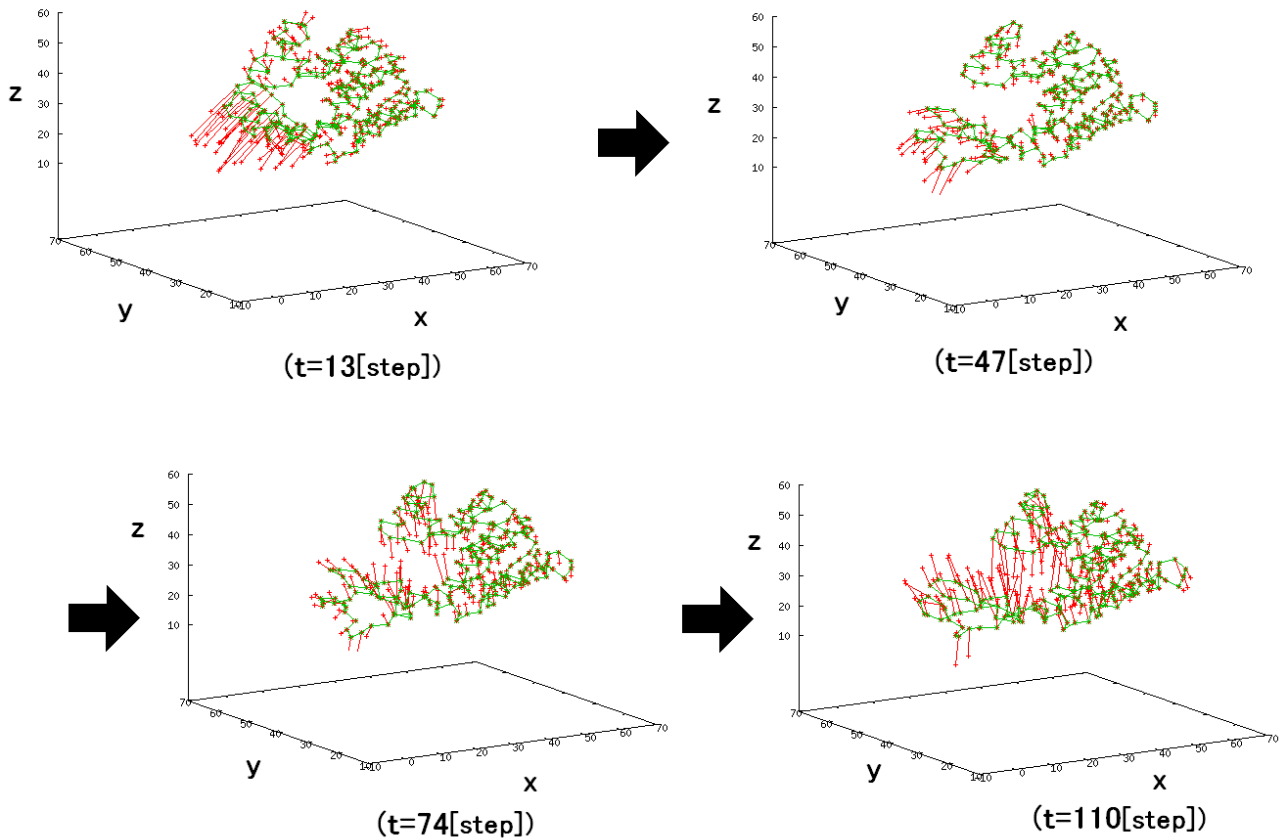


図 84: AK の動き (trj03)

trj03 における代表的な場所相関の 4 例を挙げた。  $t = 13[\text{step}]$  では AK が開き始める前の動きをみており、  $t = 110[\text{step}]$  では AK が開いた後の動きを見ている。 trj03 の  $t = 110[\text{step}]$  と trj01 の  $t = 110[\text{step}]$  の場所相関の図と比較すると、 trj03 の方が集団運動がばらけていることが分かる。これは、 trj03 の方がより熱平衡状態に落ち着いているということを意味しているのではないだろうか。  $t = 13[\text{step}]$  から  $t = 47[\text{step}]$  にかけて、集団運動性が強くなっていることが分かる。これは、慣性テンソルを見たときにも確認できたことである。

trj01 と trj03 において、慣性テンソルが著しく大きくなっていった時間領域、つまり、クローズ型構造からオープン型構造へ転移する過程で、場所相関の指標を用いることにより集団運動が起きていることが確認できた。これは、トラジェクトリにかかわらず、クローズ型構造からオープン型構造へ変移する際に、分子内で集団運動が起こっていることを示唆している。

## 4.8 まとめ

本章では、AK がクローズ型構造からオープン型構造へ転移する際のプロセスについて調べた。trj01 と trj03 という 2 つの trajectory について wavelet 変換と多項式フィッティングを用いて解析を行った。これら 2 つの trajectory の共通点と相違点についてまとめる。

### 1. 共通点

- (a) 多項式フィッティングから得られた慣性テンソルの固有値と trend から得られた慣性テンソルの固有値はどちらも傾向が一致していた。
- (b) 多項式フィッティングと trend の軌跡において、多項式フィッティングの軌跡はある程度 trend の軌跡の傾向を捉えていることが確認できた。
- (c) 集団運動性の指標をみると、trj01 と trj03 において構造転移が起こる際に、指標の値が大きくなっていることが確認できた。
- (d) 場所相関を求めると、trj01 と trj03 の両方において、構造転移が起こる際に、集団運動が起こっていることが確認できた。

### 2. 相違点

- (a) 多項式フィッティングと trend の軌跡において、運動エネルギーが大きい値をとる際に、trj03 では trend の軌跡が多項式フィッティングの軌跡に沿うようになっていたのに対し、trj01 ではある程度のずれが見られた。この原因としては、trj01 における停滞期の存在が挙げられるであろう。
- (b) 集団運動性の指標をみると、trj01 では 2 次関数と 3 次関数で大体の一致が見られるのに対し、trj03 では 2 次関数と 3 次関数でずれが見られた。これは、trj03 における局所的なフィッティングの必要性を示唆しているのではないかと考えられる。

## 5 総括と展望

本研究では、wavelet 変換と多項式フィッティングを用いて分子動力学計算により求めた AK の炭素の座標の時系列データに対して時系列解析を行った。これらの解析方法を用いることにより、AK の 2 つの dynamics において集団運動を抽出することに成功した。

オープン型構造の AK については、定常的な集団運動を取り出すことに成功した。また、100[psec] 程度の時間スケールでは熱平衡的なゆらぎが確認できた。この結果は、wavelet 変換が定常的な振動を取り出すのに有効であることを示唆している。しかしながら、慣性テンソルの固有値を見た際に 1[nsec] 程度の時間スケールで確認できる非定常な動きについては解析が行えていない。十分な長さの時系列データがあれば、これも可能になるであろう。

続いて、クローズ型構造からオープン型構造へ構造転移する AK については、非定常な方向性を持った集団運動を取り出すことに成功した。また、異なる初期条件の軌跡で、開いていくプロセスにおいて集団運動が形成されていることを確認した。本研究では、trj01 と trj03 という初期条件が異なる 2 つの trajectory に対して解析を行ったが、trj03 は初期段階で構造転移が見られるのに対し、trj01 は構造転移が起こるまでに時間がかかるものであった。しかしながら、構造転移が起こる時間スケールはこれら 2 つの trajectory で共通しており、trj01 における停滞期の存在が、AK が開いていくプロセスの律速過程にまさに影響を与えているということが考えられる。ただし、これについてはより多くの trajectory について解析を行う必要があるといえる。

今後の課題としては、第一に、オープン構造の AK において確認できた 1[nsec] 程度の時間スケールでの非定常的なゆらぎについての解析の必要性が挙げられるであろう。第二に、AK が開いていくプロセスにおいて集団運動が励起されるメカニズムの解明が挙げられるであろう。また、構造転移の速度を決めているメカニズムの解明も必要であろう。ところで、本研究においては、場所相関など 1 次の統計量についての解析であった。2 次の統計量、例えば、レベル間での相関などについても考える必要がある。

AK の集団運動について考え、AK 内のどの部分とその機能に関係しているのかということが解明できれば、AK のコントロールひいては生化学反応のコントロール可能性についても考え得るであろう。

## 6 謝辞

# 謝辞

まずはじめに、指導教官である戸田幹人先生に厚く御礼申し上げます。「研究とは何か？」ということを実践して学べたということに非常に感謝しております。これからの「生きる糧」になったと確信しております。有難うございました。

また、貴重なデータを頂きました横浜市立大学大学院の木寺詔紀氏と淵上壮太郎氏には、研究を進めて行く上で様々なアドバイスなども頂き、大変感謝しております。有難うございました。

ゼミや授業、そして研究室内のPCの扱いなどにおいて、上江洌達也先生や狐崎創先生にも大変お世話になりました。有難うございました。

研究室内の方々には、ゼミや研究室内で様々なアドバイスや励ましを頂き、本当に感謝しております。有難うございました。

この場をお借りして、皆様に厚く御礼申し上げます。

## 7 Appendix A

### 7.1 慣性テンソル

回転している剛体の質量中心のまわりの相対的回転の運動エネルギー  $T$  は、 $m$  を質量、 $\vec{v}$  を相対速度ベクトル、 $\vec{\omega}$  を相対回転角速度ベクトル、 $\vec{r}$  を相対位置ベクトルとすると、

$$\begin{aligned} T &= \sum_i \frac{1}{2} m_i v_i^2 \\ &= \sum_i \frac{1}{2} m_i \{ (\vec{\omega} \times \vec{r}_i)_x^2 + (\vec{\omega} \times \vec{r}_i)_y^2 + (\vec{\omega} \times \vec{r}_i)_z^2 \} \\ &= \frac{1}{2} (\omega_x \quad \omega_y \quad \omega_z) \begin{pmatrix} \sum_i m_i (y_i^2 + z_i^2) & -\sum_i m_i x_i y_i & -\sum_i m_i z_i x_i \\ -\sum_i m_i x_i y_i & \sum_i m_i (z_i^2 + x_i^2) & -\sum_i m_i y_i z_i \\ -\sum_i m_i z_i x_i & -\sum_i m_i y_i z_i & \sum_i m_i (x_i^2 + y_i^2) \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \\ &= (\omega_x \quad \omega_y \quad \omega_z) A \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \end{aligned} \quad (7.1)$$

とあらわすことが出来る。この行列  $A$  の固有値を求める数値計算を 1ake と 4ake に対して行った。この際用いたデータは、分子動力学により求めた 1ake と 4ake における炭素の座標の時系列データである。ここで、今回は炭素の骨組みだけを見ているので、 $m_i = 1$  ( $i = 1, 2, \dots, 214$ ) とした。

## 8 Appendix B

### 8.1 エネルギー (1ake)

本文では 4ake に対してのみ載せたが、本章では 1ake の trj01 と trj03 に対して求めたレベルごとのエネルギーの時間発展

$$E(l, t) = \sum_{i=1}^{214} |\mathbf{v}_i(l, t)|^2 \quad (8.1)$$

の結果を載せる。ここで、 $l$  はレベルの番号を表し、 $i$  はアミノ酸残基の炭素の番号を表す。また、ここでの wavelet 変換は DAUB20[2] を用いた。AK が開いていく過程でエネルギーの値が大きくなっていることに注目して見てもらいたい。今後、開いていく過程でエネルギーが大きくなっていることに対する解析も必要であろう。

### 8.1.1 エネルギー (1)

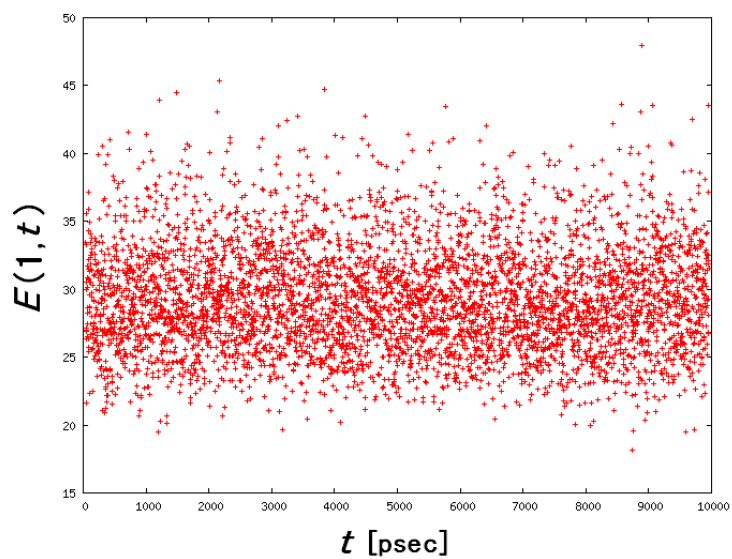


図 85: trj01:レベル 1

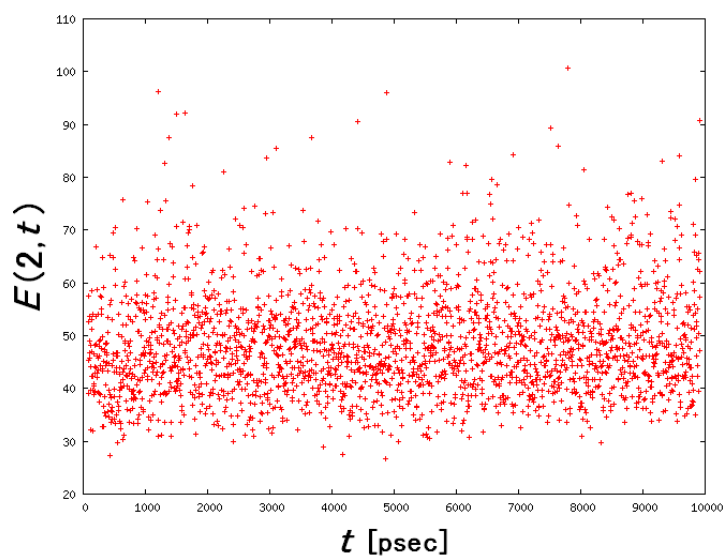


図 86: trj01:レベル 2

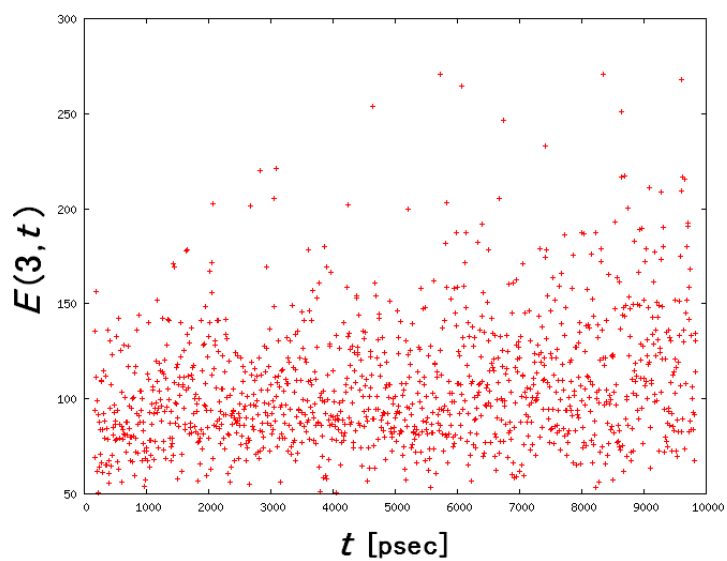


図 87: trj01:レベル 3

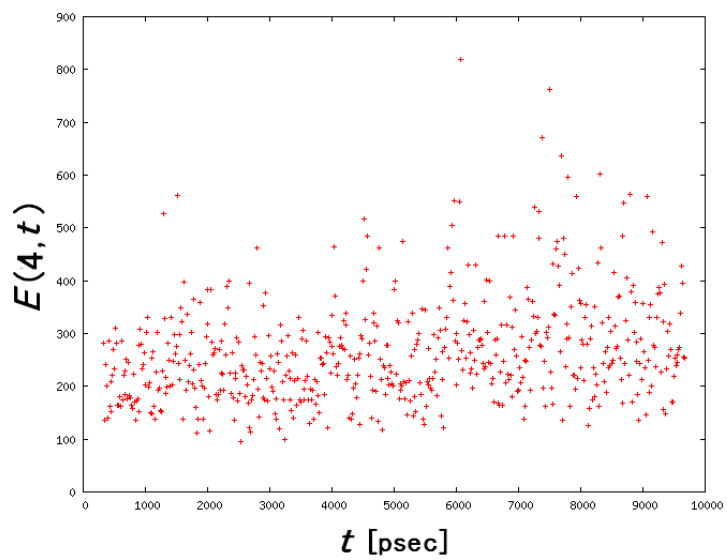


図 88: trj01:レベル 4



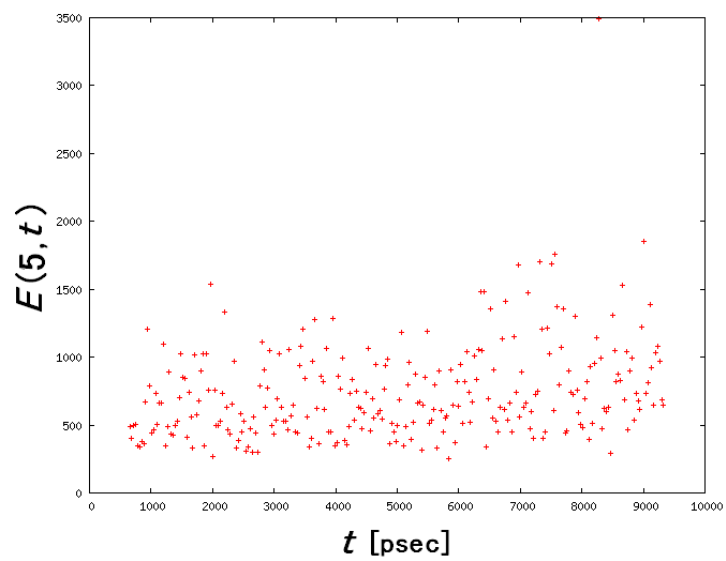


図 89: trj01:レベル 5

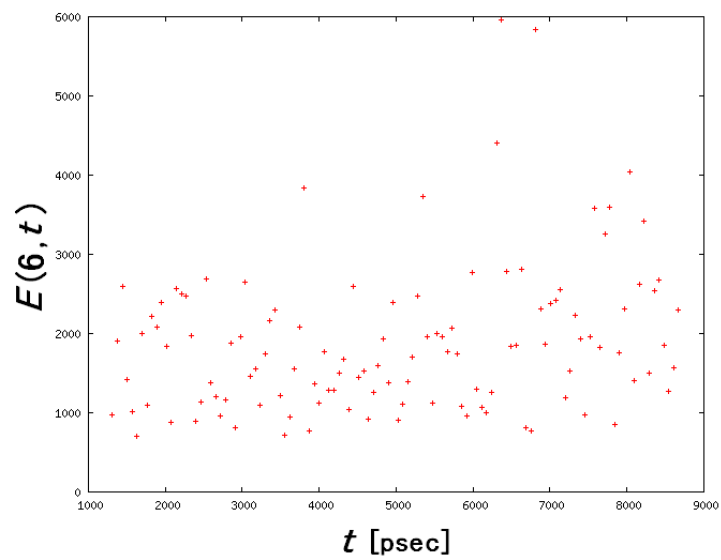


図 90: trj01:レベル 6

### 8.1.2 エネルギー (2)

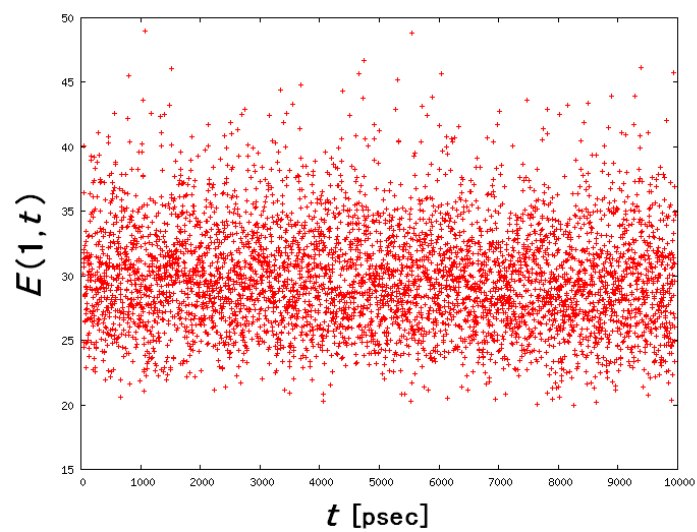


図 91: trj03:レベル 1

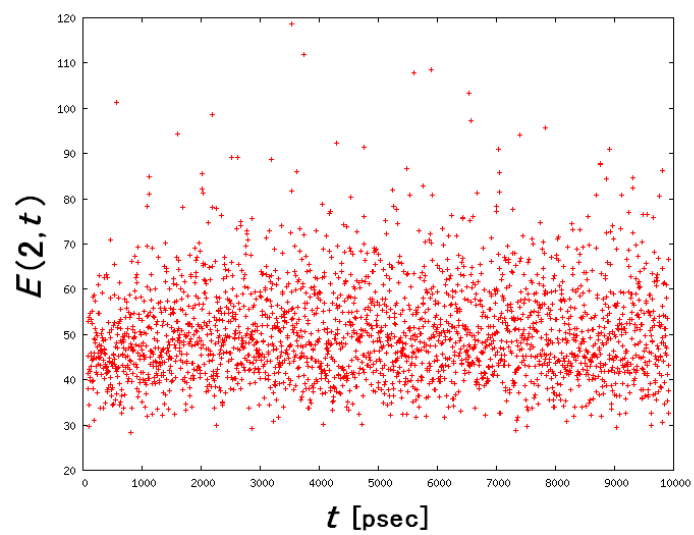


図 92: trj03:レベル 2

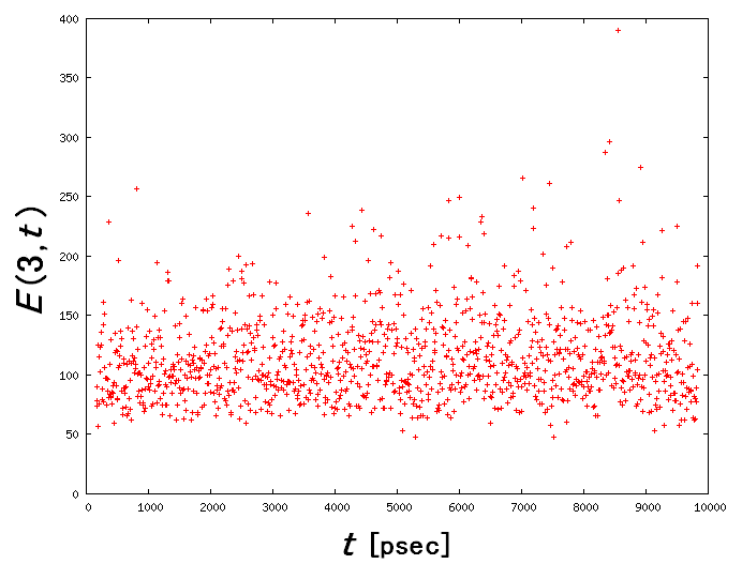


図 93: trj03:レベル 3

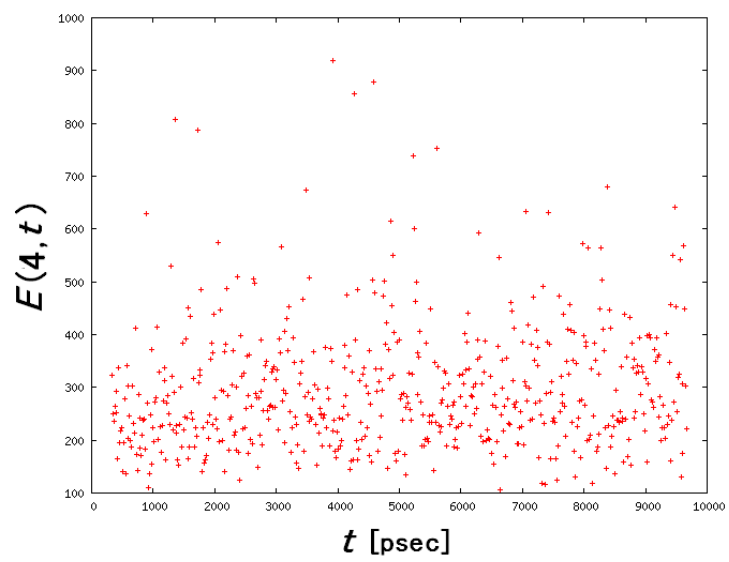


図 94: trj03:レベル 4

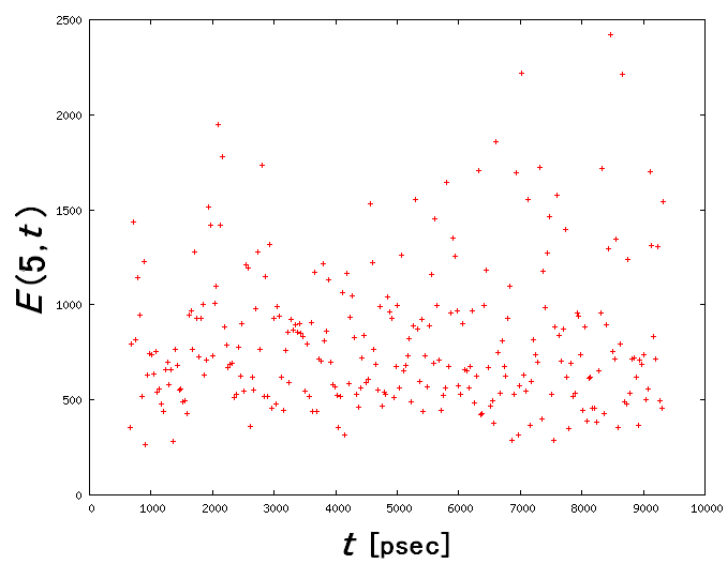


図 95: trj03:レベル 5

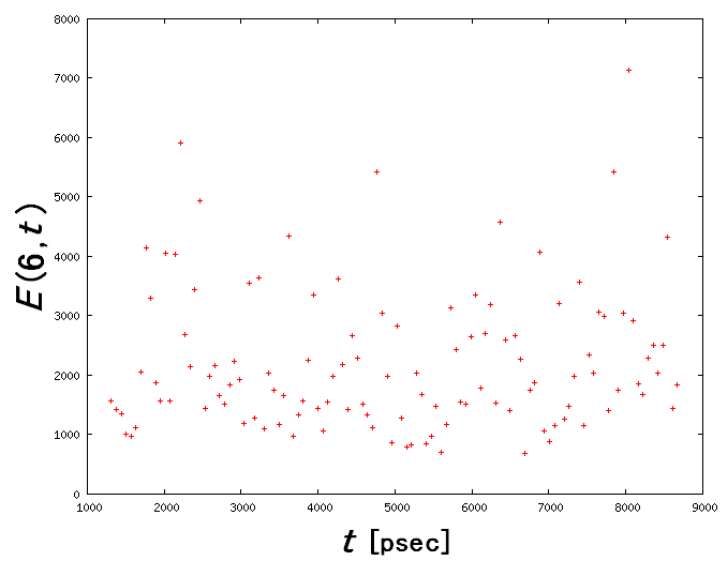


図 96: trj03:レベル 6

## 参考文献

- [1] 『ウェーブレット10講』  
I. ドブシー 著、山田道夫 / 佐々木文夫 訳  
シュプリンガー・フェアラーク東京, 2003
- [2] 『NUMERICAL RECIPES in C The Art of Scientific Computing Second Edition』  
(特に、13.9を参考)
- [3] 『数理科学 ウェーブレットビギナーズガイド』  
榊原進 著  
東京電機大学出版局
- [4] 『ウェーブレット解析の基礎理論』 新井康平 著  
森北出版株式会社
- [5] 『Trend assessment in a long memory dependence model using the discrete wavelet transform』  
Peter F. Craigmile, Peter Guttorp and Donald B. Percival  
Environmetrics 2004; 15: 313-335
- [6] 『Conformation, Energy, and Folding Ability of Selected Amino Acid Sequences』  
M Sasai  
Proc. Natl. Acad. Sci. USA Vol 92, pp.8438-8442, August 1995 Biophysics
- [7] 『The effects of solvent on the conformation and the collective motions of protein: normal mode analysis and molecular dynamics simulations of melittin in water and in vacuum』  
Akio Kitao, Fumio Hirata and Nobuhiro Go  
Chemical Physics 158(1991)447-472
- [8] 『蛋白質・水系の階層的ダイナミクス』  
北尾彰朗  
物性研究 60-3 (1993-6)
- [9] 『Vibrational Energy Transfer in a Protein Molecule』  
Kei Moritsugu, Osamu Miyashita, and Akinori Kidera  
Physical Review Letters, 2000

- [10] 『蛋白質機能発現のダイナミクスの側面』  
木寺詔紀 / 池口満徳  
日本物理学会誌 Vol.58, No.4, 2003
- [11] 『Protein Structural Change Upon Ligand Bindin: Linear Response Theory』  
Mitsunori Ikeguchi, Jiro Ueno, Miwa Sato, and Akinori Kidera  
Physical Review Letters, 2005
- [12] 『生命科学史』  
遠山益 著
- [13] 『絵とき タンパク質と遺伝子』  
池内俊彦・畠中寛 共著
- [14] 『タンパク質の生命科学 - ポスト・ゲノム時代の主役 -』  
池内俊彦著  
中公新書
- [15] 『酵素反応のしくみ - 現代化学の最大の謎をさぐる - 』  
藤本大三郎 著  
講談社
- [16] 『遺伝子の窓から見た動物たち - フィールドと実験室をつないで - 』  
竹中修 企画、村山美穂 / 渡邊邦夫 / 竹中晃子 編  
京都大学学術出版会
- [17] 『THEORY OF HEAT』  
James Clerk Maxwell
- [18] 『Maxwell's Deamon 2 -Entropy, Classical and Quantum Information, Computing-』  
H.S.Leff and A.F.Rex  
Institute of Physics Publishing Bristol and Philadelphia
- [19] 『動力的非平衡性に基づく分子機械の研究』  
野崎綾子  
2005年度修士論文

- [20] 『タンパク質のかたちと物性』  
中村春木 / 有坂文雄 担当編集委員  
共立出版
- [21] 『生物物理学とはなにかー未解決問題への挑戦ー』  
曾我部正博 / 郷信 担当編集委員  
共立出版株式会社