次世代生命体統合シミュレーションソフトウェアの研究開発

Next-Generation Integrated Simulation of Living Matter

ISLiMソフトウェア研究開発報告会

2013 年 1 月 10 日(木) - 11 日(金)

東京大学 武田ホール

文部科学省委託事業

「次世代生命体統合シミュレーションソフトウェアの研究開発」

次世代生命体統合シミュレーションソフトウェアの研究開発 Next-Generation Integrated Simulation of Living Matter (ISLiM)

ISLiMソフトウェア研究開発報告会

開催日時: 2013年1月10日(木)-11日(金) (10:00開始) (受付:9:30開始)

会場:東京大学 武田ホール (本郷キャンパス、武田先端知ビル 5F)

主催:

文部科学省委託事業「次世代生命体統合シミュレーションソフトウェアの研究開発」

協賛:

次世代スパコンの創薬産業利用促進研究会 バイオスーパーコンピューティング研究会

プログラム

| 時間 | (分) | DAY1 (1/10) | | |
|-------------|-------|---|----|--------------|
| 10:00-10:15 | (15) | ●オープニング | | |
| | | 茅 幸二 プログラムディレクター | | |
| 10:15-10:45 | (30) | プロジェクト総合報告 | | |
| | | 姫野龍太郎 副プログラムディレクター | | |
| 10:45-11:55 | (70) | 分子スケール研究開発チーム成果報告 (1/2) | | |
| | | ・MM/CG プログラム μ 2-lib によるタンパク質の大規模 サンプリング | 木寺 | 詔紀 (理研/横浜市大) |
| | | \cdot Replica-exchange interface program: REIN | 宮下 | 尚之(理研) |
| | | ・全原子分子動力学計算ソフトウエア MARBLE の開発と多剤 排出トランスポーター AcrB への応用 | 池口 | 満徳 (横浜市大) |
| | | ・タンパク質カノニカル分子軌道法 ProteinDF の京対応版 プログラムの公開 | 佐藤 | 文俊 (東大) |
| 11:55-12:50 | (55) | □ ランチタイム | | |
| 12:50-13:50 | (60) | ◆ポスター展示説明および休憩 | | |
| | | 講演会場前(5F ホワイエ) | | |
| 13:50-14:45 | (55) | 分子スケール研究開発チーム成果報告 (2/2) | | |
| | | ・粗視化分子シミュレーションソフトウエア CafeMol の開発 と遺伝子動態研究 | 高田 | 彰二(京大) |
| | | ・酵素分子の柔らかさと触媒活性:ハイブリッド QM/MM 自由エネルギー法の開発 | 林 | 重彦 (京大) |
| | | ・Platypus-QM/MM の開発と応用 | 鷹野 | 優(阪大) |
| 14:45-15:35 | (50) | 細胞スケール研究開発チーム成果報告 | | |
| | | ・細胞スケール研究開発チーム成果統括報告 | 横田 | 秀夫(理研) |
| | | ・血小板細胞シミュレーターの開発 | 後藤 | 信哉 (東海大) |
| 15:35-16:20 | (45) | ◆ポスター展示説明および休憩 | | |
| | | 講演会場前(5F ホワイエ) | | |
| 16:20-18:05 | (105) | データ解析融合研究開発チーム成果報告 | | |
| | | ・データ解析融合研究開発チーム チーム成果統括報告 | 宮野 | 悟(東大) |
| | | ・大規模遺伝子ネットワーク推定とその応用 | 宮野 | 悟(東大) |
| | | ・大規模ゲノム多型データと表現型データを関連付ける新規 アルゴリズムの開発と、妥当性、有用性の検討 | 角田 | 達彦(理研) |
| | | ・大規模タンパク質ネットワーク推定とその応用 | 秋山 | 泰 (東工大) |
| | | ・生命体シミュレーションのためのデータ同化技術の開発 | 樋口 | 知之 (統数研) |
| 18:10-19:30 | | | | |
| | | 講演会場前(5F ホワイエ) | | |

プログラム

| 時間 | (分) | DAY2 (1/11) | | |
|-------------|-------|--|----------|----------------|
| 10:00-10:15 | (15) | ● DAY2 オープニング | | |
| | | | | |
| 10:15-12:00 | (105) | 臓器全身スケール研究開発チーム成果報告 | | |
| | | ・臓器全身スケール研究開発チーム成果総括報告 | 高木 | 周 (理研 / 東大) |
| | | ・次世代超音波治療機器開発のための生体内超音波伝播 シミュレータ (ZZ-HIFU) | 沖田 | 浩平 (日大) |
| | | ・大規模並列計算に適した流体構造 / 膜連成手法 (ZZ-EFSI) | 杉山 | 和靖(理研) |
| | | ・UT-heart によるサルコメア力学から心筋細胞構造を経て 心拍動にいたるマルチスケール解析 | 鷲尾 | 巧(東大) |
| 12:00-13:00 | (60) | □ランチタイム | | |
| | | | | |
| 13:00-14:00 | (60) | ◆ポスター展示説明および休憩 | | |
| | | 講演会場前(5F ホワイエ) | | |
| 14:00-15:45 | (105) | 脳神経系研究開発チーム成果報告 | | |
| | | ・神経系の大規模シミュレーション | 石井 | 信(京大) |
| | | ・NEST による局所回路シミュレーション | 五十旗 | 嵐 潤(理研) |
| | | Multi-scale simulation of subcortical visual systems with real-time robotic visuo-motor feedback | Jan | Moren(沖縄科技大) |
| | | ・大規模数理モデル構築プラットフォームによる視覚系数理 モデルの構築 | 臼井 科大 | |
| | | 「京」による無脊椎動物嗅覚系シミュレーション –背景と現状と将来への展開 | 神崎 | 亮平 (東大) |
| 15:45-16:30 | (45) | ◆ポスター展示説明および休憩 | | |
| | | 講演会場前(5F ホワイエ) | | |
| 16:30-17:45 | (75) | 生命基盤ソフトウェア開発・高度化チーム成果報告 | | |
| | | ・ISLiM ソフトウェアの高度化 | 泰地 | 真弘人(理研) |
| | | ・MD コアプログラムの最適化と性能評価 | 大野 | 洋介(理研) |
| | | ・合成可能な化学構造および反応スキームからなる大規模仮想 ライブラリの構築 | 船津 | 公人 (東大) |
| 17:45-18:00 | (15) | ●クロージング | | |

プログラムは都合により変更されることがあります。

| ご挨拶 | | |
|--|---------------------|-----|
| 茅 幸二 プログラムディレクター | | 1 |
| プロジェクト総合報告 ISLiM:ソフトウェア開発の現状と今後 | | |
| 姫野龍太郎 副プログラムディレクター | | 3 |
| 分子スケール研究開発チーム成果報告 | | |
| MM/CG プログラム μ 2-lib によるタンパク質の大規模サンプリング | 木寺詔紀 (理研/横浜市大) | 15 |
| Replica-exchange interface program: REIN | 宮下尚之 (理研) | 21 |
| 全原子分子動力学計算ソフトウエア MARBLE の開発と多剤排出トランス ポーター AcrB への応用 | 池口満徳 (横浜市大) | 29 |
| タンパク質カノニカル分子軌道法 ProteinDF の京対応版プログラムの 公開 | 佐藤文俊 (東大) | 33 |
| 粗視化分子シミュレーションソフトウエア CafeMol の開発と遺伝子動態研究 | 高田彰二 (京大) | 37 |
| 酵素分子の柔らかさと触媒活性:ハイブリッド QM/MM 自由エネルギー法の 開発 | 林 重彦 (京大) | 41 |
| Platypus-QM/MM の開発と応用 | 鷹野 優 (阪大) | 47 |
| 細胞スケール研究開発チーム成果報告 | | |
| 細胞スケール研究開発チーム成果統括報告 | 横田 秀夫(理研) | 53 |
| 血小板細胞シミュレーターの開発 | 後藤 信哉 (東海大) | 61 |
| データ解析融合研究開発チーム成果報告 | | |
| データ解析融合研究開発チーム チーム成果統括報告 | 宮野 悟 (東大) | 65 |
| 大規模遺伝子ネットワーク推定とその応用 | 宮野 悟 (東大) | 71 |
| 大規模ゲノム多型データと表現型データを関連付ける新規アルゴリズムの開 発と、妥当性、有用性の検討 | 角田達彦 (理研) | 77 |
| 大規模タンパク質ネットワーク推定とその応用 | 秋山 泰 (東工大) | 83 |
| 生命体シミュレーションのためのデータ同化技術の開発 | 樋口知之 (統数研) | 89 |
| 臓器全身スケール研究開発チーム成果報告 | | |
| 臓器全身スケール研究開発チーム成果総括報告 | 高木 周 (理研/東大) | 97 |
| 次 世 代 超 音 波 治 療 機 器 開 発 の た め の 生 体 内 超 音 波 伝 播 シミュレータ (ZZ-HIFU) | 沖田浩平 (日大) | 103 |
| 大規模並列計算に適した流体構造 / 膜連成手法 (ZZ-EFSI) | 杉山和靖 (理研) | 109 |
| UT-heart によるサルコメア力学から心筋細胞構造を経て心拍動にいたるマル チスケール解析 | 鷲尾 巧 (東大) | 115 |
| 脳神経系研究開発チーム成果報告 | | |
| 神経系の大規模シミュレーション | 石井 信 (京大) | 121 |
| NEST による局所回路シミュレーション | 五十嵐潤 (理研) | 133 |
| Multi-scale simulation of subcortical visual systems with real-time robotic visuo-motor feedback | Jan Moren (沖縄科技大) | 139 |
| 大規模数理モデル構築プラットフォームによる視覚系数理モデルの 構築 | 臼井支朗 (理研/豊橋技 科大) | 145 |
| 「京」による無脊椎動物嗅覚系シミュレーション -背景と現状と将来への展開- | 神崎亮平 (東大) | 151 |
| 生命基盤ソフトウェア開発・高度化チーム成果報告 | | |
| ISLiM ソフトウェアの高度化 | 泰地真弘人 (理研) | 159 |
| MD コアプログラムの最適化と性能評価 | 大野洋介 (理研) | 165 |
| 創薬プラットフォーム-大規模バーチャルライブラリの開発- | 船津公人 (東大) | 169 |

ポスター一覧

| M-1: Platypus-MM/CG マルチコピー・マルチスケール分子シミュレーション法開発の基盤とな るクラスライブラリ | 177 |
|---|-----|
| M-2: Platypus-REIN レプリカ交換インターフェースプログラム | 178 |
| M-3: MARBLE 全原子分子動力学計算 | 179 |
| M-4: CafeMol 粗視化モデル計算 | 180 |
| M-5: ProteinDF 密度汎関数法に基づくタンパク質全電子波動関数計算 | 181 |
| M-6: Platypus-QM/MM-FE ハイブリッド QM/MM 反応自由エネルギー計算 | 182 |
| M-7: Platypus-QM 量子化学計算 | 183 |
| M-8: Platypus-QM/MM 量子化学計算/分子動力学連成計算 | 184 |
| C-1: RICS 細胞シミュレーションプラットフォーム | 185 |
| C-2: RICS-Pre 細胞シミュレーションプラットフォーム (RICS) プリシステム | 186 |
| C-3: RICS-platelets RICS による血小板細胞シミュレーションと力学的作用による偽足的変形に ついての検討 | 187 |
| C-4: 血小板活性化モデルのための細胞内カルシウムイオン実証実験 | 188 |
| C-5: 分子動力学法による平衡状態におけるコフィリン修飾アクチンフィラメントのエネルギー解析 | 189 |
| O-1: ZZ-EFSI 全身ボクセルシミュレーション(ボクセル構造流体連成解析プログラム) | 190 |
| O-2: 血栓形成のマルチスケールシミュレーション - 血小板の血管壁への接着に関するタンパク質間相互作用の解析 - | 191 |
| O-3: ZZ-DOSE 全身ボクセルモンテカルロ重粒子線計算プログラム | 192 |
| O-4: ZZ-HIFU 集束超音波による低侵襲治療のための HIFU シミュレータ | 193 |
| O-6: UT Heart マルチスケール・マルチフィジックス心臓シミュレータ | 194 |
| O-7: 血管内の血流シミュレーションソフトウェア開発・検証研究 | 195 |
| O-8: 筋骨格系 3 次元有限要素シミュレータの開発 | 196 |
| O-9: 血管網シミュレーター 全身循環モデルを用いた頸動脈ステント留置術に関する血流シミュレーション | 197 |
| D-1: ParaHaplo ハプロタイプ関連解析に於ける統計検定を行うためのソフトウェア | 198 |
| D-2: NGS analyzer 次世代シークエンサーの出力データを高速に解析し,個人間の遺伝的差異や がんゲノムの突然変異を高い正確さで同定する | 199 |
| D-3: ExRAT 拡張 RAT 法による 2SNP 組合せの全ゲノム関連解析ソフトウェア | 200 |
| D-4: SiGN-BN ベイジアンネットワークによる大規模遺伝子ネットワーク推定プログラム | 201 |
| D-5: SiGN-L1 L1 正則化法による大規模遺伝子ネットワーク推定プログラム | 202 |
| D-6: SiGN-SSM 状態空間モデルによる大規模遺伝子ネットワーク推定プログラム | 203 |
| D-7: SBiP データ解析融合プラットフォーム | 204 |
| D-8: LiSDAS 生命体データ同化プログラム | 205 |
| D-9: MEGADOCK 網羅的タンパク質ドッキング解析プログラム MEGADOCK の開発と応用 | 206 |
| B-1/B-2: NEST による局所回路シミュレーション | 207 |
| B-3: VSM 全視覚系モデルによる視覚情報処理の解析 (視覚系シミュレーションのための共有プ ラットフォーム) | 208 |
| B-4: NeuroMorphoKit 神経細胞形態シミュレーションキット | 209 |
| B-5: IOSSIM カイコガ運動指令形成回路のシミュレーションの構築の試行 | 210 |
| B-5: IOSSIM 昆虫嗅覚系全脳シミュレータ | 211 |
| H-1: cppmd 大規模並列用 MD コアプログラム | 212 |
| H-2: LSV 大規模データ可視化ソフトウェア LSV の開発 | 213 |
| H-3: SPHERE Simulation Framework for Large scale Coupled Analysis | 214 |
| H-4: VLSVL 合成可能な化学構造および反応スキームからなる大規模仮想ライブラリ | 215 |

ご挨拶

茅 幸二 次世代計算科学研究開発プログラム プログラムディレクター



発表者紹介

昭和41年3月東京大学大学院理学研究科博士課程修了 昭和45年9月東北大学理学部化学科助教授 (昭和48年9月~昭和49年9月 米国ベル電話研究所研究員) 昭和56年4月慶應義塾大学理工学部教授 平成11年4月岡崎国立共同研究機構分子科学研究所長 平成11年4月慶應義塾大学名誉教授 平成16年4月独立行政法人理化学研究所和光研究所・中央研究所長 平成18年10月 次世代計算科学研究開発プログラムプログラムディレクター(兼) 平成20年4月 次世代スーパーコンピュータ開発実施本部 副本部長 平成20年5月 基幹研究所特別顧問(兼) 平成22年4月社会知創成事業次世代計算科学研究開発プログラム プログラムディレクター(兼)

研究分野

ナノ物質化学

ご挨拶

茅 幸二

次世代計算科学研究開発プログラム プログラムディレクター

京コンピュータが世界最速の記録を達成し、一年あまりとなり、すでに米 国では計算速度で京を超す計算機が完成している。この現況で、もっとも重要 なのは、大型計算機時代における、計算機利用をいかに深く突き詰めるか、そ してその成果を社会に還元できるかという点にある。

「次世代生命体総合シミュレーションソフトウエアの研究開発」は、発足以 来7年目となり、2013年3月に終了する。生命科学の広範な分野の我が国を代 表する計算科学の研究者が集い、京の利用による、生命科学における、医療、 あるいは生命機能の複雑なシステムの理解に向けた日夜を問わない熱い挑戦を された結果が、30本あまりのソフトウエアの大半がペタスケールの計算を可 能とする結果となった。 本プログラムに参加された多くの研究者そして、と くに学術論文への投稿の機会を逸しながら、きわめて高度なソフト開発に従事 された若手研究者の熱意と努力に心からの敬意と感謝を表する。

本報告会では、それらのソフト開発の成果を纏め、これからの高度計算によっ て、生命科学における定量化された成果がいかに意義あるものであるかを報告 し、計算科学でのライフサイエンス分野の今後の寄与に、生命科学のみならず、 多くの分野の方々に御理解と支持をいただくことを目的としている。

ISLiM:ソフトウェア開発の現状と今後

姫野 龍太郎

次世代計算科学研究開発プログラム 副プログラム・ディレクタ



発表者紹介

1979年3月 京都大学大学院工学研究科電気工学第二専攻修士課程修了

1979年4月 日産自動車株式会社入社、中央研究所、車両研究所勤務

主に車の空力シミュレーションの研究開発に従事

1998年1月 理化学研究所入所 情報環境室長

2003年10月 情報基盤センター長

2006 年 4 月-2008 年 8 月 次世代スーパーコンピュータ開発実施本部開発グループ、グループ・ディレクタ

2006年10月から、次世代計算科学研究開発プログラム、副プログラム・ディレクタ

研究分野

高性能計算、計算生体力学、計算流体力学

ISLiM:ソフトウェア開発の現状と今後

姫野 龍太郎 次世代計算科学研究開発プログラム 副プログラム・ディレクタ

1. はじめに

次世代生命体統合シミュレーションソフトウェアの 研究開発プロジェクト(ISLiM)は次世代スーパーコン ピュータ開発実施プロジェクトに含まれるライフ分野 のグランドチャレンジとして2006年10月から開始した。 このプロジェクトは理化学研究所を中核拠点として全 国13の機関から、およそ200名の研究者が参加した。 (13機関:東京大学、大阪大学、京都大学、東北大 学、北陸先端大学、東海大学、千葉大学、慶応大学、 横浜市大、東工大、広島大学、統計数理研究所、理 化学研究所)

ソフトウェア開発は、分子スケール、細胞スケール、 臓器・全身スケール、脳神経系、データ解析融合、 HPC の6チームに別れ、表1に示すように34本のソフ トウェアを開発した。HPC チームは他の5チームで開 発しているソフトウェアの高速化を援助、特に2011年 4月からの京へのチューニングでは大きな役割を担 った。

2. 開発の現状

我々の京の利用可能期間は京の供用開始まで ということになっていた。この供用開始は当初 2012 年 11 月末を予定していた。しかしながら、 この供用開始が9月中に早められた結果、我々の 利用可能期間が2か月短縮されてしまった。大部 分のソフトウェアでは京での大規模計算が可能 にはなったが、当初予定していた研究成果を出す ための計算を行う時間がとれなくなってしまった。非常に残念である。

図1は2006年からの開発を通して設定したマ イルストーンで管理した進捗状況である。開発を 3期に分け、第一期を計算アルゴリズムのチェッ クと並列化のテスト、第二期を理研で所有する 1028ノード8192コアのクラスタを利用した高並 列のテスト、第三期は京を使ったテストに位置づ けている。このようにして3ヶ月毎に進捗をチェ ックしてきた。その結果、34本中13本のソフ トウェアで1万ノードを超えるところまで性能 が向上する高性能さを実現できた(図2)。また、 cpmd, UT-Heart, ZZ-EFSIの3本は京全体を使 ってテストを行い、約3PetaFLOPS、ZZ-EFSI では 4.5PetaFLOPS という性能なソフトウェア に仕上がった(図3)。

3. 今後

現在、各ソフトウェアの公開に向けてマニュア ル(日本語・英語)やダウンロードサイト等の整 備を進めている。一部ソフトウェアに関しては一 般課題等として京の計算時間を確保し、引き続き 京での計算を行い、研究を継続している。今年度 末、このプロジェクトは終了するが、来年度以降 も引き続き、開発したソフトウェアが利用できる よう、整備し続ける予定である。



図3京全体を使ったテストでの性能

表1開発ソフトウェア一覧(全34本)

| _ | 122-1717支件 | 12 M |
|---|-------------------|---|
| | Farmer Lakibo | マムチョニー・フルチスターム分子をなしい~ ションは開発の単数上でもフラスライリラリ |
| | Phone NEW | ムゴリカを取りす数ださいますっさーフェイス |
| | MATHULE | 運動主任子動力学計算 |
| 1 | (SeliMu) | 和供生生产作的原 |
| 1 | | 医電影響和語に基づくすンパジ数を電子計算 開始計算 |
| | Fergin GMIMMINT | 1++21-+24 MARS. 8 81 5 - 818 |
| | Philippine GM | ALCAN A |
| | Planame CAN MAN | · 查卡卡学計算:"计子数大学計算 |
| | Painter - CBI-RBI | 副病也在外从引至 后子数大量計算 |
| 쁥 | HETE . | 補助シネコレーショングラオフォーム |
| | 12-018 | 検察市営営市5/ELL/一5/45/10学校市場建立 構造成局をつわせると |
| H | ti-mt | ※小線理レスメレーラ 1月の込み線表面によく毎小板間30251 |
| 1 | CC:0011 | 常料手能力者などもなーション |
| 2 | Hart. | ゆうちゃ 行き きをあつのうちん |
| F | 12-14/04 | 新加加速·後期得つきょいいう(100) |
| | Arman - | マルテスケール・マルキフィニ・クスの聞いてき ネージのロ |

| · 百利 |
|--|
| 1-3034222336110176四戸株字町1735- 第四5275027 |
| スタイシークエンス解剖プロプラム |
| 総合いいほど3-5200円目前やむタテノム目出版 新ジントウィア |
| · REREGIANIANO-74220734 |
| ・ うみがごかせれによる本自的にその大規模キット コンーク用きづけまうよ。 |
| M ダ気変動モデルによるMALのデーフかんが途出干 キットワーう地震プログラム |
| F-144000250034-6 |
| 2.001-1-100/04 |
| MARSHITA-AUPMENTEL |
| Notest Ethiotetest Full |
| Contrast Management Descenared on HEST |
| 金視覚系もそんによら視覚情和影視の能利(現象 系ジミュレーンがの75的分片有ブラットファーム) |
| 2 時間間の目的にない、シスンキット |
| 記念機関系を描いたメレーす |
| 決議議会員用地は13720754 |
| 計算書無大政理ドータ関係センスをム |
| マナリナーションコドルワルア |
| the set of |
| |







ISLiM ソフトウェア研究開発報告会(2013.1.10-11)





| | 整理後 ^{飯期目載} | 長期目標 | | | | | 」 他チームとの連携によ る効率化 | Q | G | - 算(CG) 本総 | くにしなげ、笛を思え、雷時酸代謝 | lc. tho | いヨンを実現 | | | | ストを続けな | 福 |
|-----------------------------------|---|---|---|---|--|---|--|---|------|---------------------|--|---|--|--|---|--|--|----------------------------|
| (2008年での整理 | 会議会会には、 「一般」の「日本のので、 「一般」の「「一般」の「日本のので、 「一般」の「一般」の「日本のので、 「一般」の「一般」の「日本のので、 「一般」の「一般」の「日本のので、 「一般」の「一般」の「日本のので、 「一般」の「一般」の「一般」の「一般」の「一般」の「一般」の「一般」の「一般」の | 1 回避セシュレーション 第4日前の株法サンプリング技術 動物が3-リーション統合プラかりフォーム | 田田市小葉 (株) (株) (株) (株) (株) (株) (株) (株) (株) (株) | 調整には、「「「「」」」」」」」」」」」」」」」」」」」」」」」」」」」」」」」」」 | 今 全身本ウセルデータ・全身力学モデル は音波伝搬シミュレーション 第 | マロロ 「日本の」 | 「「「」」」、「」」、「」」、「」」、「」」、「」」、「」」、「」」、「」」、 | a the second | 研究開発 | 寛(MM)・粗視化モデル計 | 発現過程のシミュレーション 発見過程のシミュレーション :多剤排出トランスポータ、 | デムの開発を進めるとと ! | マルチスケールシミュレー | 相視化モデル計算 | より大きく より大きく より 大き | より近似的 | 方法をプログラム化し、テ ^{はっている。} | ^{発報告会} 2011.1.23時 |
| 組む課題 | 共通的な開発するの設置 | | R 衣 家 の 画 | <u>「</u> 「 「 「 「 」 「 」 「 」 「 」 「 」 | ····································· | | L L L L L L L L L L L L L L | imy Jr.P.I. F. P. P. S. B. | スケール | 子動力学計 | 着売の あま の の ーゲット | ールでプログ | いたって | 分子動力学計算 | | | いぞれで計算 が出るまでにな | iMソフトウェア研究開 |
| チームで取り | GM/MM 全電子量子化学計算 ハイブリッドQM/MM法 分子動力学計算法 | 粗視化シミュレーション 蛋白質の構造サンプリング技術 細胞シミュレーション統合プラットフォーム | 肝 歯 昭(た 饕) 膵臓B 箱 昭 人 イソナャ サ テ | 日本 | く モン・ノーン・シート・シート・シート・シート・シート・シート・シート・シート・シート・シート | 至身木りでルトェー9・至身刀字モテノル 起音波伝搬シミュレーション 肺シミュレーション | 大規模遺伝子ネルワーク探索ソフトウェ 生命体データ同化プログラム タンパク質ドッキング解析フログラム | 大規模SNP解析リフトウェア HSL | 小小 | -化学計算(QM)·分 | F開発、タンパク質や う有機的な連携を図る | 、MM、CGの各スケ | たCM/MM、MM/CO | 量子化学計算 | | | MM/CGの3階層それ 並列程度まで性能が | ISI |
| | 整理前 計算 話 第 第 | 1 | 1111日11日11日11日11日11日11日11日11日11日11日11日1 | 4 | 計算対象 | | プログラム | 4 □ | Ń | 開発目標:量子 | 合代する技術を 胞スケールとの 素の応 | 来べた 開発計画: QM | の手法を結合しするためのプロ | | v | | <u>進捗状況:</u> QM/I がら開発、8000 | |
| | | | | | | | | | | | | | | | | | | |
| C. | 平成24年度 (2012) | 住能 チューー ソグ | | | | | | 計画 ⁵ | Ç | NEWR | 神経系 | ラメータ推 (モデル化) | 森口塑柱・ 柱・イオンチ オテ | 胞プラットフ 一ム供給 | 胞の変形 | 胞から高次 能までの一 | 析·高速化 | |
| C | 年度 平成23年度 平成24年度)) (2011) (2012) | 画・製造 チューニ ソグ | 東部 | | 楽 | | |)年度時点での計画 ⁵ | C | 超 | 器全身スケ 脳神経系 | 缘処理パラ <mark>パラメータ推</mark> −9推定 定(モデル化) | 栓形成 神経可塑性・ 変性・イオンチ ャネル | 数細胞から 細胞プラットフ 器へ・血栓 オーム供給 戎 | 環器系 細胞の変形 | 動制御 細胞から高次器制御 機能までの一 | minut ji | |
| 置づけ ジェクトの一部 | 平政21年度 平政22年度 平成23年度 平成24年度 (2009) (2010) (2011) (2012) | 試作・評価・製造 テューニ - プ | · · · · · · · · · · · · · · · · · · · | 第2回 中間評価委員会 | 通 | | 経験 | ^{駿告会} 註: 2010年度時点での計画 ⁵ | | - 4 间連携課題 | 胞スケール 職器全身スケ 脳神経系 ール | 、体分子ネッ 画像処理パラ パラメータ推 フークモデル メータ推定 定(モデル化) | での輸送モ 血栓形成 神経可塑性・ ル 変性・イオンチ ャネル | 細胞 複数細胞から 離胞ブラットフ 職器・ 自格 オーム供給 形成 | 流・分泌物 循環器系 細胞の変形 送 | 長・変形モ 運動制御 細胞から高次 ル 職器制御 機能までの一 | 析·高速化分析·高速化分析·高速化 | 報告会 |
| 研究開発の位置づけ コン開発プロジェクトの一部 | 平成20年度 平成21年度 平成22年度 平成23年度 平成24年度 (2008) (2009) (2010) (2011) (2012) | 详細設計 | 開発・製作・評価 | ·開評価委員会 第2回 中間評価委員会 | 開発:評価 実証 | 確設 | EX I WE EX | いフトウェア研究開発報告会 註:2010年度時点での計画 5 | | ダモしたナーム 同連携課題 | 子スケール 細胞スケール 臓器全身スケ 脳神経系 ール | ータ同化技 生体分子ネッ 画像処理パラ パラメータ推 トワークモデル メータ推定 定(モデル化) | 謝・膜 膜での輸送モ 血栓形成 神経可塑性・ デル 変性・イオンチ ャネル | 謝反応・膜 肝細胞 複数細胞から 細胞ブラットフ の物質輸送 一番 - 一本校 オーム供給 | —今提供・ド 血流・分泌物 循環器系 細胞の変形 ッグデリバリ 輸送 | 塑性や神経 成長・変形士 運動制御 細胞から高次 性に関する デル 職業制御 機能までの一 | | 「シートウェア研究開発報告会 |
| この研究開発の位置づけ 次世代スパコン開発プロジェクトの一部 | 18年度 平成19年度 平成20年度 平成21年度 平成22年度 平成23年度 平成24年度 7.2012) (2011) (2012) (2012) (2012) | 概念設計 詳細設計 | 開発・製作・評価 実証 | 第1回中間評価委員会 第2回 中間評価委員会 中間評価委員会 | 調発・評価 | 25.44 26.45 27.45 | 酸計 建設 | ISLMソフトウェア研究開発報告会 註:2010年度時点での計画 ⁵ | | 008年度に設定したナーム间連携課題 | | iがんと薬 データ同化技 生体分子ネッ 画像処理パラ パラメータ推 術 トワークモデル メーダ推定 定(モデル化) | ンパク賞-リ 代謝・膜 アル デル デル ポキャー 神経可塑性・ 変性・イオンチ オネル | (体分子ネッ 代謝反応・膜 肝細胞 複数細胞から 細胞ブラットフ フークデータ での物質輸送 藤器へ・血栓 オーム供給 アータ データ 形成 | (療画像デー データ提供・ド 血流・分泌物 循環器系 細胞の変形 ラッグデリバリ 輸送 | 経細胞デー 可塑性や神経 成長・変形モ 運動制御 細胞から高次 ・局所回路デ 変性に関する デル 聴発制御 機能までの一 | タ データ データ Wet and a manual m Manual manual manu | ISLMソフトウェア研究開発報告会 |
| この研究開発の位置づけ 次世代スパコン開発プロジェクトの一部 | 平成18年度 平成19年度 平成20年度 平成21年度 平成23年度 平成23年度 平成23年度 平成24年度 (2012) (2008) (2008) (2011) (2012) | | #代ナノ鉄合 ユレーション ユレーション | \$1回中間評価委員会 \$44±会比益会 第2回 中間評価委員会 | ヨバエのみを取らって、「「「」」、「」」、「」、「」、「」、「」、「」、「」、「」、「」、「」、「」 | 幕機模 | 资禁 授計 通過設 | ISLIMソフトウェア研究開発報告会 註:2010年度時点での計画 ⁵ | | 2008年度に設定したナーム间連携課題 | 受 データ解析融 分子スケール 細胞スケール 臓器全身スケ 脳神経系 E供 合 ール ール <td>データ解析融 <mark>肺がんと薬 データ同化技 生体分子ネッ 画像処理パラ パラメータ推</mark> さ もワークモデル メータ推定 定(モデル化)</td> <td>3子スケール タンパク賞・J 代謝・膜 膜での輸送モ 血栓形成 神経可塑性・ ガンド デル デル オント</td> <td>田池スケール <u>生体分子ネッ</u> 代謝反応・護 肝細胞 複数細胞から 細胞ブラットフ トワークデータ での物質輸送 聴器へ・血栓 オーム供給 データ</td> <td> </td> <td>当神経系 神経細胞デー 可塑性や神経 成長・変形モ 運動制御 細胞から高次 タ・局所回路デ 変性に関する デル 職業割御 機能までの一</td> <td>一タ データ データ データ 1000000000000000000000000000000000000</td> <td>ISLIMソフトウェア研究開発報告会</td> | データ解析融 <mark>肺がんと薬 データ同化技 生体分子ネッ 画像処理パラ パラメータ推</mark> さ もワークモデル メータ推定 定(モデル化) | 3子スケール タンパク賞・J 代謝・膜 膜での輸送モ 血栓形成 神経可塑性・ ガンド デル デル オント | 田池スケール <u>生体分子ネッ</u> 代謝反応・護 肝細胞 複数細胞から 細胞ブラットフ トワークデータ での物質輸送 聴器へ・血栓 オーム供給 データ | | 当神経系 神経細胞デー 可塑性や神経 成長・変形モ 運動制御 細胞から高次 タ・局所回路デ 変性に関する デル 職業割御 機能までの一 | 一タ データ データ データ 1000000000000000000000000000000000000 | ISLIMソフトウェア研究開発報告会 |

ISLiMソフトウェア研究開発報告会(2013.1.10-11)

| ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● | 開発目標:ペタフロップス級の計算量を用いて、数時間で実行可能な職器全身シミュレーションの構築を目指す。これにより、病態予測や治療予測等の医療支援ツールの開発に つなげる。 | 開発計画: 血管網、各種臓器、全身を3次元的に再現した臓器全身モデルを構築し、 病態予測や治療予測を行う臓器全身モデルを次世代スパコンに実装、循環器系疾患 を始めとする種々の病態のシミュレーションと、ガンなどの治療支援シミュレーション(放 射線・集束超音波など)につなげる。 | 第二、1、1、1、1、1、1、1、1、1、1、1、1、1、1、1、1、1、1、1 | 進捗状況・ 1mm分解能の全身ボクセルモデルを開発すると同時に新しい構造流体連成手法と多媒質内の超音波伝搬解析手法を開発した。これらにより、外科手術シミュレーションや収束超音波・重イオンビームによるガンの治療等へつなげる基盤を構築した。 2011.1.23時点 10 | シ 脳神経系研究開発 | 開発目標:脳全体レベルでの入出力関係を再現できるモデルとシミュレーションソフトウェアを開発、次世代スーパーコンピュータの性能を引き出して、脳の柔軟な環境適応能力である脳の発達と学習の再現を目指す。 | 開発計画: 神経細胞と局所回路のシミュレータ、および、脳のシミュレータとして昆虫の匂い情報 処理回路モデル、ほ乳類の網膜モデルとを開発、次世代スーパーコンピュータに適したプログラ ムとすることで、昆虫の嗅覚系情報処理では実時間で、ほ乳類の視覚系情報処理は60倍の時 間で計算可能とする。 | 神経マルチフィジクスシミュリータ 高所回路シミュリータNEST | ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● | 機能種性形成シミュレーション 進捗状況:本チームは2008年10月に発足、昨年度は神経マルチフィジックスシミュレータでは神経 細胞の極性形成について、分子レベルからのシミュレーションができた。また、局所回路シミュレー タNESTでは105個の神経細胞、105個のシナブスからなる大規模局所回路のシミュレーションがで きた。並列性能に関しては32000並列までのテストを終了 2011.1.23時点 ¹² |
|---------------------------------------|---|--|---|---|-------------------|--|---|--|---|--|
| ● 第七人一と母党開発 | 開発目標:ペタフロップス級の計算力を活かして細胞を100万のボクセルに区画、そこに 包括的な実証データを組み込んだ肝細胞・肝小葉のシミュレーションの実現を目指す。薬 剤等の作用や副作用の予測を行うシミュレーションの開発につなげる。 | 開発計画:生物の空間的情報を入れ込むことが出来るプラットフォームの開発と平行して、 プラットフォームに実装するためのデータ収集とモデルの構築を進める。また、プラット フォームと各種データ・モデルとの有機的連携を図る。 | | <u>進捗状況:</u> 細胞を約3万個のボクセルに分割、複数のオルガネラを表現し、移派抵散万程式による物質移動をシミュレーションできる細胞シミュレーションプラットフォームを開発、E-CELL3をベースにした代謝反応がオルガネラ間も含めてシミュレーションできるようになった.現在、1024並列までの性能テストを終了 2011:1:23時点 。 | データ解析融合研究開発 | ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ | ケーンノル国は11米キの天がにしゅこの。 開発計画:遺伝子ネットワーク推定・タンパク質ネットワーク推定・ゲノム多型解析技術等の研究、ならびにデータ同化技術の開発を進め、開発した解析法やデータ同化技術を次 世代スパコンに実装し、データ解析とシミュレーションを融合した技術を開発する。 | シミュレーションモデルの構築 関連するデータをモデルに同化 病気のシステム的理解による医療開発イノベーション | | (B3260)RE18(H)つ7CU/847/W) (B3220)RE18(H)つ7CU/847/W) (B3220)RE18(H)の10000000 (B3220)RD100000000000000000000000000000000000 |

| 🔊 開発ソフトウェア(全34本) | レントウェア名称 説明 レントウェア名称 説明 Platvouus-MM/CG マルチコビー・マルチスケール分子シミュレー ParaHaplo ハンロタイフ園連解析に於ける統計検定を行うた | mathematical ション法開発の基盤となるクラスライブラリ NGS analyzer 次世代シークエンス解析プログラム Platypus-REIN レブリ功交換分子動力学計算インターフェイス KERAT ボ州フトウェン ボ州フトウェン A MARBLE 金原子分子動力学計算 FLAT ボリフトウェア ボリフトウェア A MARBLE 金原子分子動力学計算 1 ボリフトウェア ボークェンス解析フログラム A MARBLE 金原子分子動力学計算 5 Sick-BN (Sick) 大規模違低子手術調ネットワーク推定プログラム ProteinDF 密度汎関数法に基づくタンパグ資金電子波動 第 Sick-BN (Sick) 大規模当小 () ProteinDF 國政計算 第 Sick-Li (Li CN) 7 | Platypus-OM/MM-FE ハイブリッドOM/MM 反応自由エネルメーギャン 10 5:001-5:05 (Bit Platy) - 501:00 (Bit Platy - 501:00 (Bit Platy) - 501:00 (Bit Platy - 501:00 (Bit | 身 | ◆ それぞれの成果アピール時期 | | | 第一走者 のpmd(分子動力学) 多剤排出レランスポーカ 神経系統合シミレータ のpmd(分子動力学) MARBLE 緒胞から臓器へ ProteimDF Platypus 分子から縮脂へ CateMol(ヒストンDNA) RICS:肝細胞 UTHeart HFU ZZ-DOSE重イナン粒子 | ZZ-EFSI SiGN SSM L1GNE ParaHaplo LiSDAS ParaHaplo MEGADOCK NEST(神経回路網) MEGADOCK Sid% 5*7cmu AFD/A た成後すぐに成果 完成後、1年以内に たアピール 売までの | |
|---------------------|--|--|--|--|-----------------|---|--|---|---|---|
| ★ 生命体基盤ソフトウェア開発・高度化 | 開発目標 :プロジェクトで開発される各種ソフトウェアの高性能化、特に次世代スーパーコンピュータ向けの最適化を行い、次世代スーパーコンピュータの生命科学への応用を促進すると同時にグランドチャレンジを実現する。 | 開発計画 :コアソフトウェア・基盤ライブラリの開発と他チームアプリケーションの調査・ 高速化に向けた検討を進め、次世代スーパーコンピュータ実機でのアプリケーションの 高速化、チューニングを行う。 | ^{住能} アムダールの法則 並列化度が上がるにつれ、わずかな未並列化でも 性能に大きな影響を及ぼす 2.3.1.cg | 進捗状況:本チームは2007年11月に発足、これまでに分子動力学計算高速コアソフトウェア (cppmd)の解析・開発、量子化学計算ソフトウェアの高性能化を行うと同時に、基盤ライブラリ・可祝 化ソフトウェアの調査・開発を進めた。Cppmdでは8000並列、可視化ソフトでも100並列での動作テ ストを行った。 | SLiMのアプリ開発状況 | Phase I < 256 Phase II < 8192 Phase II on K II-0 II-1 II-2 1024nds 12288nds 82944nds | Playpus-MMVCG Playpus-REIN MRRLE CateMo ProteinOF Playpus-CMMMM Playpus-CMMMM Playpus-CMMMM | ZZEF8 ZZHWG ZZHWG UTHean Prankapio NERAT SGNAS | MESANDOX NEST Nest COUNN VGSM Neurokophoki VGSM VLSV 0 1 2 3 4 5 6 7 | cppmd, ZZ-EFSI, UT-Heartの三つは2Peta FLOPSを超えた ¹³ |

ISLiMソフトウェア研究開発報告会(2013.1.10-11)





分子スケール研究開発チーム成果報告

分子スケールチーム

木寺詔紀

次世代計算科学研究開発プログラム 分子スケール研究開発チーム チームリーダー



発表者紹介

- 1982年3月 京都大学工学研究科高分子化学専攻博士課程退学
- 1982 年 8 月 Cornell 大学 化学科 博士研究員
- 1986年2月 京都工芸繊維大学 工芸学部 助手
- 1988年4月 蛋白工学研究所 主任研究員
- 1996 年 4 月 京都大学大学院 理学研究科 助教授
- 2001 年 4 月 横浜市立大学大学院 総合理学研究科 教授
- 2006 年 10 月 理化学研究所次世代計算機科学研究開発プログラム チームリーダー (兼 任)
- 2011 年 4 月 理化学研究所 HPCI 計算生命科学推進プログラム 副プログラムディレク ター (兼任)

研究分野

計算生物学、生物物理学

分子スケールチーム

木寺記紀 次世代計算科学研究開発プログラム 分子スケール研究開発チーム チームリーダー

1. 目的

生命現象の最下層にある生体分子の機能発現過 程のシミュレーションは、それ自体として時間、 空間の階層性を持っており、それぞれの階層に対 応したシミュレーション方法を必要とする。短時 間の比較的少数の原子が関わる化学反応を含む 高精度の相互作用のレベルでは量子化学計算

(QM)を用いる。より長時間のタンパク質の全 原子の動きのシミュレーションは、環境として存 在する水や脂質分子をあらわに取り入れた系で、 経験的な力場(MM)に基づいた古典分子動力学 計算を用いる。生理学的なレベルに近づく時間ス ケールでの現象をより巨大な系においてシミュ レーションする場合は、MM をさらに粗視化した モデル (CG) に基づいて、環境を陰的に考慮す る方法によって行う。分子スケールでは、生体分 子に関わるあらゆる現象を対象としたシミュレ ーションを可能とするために QM-MM-CG とい う3階層にわたる方法論を開発する。さらに、そ れらの階層を接続し、それぞれの方法の限界を超 えるマルチスケール法として、QM と MM とを連 成した QM/MM 法、さらに MM と CG を練成し た MM/CG 法を開発する。これらの方法論による シミュレーションをスーパーコンピュータ京の 性能を十分に生かして行うことを可能とするソ フトウェアを開発する。さらに、実証研究の対象 として、多剤排出トランスポーター、脂質代謝多 酵素複合体を分子スケールの共通ターゲットと して選び、スーパーコンピュータ京上で多階層で のシミュレーションを行うことで、これらのソフ トウェアの有用性を示す。

3. 現時点での成果

2.1 ProteinDF (QM)

巨大複雑系であるタンパク質の状態を最も精密 なレベルでシミュレートする目的で、タンパク質 全体を含んだ全電子カノニカル分子軌道計算を 行う。計算法には基底関数にガウス型関数を使用 する密度汎関数法(DFT)を採用した。実績のあ る汎関数 B3LYP を実装し、高度並列化を行って いる。現時点で 8,000 並列を達成している。適用 例として、インシュリン 6 量体、ミオグロビン、 チトクローム c などのカノニカル分子軌道を求め、 カノニカル分子軌道の重要性を示し、分子軌道と 分子機能との関わりを明らかにした。現在、光合 成アンテナ蛋白質部分構造の計算を行っている。

2.2 Platypus-QM/MM (QM/MM)

QM と MM を連成することで、QM/MM 分子動力 学シミュレーションを高精度で高速に実行する ことができる。*ab initio* 量子化学計算として RHF. UHF, R-DFT, U-DFT, CIS, CIS(D), CIS-DFT, MP2, CASSCF を組込み、基底状態だけでなく励 起状態解析にも対応し、電子移動や生化学反応を on-the-fly 計算によりシミュレートできる。さら に、Chain-of-state 法を利用して、反応の最小自 由エネルギー経路が得られる。QM 計算部は最大 32,768 コアの並列(並列化率 99.9888%)を、 Chain-of-state 法による QM/MM 計算では 98.304 コアの並列(Weak Scaling 効率 93.36%) を達成している。また、遠距離力を高精度で計算 できる Zero-Dipole Summation 法を高度化した。 適用例として発光蛋白質 Sirius の発色団の励起 状態解析による発光機構の解明と RNase HI の加 水分解反応機構の解析を行っている。

2.3 Platypus-QM/MM-FE (QM/MM)

生体分子機能に関わる化学反応をハイブリッド QM/MM 法を用いて解析し機能発現の分子機構 の解明を可能にする方法論を開発した。 QM/MM-FE 法では、MM 分子力場を用いた MD 計算による MM 領域の構造サンプリング を行い、その MM 分布によって定義される自由 エネルギー曲面上で QM 領域の反応経路探索を 行うことで、MM 領域のゆらぎの中で起こる酵素 反応を定量的に評価することができる。これまで に 700 ns の MD サンプルによる自由エネルギ ー曲面上での QM 構造最適化に成功している。 現在、京で 4000 並列を達成している。α-アミラ ーゼの酵素反応機構に適用し、タンパク質の構造 揺らぎの反応自由エネルギーへの寄与を明らか にした。

2.3 Marble (MM)

全原子分子動力学法のソフトウェアである。空間

分割、PME 法、シンプレクティック剛体法など の高効率・高精度のアルゴリズムを用いており、 現在、京で 8000 並列超を実現している。応用と して MD と NMR、X 線溶液散乱実験との連携技 術の開発を行った。また、共通ターゲットである 多剤排出トランスポーターの分子動力学シミュ レーションを行い、プロトン移動による構造変化 誘導,トランスポーターに対する薬剤の結合・排 出機構についての全原子シミュレーションを実 施した。

2.4 Platypus-REIN (MM/CG)

レプリカ交換分子動力学法を行うためのソフト ウェアである。レプリカ交換分子動力学法は異な った環境変数で行われる分子動力学計算で駆動 される多数のレプリカが確率的に他の環境変数 に移行することによって、通常の分子動力学計算 では到達できない広範囲の構造アンサンブルを 精密に求めることができる。ここではさらに、多 次元の環境変数を選択することでさらなるサン プリング空間の拡張をめざす多次元レプリカ交 換法を採用している。京で3万並列超を実現する とともに、溶液中の糖鎖構造のモデリングや蛋白 質によるペプチド認識機構の解明などに適用し た。さらに細胞膜中でのアミロイド前駆体のシミ ュレーションを行うことで、アミロイド形成過程 の初期のモデリングを行っている。「京」に資す るソフトウェアとして、ISLiM ホームページ上で ソースコードも含めて公開を行った。

2.5 Platypus-MM/CG (MM/CG)

MM と CG の二つのモデルを共役項で制約させな がらひとつのシミュレーションとして行う。それ によって、MMの精度とCGの効率を両立させる ことを目指している。主要な方法論は、マルチス ケール拡張サンプリング(MSES)法、On-the-fly ストリング法である。MSES 法は、上記のシミュ レーションをそのまま分子動力学計算として実 施し、最終的に CG からの影響のない MM の構造 アンサンブルを得るために、共役項の強度を変数 とする Hamiltonian 交換法を用いている。 On-the-fly ストリング法は、同様なシミュレーシ ョンを経路探索の問題として行う。CG のポテン シャルを flat にして、共役項の強度を大きくする ことによってCG自由度の上でのMM ポテンシャ ルの平均力場を求めることができ、最小自由エネ ルギー経路を得ることができる。MSES 法は天然 変性タンパク質 Sortase A の disorder-order 転移、 Barnase-Barstar のタンパク質間相互作用に、 On-the-fly ストリング法は Adenylate kinase の

構造変化に適用した。

2.5 CafeMol (CG)

粗視化ポテンシャルを用いて、タンパク質、核酸、 脂質を含んだ系の大規模、長時間の分子動力学シ ミュレーションを行った。これまでに、蛋白質、 DNA および脂質それぞれの単体系および核酸-蛋白質、脂質-蛋白質の複合体の粗視化モデルの 開発および実装を完了した。現在、京への実装と、 レプリカ交換法によって 98000 コアの並列を達 成している。これまでに、タンパク質のフォール ディング、キネシンの微小管上の歩行、多剤排出 トランスポーターの薬剤排出、ヌクレオソームの 動態などのシミュレーションに適用し、実験また は既存のシミュレーション手法では困難であっ た、それぞれの運動の機構を明らかにしてきた。 特に、多剤排出トランスポーターでは、機能的回 転の機構および薬剤取り込みの経路についての 知見を与えた。

2. プロジェクト終了時の達成目標

それぞれの方法論によるシミュレーションを京 の性能を十分に生かして以下の実証研究の目標 を以下にあげる。

ProteinDF

光合成反応中心、光合成アンテナ蛋白質のカノニ カル全電子計算

Platypus-QM

数万電子系の基底状態/励起状態の電子構造の計 算

Platypus-QM/MM

数千電子系とそれを取り巻く数 10 万古典原子に よる反応自由エネルギー地形の計算

Platypus-QM/MM-FE

多剤排出トランスポーターのプロトン移動の自 由エネルギー計算

MARBLE

多剤排出トランスポーターの薬剤排出過程のシ ミュレーション

Platypus-REIN

タンパク質の構造変化、構造予測および、タンパ ク質間の結合自由エネルギープロファイル(数万 ~十万粒子系)

Platypus-MM/CG

多剤排出トランスポーターの構造変化経路計算 脂質β酸化多酵素複合体の構造変化計算 CofeMal

CafeMol

ヌクレオソーム (DNA ヒストン複合体) 動態の シミュレーション

以上

ISLiM ソフトウェア研究開発報告会(2013.1.10-11)



18





Replica-exchange interface program: REIN

宮下 尚之

理化学研究所 生命システム研究センター 分子機能シミュレーション研究チーム 研究員



発表者紹介

2003 年 3 月 総合研究大学院大学数物科学研究科機能分子科学専攻博士課程修了

2003 年 4 月 東京大学 分子細胞生物学研究所 学術研究支援員

2005 年 4 月 ボストン大学 化学科 Research Associate

2007 年 4 月 岐阜大学 人獣感染防御研究センター 助教

- 2007 年 12 月 理化学研究所 次世代計算科学研究開発プログラム 研究員
- 2011 年 4 月 理化学研究所 生命システム研究センター 研究員
- 2011 年 4 月 理化学研究所 次世代計算科学研究開発プログラム 兼務
- 2011 年 4 月 理化学研究所 計算科学研究機構 兼務

研究分野

物理(理論)・生物物理・分子シミュレーション





ISLiM ソフトウェア研究開発報告会 (2013.1.10-11)



2年12月13日木曜日

node

\$

node

→ Hybrid parallelization MPI/OpenMP 1

ŧ

node

Language/ parallelization

MD Packages

Fortran/MPI Fortran/MPI

Ξ 2 $\overline{\mathbb{C}}$ 4 2 0 E

CHARMM

AMBER NAMD C/MPI/OpenMP

MARBLE TINKER

12年12月13日木曜

Fortran/MPI

C/MPI/OpenMP C/MPI/pthread

GROMACS DESMOND



| Very easy execution three steps | | | Example: In | put file of Inp | ut builder | |
|---|--|--|--|--|--|---------|
| <pre>1) Input_builder → rein.inp #input_builder ib.inp</pre> | | # control paramet rein_input | rs in input_builder = rein.inp # rein input file | lame | Temperature - distance 2D-REMD | |
| 2) <mark>Batch_builder</mark> → rein.sh #batch_builder bb.inp | | (CMBA) 1962ixe 100-2011 100-2011 100-2011 100-2011 | = temperature # axis (TBMPENATUR = 8 # 0f replicas fo = 300.0 300.0 # Lowest, highest = exp # (EOUTVER/CYCLE | LE/HARPHONIC) or the axis value | ex) Alanine Dipeptide REIN/NAMD, NVT MD steps: 1000 steps (2ps) | |
| <pre>3) submitting of job # pjsub rein.sh</pre> | | akis602 nulre91Lcas602 range602 divis10n602 springe02 springe02 | = harmonic = 0.08.0 = 2.08.0 = 40P Ct 1-1 ADP Ct 1-1 = 2.0 | | Exchange: 100 exchanges Number of Replicas: Temperature : 8 replicas Distance : 8 replicas Tetel: 8 v 8 - 64 replicas | |
| Restart #batch_builder bb.inp restart | | 100) nd_program nd_exe_file nd_temperature nd_sttps system_size_x | <pre>= namd # md program (NAMG = ./namd2 # md eve file = 300.0 # md temperature = 1000 # md steps / exch = 23.4 # system sizeau</pre> | VMARBLE) inge itomatically create pme grids | Range of Distance : 2.0-8.0 Range of Distance : 2.0-8.0 | oK Å |
| <pre>rein_converter → creating the trajectories #rein_converter rc.inp 13</pre> | series | system_size_y system_size_z | = 133.2 = 23.9 # | 4 | \$ input_builder ib.inp 実演 | |
| r:139:18:14:18:19 Fyamnle- innut file of Batr | a builder | 12年12月13日大曜日 | C | utout format | | 4 |
| | REIN: 32 core (4 node) | Replica#, | condition#, Temperature | (Distance etc), T | otal, Kinetic, Potential, Extra | |
| # control parameters in batch_builder rein_input = rein.inp # REIN input file name [BATCH] batchfile_name = rein.sh # batch filename batchfile_name = rein.sh # batch filename nomen meame * (ISMAL) AFOB: AFOB | Thread: 1 (flat MPI) Exchange steps : 100 exchanges | # -1 41 60 45 10 | EXCHANCE_STEP = 1 1 290.14150 2 314.56290 3 313.3160 5 3013.9160 5 | -1624.33470 -1624.33470 -1657.17940 -1658.91640 -1652.479160 | 370.15570 -1994,49040 -1994,49040 401.31190 -2010.25580 401.31190 -2060.225950 -388.59900 -2031.47760 | |
| <pre>nu_thread:_ind = 8 # number of threads in ad nu_timplerc_ind = 4 # # of mpl process for ad cpu_timp cpu_tim = 01:00:00 # cpu timp num_core_pu = 8 # number of core in cpu, for k = 8, for fx10 = 16, num_core_pu = 8 # number of core in cpu, for k = 8, for fx10 = 16, num_core_pu = 8 # number of core in cpu, for k = 8, for fx10 = 16, num_core_pu = 8 # number of core in cpu, for k = 8, for fx10 = 16, num_core_pu = 8 # number of core in cpu, for k = 8, for fx10 = 16,</pre> | K (fx10) or SGE(UGE) type batch script can be selected. Batch_builder create batch script for submission the job to job scheduler. \$ batch_builder bb.inp \$ batch_builder bb.inp restart | 0 - 0 - 0 - 0 - 0 - 0 - 0 - 0 - 0 - 0 - | 6 239, 4776 8 289, 4420 8 288, 6444 2 306, 7844 4 310, 8750 4 310, 8750 6 284, 3490 5 284, 3490 6 284, 3490 5 284, 3490 5 284, 3490 5 294, 5500 5 296, 5520 7 299, 5500 | 1655, 760.70 1655, 760.20 1653, 471.40 1664, 285.40 1652, 92830 1652, 92830 1652, 98420 1652, 98420 1652, 98420 1668, 18320 1668, 18320 1662, 94080 | 332.1296.30 - 2031.0800 383.22022 - 2067.791.30 386.2202 - 2067.791.30 384.5210 - 2056.65.340 396.5520 - 2056.65.30 396.5520 - 2056.5630 375.5560 - 2015.2156 375.5560 - 2015.21560 377.3510 - 2056.2520 378.17260 - 2016.45900 378.17260 - 2016.01340 | |
| 15 | | 미미나 다 아파 (1999) 다 아파 (1999) | | 9 | | ą |

[BATCH] batchfile_name job_scheduler queue_name

num_threads_md num_mpiproc_md cpu_time num_core_cpu

12年12月13日木曜日



ISLiM ソフトウェア研究開発報告会 (2013.1.10-11)



Summary

Replica-Exchange Interface Program (REIN)

•



全原子分子動力学計算ソフトウエア MARBLE の開発と

多剤排出トランスポーターAcrB への応用

池口 満徳 横浜市立大学 大学院生命ナノシステム研究科 分子スケール研究開発チーム



発表者紹介

- 1994 年 3 月 東京大学大学院農学系研究科応用生命工学専攻博士課程修了
- 1994年4月 東京大学農学部助手
- 2001年4月 横浜市立大学大学院総合理学研究科助教授
- 2009年4月 横浜市立大学大学院生命ナノシステム科学研究科准教授

研究分野

生物物理学、分子シミュレーション








タンパク質カノニカル分子軌道法 ProteinDF の京対応版プログラムの公開

佐藤 文俊

東京大学生産技術研究所 革新的シミュレーション研究センター 教授



発表者紹介

- 1990年 3月 慶應義塾大学大学院理工学研究科物理学専攻博士課程修了
- 1990年 4月 科学技術庁特別研究員
- 1991年 3月 九州工業大学情報工学部助手
- 2002 年 11 月 東京大学生産技術研究所客員助教授
- 2004 年 12 月 東京大学生産技術研究所助教授
- 2008年 4月 東京大学生産技術研究所教授

研究分野

計算生体分子科学、量子分子生物学

😽 東京大学 😽 惠京大学 基本的な計算機能は「イノベーション基盤シミュレーションソフトウェアの研究開発」 10 of Fe(II)⇔Fe(III)で電子移動 cvtochrome c 本プロジェクトでは、「京」対応のために新しい動作形式を研究開発 85.2%: Feのd軌道 機能・性能がどこまで達成しているか ・ MPI / OpenMPでのハイブリッド並列で動作確認 ・ ブランチを作らずに、様々なハードウェアに対応 - 分散並列計算機における数チハイブリッド並列計算 ISLIMソフトウェア研究開発報告会 背景·日的 - 「京」での演算性能向上・チューニング作業中 分子積分・数値積分の単体性能の向上 ProteinDFIこよるタンパク質の分子軌道計算 ProteinDFの京対応版プログラムの公開 MbCO MO of 1024x8並列まで動作確認 99.9%: Heme+近位His Fe(III)で失活 プロジェクトで開発(済み) 94.7%: Feのd軌道 - 「京」での正常動作 - バージョン徳田 性能 弯星 回咒 機能 SL UN 😽 東京大学 🍾 東京大学 - タンパク質のカノニカル分子軌道を計算する超大規模量子化学計算プログラム タンパク質カノニカル分子軌道法ProteinDFの A-branch 京対応版プログラムの公開 大規模SCF計算を成功させるための高精度初期値作成法 革新 的シミュ レーション 研究 センター 分子積分にKI法やKT法アルゴリズムなどが選択可能 東京大学生産技術研究所 既要・アプローチ ISLIMソフトウェア研究開発報告会 A-branch 佐藤文俊 教授 光合成反応中心タンパク質の分子軌道 ・アクセサリーバクテリオクロロフィル HE, DFT(含ハイブリッド汎関数) ・Pから4Å以内のアミノ酸残基 ・、、クテリオフィオフィチン エネルギー勾配計算 - C++ドロードイング - ガウス型基底関数

ISLIMソフトウェア研究開発報告会

ISLIMソフトウェア研究開発報告会

・スペシャルペア

主な機能と特徴

I

ProteinDFとは

ISL AN

ISL UN



| 4 (C) | 担当 超々並列 プレ・ポスト処理 複雑計算収束 自由エネルギー・構造作成 超大規模用交換相関積分 | o |
|---------|---|------------|
| 参考文献 | 描画 H19-20 H19-23 H20-23 H21- H21- | 了研究開発報告会 |
| · 提耀 | 「「」 「」 < | ISLIMVZFウェ |
| NY: Yaa | 参平西上恒松 ● 御野村村川田 京秋康典直潤 速子を 予報 「 | |

粗視化分子シミュレーションソフトウエア CafeMolの開発と遺伝子動態研究

高田彰二

京都大学大学院 理学研究科生物科学専攻



発表者紹介

- 1990年3月 京都大学理学研究科化学専攻修士課程修了
- 1991 年 3 月 岡崎国立共同研究機構分子科学研究所技官
- 1998年4月 神戸大学理学部化学科講師·助教授
- 2007年3月 京都大学理学研究科生物学専攻准教授(現在に至る)

研究分野

理論生物物理学



ISLiM ソフトウェア研究開発報告会(2013.1.10-11)



ISLiM ソフトウェア研究開発報告会(2013.1.10-11)



酵素分子の柔らかさと触媒活性:

ハイブリッド QM/MM 自由エネルギー法の開発

林重彦

京都大学大学院理学研究科 分子スケール研究開発チーム 開発責任者



発表者紹介

- 1997 年 11 月 京都大学大学院理学研究科化学専攻博士課程修了
- 1998年4月 日本学術振興会特別研究員(名古屋大学大学院理学研究科)
- 2000年4月 米国イリノイ大学アーバナシャンペーン校 博士研究員
- 2003年6月 京都大学福井謙一記念研究センター 研究フェロー(博士研究員)
- 2003 年 10 月 科学技術振興機構 PRESTO さきがけ研究者
- 2005年5月 京都大学大学院理学研究科化学専攻 准教授 現在に至る

研究分野

理論化学、生物物理学



ISLiM ソフトウェア研究開発報告会 (2013.1.10-11)

| アシンシンシンシンシンシンシンシンシンシンシンシンシンシンシンシンシンシンシン | Contention of a continuum electron density Conventional: hard to implement for continuum electron density Content electron density Content electron density Content electron density Convention electron density Content electron density <l< th=""></l<> |
|---|--|
| Mond A MM の計算時間の大きなギャップ 100 days with 1 fs / step = 8.64 Ms 100 days with 1 fs / step = 8.64 Ms Trajectory time Computational time/ step 10 ps 864 s = 14.4 m 0 ps 864 s = 14.4 m 10 ps 864 s = 14.4 m 10 ps 865 computational cost 1 ns 8.6 s computational cost 0.86 s 0M>>>Mn / step 1 ns 0.86 s 0M>>>>Mn / step 1 ns 120 step / s Anton 1 ms 120 step / s | $ \begin{array}{c} \hline & \\ \hline \\ \hline$ |





| まとめ ため <l< th=""><th> GMMMM 目由エイルナー法でののGMMMM KWFE-SUF 法 を開発した。 MD シミュレーションによって得られる MM サンプルによる 自由エネルギー面上 での反応経路解析が可能となった。 QMMM 計算と MD サンプリングが完全に分離: MD シミュ レーションにどんなプログラムを使っても良い (AMBER, レーションにどんなプログラムを使っても良い (AMBER, CHARMM (クラスタ版) 分子 力場使用可)。 「京」上で 4000 並列まで動作検証済み。現在、もう少し大き な茶へのチューニングを実施中。 タンパク質の特徴的な揺らざは、酵素活性に重要 (かも?) </th><th>ISLiMソフトウェア研究開発報告会</th></l<> | GMMMM 目由エイルナー法でののGMMMM KWFE-SUF 法 を開発した。 MD シミュレーションによって得られる MM サンプルによる 自由エネルギー面上 での反応経路解析が可能となった。 QMMM 計算と MD サンプリングが完全に分離: MD シミュ レーションにどんなプログラムを使っても良い (AMBER, レーションにどんなプログラムを使っても良い (AMBER, CHARMM (クラスタ版) 分子 力場使用可)。 「京」上で 4000 並列まで動作検証済み。現在、もう少し大き な茶へのチューニングを実施中。 タンパク質の特徴的な揺らざは、酵素活性に重要 (かも?) | ISLiMソフトウェア研究開発報告会 |
|--|--|-----------------------|
| Water Molecules Are Squeezed Out < | | ISLIMソフトウェア研究開発報告会 13 |

Platypus-QM/MMの開発と応用

鷹野 優

次世代計算科学研究開発プログラム 分子スケール研究開発チーム



発表者紹介

- 2002年3月 大阪大学大学院理学研究科化学専攻博士後期課程修了 博士 (理学)
- 2002 年 4 月 日本学術振興会海外特別研究員
- 2004年4月 日本学術振興会特別研究員
- 2005年6月 大阪大学蛋白質研究所 助手
- 2007年4月 大阪大学蛋白質研究所 助教

研究分野

計算化学、量子化学、生物物理学、蛋白質計算科学





| 🚵 💦 RNase HI | MMDを利用した破結合(chain of state) 超並列シミュレーションによる酵素反応機構研究 の mbe fly QM/MD の mbe fly QM/MD 0 | 注回研究者 中村春木(阪大蛋白研 中村春木(阪大蛋白研 米澤康滋(阪大蛋白研 山中秀介(阪大蛋白研 山中秀介(阪大蛋白研 山中多介(阪大蛋白研 |
|--------------|---|---|
| ◎ 線形応答関数解析 | (法: このいいいいいいいいいいいいいいいいいいいいいいいいいいいいいいいいいいいい | 第二十二章 中国 第二十章 中国 < |

細胞スケール研究開発チーム成果報告

細胞スケール研究開発チーム 成果統括報告

横田 秀夫

次世代計算科学研究開発プログラム 細胞スケール研究開発チーム チームリーダー



発表者紹介

1993 年 3 月 日本大学大学院農学研究科畜産学専攻修士課程修了

1999 年 3 月 東京大学大学院工学研究科精密工学専攻 博士授与

1999 年 4 月 理化学研究所素形材工学研究室協力研究員

2003 年 10 月 理化学研究所 VCAT 開発チームリーダー

2006年10月理化学研究所細胞スケール研究開発チームリーダー兼務(現在に至る)

2011 年 10 月 理化学研究所基幹研究所生物情報基盤構築チームヘッド(現在に至る)

2012 年 4 月 理化学研究所計測情報処理研究チーム副チームリーダー(現在に至る) 北海道大学大学院 工学研究科 客員教授

神戸大学 大学院工学研究科 システム情報学研究科 客員教授 東海大学 総合医学研究所 循環器内科学 客員教授

東京農工大学大学院 工学府 客員教授

研究分野

バイオイメージング 細胞内画像処理 バイオメディカルシミュレーション 細胞スケール研究開発チーム成果統括報告

横田 秀夫

次世代計算科学研究開発プログラム 細胞スケール研究開発チーム チームリーダー

1. 目的

細胞や臓器の部位特異的な生物学的情報を系 統的に実装させて生命現象の予測や実験仮説構 築に利用できるモデル開発に必要なソフトであ る細胞シミュレーションプラットフォームを完 成させ、これを超高速計算機で稼動させること により、細胞や臓器単位の機能的形態学的応答 の再現や病態メカニズムの理解や制御法の開発 に資するシミュレータを開発することを目的と する。細胞生物学の目的は細胞内の現象を理解 し説明することにあるが、細胞内の現象は未だ 理解されていない現象が多く残されており、定 量的な解析が可能な現象はごくわずかである。 一方、シミュレーションはある規則に則り、数 値解析して現象を再現することに他ならず、現 象の理解、数理モデルの構築無しに細胞のシミ ュレーションの実現はあり得ない。この圧倒的 に情報が少ない『細胞』を対象としたシミュレ ーションに於いては、計算科学としてのアプリ ケーションソフトウエアの開発だけでは、意味 のある細胞現象のシミュレータを構築する事は 不可能である。そこで、シミュレータの開発と 同時にシミュレーションのための実験を行う事 により、新たな細胞シミュレータの開発を目的 とした。さらに、近年進展が著しい各種イメー ジング技術により取得した細胞内における物質 の局在の定量データを用いることにより、細胞 シミュレーションの解析結果の検証を行い、実 際の細胞内で生じている現象を再現出来るシミ ュレータの開発を目指す事とした。

これまでの細胞シミュレーションでは、細胞 内生じている現象のモデル化と個別の現象のシ ミュレーションが行われてきた。これらの研究 の成果により、細胞の膜機能(物質の透過)、物質 の拡散、生化学反応について、その数理モデル の構築が実施されてきた。しかしながら、従来 の研究では、細胞内や膜の場所を考慮すること は不可能であった。一方、細胞の中ではさまざ まな現象が複雑に影響し生命が成立している。 細胞の中では、物質の濃度勾配やオルガネラ等 によってその機能の役割分担がされていると共 に、臓器内部における環境によって同じ細胞で も挙動が異なっている。疾患時においては、細 胞の応答やその周囲の環境の異常を再現するこ とが必要であり、これらの現象の再現のために は、複数の事象と空間(場)を考慮したシミュ レーションを連成して解析することが必要であ る。我々は、正常時の細胞の挙動や異常な挙動 である疾患の再現、さらにはその治療法に関す るシミュレータの開発を目指して研究を推進し ている。なお、具体的課題として、ターゲット を赤血球と肝臓細胞、血小板、膵細胞に限定し、 細胞内外の代謝、膜透過、物質の拡散の現象を 連成して解析することを目的とした。また、こ れらの対象は、細胞内のオルガネラを考慮した 細胞内現象のシミュレーションと複数細胞計算 が必要なことに加えて、将来的な計算機能力の 向上をにらんだ臓器や全身レベルでのシミュレ ーションの開発を目指している。

1. ソフト研究開発の概要・アプローチ

2.1 細胞シミュレーション統合プラットフォ ーム(RICS)の概要

細胞シミュレーション統合プラットフォーム (Riken Integrated Cell Simulator: RICS)は、細胞 内の場や複数のシミュレーションの連成を考慮 した共通基盤である。固定格子の空間内に細胞 の形状モデルや反応モデルを構築し、様々な細 胞反応のシミュレーションを可能にする。具体 的には細胞を100万のボクセルに区画し、区 画したボクセルに実測データより得られた細胞 内の物質量・移動量などの情報を取り込み、代 謝のパスウェイを設定することで細胞内現象を シミュレーションする。

2.2 RICS のアプローチ

RICS は、細胞内の生化学反応と物質拡散の連 成解析、物質の膜透過、膜電位計算、移流拡散 を弱連成の形式で連成解析する。RICS で計算可 能な形状を作成するには 2 つの方法があり、一 つは CAD を用いてポリゴンの細胞・オルガネラ 形状を作成し、RICS プリシステムを使用してボ クセル化する方法、もう一つは実際の顕微鏡画 像から得られる3次元断層画像を再構築し、オ ルガネラ等の領域のラベル化を行った後、ボク セル化する方法である。CAD と顕微鏡画像を組 み合わせて形状を作成することも可能で、特殊 な状況を作り出し、シミュレーションすること が可能である。細胞の反応にはこれまでの細胞 シミュレーションにおいて実績のある E-Cell3 システムを使用し既存の生化学反応モデルを使 用できるだけでなく、SBML2.0 などの他の言語 で書かれたものも既存のコンバータを介して利 用することを可能とした。 RICS は大規模並列 計算機による実行を考慮して作成されており、 PCクラスターから、「京」を含んだ多くの Linux ベースの高速計算機で実行可能である。これま でに 8192 並列までのスケービラリティを確認 しており、現在「京」での超高並列計算に向け たチューニングを進めている。

- 2. 現時点でのソフト研究開発成果
 - 3.1 RICS の開発状況

作成したシステムを用いて実際の顕微鏡デー タから得た細胞形状を用い、内部でのカルシウ ムイオン(Ca²⁺)の移動、反応、膜輸送担体をシミ ュレーションした。細胞形状はヒト肝臓由来細 胞である HepG2 細胞株を用い共焦点レーザー顕 微鏡を使用して取得した細胞形状および核とミ トコンドリアの形状からなる。顕微鏡画像から 3 次元 Volume データを作成し、細胞内の反応と して 24 種類の物質と、11 種類の生化学反応を 設定した(図 1)。



図1:共焦点レーザー顕微鏡から RICS プリシス テムを使用して再構築したヒト肝臓細胞の形状



図 2: RICS を用いて計算した細胞内 Ca²⁺の移動

Ca²⁺だけを通過させるチャネルを細胞膜の一部に局在化させて Ca²⁺を流入させて解析した (図 2)。細胞内の Ca²⁺の緩衝反応を設定した場 合は、生化学反応が無い場合に比べ細胞内 Ca2+ のみかけの拡散速度の低下がみられ、実際の細 胞内に近い Ca²⁺の動態が再現された。

3.2 肝細胞シミュレーション

肝細胞の解糖系、糖新生、グリコーゲン生成と 消費、フルクトース代謝、尿素回路、ミトコン ドリアクエン酸回路、酸化的リン酸化、脂肪酸 合成と分解、アルコール代謝、グルタミン/グ ルタミン酸代謝、メチオニン代謝、プリン核酸 代謝など、糖とアミノ酸を中心とした代謝経路 を調査し、個々の酵素反応速度に基づいて大規 模な代謝モデルを構築した。実証実験により、 さまざまな環境下での代謝物の定量計測を行い、 その数値情報を基に代謝モデルを構築した。現 在、485変数、582の反応からなる詳細な肝細胞 の代謝モデルを構築することに成功した。肝手 術における虚血とその後の再灌流では、クラン プと再灌流による代謝機能低下が重篤な不良を 起さない限度として、15分というクランプ時間 が慣習的に使用されている。そこで、細胞外の 酸素濃度と栄養としての代謝物質濃度、細胞膜 を介した膜輸送が起こる代謝物質濃度について 一時的に 15 分間 50%、10%、1%に低下させる摂 動を設定し、代謝変動を計算した。また肝臓の 低酸素・低グルコース応答に必要な代謝システ ムから、糖代謝、ミトコンドリア機能を実装す ると共に、血管の上流、下流の細胞モデルを構 築して、酸素やエネルギー代謝の heterogeneity を勘案した類洞モデルをほぼ完 成させた。門脈型(P型)と中心静脈型(C型) の細胞シミュレーションを連結することにより、 虚血および再灌流に対する主要代謝物の濃度変 動が顕著に buffering されて robustness が獲得 できるようになることが明らかになった。さら に、細胞レベルの精緻さを保持した臓器代謝シ ミュレーションの妥当性を検証するために先端 的バイオイメージングを利用した肝細胞・類 洞・肝小葉レベルの代謝情報収集を研究室で独 自に開発した2光子レーザー顕微鏡を用いて実 施し、肝細胞索の画像情報を収集し、RICSによ るモデル構築に利活用した。また肝臓微小循環 系の上流と下流にあたる門脈終末枝と中心静脈 における血流速度計測と酸素分圧計測を大腸が ん肝臓転移モデルにおいて実施し、門脈域に形 成された微小転移病巣の大きさと酸素消費の間 の相関関係を明らかにした。

3.3 血小板シミュレーション

細胞シミュレータープラットフォームに、実証 実験により取得した血小板細胞内のミトコンド リア、ゴルジ体、小胞体、濃染顆粒、血小板膜 特異的受容体蛋白などの局在情報を取り入れた 血小板細胞基盤モデルを作成した。刺激を受け 活性化した血小板にて起こる濃染顆粒の放出機 構の再現を目指した。作成したモデルにより、 直径 2 µm の血小板細胞表面の数十 nm2 程度の受 容体刺激の血小板細胞局所の細胞応答(細胞内 シグナル産生)の細胞内拡散による血小板細胞 全体の活性化過程(GPIIb/IIIaの活性化構造転 化)と、活性化の中途にて濃染顆粒の放出が起 こり、放出された ADP による P2Y12 ADP 受容体 刺激が加算された時の血小板細胞全体の活性化 過程の差異を見いだした。このことにより、放 出反応を介する活性化刺激伝達促進効果の論理 的予測が可能となった。血管壁の損傷部位にお ける、止血に向かう血小板細胞の positive feedback 機構を論理的、定量的に予測、説明す る初めてのモデルとなった。



3 D 血小板モデル 血小板内物質の局在 3.4 小胞動態モデル

全反射蛍光顕微鏡による膵臓β細胞のインスリ ン顆粒動態の画像解析から、グルコース刺激に よるインスリン分泌では、restless newcomer が主様式であることが判明した。その動態を元 にインスリン顆粒が細胞膜に融合する過程をモ デル化した。1)細胞内部では顆粒がゴルジ体 から分離し、細胞膜に接近するまでは細胞骨格 に沿って移動する。2)細胞膜付近ではブラウ ン運動に近い挙動を示し、細胞膜に融合した後、 インスリンを分泌する。これらの顆粒動態の違 いから、細胞内を外部層と内部層の2つに分割 してモデルを構築した。構築したモデルは、全 反射顕微鏡(TIRF)による分泌とほぼ同様の傾向 を示した。現在、このモデルをRICSに搭載する ことで小胞の3次元解析を計画している。

3. ソフト研究開発の達成目標

基本アプリケーションである RICS の機能開発 は終了し、現在「京」での動作確認と計算性能 向上を図っており、平成 24 年度には、当初目的 の計算を実現できる見込みである。さらに、開 発したアプリケーションを用いて、生命科学の 実問題を計算することに注力している。具体的 な課題は、上記に述べた肝臓、血小板、膵臓の 細胞内外の物質の反応と挙動である。これらの モデル化を完成すると共に「京」での大規模計 算に適した計算モデルに改変する予定である。 この作成したモデルにより、臓器内における細 胞の役割、機能分担を再現することが出来ると 共に、疾患時における細胞内の状況、薬物投与 時の反応を再現することが可能となり、治療や 薬物開発に貢献出来ると考える。さらに、血小 板や膵臓β細胞、腎臓細胞、脳神経細胞など局 所的な反応が全体に強い影響を及ぼす計算を行 いたいと考えている。

開発したソフトウェアとモデルは、プログラム 終了までに無償公開を計画している。現状の大 規模計算システムでは、情報科学の研究者が生 命現象を解析しているが、研究の裾野を広げる ためには、生命科学の研究者が「京」を使いこ なして細胞のシミュレーションを実施する環境 を整えることが必要である。そこで、細胞スケ ール研究開発チームとして、RICS の利用範囲を 拡大することを目的に周辺アプリケーションの 開発を進めている。具体的には、シミュレーシ ョンの形状モデルの構築、物質分布の配置、オ ルガネラの場所と反応式の対応などのシミュレ ーションセッティングを設定するプリポストシ ステムを開発している。このシステムは、関連 する外部のアプリケーションである、VCAT(形 状処理)、E-CELL IDE (反応式)、V-Sphere∷CBC (流体解析)、V-Isio(可視化)を連結してシミ ュレーションのモデル構築をGUIのみで実現 するものである。これらのアプリケーション開 発と共に、細胞シミュレーションモデルの整備 と公開に加えて、RICS ユーザー会を設立して、 細胞シミュレーションに関連する研究者コミュ ニティ作りを実施する予定である。プロジェク ト終了後には、開発し公開する基本細胞シミュ レータを元に、個別の臓器に分化したシミュレ ーションに関する研究が花開くことを期待する。 5. 謝辞

本発表の結果は、理化学研究所が実施してい る京速コンピュータの試験利用と理化学研究所 情報基盤センター(RICC) RIKEN Integrated Cluster of Clustersの計算による。



ISLiM ソフトウェア研究開発報告会(2013.1.10-11)



| Summary | PICS (Riken Integrated Cell Simulation Platform) |
|-----------------------|--|
| Liver cell simulation | Participants of the second sec |
| SCL MALLER | Perisinusoid Capillary Portal vein Hepatic artery |



C.

血小板細胞シミュレーターの開発と血栓 シミュレーターへの応用

後藤信哉

東海大学医学部内科学系循環器内科学 教授 東海大学バイオ研究医療センター代謝疾患研究センター長



発表者紹介

- 1986年3月 慶應義塾大学医学部卒業
- 1990 年 3 月 慶應義塾大学大学院博士課程内科学専攻単位取得退学
- 1992年3月 博士(医学:慶應義塾大学)
- 1992 年 7 月 スクリプス研究所分子実験医学部門博士研究員
- 1996年4月 東海大学医学部内科学 | 教室助手
- 1998年4月 東海大学医学部循環器内科講師
- 2002 年 4 月 東海大学医学部内科学系循環器内科部門助教授
- 2007 年 4 月 東海大学医学部内科学系(循環器内科学)教授
- 2008 年 4 月 東海大学大学院バイオ研究医療センター代謝疾患研究センター長兼任 東海大学総合医学研究所代謝システム医学部門長兼任

研究分野

循環器内科学、血小板生物学、システム医学、薬剤開発

| ₹ | ・ 背景 「血小板細胞」は、分裂、増殖せず、生理機能が止血、血栓形成に特化している。血小板細胞の活性化は、細胞度上の受容体刺激開始後数分の時間経過にて完結する。シミュレーンョンすべき生理機能は単純、かつモデル化が容易である。直径-5 µmの血小板細胞数個を対象としたマイクロメートルスケールの実証実験により精緻化できる。 しかし、「血小板細胞」は心筋梗塞、脳梗塞などの動脈硬化血栓性疾患の発症において 重要な役割を満じる。アスピリン、クロビドグレルなどのサノメールムンテーレの分子としての抗血小板薬により、筋梗塞再発予防効果は臨床的に証明されている。 しかし、「血小板細胞」は山球細胞として、血管内を常に全身循環している。血管壁損傷部位への接着、集積したの以上を対象とする臓器全身スケールのイベントである。 ・ ロかし、「セントレスケールの「血小板細胞」に注目し、血小板細胞を構成する。 ンメートルスケールの分子の細胞内局在変化により、活性化に伴う機能変化を再現する「血小板細胞」代謝、シグナルンニーションを作成する。 ・ 他小板細胞」を中心に、血管機能の構成論の構成論的理解を目指す ・ 血小板細胞」を中心に、血管機能の調節機構の構成論的理解を目指す | 現在までの研究開発成果 現在までの研究開発成果 現時点の開発ソフトウェアについて ・ オルガネラの局在を組み込んだ「血小板細胞」形状モデルは確立できた。また、実証実験をした。 * はいちがら、実証実験をしいて、血胞内の現象と同時に再現するために、RICSにおいておが、生物学的に安当な条件では明らかな計算エラーがでる。 * によいで計算は可能できた。 * において計算は可能できた。 * においた計算を行なうことの。 * においた計算に可能できた。 * においた計算と行なっている。 * においた計算エラーがでる。 * においた計算に可能できた。 * においた計算に可能できた。 * においた計算エラーがでる。 * においた計算エラーがでる。 * においた計算に可能できた。 * においたが、 * においかの移活拡散計算のためににおみ込んだ * いたいかい、 * たいかうい。 * におすののやのが活動説はない、 * たいかいかい * におすののやいたいない、 * たいたいない。 * たいたいない、 * たいたが、 * はないたいない。 * はないたがたます * はないたいない。 * はないたがたます * はないたいない。 * はないたたいない。 * はないたたいない。 * はないたたいない。 * はないたたが、 * はないたたたいない。 * はないたたいない。 * はないたたいない。 * はないない。 * はないないないないないないないないないないないないないないないないないないない |
|---|---|--|
| | <section-header></section-header> | ・アプローチ ・ 建造プログラムとしてRicken Cell Simulator Platform (RICS)を用いる ・ 基盤プログラムとしてRicken Cell Simulator Platform (RICS)を用いる ・ 基盤プログラムとしてRicken Cell Simulator Platform (RICS)を用いる ・ 美証実験にて取得した血小板細胞の構造情報をRICSに入れ込み、「血小板細胞」形状モデルを 体成する ・ 美証実験にて取得した血小板細胞」形状モデルは ・ 「血小板細胞」を供成し、複数面小板の接着 ・ 「血小板細胞」を作成し、複数面小板の接着 ・ 「血小板細胞」と作成し、複数面小板の接着 ・ 「血小板細胞」に、複数面小板の接着 ・ 「血小板細胞」と作成し、複数面小板の接着 ・ 「血小板細胞」と作成し、複数面小板の接着 ・ 「血小板細胞」には大する ・ 「血小板細胞」に、「酸、活性化にデルと「血小板細胞」を行 ・ いっ連成計算を行 ・ 1.00連成計算を行 |

データ解析融合研究開発チーム成果報告

データ解析融合研究開発チーム

宮野 悟

次世代計算科学研究開発プログラム データ解析融合研究開発チーム チームリーダー



発表者紹介

- 1977 年 3 月 九州大学理学部数学科卒業
- 1979 年 3 月 九州大学大学院理学研究科修士課程数学専攻修了
- 1979 年 6 月 九州大学理学部助手
- 1985 年 4 月 Alexander von Humboldt 財団研究員
- 1987年4月 Universität GH Paderborn 助手
- 1987年12月 九州大学理学部助教授
- 1993年3月 九州大学理学部教授
- 1996年4月 東京大学医科学研究所教授

研究分野

計算システム生物学、バイオインフォマティクス、メディカルインフォマティクス
データ解析融合研究開発チーム

宮野 悟

次世代計算科学研究開発プログラム データ解析融合研究開発チーム チームリーダー

1. 目的

「計測技術の大規模化・精緻化・簡便化により、 遺伝子から環境因子にわたる生命システムに関 するデータが、超高次元化・超ヘテロ化・超膨大 化した。しかも、不観測性・欠損の問題も同時に 有している。こうした中、データ解析はデータの 多様化と増大化に急速に引き離され、シミュレー ションは生命体個別の現実データを反映できず 予測能力に限界がある。方法論にパラダイムシフ トが必要となった。」これが、2006年、研究の準 備段階での研究推進のメッセージあった。それか ら7年を経過した現在、次世代シークエンサーの 発達や遺伝子発現データ解析の高精度化・低コス ト化が大きく進み、がんなどの病態の分子システ ムの理解と病気の原因遺伝子の解明に向けたデ ータ解析及びシミュレーション技術の研究は全 世界的潮流になった。本チームの目的設定が極め て的確であったといえる。ただ、京の利用期間が 想定していたよりも短かったこと、及び、現時点 での京の運用がライフサイエンス研究に最適に なっていなことが残念なところである。

本研究チームは、ペタスケールの計算を前提に して、超高次元大規模へテロデータ解析と生命体 シミュレーションを融合し、生命体システムに対 する予測と発見の基盤情報技術を構築すること を目的として研究開発を行ってきた。4人の PI が、図1の文脈で、次の3つの技術開発を目標に 研究を行った。



① ゲノムワイド関連解析により疾患や薬物反応

図1

性などの表現型に関連する遺伝子の解明と、 個人の表現型をゲノム情報と環境情報により 予測する技術

- 理化学研究所ゲノム医科学研究センター・角田達彦(2011.4~2013.3);鎌谷直之(2006.10~2011.3)
- ② 大規模な生体分子のネットワークを推定する 技術を開発し、これを「地図」として薬物・ 疾患に関与する遺伝子群を探索する技術
 - 東大医科学研究所・宮野悟
 - 東京工業大学情報理工学研究科・秋山泰
- ③ データ同化技術の活用による一般のモデルか ら個のモデルを創出する技術
 - 統計数理研究所 · 樋口知之

2. 成果

2.1 「肺がんと薬」を共通軸とした研究と、その 広がり



本チームの研究開発は 2006 年に開始したが、 2008 年に実施された研究の進捗状況についての 中間評価において、チームとして共通の研究軸を もつことによりチーム内での相乗効果を発揮さ せることが指示された。そこで、本チームのメン バーが実験データの取得も含め対応可能なもの として「肺がんと薬」を共通のテーマとしたが、 優れた技術が開発されたため、社会的・科学的イ ンパクトの大きな成果を出すこともできるよう、 多少研究に広がりを持たせた。図2のもと、図3 のように技術開発の方向を明確化したが、肺がん とその薬に限らず、本チームが開発したソフトウ ェアアプリケーション群は有効であることが示 された。

2.2 開発ソフトウェアと開発責任者

チームの構成と役割を図4に示している。チーム リーダーの宮野悟は「大規模遺伝子ネットワーク 推定とその応用」を業務委託課題として担当し、 ヒト全遺伝子規模の遺伝子ネットワークを推 定・探索するソフトウェア群を開発した。この全 体パッケージを SiGN と称している。ベイジアン ネットワークにノンパラメトリック回帰を組み 合わせた方法(SiGN-BN)、状態空間モデルに次元 圧縮を組み合わせた方法、構造方程式もでるなど に L1 正則化法を駆使したネットワーク推定法群 (SiGN-L1)などのアプリケーション群がパッケー ジに含まれており、すべて京コンピュータでの高 並列化を達成した。トランスクリプトームのネッ トワークを推定することに用いることができる が、ソフトウェアとしては生命科学に限らず、汎 用的なものである。ノード間の通信量が京をもち いても効率よくスケールさせることが困難であ ることが判明した「最適ベイジアンネットワーク 推定ソフトウェア」では、32ノードの最適ベイジ アンネットワークの推定を実現し、これは世界記 録となっている。



図3



図4

秋山泰は、1000×1000 規模のタンパク質相 互作用を網羅的に推定することでタンパク質ネ ットワークを構築できるソフトウェア MEGADOCK を開発し、京コンピュータでの高 並列化を達成した。その能力は1000×1000を超 えるものとなった。網羅的なタンパク質間相互作 用予測システムを京コンピュータの上で効率的 に動作する並列ソフトウェアパッケージとして まとめ、前処理・後処理のツール群も京コンピュ ータ周辺環境で動作することを実現した。 MEGADOCK の開発により、予測されたタンパ ク質相互作用情報をトランスクリプトームネッ トワークと合わせてシステムを理解することを ことが京コンピュータの計算リソースが十分に 利用することができれば、様々疾患においても、 分子ネットワーク・薬剤応答ネットワークを解析 し、その分子病態を描出することを狙うことが可 能になり、今後に期待が大きい。

角田達彦は、前任者の鎌谷直之の開発を引き継 ぎ、ParaHaploというソフトウエアを開発した。 京コンピュータでの高並列化を達成し、Wellcome Trust のデータを用いてその優れた点を確かめて いる。これは肺がんに限ったものではなく、患者 群と対照群の全ゲノム上の SNP を用いハプロタ イプ単位のゲノムワイド関連解析を行い、疾患関 連遺伝子を網羅的に探索するためのソフトウェ アである。Nature Genetics に掲載されている論 文の 55%が GWAS を用いて、疾患遺伝子を探索 しており、GWAS は理化学研究所の中村祐輔 (現・シカゴ大学教授)の率いる遺伝子多型研究 センター(当時の名称。その後、ゲノム医科学研 究センターに改組)が世界で最初に開発し、その 有効性を証明したものである。これから時間をか けて世界において HaraHaplo の有効性が確認実 証されることを期待したい。また、角田はこれに 関連して、拡張 RAT 法による 2 SNP 組合せの全

ゲノム関連解析ソフトウェア(ExRAT)を開発し ている。これは、病気へのリスクを上げる原因と なるような複数の遺伝子による相乗効果を大規 模体系的網羅的に見つけ出すソフトウェアであ り、京コンピュータのアーキテクチャに合わせ並 列度を上げ、京コンピュータ上で実データを用い て実行させて、未知の複数要因による新たな疾患 発症機序の発見と、未だ解決してない遺伝力の問 題を解く成果を狙っている。京コンピュータにお いてかなりの高並列化を達成している。また、次 世代シークエンサーデータ解析プログラム (NGSanalyzer)は、がんを体細胞変異のゲノム 学と遺伝統計学により解明するために、次世代シ ークエンサーからの人間のがん細胞と正常細胞 の全ゲノムの膨大なデータを網羅的・正確・高速 に解析するためのソフトウェアである。がんの種 類や対象とする病気により、解析パイプラインの パラメータの調整が実験による検証のプロセス を経ながら調整する必要があるため、パイプライ ンの並列度を上げ、オープンソースでコードを配 布し、様々のゲノム解析研究において京コンピュ ータ上で利用されていくことが期待される。

データ同化技術は、個別データをモデルに同化 させ、個別化モデルを作り、個に対するシミュレ ーションと予測をするものである。LiSDAS は、 階層性を考慮した粒子フィルタアルゴリズムな どを京コンピュータ上で超高並列化したソフト ウェアであり、京コンピュータにおいて高並列化 を達成している。オミックス計測技術から生成さ れるデータを利用して、生体内分子相互作用ネッ トワークのシミュレーションモデルを構築する ことに応用している。

最後に、開発したソフトウェアを連携して使う ことができるSBiP(Systems Biology integrative Pipeline)というデータ解析プラットフォームを 宮野のグループで開発した。SBiP は京コンピュ ータで走らせるソフトウェアではなく、ユーザの コンピュータにインストールし、高機能 GUI に より、SBiP から京コンピュータの上で走るソフ トウェアをつないで解析パイプラインを作成で きるものである。京コンピュータの運用が開始さ れてから歯科開発に着手できない部分が有り、ま た、東大医科学研究所ヒトゲノム解析センターの スーパーコンピュータを使って考案した解決策 も京コンピュータの運用方針に合わないとのこ とで、苦難と膨大な努力の末、SBiP から SiGN 関係のソフトウェアを京コンピュータで実行し その結果をもどし、人手では実行不可能な手順の 大規模データ解析(遺伝子ネットワーク解析を中 心)ができるものになりつつある。ドキュメンテ ーションもきちんとしており、オープンソースで すでに公開している。

3. プロジェクト終了時の達成目標

ネットワーク解析による薬のターゲット探索 やがんなどの病態を理解する情報技術、並びにゲ ノム情報に基づいた個人差を考慮した医療のた めの基盤情報技術を創出することができ、目標を 達成することができた。本プログラム期間内に得 られた個々の成果については、各 PI から報告が ある。

2012 年になり、半導体チップで DNA を読む 革新的なシークエンサーの実用化された。さらに、 2013 年には、高集積半導体チップがリリースさ れ、誰もが自分の全 DNA 情報を 10 万円程度で 丸ごと手にいれられることが確実になった。さら に、1万円ゲノムを実現するナノポア半導体シー クエンサー技術も実用化されている。これまで主 に生物や病気の「研究」のために行っていたシー クエンスに対し、がんや患者さんの全 DNA 情報 や遺伝子発現情報を得て、臨床的に翻訳・解釈し、 治療として「患者さんに戻す」臨床シークエンス の研究が米国やカナダでは既に始まり、国を上げ て個別化医療を推し進めている。そして、医療・ 生命科学のビッグデータが誕生する。エピゲノム、 トランスクリプトーム(シークエンス、及び定量 データ)もこの次世代のシークエンサーで解析で きるものである。これまで遺伝子発現プロファイ ルデータの取得には、マイクロアレイを使ってき たが、これが次世代シークエンサーによる RNA-seg という方法に置き換わり始めている。

本チームが開発したソフトウェアアプリケー ションは、まさにこの医療・生命ビッグデータ時 代に挑戦できるものであり、京コンピュータ、並 びにその他のスーパーコンピュータを用いてい た活用に大きな期待を寄せている。本チームの研 究開発は、単に目標達成という「終わり」ではな く、個別化医療新時代の幕を開けてくれるものと 考えている。



| A等 描述 目前 日本 | 格称 担当PI | 角田達彦 zer 角田達彦 | 角田達彦 | 呈现 zign-ru | 宮野 悟 スパコンには実装Uaiv | 樋口知之 | X 秋山 泰 |
|---|--|------------------|-------------------------------------|---------------------------|----------------------|---------------|--------------------------|
| raianapiu | olachered | NGS analyz | ExRAT | Sign-BN, sign-5 | SBiP | LISDAS | MEGADOC |
| 次世代シークエンサーデータ解析プロ グラム | ハブロタイブ関連解析に於ける統計検 定を行うためのソフトウェア 次世代シークエンサーデータ解析プロ グラム | | 拡張RAT法によるZSNP組合せの全ゲノ ム関連解析ソフトウェア | 大規模遺伝子制御ネットワーク推定プ ログラム | データ解析融合プラットフォーム | 生命体データ同化プログラム | 網羅的タンパク質ドッキング解析プロ グラム |























大規模ゲノム多型データと表現型データを関 連付ける新規アルゴリズムの開発と、妥当性、 有用性の検討

角田 達彦

理化学研究所 ゲノム医科学研究センター 統計解析・技術開発グループ グループディレクター



発表者紹介

1995年3月 東京大学大学院 工学系研究科 博士課程修了

1995 年 4 月 京都大学大学院 工学研究科 助手

1997 年 4 月 東京大学医科学研究所 ヒトゲノム解析センター リサーチアソシエイト

1998 年 4 月 東京大学医科学研究所 ヒトゲノム解析センター 助手

2000 年 4 月 理化学研究所 遺伝子多型研究センター チームリーダー

2008 年 4 月 理化学研究所 ゲノム医科学研究センター チームリーダー

2011 年 4 月 理化学研究所 ゲノム医科学研究センター グループディレクター

医学博士・工学博士

研究分野

ゲノム医科学、医科学数理

受賞歴

1993 年 International Joint Conference on Neural Networks Student Award

1994年 情報処理学会第48回全国大会奨励賞受賞

2004 年 The 13th Takeda Science Foundation Symposium on Bioscience on Genome Analysis and Medicine Poster Award

2009 年 日本人類遺伝学会奨励賞受賞

| 背景·目的 | 背景: 全くの未知のものも含め, 疾患の原因を探索するには, 全ゲノム上で患者群と対照群との間でゲノムDNA配列を比較する, ゲノムワイド関連解析 (GWAS)が極めて有効であり, われわれ理化学研究所ゲノム医科学研究センターでは, 2002年に世界に先駆け初めてのGWASを実現してから, 世界を牽引してきた. この方法論をさらに劇的に推進するために, は点の解析だけでなく, 近傍の, あるいは遠距離にある複数点の解析 | を11.2/こかで4.6-4/Falanahabacexive10、約家とりのメールーで、これまでの回たと4.6 たセットから拡張し、次世代シークエンサーにより全ゲノム配列への探索を行ったり (NGS analyzer)するなどの新たな方法論を導入する必要がある、これらには、スー パーコンピューティングを駆使することが必要不可次になる、 | 目的: - ParaHaplo: GWASの方法論を拡張し、単点ではなく連続する複数点に対して超並 列に行うことによって、飛躍的な検出力の向上と高速性能を狙う. - ExRAT: 複数の遺伝子の相互作用により発症する可能性を、全ゲノム上で網羅的 に起並列に探索する. | - NGS analyzer: 次世代シークエンサーの爆発的データを超並列に処理し, ゲノム多様性を高精度かつ包括的に解析する. | ISLIMソフトウェア研究開発報告会 2 | 研究開発成果 | 開発ソフトウェア ParaHaplo について 目者ソフトウェア ParaHaplo について - ParaHaplo ver. 4.0 β, phase IV, 50% - 富士通コンパイラ (RICC, FX1, 「京」、の全て)対応済, 実行成功 - アビクラスタで8,000ノードの並列性能を実証. - 「京」ではコア並列(ハイブリッド並列)も実装している. - 一京」ではコア並列(ハイブリッド並列)も実装している. - 現在は、一層の高速化を目指LSIMD化を行っている. - シーズコード公開済. - シーズコード公開済. - 1 論文出版: Misawa K, Kamatani N (2011) ParaHaplo 3.0: a program package for imputation and haplotype-based whole-genome association study using parallel computing. Source Code Biol Med. | |
|-------|---|--|---|--|----------------------|---|--|--|
| | • | | • | | | | • | |
| | 大規模ゲノム多型データと表現型データを関連付ける新規アルゴリズムの開発と、妥当性、有用性の検討 | 理化学研究所 | ゲノム医科学研究センター グループディレクター 角田 達彦 | | ISLIMソフトウェア研究開発報告会 1 | いまた しんしょう しんしょう しんしょう しょうしょう しょう | 研究開発コードの概要 日空間発コードの概要 日空間目的:人のゲノム全体にわたる遺伝的な相違点を、近傍にある複数点を同時にみたいプロタイプという単位で、患者群と対照群とで比較することにより、疾患の道伝的原因を探るための統計検定計算 ExRAT:遺伝子間相互作用が発症リスクを変化させる疾患関連遺伝子の組合せを をゲノムで探索する.2SNP間の全組合せを超述剤に行う方法と、SNP間の連鎖不 報償(相関)も考慮した、より精密な方法の2種類を実装、前者で全組合せをスクリー ニングし、後者で経験的p値を求める手順を想定している。 NGS analyzer:次世代シークエンサーの出力データを高速に解析し、個人間の遺 低的差異やがんゲノムの突然変異を高い正確さで同定する. NGS analyzer:次世代シークエンサーの出力データを高速に解析し、個人間の遺 低的差異やがんゲノムの突然変異を高い正確さで同定する. ParaHaplo: Haplotpeig度を用いたTypel errorの確率計算を、ハイブリッド並列 で実装したマルコフ連鎖モンテカルロ法で行う. ExRAT: ErRAT(Rapid Association Test)を、データ分割法で実装したインボータンス サンブリング法で行う. NGS analyzer: LF標準ゲノム配列におする マッピングと確率計算に基づいた多様性検出を、領域分割で実装した直接法による密度行列の対角化により行う. | |



1200

000

800

009

400

200

0

5

ファ確保が原因だったため、通信処理を見直し、処理完了待ちのバッファの大きさを

最小とすることで対処した.

テスト運用の結果から問題点の洗い出しおよび改善方法の検討を行ってきた。

WTCCCの実データを解析し, 評価と結果を出した.

ISLIMソフトウェア研究開発報告会

ノード数 ISLiMソフトウェア研究開発報告会









タンパク質間相互作用ネットワークの推定と その応用に関する研究

秋山 泰 東京工業大学 大学院情報理工学研究科 計算工学専攻 教授



発表者紹介

1990年3月 慶應義塾大学大学院理工学研究科電気工学専攻博士課程修了(工学博士)

1990 年 4 月 通商産業省工業技術院電子技術総合研究所 研究官

1992 年 4 月 京都大学 助教授(化学研究所)

1996年4月 技術研究組合新情報処理開発機構 研究室長(並列応用つくば研究室)

2000 年 4 月 通商産業省工業技術院電子技術総合研究所 主任研究官 (生命情報科学 研究センター検討チーム長)

- 2001年4月 独立行政法人産業技術総合研究所 生命情報科学研究センター長
- 2007 年 4 月 東京工業大学大学院情報理工学研究科計算工学専攻 教授

2011 年 12 月 東京工業大学情報生命博士教育院 教育院長 (兼任)

研究分野

バイオインフォマティクス、並列処理応用









生命体シミュレーションのための データ同化技術の開発

樋口知之 次世代計算科学研究開発プログラム データ解析融合研究開発チーム



発表者紹介

- 1989年 3月 東京大学理学系研究科地球物理学博士課程修了
- 1989 年 4月 統計数理研究所予測制御研究系予測理論研究部門助手
- 1994 年 12 月 統計数理研究所予測制御研究系予測理論研究部門助教授
- 2002 年 7月 統計数理研究所予測制御研究系システム解析研究部門教授
- 2004 年 4月 統計数理研究所モデリング研究系教授,予測発見戦略研究センター 副所長(兼務)(~2010.3)
- 2011 年 4月 統計数理研究所 所長

研究分野

ベイジアンモデリング、データ同化













臓器全身スケール研究開発チーム成果報告

臓器全身スケール研究開発チーム 成果総括報告

高木 周

次世代計算科学研究開発プログラム 臓器全身スケール研究開発チーム チームリーダー



発表者紹介

1995年3月 東京大学大学院工学系研究科機械工学専攻博士課程修了 1998年4月 東京大学大学院工学系研究科 講師 2001年4月 東京大学大学院工学系研究科 助教授 2007年4月~現在 理化学研究所次世代計算科学,チームリーダー 20010年4月~現在 東京大学大学院工学系研究科 教授

研究分野

計算生体力学,流体力学,マルチスケール計算力学

臓器全身スケール研究開発チーム成果総括報告

高木 周 次世代計算科学研究開発プログラム 臓器全身スケール研究開発チーム チームリーダー

1. 目的

臓器全身スケール研究開発チームでは、CT, MRI,超音波画像装置等により取得された各個人 の人体ボクセルデータを基にして、コンピュータ 上に人体を再構築し、各種病態の予測や治療の支 援を行なうためのシミュレーションツールの開 発を行うことを目的としている.特に、日本人の 死因で大きな比重を占める癌の治療、循環器系疾 患に関連する計算手法の開発に重点を置き、病態 のメカニズムの解明と薬効の評価、低侵襲治療法 の支援を行なうためのシミュレーションツール を開発している.

3. 現時点での成果

2.1 医用画像データに適した計算手法の開発(オ イラー型流体構造連成手法: ZZ-EFSI)

本研究では、従来は境界適合メッシュを用いた 有限要素法による解析が主流であった流体・構造 連成問題に対して、メッシュ生成のプロセスなし に画像データから直接計算を行なうオイラー型 流体構造連成手法を新たに開発した.本手法は、 医用画像データのみならず、一般に画像データを 基にして時々刻々形状が変化していくような問 題全般に適用可能であり、固定された構造格子の 上を自由に変形する物体の運動を計算すること ができ超並列の計算に特に適した定式化を行っ ている.すでに、フルコアを用いた計算で実効性 能 4.5Peta Flops の計算速度を記録している.

2.2 血小板上と血管壁のタンパク質間相互作用と 流れの計算を連成させた血球細胞・流体連成手法 (マルチスケール血栓シミュレータ: ZZ-Throm)

ZZ-EFSI 法を膜と流体の連成問題へまで発展 させ、赤血球や血小板などの血球細胞を大量に含 む血流の計算を可能にさせた手法である.さらに、 血小板上の糖タンパク GP-1baと損傷血管壁上の vWF 分子との分子間相互作用である Receptor-Ligand 反応をモンテカルロ法により計算するこ とにより、分子間結合による吸着力と、流体運動 により引き剥がす流体力を連成させながら解く 計算手法を考案した.赤血球の有無が、血小板吸 着に与える影響について調べ、血小板吸着にとっ て赤血球がもたらす血漿側のゆらぎが本質的な 影響を与えることを再現するのに成功している. さらに,赤血球の体積率を変化させ,実験データ との定量的比較を行っている段階.

2.3 心筋細胞の詳細な挙動から心臓全体を再現す る心臓シミュレータ(マルチスケール・マルチフ ィジックス心臓シミュレータ:UT-Heart)

ISLiM プロジェクトで開発されたプログラム の中で唯一,「京」の優先課題として選ばれたソ フトウェア.東京大学,久田らにより10年以上 の年月をかけて開発されてきた心臓シミュレー タである. このシミュレータでは、細胞膜のイオ ンチャンネル, ポンプ, トランスポーターなどの 作用に基づく細胞内イオン環境のダイナミック スを記述する電気生理学モデル,並びにイオンに よって制御される細胞内収縮タンパクであるア クチン・ミオシン間のクロスブリッジ運動を記述 する興奮収縮連関モデルから出発し、細胞、組織、 臓器を経て最終的な血液拍出に至る心臓機能の 全過程を各レベルにおける要素間の相互作用を 含めて合理的に数理化し、心臓の振る舞いを再現 できる.本年度は、サルコメ←→心筋細胞←→心 臓全体の3レベル階層統合に成功し、「京」での 超大規模計算に成功した.

2.4 国産初の治療器開発に向けた超音波治療シミ ュレーション(強力集束超音波シミュレータ: ZZ-HIFU)

患者個々の人体ボクセルデータを用いて強力 集束超音波(HIFU)により腫瘍焼灼を行なうシミ ュレーションを行う方法の開発を終えた.現時点 において、98304 コアを用いた計算で、実行性能 21%を達成している.現在、国産初の超音波治療 器を目指して、国内の医療機器メーカーと連携し て乳腺治療を対象として実機の設計支援のため のシミュレーションを開始している.具体的には、 実際の乳房の画像データを基にし、焦点制御のシ ミュレーションを実施し、位相制御の重要性を確 認している.

2.5 (モンテカルロ重粒子線治療シミュレー タ:ZZ.DOSE)

重粒子線治療の手術計画に利用可能なモンテ

カルロ計算によるシミュレータを開発している. 現時点では、プログラムのハイブリッド並列化が 済み、大規模計算の準備をしている.また、理研 の仁科加速器センターとの共同研究により、計算 結果の検証用の実験データを取得している.現時 点では、計算負荷の軽い系に対して、実験との比 較を行い、モデルの検証を進めている.

2.6 その他(0D-1D-3D 全身血管網シミュレータ: ZZ-Vasc)

大規模3次元血流計算に対する境界条件を与え るための計算を行なうソフトウェアである.現時 点では、東京大学の大島まりの研究グループと連 携をとり、動脈硬化を起こした総頸動脈および脳 血管系であるウィルス動脈輪に対して、患者の画 像データから得られた情報を基に3次元計算と 0D-1D計算の連成を行い、術後の血流量の分配予 測のシミュレーションなどを行っている.

3. プロジェクト終了時の達成目標

1) ZZ-EFSI

上述の通り、ZZ-EFSI はすでに京のフルノード 約 65 万コアを用いた計算で実行性能 4.5Peta Flops を達成している. プロジェクト終了時まで に、ソフトウェアの最新版の公開と随時更新を行 なっていく.本手法は、医療応用だけでなく、「も のつくり」分野も含めた幅広い分野への適用が可 能なソフトであり、公開を通して多くのユーザー を募り、プロジェクトの成果の有効活用を図る.

2) ZZ-Throm

損傷血管壁への血小板の吸着に伴う,血栓成長の初期過程を再現するシミュレータを完成させる.これにより,分子レベルでのReceptor-Ligand結合から,多数の変形赤血球・血小板を含む大規模な流動構造までを連成して解くマルチスケールシミュレーションを達成し,Receptor-Ligandの分子間結合が血栓成長に与える影響を調べる. また,東海大学医学部,後藤らにより取得される実験データとの比較により,計算結果の妥当性を検証し,さらに実験データと計算結果の相補的な利用により実験で解析が困難な膜上における糖タンパク GP1b-αの動的振る舞いなどについて知見を得る.

3) 心臓シミュレータ (UT-Heart)

筋原線維を構成するアクチン・ミオシンフィラ

メントにおける機能タンパク個々の運動やそれ を支えるZ帯などの微小構造を具体的にモデル 化し、そこから心臓の拍動までをシームレスに接 続し、ミクロとマクロの因果関係を明らかにすた めのシミュレータ開発をさらに進めていく.この ようにタンパク質レベルの確率的運動、複合体と しての細胞レベルの収縮運動、および臓器レベル の拍動運動の3つのスケールを合理的に結び付 けるシミュレーションを達成することにより、心 筋細胞内で起こる分子レベルの物質輸送から心 筋の収縮を再現し、新たな医学的・生理学的知見 を得ることが可能となる.

4) ZZ-HIFU

国産初の治療器の設計に向けた大規模計算を 引き続き行う.本年度は実際に乳房のデータを基 にして計算を実施しており,このデータに対して さらにパラメータを振った計算を進める.特に, 治療の対象とする部位への超音波集束の制御を 考慮した場合に必要となる超音波パネル(アレイ トランスデューサー)の数と配置,リアルタイム モニタリングの方法について,実験データとの比 較を行いながら機器設計を進めて行くシステム を構築する.これにより,大規模計算の結果を利 用して,機器の設計および実機へ搭載していくた めのソフトウェアの開発を進めることができる.

5) ZZ-DOSE

仁科加速器センターで新たに取得される実験 データにより、モンテカルロ法による計算結果の 比較を行い、開発された計算手法の妥当性の検証 を行う.加えて、「京」での計算に関して、ハイ ブリッドプログラムの改良を行い、並列化効率を 上げる.

6) ZZ-Vasc

本年度構築された 0D-1D-3D の連成手法に関し て,患者の血管データを使用したシミュレーショ ンを引き続き実施し,計算結果の検討および連成 手法の安定性の検証を行う. ZZ-VASC は、3 次元 医用画像データを用いた計算との連成以外にも、 0D-1D 連成の部分を用いてさまざまな共同研究 を行ってきた.その一つとして,心臓周りの動脈 硬化を調べる新たな電子血圧計の開発に協力し 製品化に成功したが.さらに、シミュレーション の結果を用いて、新たな動脈硬化判定指標の提案 を行う.

以上






次世代超音波治療機器開発のための 生体内超音波伝播シミュレータ (ZZ-HIFU)

沖田 浩平

日本大学生産工学部機械工学科 准教授 理化学研究所 次世代計算科学研究開発プログラム 臓器全身スケール研究開発チーム 客員研究員



発表者紹介

- 2002 年 3 月 大阪大学工学研究科機械物理工学専攻博士課程修了
- 2002 年 4 月 東京大学 IML 研究員
- 2004 年 4 月 東京大学 21 世紀 COE 機械システムイノベーション 特任研究員
- 2006 年 4 月 理化学研究所 VCAD システム研究プログラム 研究員
- 2007 年 4 月 理化学研究所 VCAD システム研究プログラム 上級研究員
- 2011 年 4 月 東京大学大学院工学系研究科機械工学専攻 特任研究員
- 2011 年 9 月 日本大学生産工学部機械工学科 准教授

研究分野

流体工学、数值流体力学、混相流





| Results of the simulation assisted focusing | い おお お お お お お か お か お か お か お か お か い か い |
|---|--|
| 「「」」」 「」」 「」」<th>ISBMILTablaGe@288bilia@8Bonfige.ox1d ation atio</th> | ISBMILTablaGe@288bilia@8Bonfige.ox1d ation atio |
| 新月のMRI画像を利用したHIFUシミコレーション からし、 からし、 がらし、 がらし、 | A Contraction during Sonothrombolysis From Cavitation during Sonothrombolysis From Cavitation during Sonothrombolysis From Cavitation during Sonothrombolysis Contraction during Sonothrombol |
| * 嗎見ら、日本超音波治療研究会, 2012 | "Hope et al. J. Acoust. Soc. Am. (accepted) |

ISLiMソフトウェア研究開発報告会(2013.1.10-11)



大規模並列計算に適した 流体構造/膜連成手法 (ZZ-EFSI)

杉山 和靖

独立行政法人 理化学研究所 情報基盤センター 技術開発ユニット ユニットリーダー



発表者紹介

2000 年3月東京大学大学院工学系研究科機械工学専攻博士課程修了
2000 年4月~2001 年3月東京大学IML研究員
2001 年4月~2005 年3月 独立行政法人海上技術安全研究所研究員
2005 年4月~2007 年9月 Twente University (オランダ)研究員
2007 年10月~2008 年3月東京大学大学院工学系研究科 リサーチフェロー
2008 年4月~2012 年9月 東京大学大学院工学系研究科特任准教授
2012 年10月~現在 理化学研究所情報基盤センター

研究分野

流体力学, 計算力学, 生体力学

| ● 「「「」」」 ● 「「」」 ● 「」」 ● ● 「」」 ● ● 「」」 ● ● ● 「」」 ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● | 日的・方策 日的・方策 日的・方策 ・大規模並列計算に適した流体構造/膜連成手法 (ZZ-EFS)の開発・ ・大規模並列計算に適した流体構造/膜連成手法 (ZZ-EFS)の開発・ ・防存の手法のチューニングではなく、新たに、実装すべき式を定め、 ・防存の手法のチューニングではなく、新たに、実装すべき式を定め、 ・大ーム、アルゴリズムを作る ・はのに型有限差分法 ・ま行効率の高い解析法を1から作る) ・回回に型有限差分法 ・目のに型有限差分法 ・この「この「「「、「、「、「、「、「、「、「、「、「、「、「、「、「、「、「、「 |
|---|--|
| ◇ <i>大規模並列計算に適した</i> <i>流体構造/膜連成手法</i>(ZZ-EFSI) 選化学研究所情報基盤センター技術開発ユニット <i>能山和靖</i> (五人・五、理研)、松本洋-郎(東大・エ) ◆ (大・五、理研)、松本洋-郎(東大・エ) | |



C.

ISLiM ソフトウェア研究開発報告会 (2013.1.10-11)





UT-heart によるサルコメアカ学から 心筋細胞構造を経て心拍動にいたる マルチスケール解析

鷲尾 巧

東京大学大学院 新領域創成科学研究科



発表者紹介

- 1990年3月 大阪大学理学部数学科修士課程修了
- 1991 年 4 月 日本電気(株)入社 コンピュータ技術本部にて HW 開発に従事
- 1993 年 4 月 同社 C&C 研究所にて高速計算アルゴリズムの研究開発に従事
- 2004 年 4 月 東京大学新領域創成科学研究科・(独)科学技術振興機構 CREST 研究員として心臓シミュレータの開発に従事
- 2009 年 4 月 同研究科 特任研究員として同シミュレータの開発を継続

研究分野

数值解析,計算生物学

| ・ 背景 ・ 背景 ・ 背景 ・ 背景 「ヒトゲノムプロジェクト」の完成を受けて生命科学の焦点は遺伝情報の解明からたんぱ | 「フィジオーム」とは、たんぱく質から細胞内小器官、細胞、組織を経て臓器、そして個体までの各階層に属する膨大な数の機能モジュール間の相互作用として表現される生命現象を計算機シミュレーションによってモデル論的に再現しようとする研究領域であり、ポストゲノムの医学・生物学の研究の中心課題のひとつとなっている。 目的 近年発展の著しい分子生物学により次々と得られるミクロな知見をマクロな臨床医学くと合理的に接続し、医療へ貢献すると共に、新たな学術領域[計算医学]を創出する。 マクロな心機能を細胞下レベルの電気化学・力学的なミクロ事象からありのままに再現するマリレテスケール・マルチフィジックス心臓、ニュレータ UT-Heart を実現し、ミクロレベルの異常といた # a ling 医学・氏 # a f a f a ling に まか a ling に まか a ling に まかし | 18LIMソフトウェア研究開発報告会 2 | → 現在までの研究開発成果 | 現時点の開発ソフトウェアについて ミクロスケール(心筋細胞複合体)とマクロスケール(心室)のマルチスケール解析を 均質化法により実現。 克服すべきであった問題点 均質化法における非圧縮性の取り扱い 分子モデルの確率的振る舞いと連続体モデルの有限要素解析とを結び付ける計算 手法の確立 解決策 | 非圧縮や料を扱うことの CF のり見しばを用来 ミクロとマクロでエネルギー的に矛盾の生じないロバストな連成法を開発 - 計算事例 問題サイズ:5万自由度細胞モデル×63万細胞(計:約315億自由度) 京のフルノードを使用し、正常な一心拍を再現 | ISLIMソフトウェア研究開発報告会 |
|---|--|----------------------|---------------|--|--|----------------------|
| UT-heartによるサルコメア力学から い防細胞構造を経て心拍動にいたる マルチスケール解析 | 東京大学新領域創成科学研究科 特任研究員 鷲尾巧 鷲尾巧 | ISLMソフトウェア研究開発報告会 | ☆ 概要・アプローチ | 研究開発コードの概要 研究開発コードの概要 本シミュレータは、イオンチャネルや収縮タンパクの電気化学・力学現象を再現した心筋細胞有限要素モデル、並びに医療画像(CT)から精密に作られた全心臓有限要素モデルから成る。後者の各有限要素に可者の心筋細胞有限要素モデルを埋め込み、両者を独自に考案した非線形均質化アルゴリズムにより連成させることにより、細胞下のミクロ事象から心臓の拍動や血液の拍出などのマクロ事象を再現することを可能とする。 シ以上により、これまで推測の域を出なかったミクロ事象(例えばデスミンなど機能タンパクの異常)とマクロ事象(例えば拡張型心筋症、心肥大症など)の関係を合理的に解明し、従来にない観点から医学・医療に貢献することを目指す。 | アプローチ 大規模な電気・化学・力学現象のマルチフィジックス問題を、流体構造連成解析などの 大規模な電気・化学・力学現象のマルチフィジックス問題を、流体構造連成解析などの 有限要素法をベースとする最先端の科学計算手法により正面からシミュレートする。 形ある細胞モデル(有限要素法により内部構造を再現した3次元数値細胞)を開発し、 均質化法をベースにミクロ現象とマクロ現象をシームレスにつなぐマルチスケールシミュ レーションを達成する。 | ISLIMソプトウェア研究開発報告会 3 |

| *! *! | | |
|----------------|---|-------------------|
| プロジェクト終了時の達成目標 | 約束 一節原線維サルコメア内のたんぱく質レベルの確率的運動、複合体としての細胞モデル収縮運動、および臓器レベルの拍動運動の3つのスケールをシームレスに結び付けるシミュレーションを実現し、たんぱく質レベルの機能と臓器レベルの拍動性能の関連について新たな知見を得る。特に、心疾患の事例として肥大型心筋症を取り上げる。 | ISLMソフトウェア研究開発報告会 |

脳神経系研究開発チーム成果報告

神経系の大規模シミュレーション

石井 信

次世代計算科学研究開発プログラム 脳神経系研究開発チーム チームリーダ



発表者紹介

- 1986 年 3 月 東京大学工学部 卒業
- 1988 年 3 月 東京大学大学院工学系研究科 修士課程 修了
- 1988年3月 (株)リコー中央研究所 研究員
- 1997年3月 工学博士(東京大学)
- 1997 年 4 月 奈良先端科学技術大学院大学情報科学研究科 助教授
- 2001年4月 同上教授
- 2007 年 7 月 京都大学大学院情報学研究科システム科学専攻教授
- 2008 年 10 月 理化学研究所次世代計算科学研究開発プログラム チームリーダ

研究分野

システム神経生物学、バイオインフォマティクス、統計的学習

神経系の大規模シミュレーション

石井 信 次世代計算科学研究開発プログラム 脳神経系研究開発チーム チームリーダ

1. 目的

脳神経系は、物質基盤の上に情報基盤を構築し た点で特異な生命システムである。また、遺伝 子・分子から脳全体、個体集団に至る生命の階層 の中で情報処理基盤が構築されるため、必然的に 多階層をなす複雑システムである。脳神経系の情 報処理は、二つの異なる時定数を持つ様式に大別 される。第一は、ミリ秒から秒の時定数を持ち、 外部からの入力刺激に基づき外部への働きかけ (運動)を実施する刺激運動変換過程であり、第 二は、分から時、日といった長い時定数を持ち、 環境に依存して、刺激運動変換自体を変化させる 適応(学習)過程である。前者は、神経回路レベ ルでの電気生理学的な現象(神経発火)がメイン プレーヤであり、後者は、神経発火のみならず神 経細胞内のイオンや酵素、受容体などの分子がメ インプレーヤとなる。脳神経系の情報処理は、常 に脳をとりまく物理世界の時定数(すなわち実時 間)において実現される。したがって、入力から 出力までをつないだ刺激運動変換の実時間での 再現、また、実環境において外的刺激にさらすこ とで適応過程を再現する必要がある。以上のこと から、脳全体レベルでの入出力変換、およびその 学習をシミュレーションのターゲットとし、その ための構成要素である神経細胞および神経回路 の研究(神経細胞シミュレーション、神経回路シ ミュレーション)を並行して実施している。前者 について、200 億個を超える細胞からなるヒトの 脳全体のシミュレーションは、「京」であっても 困難であるので、対象を限るものとし、ほ乳類の 視覚系と無脊椎動物(特に昆虫)の嗅覚系を対象 として、入出力にわたる回路のシミュレーション 研究を実施している(視覚系シミュレーション、 昆虫嗅覚系シミュレーション)。

2008年10月に発足した脳神経系研究開発チームでは、石井(京大、理研)が神経細胞シミュレーション、深井、Diesmann(理研)が神経回路シミュレーション、黒田(東大)、臼井(理研)、銅谷(沖縄科学技術大)が視覚系シミュレーション、神崎(東大)が昆虫嗅覚系シミュレーションを主担当することとして、7 PI 体制でスタートした。2010年4月からは、神経細胞シミュレーシ

ョンに市川 (東大医科研) を加えた 8 PI 体制になっている。また、Diesmann は 2011 年 9 月より ドイツ Juelich Research Center に転出したため、 現在は、国際協力体制で研究開発を行っている。

2. 現時点での成果

2.1 神経細胞シミュレーション

石井(京大、理研):発達期の神経細胞における 特徴的な局所構造の獲得や、軸索伸長などの動的 な形態変化の理解を目指し、神経細胞形態形成シ ミュレーションプラットフォーム NeuroMorphoKit を開発してきた。細胞形態形成 に関わるシグナル分子の反応拡散現象、細胞骨格 の動態(主にアクチン系)、細胞膜の変形を含む マルチフィジックスモデルにより汎用的な細胞 形態変化のシミュレーションを実現し、これまで に、アクチン線維により駆動される葉状仮足の自 己組織化、化学走性などを再現した(Nonaka, et al., Neural Networks, 2011). NeuroMorphoKit O 高度化のため、大規模並列環境におけるパフォー マンスチューニング、神経系への適用を可能にす るためのシミュレーションモデルの拡張を行っ た。前者について、MPI による並列化アルゴリズ ムの改善を推し進め、並列実行時に特にボトルネ ックとなっていた、重合・脱重合・分岐・切断・ キャッピングなどアクチンキネティクスの確率 シミュレーションにおける分子濃度情報の更新、 および細胞膜の変形シミュレーションにおける エネルギー計算で発生するルートノードへの通 信量を大幅に削減した。これにより、京上での初 期アクチン繊維数約120万本のシミュレーション におけるストロングスケールは、従来のα =0.91(512)からα=0.99(8192)に改善した。後者に ついて、力学的な曲げと結合タンパク質による結 束を考慮したアクチンフィラメントのバネネッ トワークモデル化、および膜モデルの三次元化を 実施し、現在統合を進めている。これらのモデル 拡張により、三次元でのアクチンフィラメントの 挙動を考慮したシミュレーションを行うことが 可能となった。このことは、軸索伸長時に軸索先 端部に見られる特徴的な構造である成長円錐の シミュレーションなどにおいて特に重要である。 市川(東大医科研): A-Cell は形態を有する神経 細胞について、電気生理学反応と細胞内分子反応 を同時に計算できるシミュレータである。これま での A-Cell は細胞の形態変化に対応できなかっ たので、NeuroMorphoKit との間で形態変化、局 所膜電位やコンパートメント内分子濃度などを 共有できるためのデータ交換形式 netCDF-neuronを規定しその基本動作を確認し た。また、記憶形成を修飾する転写因子 NF-κB の振動について約 40 万コンパートメントを有す る 3D モデルを構築(Ohshima, et al., PLoS ONE, 2012)し、MPI 化された A-Cell によって 99.99% の並列化率を達成した。

2.2 神経回路シミュレーション

Diesmann (Juelich Research Center):神経回路 の京上での大規模シミュレーションを可能とす るソフトウェア環境 NEST(NEural Simulation Tool)を開発してきた。NEST は主に点神経細胞 (形態を有しない神経細胞)からなる回路のシミ ュレーションを得意としているが、形態のある細 胞も対応可能である。神経の回路内情報伝達は神 経発火による。発火ごとにそのイベント情報を全 細胞に伝達するイベント駆動方式は、正確である ものの効率が悪い。NEST では、時間軸を一定の 間隔に区切り、その間隔内では予測に基づき細胞 ごとに積分を行った上で、間隔内で発生した発火 を全細胞に伝え、発火については時間をさかのぼ って各細胞の積分の更新を行うという時間駆動 型のアプローチをとることで、効率の良い計算を 行っている。この際にカルマンフィルタなどの精 度の良い予測法を用いることで精度を確保する (Plesser, et al., Neural Computation, 2009; Hanuschkin, et al., Frontiers in Neuroinformatics, 2010)。神経回路シミュレーションの難しさは、 原理的には全細胞間が接続し得るため、空間局所 性など並列化できるてがかりが少ないことであ る。しかし、これまでに NEST は、京の上で、 5×10^8 個の細胞、3×10¹²個のシナプスからなる、小型の ほ乳類の脳全体に匹敵する巨大回路のシミュレ ーションを実施した。24,576 ノード(196,608 コ ア)までの hybrid 並列実装による実行確認ができ ており、上記回路でのシミュレーションを行った ところ、196,608 コアを用いて、1 秒の生物時間 を1.544 秒でシミュレーションできている。各コ アの1神経細胞あたりの計算速度で比較すると、 Juelich センターの Blue Gene P(JUGENE)よりも 数倍程度高速である。このような、また、さらに 規模を拡大した回路シミュレーションを行うた め、新たにデータ構造の見直しを行い、効率的な メモリ利用を実現した(Helias, et al., Frontiers in Neuroinformatics, 2012)。NEST が目指すのは、 大型のほ乳類の脳全体レベルとなる大規模回路 のシミュレーションである。現在、京の全ノード が使用できれば、10⁹ 個の細胞からなる大規模回 路のシミュレーションをテストできるまで開発 が進んでいる。今のところ京の利用機会が得られ ていないが、今後機会を得て、世界最大規模の回 路シミュレーションを行い、ボトルネックとなる 可能性のあるコミュニケーション性能や回路規 模についての性能評価を行うことを目標として いる。また、NEST において部位局所的に遅い時 定数で作用する神経修飾物質の効果を取り入れ、 ドーパミンによる時間差分学習が可能であるこ とを示した(Potjans, et al., PLoS Computational *Biology*, 2011).

深井(理研):大脳新皮質で見られる特徴的な階 層構造が神経活動とどのような関係を持つのか、 いかなる情報処理を実現しているかを明らかに することを目指して、神経回路シミュレーション の研究を行っている。視覚野では、外界から得ら れる膨大な視覚情報から、視野内にある注目すべ き情報に焦点を当て、それを重点的に処理してい る(視覚的注意)。電気生理学的・解剖学的知見 に基づく詳細なネットワークを持つ視覚皮質の 層構造局所回路モデルを NEST により構築 (Potjans and Diesmann, Cerebral Cortex, 2012) し、視覚的注意が視覚皮質層の動態に与える影響 をシミュレーションにより検討した。Bottom-up 的な視覚情報と Top-down 的な注意入力が局所回 路内で作用しあうことを示し、視覚的注意が皮質 層の視覚処理に果たす役割を予測した (Wagatsuma, et al., Frontiers in Computational Neuroscience, 2011)。また、局所回路の数を増や し、視覚野内の局所領域をより忠実に再現するシ ミュレーションを行った結果、空間的注意と特徴 に基づく注意による、神経細胞応答の変調効果が、 定性的かつ定量的に良く説明されることがわか った。

2.3 視覚系シミュレーション

黒田(東大):これまでに、大脳皮質の錐体細胞 を対象として、カルシウムシグナル関連分子を含 む詳細学習(Spike Timing Dependent Plasticity; STDP)モデルを構築、また、カルシウムの上下流 の分子を中心とした簡素化モデルを構築、それを 用いて、オタマジャクシの視蓋における方向選択 性機構の獲得をシミュレーションにより明らか にした(Honda, et al., Journal of Neuroscience, 2011)。この簡易 STDP モデルを NEST のプラグ インとして実装し、京上で動くようにした。さら に、STDP の制御モデルとしての再モデル化を行 い、このモデルにより、カルシウムダイナミクス とリン酸化ダイナミクスの異なる役割が明白と なった。

臼井(理研):ヒトの視覚系は、眼光学系、網膜、 視覚皮質、脳幹(眼球運動駆動部)などで構成され る多階層複雑系であり、その機能の理解には、構 成要素のモデル化とともに、それらを統合した大 規模シミュレーションが必要である。これまでに、 計算機環境に依存せずモデルの構築や統合が可 能なモデル開発共有プラットフォーム環境 (PLATO)の開発(Kannon, et al., Neural Networks, 2011)を進め、また、モデル結合に用いる共有デ ータフォーマットならびにその利用支援ライブ ラリをRICC ならびに京上で MPI 動作可能なよう に実装を進めた。視覚系を構成する各構成要素モ デルの状況は以下のようである。(1) 眼球運動系 は、コンダクタンスベースの細胞モデルとして記 述し、物体注視時に生じる微小眼球運動ならびに 視野上の物体を反射的に捉える急速眼球運動の 生成を可能とした (Inagaki, et al., Neural Networks, 2011; 京上で 30 並列を実現)。(2) 眼光 学系は、眼球形状の特性、瞳孔系の特性などの知 見を追加して構築を進め、1000x1000の網膜像の 生成を可能とした(400 並列を実現)。(3) 網膜は、 視細胞、水平細胞、双極細胞、ならびに神経節細 胞から構成されている。視細胞について、約3600 万個の細胞から構成されるモデルの構築を行い、 それらのうち 150 万個の錐体細胞からなる Cone Mosaic と呼ばれる L 型、M 型、S 型の視細胞配 置による応答を再現し、1,504 並列での動作(スト ロングスケールα=0.89)を確認した。双極細胞層 および神経節細胞層は、それぞれ 150 万個、40 万個の細胞から構成されるモデルを構築し、それ ぞれの細胞応答を再現するとともに1,024 並列で の動作を確認した. また、これまでに開発した 共有データフォーマットならびにその利用支援 ライブラリ PLATONIC を用いて、眼球運動モデル、 眼光学モデル、網膜視細胞モデル(Cone Mosaic) ならびに双極細胞層・神経節細胞層モデルの統合 を進め、自然画像、錯視画像を用いたシミュレー ションを RICC ならびに京上で並列実施すること に成功した。

銅谷(沖縄科学技術大、京大、理研):視覚系の 出力を司る眼球運動系、すなわち、網膜-上丘-眼球運動制御系のモデルを NEST 上に実装した。 特に上丘は多階層から成るスパインキングニュ

ーロンの回路モデルとして実装した。網膜入力に 基づきサッカード眼球運動が誘発される際に、上 丘回路の各層では特徴的な時空間パターンが形 成される。例えば、サルの神経生理学実験によれ ば、上丘の出力を与える中間層のバースト細胞は、 ベル型の発火パターンを持ち、深層では活動の時 空間的な拡張が生じることが分かっているが、シ ミュレーションはこれらを良く再現し、その新た な機能的な意味付けを与えた。また、上丘により 駆動される眼球運動は、サッカード角度に対して ピーク速度は特徴的な非線形カーブを描くが、こ の性質もシミュレーションにより良く再現でき た。シミュレータは現在 RICC 上で並列動作して いるが、今後京上で動作確認を行う。また、視覚 運動系を中心とした閉回路シミュレーションを 実施した。このシミュレーションはATR(京都) に置かれたロボット頭部、RICC(和光)上の上 丘シミュレータ、それらをネットワークでつなぐ クラスタモジュール(MUSIC を用いて実装)から なる。ロボットの眼球にあたるカメラ画像から特 徴点を抽出し、その入力に対する上丘モデル (RICC の 256 並列で動作)の回路動作に基づき、 ロボットカメラが特徴点位置に跳躍眼球運動を 実行する。システム全体の遅れ時間は 120ms で あるが、これは主にネットワークの遅れによる。

2.4 昆虫嗅覚系シミュレーション

神崎(東大): モデル生物であるカイコガを対 象に 10⁵程度の神経細胞からなる脳全体レベルの 嗅覚情報処理、すなわち、脳における入力(匂い 受容)から出力(匂い源探索行動)までの情報処 理のシミュレーションを目指している。これまで の成果は以下のようである。(1) カイコガ神経デ ータベース(BoND)の整備を進め、約 1,600 個 のカイコガの脳内神経の三次元形態・生理データ を収集、解析などに利用可能にした (Kazawa et al. **2008**)。(2) 嗅覚系一次中枢である触角葉(哺乳 類脳の嗅球に対応)、および匂い源探索の運動パ ターンを生成する前運動中枢である側副 葉 (LAL-VPC)において, 各領域また各領域を連結 する神経細胞について細胞レベルで構造と機能 のデータを解析した(Namiki and Kanzaki 2008, Kazawa et al. 2009, Iwano et al. 2010, Namiki and Kanzaki 2011)。 側副葉においては、 持続的な ステアリング信号を出力する側副葉からの出力 である下降性神経が側副葉の特定の領域に分枝 していること、側副葉の入力細胞の嗅覚応答が一 過的であること、また側副葉は両側の触角情報を 統合すること、さらには、行動指令信号の長期的 応答が側副葉で生成されること等を明らかにし

た。(3) 脳から胸部運動系に下降する神経細胞を 選択的にラベルし, 選別したニューロンをスライ ドガラス上に展開し培養することで、電位依存性 の電流を計測する手法を確立した。同時に触角葉 や側副葉を構成する神経細胞の電位や電流クラ ンプ実験を進め、触角葉から上位中枢への出力神 経や側副葉神経細胞のスパイク発火の特性情報 を得た。(4) 側副葉を構成する神経細胞の膜電位 データと形態データから、側副葉の神経回路を構 成する両極性神経(BLN)や局所介在神経(LN) をさらに分類し (Iwano et al. 2010)、神経細胞の 平均応答波形を作成した。この平均応答を作り出 す側副葉神経回路のシナプス結合強度の推定を 行い、BLN が抑制性であるという条件下では一意 に近い解を得た(Yamagishi et al. 2012)。現在こ の結果を基に、側副葉回路モデルを構築中である。 (5) 多数の脳の形状に対して一種の平均操作を加 えることで、カイコガの「標準脳」を作成した。 神経細胞をこの標準脳座標系に 10μm 以下の誤 差でマッピングするプロコトルを確立した (Ikeno et al. 2012)。(6) 神経細胞の三次元形態か ら神経形態を抽出する技術を試行し (Nakajima et al. 2008, 2009)、神経細胞の形態を効率よく取 得するソフトウェアを開発した(Ikeno et al. 2012)。(7) 神経細胞プラットホームシミュレータ である NEURON を京に移植し、チューニングを 行い、単体性能で 1.54GFLOPS/core を達成した。 これは10,000 コンパートメントの神経細胞を32 コア並列でシミュレーションした際にリアルタ イム実行するのに必要な速度であり、シミュレー ション時のコンパートメント数を節減すればリ アルタイムシミュレーションが可能である。さら にシナプス間の通信をノード配置最適化したポ イント間通信に改良することによって通信コス トを2倍程度以上に改良した。現在は196,608 コ アの並列実行で187TFLOPSの計算性能を達成し ている。(8) 実数値遺伝的アルゴリズ ム (RCGA) を用いて、単一細胞の膜特性を推定 するソフトウェアを作成した。現在までに、13 万コア未満の並列実行について良好な並列性能 を得た。このプログラムを触角葉出力神経に適用 し、入力信号(電流注入)として適当な正弦+三 角波を入力することで早い収束が得られること を示した。

2.5 学術的成果とアウトリーチ

本研究に関わる国際学術専門雑誌掲載論文の 数は、2008 年(10 月から 12 月)は 5 件、2009 年は 10 件、2010 年は 9 件、2011 年は 24 件、2012 年は 22 件と順調である。その一部は、Nature Cell Biology, Molecular Systems Biology, Nature Neuroscience, Neuron, Journal of Neuroscience, PNAS など分野を代表する一流雑誌に掲載された。 これらはシステム生物学研究、大規模データ解析 などによるもの(いわゆる計算神経科学に関する もの)であり、スパコン利用を前提としたものと は限らないが、本プロジェクトに関わるメンバー の学術活動のレベルの高さを示している。京によ るシミュレーションが含まれるものとして、 Helias, M., et al. Supercomputers ready for use discovery machines for neuroscience. as Frontiers in Neuroinformatics, 6:26 (2012) が誌上 発表され、また、国際会議 SC12 においても、 Miyamoto, D. et al. Neural circuit simulation of Hodgkin-Huxley type neurons toward peta scale computers が発表された。また、プロジェクト期 間内で、関係研究者および学生の受賞は、稲垣の **CNS2011 poster award** など 16 件、メディアに取 り上げられたのは、NHK サイエンスゼロ(神崎グ ループ、2009年9月19日)、NHK スペシャル(石 井、2012年6月3日)を含め30件である。この ように、研究活動およびアウトリーチ活動は良好 である。

3. プロジェクト終了時の達成目標

3.1 神経細胞シミュレーション

神経細胞におけるマルチフィジックス(骨格系、 細胞膜、細胞内分子の反応拡散)を統合的にシミ ュレーションするためのプラットフォーム NeuroMorphoKit により、神経伸長の方向決定性 に関するフィラメントベースの三次元シミュレ ーションを実行する。京上での10,000 コア並列 を目標とする。並行して、膜電位および分子の反 応拡散の計算を可能とし、可視化など各種インタ ーフェースを備える A-Cell の MPI 化を進め、 NeuroMorphKit との動作統合のためのデータ交換 法を開発する。

3.2 神経回路シミュレーション

NEST のコミュニケーションボトルネックについてアルゴリズムの改良をさらに進め、京上での実行時間が確保されれば、世界最高水準となる、10⁹ 個の細胞、10¹³ 個程度のシナプスからなる回路シミュレーションを行うことで、小型の霊長類の脳全体レベルの回路動態の再現を目指す。 NEST による複数カラムモデルシミュレーションを実施し、特徴注意と空間注意の違いの説明を試みる。また、電気シナプスを考慮した並列計算技術を NEST に実装し、既存のシミュレータの技術的な困難を排除することで、より現実的な局所回路モデルの高速シミュレーションを可能にする。

3.3 視覚系シミュレーション

視覚系の入力部に相当する眼光学-初期視覚 系に関する統合シミュレーション環境 VSM に対 して、脳の情報処理の実時間性を考慮した並列化 を進め、RICC 上で高並列実行するとともに、FX (京)上でのチューニングを進める。FX 上では 2000 コア並列以上を目指す。また、VSM を用い て、錯視現象のモデルに基づく機構解明を目指す。 網膜モデル、上丘モデル、眼球運動モデル、視覚 空間モデルを MUSIC により結合した視覚系の入 出力閉回路シミュレーションを、FX 上で、実時 間と同程度の時間で動作させる。

3.4 昆虫嗅覚系シミュレーション

詳細な生物学実験に基づいた昆虫の嗅覚系神経回路および細胞に関するデータベースと、それ に基づいて形態抽出や標準脳マッピングを行う ことでシミュレーションにつなげるシミュレー タ IOSSIM の開発を完了する。京上の 80,000 ノ ード実行で1ペタに近い実効性能の達成を目指す。 神経回路レベル・単一細胞レベルの生物的パラメ ータ探索とその評価を行う。匂い情報入力から匂 い源探索行動の行動指令を生成する中枢回路 (LAL-VPC)のシミュレーションを、京上で高並列 実施する。大規模マルチコンパートメントモデル については、京上で 2~3 万コア並列程度により LAL-VPC 神経回路の特性の時間計算を達成する。 実時間動作させた昆虫脳シミュレーションを移 動ロボットとリアルタイムで通信させる。

3.5 将来に向けて

脳神経系研究開発チームの研究は、長期的には、 脳神経系における発達障害の機序の解明、最適な 学習(教育)過程の設計、脳を模した知的情報処 理デバイスの開発などの社会・産業応用、また、 ブレインマシンインターフェースの高精細化な どの医療応用に貢献することを目指している。欧 米では、脳神経系シミュレーション研究はこれま でスパコン開発と足並みを揃えて発展しており、 Blue Gene/Qのターゲットの一つは神経回路シミ ュレーションとされている。また、欧州では Human Brain Project というメガサイエンスの動 きがあるが、ここでもスパコン開発が重要となっ ている。われわれは、次世代の Exa スケール計算 機と NEST の拡張により、ヒトの脳全体レベルの 回路シミュレーションが可能と考えているが、そ のためには脳のマクロ、ミクロのレベルでの回路 パラメータを調べる基礎研究も重要である。また、 大脳皮質の運動などの機能の理解には、複雑な形

態を持つ細胞からなる大規模回路の超高並列シ ミュレーションが必要となる。後者について、 ISLIM プロジェクトでは回路(通信)を得意とす る NEST と、精緻な細胞を得意とする NEURON の両者について、京上でチューニングを行ってき たので、今後はそれらを統合することのできるソ フトウェアの開発も必要であると思われる。

なお、京上の性能に関しては京速コンピュータ 京の試験利用および特別運用による。RICC 上の 性能に関しては理化学研究所情報基盤センター のPC クラスタ RICC による。京上の実装につい ては ISLiM 高度化チームの強力なサポートも得た。 関係各位に感謝申し上げる。



| ◎ 「「「」」」」」) | guided by neuroscientific goal to combine microscopic and macroscopic structure | Step Description Network Neurons Synapses Status Dec 2012 type | 1 reduced primate random at 3·10 ⁶ 3·10 ¹⁰ technically visual cortex micro and 3·10 ¹⁰ technically achieved on K on (32 areas, each 1 macro scale 28. Aug 2011 28. Aug 2011 | 2 reduced primate spatially 10 ¹² random network of visual cortex visual cortex organized at respecting relative 10 ¹² synapses 10 ¹⁰ synapses size of areas (32) actieved on K on for 10 ¹² synapses actieved on K on for 15. Sep. 2012 | 3 human brain model spatially max 10 ¹¹ max 10 ¹⁵ (approx. 100 areas) organized | Expected results: Step 1: investigation of reduced visual cortex model now possible on K characterization of stationary state (rates, irregularity) comparison of low-frequency EEG and LFP to experiments Step 2: models with closed functional circuits capability to simulate full scale model of a functional unit at microscopic resolution Step 3: full human brain simulation | assessment of capability to simulate full scale brain model b | ● B A A B B A A A A A A A A A A A A A | ^{80,000 LF-neurons} 1) 層構造をもつ視覚野局所回路の構築 Potjans and Diesmann and Fukai (2012) Cereb Cortex 2) それを用いた視覚的注意の2カラムモデル Wagatsuma, Potjans, Diesmann and Fukai (2011) Front Comput Neurosci 0. chell いき、(ト、)、小士体いいき、(ト、)、co・カーバ・エーニ | by Potjans and Discremin 3) 江田江忠(L) C村取江品(P) UOUJUCTJU Bibliggigt 9 亀気珪環線テータ 「白田山山山山山山山山山山山山山山山山山山山山山山山山山山山山山山山山山山山山 | A) 副所回路十丁儿/O | ためのデータ解析 ためのデータ解析 • 大脳皮質回路と海馬(Colein et al. Nature 2009) の計算原理の共通性 を発見(シータ波-ガンマ波の結合 波) Mcdams & Maursell, 1999; 加, Mcdams & Maursell, 1999; (Matine-Trujillo & True, 2004; (Matine-Trujillo & True, 2004; (Matine-2004; (Mat |
|--|---|--|--|--|---|--|---|---|---|--|--|--|
| 🜺 - Cellによる転写因子NF-ĸBシミュレーションの高速化 👉 💮 💽 | 背景と目的 NF-kBはCNSに発現しており、活性化がシナプスや樹状突起で進行。 NF-kBはCNSに発現しており、活性化がシナプスや樹状突起で進行。 | - Outainter Other Note Structure Store 「成本土井IAC Variation 1.7。 - NF-48の活性化は細胞接着タンパク質、NOS、BDNF、CamKII8の遺伝子が発現。 - NF-48及損マウスでは放射迷路学習に異常(図1)。 | 核内NF-rBが振動することが知られるが、これまで3Dシミュレーションは存在しなかした。 そこで空間パラメータの影響を調えるためにA-Cellによる3Dシミュレーションを行う。 七 七 は も は間辺の空間パラメータを明らかにするため、特徴細胞の細胞体を | 様した球形細胞で検討(コンパートメント数:約400,000、図2)。 - IKK活性化→IkB分解→NF-kB活性化→ NF-kB液内移行 →IkB産生→ NF-kB済軟い移行をモデル化(図2)。 - A-Cell画递化のためにこれまでのMPIバージョンを改良。 - かしモデル手液存的なMPI点速化を実施。 - 44目 | nack − 核/細胞質体積比、拡散定数、核膜輸送量、IkB翻訳場所の各空間パラメータすべてが振動パターンに影響した(図3)。 | MPI parallelization of A-Cell>99.99% MPI parallelization of A-Cell>99.99% 一 WPI parallelization of A-Cell>99.99% 一 WPI parallelization of A-Cell>99.99% - より現実的な細胞形態やオルガネラの影響を明らかにし、かつ並列化率のさらなる向上を図る。 | 5 | SNESTの京上での性能 R | (1) Runtime improvements speedup α = 0.59 @ 65536 cores (strong scaling) hybrid MPI + OpenMP code (simulation and wing) achieved phase VI on K | (2) Memory improvements storage of neurons and synapses with heterogeneous plasticity rules employing sparse representations enables efficient distribution on boots | Milliouk cores (3) production network 5-10 ⁸ (4) production network 5-10 ⁸ (5) production n | (4) Next milestone: N_{max}=10° theory of memory usage possible with present code's memory consumption on full machine size of K (end of red broken line) |



| 学術的成果 😽 💮 🔝 | 2012-2008(OCT-) 2009 2010 2011 2012- | 神 篠笛 話 2 3 2 9 5 シミュレーション 9 5 | | 全視覚系 0 0 0 0 0 2 2 2 2 2 3 1 1 5 4 5 5< | 昆虫嗅覚系 3 3 1 5 4 シミュレーション 3 1 5 4 | i文掲載いイライト Helias, et al. (2012). Frontiers in Neuroinformatics, 6(26). Nishiyama, et al. (2011). Nature Cell Biology, 13(6). | Linden, et al. (2011). Neuron, doi:10.1016/j.neuron.2011.11.006. Honda, et al. (2011). Journal of Neuroscience, 31(4). Fujii, et al. (2011). Proceedings of National Academy of Science USA, 108(44). Trainana et al. (2010). Molecular Science Biolocy, 6(304). | ronyama, et al. (2009). Nature Neuroscience , 12. | NeuroMorphoKit の概要・アプローチ 💮 💽 | モデル モデル パダの容 パクテンプ・レ伝達は、二次元 ディンホンローに達は、二次元 オーレンパートントに森在する各種 アンパートントに存在する各種 アクニンプ、一トントに確定して確率的に重 オーレビング、 | ・脱重合、キャッピング、分岐を むイベントを生成する。 脳膜に膜、ードを頂点としたポ ゴンで表現、エネルギー最適化 | より形態を来のる。 日する計算法 並列化の方法 開発言語と主なライブラリ 最急降下法 ・ 膜エネルギー最適化、アクチン ・ C/O+ Runge-Kutta法 OpenMP, MPIIによるハイブリッ ・ MPI, OpenMP, GSL, NetCDF ド並列化 | 16 |
|-----------------------|--------------------------------------|---|---|--|--|--|---|--|---|--|---|---|----|
| N | Ī | 14 11 | | | | """" | | | | 計 | ・も能ご | <u>5</u> 使 | |
| C. () () | ソフトウェアの最 ナュレー | | | アクチン緑維数 1.2×10 ⁶ | 雒問教5×10⁸ シナプス教 3×10¹² | 視細胞(錐体)数 1.5×10 ⁶ | □ソパートメソト 数10 ⁴ 雒튭数8×10 ⁵ | 13 | _{7 7} 👧 | ig, et. al., Nat Cell Biol, 2001 ig: fringing in theolast viat tab | て、柔軟に変化 :件 Growth cone at axonal tip | by Dr. Kaoru Katoh | 15 |
| 或状況 <mark><</mark> | 京上のストロン | シンケナード(滅) デコア(数) | | 0.994(8192) Hybrid | 0.59(65536) Hybrid | 0.89 (RICC 1024) MPI | 0.83(196603) Hybrid | | レチフィジッ 8/膜~ | Solution and Solution | 各の再編に依存し と学反応の境界条 ァッチン ^{重合阻害} | Biophys J. 2004 | |
| こアの達 | 京上の達成並 | | | 8192 | 196608 | 1504 | 196608 | | におけるマ) ル伝達/骨‡ | | 笛 時 岐 - | Gerisch, et. dl | |
| ソフトウュ | RICC上の 株型 (一位) 業 | | | 4096 | 8000 | 2460 (whole system 1024) | 8192 | | のののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののののの | 2 達 を情報処理 | イ 茶 参 御 御 御 御 御 御 御 | わら体から放射状に存在 皆要素を統合して | |
| ar in | | | 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - | 事業 番翫 シミュノーション NeuroMorphoKit | 画 字回路 ☆ミュレーション NEST | 全視覚系 シミュレーション VSM | 昆虫嗅覚系 シミュレーション IOSSIM | | いた。 | 油 昭 内 ン グ ナ ノ 伝- - < | 41日 41日 | 細胞膜直下に多く存在 | |



NEST による局所回路シミュレーション

五十嵐 潤

次世代計算科学研究開発プログラム 脳神経系研究開発チーム 特別研究員



発表者紹介

- 2007 年 3 月 九州工業大学 生命体工学研究科 脳情報専攻 博士課程修了
- 2007 年 4 月 九州工業大学 生命体工学研究科 脳情報専攻 研究員
- 2008 年 7 月 ホンダ・リサーチ・インスティチュート・ジャパン 客員研究員
- 2009 年 4 月 次世代計算科学研究開発プログラム 特別研究員

研究分野

計算論的神経科学、大規模神経回路シミュレーション、神経科学










Multi-scale simulation of subcortical

visual systems with real-time

robotic visuo-motor feedback

Jan Morén Brain and Neural Systems Team, Integrated Simulation of Living Matter Group, RIKEN Computational Science Research Program



発表者紹介

| 1995-11 | MSc Computer Science, Lund, Sweden |
|---------|--|
| 2002-09 | PhD Cognitive Science, Lund, Sweden |
| 2002-06 | Research Assistant, Lund, Sweden |
| 2003-11 | Post-Doc, CRL/NICT, Kyoto, Japan |
| 2005-05 | Visiting Researcher, NICT, Kyoto, Japan |
| 2006-09 | Researcher, ATR, Kyoto, Japan |
| 2009-02 | Researcher, Kyoto University, Kyoto, Japan |
| 2012-11 | Researcher, OIST, Okinawa, Japan |

研究分野

Integrated Systems Biology Laboratory, Department of Systems Science, Graduate School of Informatics, Kyoto University











Conclusion

- We have built a large-scale, spiking neuron-level model of the Superior Colliculus
- Reproduces burst neuron output profiles
 - Saccade Main Sequence behaviour
- Suggest spreading activity amount (but not direction) tracks saccade end time
- Real-time distributed visuomotor control of robot by large-scale cluster model
 - Practical approach to test computationally intensive models in the real world

大規模脳数理モデル構築プラットフォーム による視覚系数理モデルの構築

臼井支朗

豊橋技術科学大学・エレクトロニクス先端融合研究所 特任教授 理化学研究所・脳科学総合研究センター

神経情報基盤センター(客員主幹研究員)



発表者紹介

- 1974年1月 カリフォルニア大学大学院博士課程単位取得満期退学
- 1974年3月 名古屋大学工学部助手
- 1974年6月 Ph.D (カリフォルニア大学)
- 1979 年 4 月 豊橋技術科学大学講師
- 1980年4月 豊橋技術科学大学助教授
- 1986年4月 豊橋技術科学大学教授
- 2002 年 4 月 理化学研究所脳科学総合研究センター・ニューロインフォマティクス 技術開発チーム チームリーダ
- 2003 年 3 月 豊橋技術科学大学退職
- 2003 年 4 月 豊橋技術科学大学 名誉教授 客員教授
- 2005 年 4 月 理化学研究所脳科学総合研究センター・神経情報基盤センター 副センター長兼任
- 2007 年 4 月 同センター長兼任
- 2012 年 10 月 豊橋技術科学大学エレクトロニクス先端融合研究所 特任教授 理化学研究所脳科学総合研究センター・神経情報基盤センター 客員主幹研究員兼任

研究分野

ニューロインフォマティクス、神経情報工学、視覚生理工学

| Image: Second Secon | 研究開発成果:眼球運動・眼光学モデル 研究開発成果:眼球運動・眼光学モデル 小体注視時の眼球運動の再現とそれを考慮した詳細な網膜像の再現 ・物体注視時の眼球運動のも現とそれを考慮した詳細な網膜像の再現 ・酸物・塩野デーク:進入(細胞・パラメータの推定 酸物・塩野デーク:こ進入(細胞・パラメータの推定 度、酸物・塩野デーク:こ進入(細胞・パラメータの推定 度、食物・酸パラメータの推定 ・酸物・酸酸、ワラメータの推定 ・酸物・酸酸、ワラメータの推定 ・酸物・酸酸、ローク ・酸物・酸酸、ローク ・酸物・酸酸、ローク ・酸物・酸、ローク ・酸物・酸、ローク ・酸物・酸、ローク ・酸物・酸、ローク ・酸物・酸、ローク ・酸物・酸、ローク ・酸物・酸、ローク ・酸、酸、酸、ローク ・酸、酸、酸、ローク ・酸、酸、酸、酸、酸、酸、酸、酸、酸、酸、酸、酸、酸、酸、酸、酸、酸、 ・酸、酸、酸、酸、酸、酸、酸、酸、酸、酸、酸、酸、 ・酸、酸、酸、酸、酸、酸、酸、酸、 ・酸、酸、酸、酸、酸、 ・酸、酸、酸、 ・酸、 ・酸、 |
|--|--|
| | (現金・アプローチ) (新会・アプローチ) (新会・アプ |





| C. | | 14 |
|--------------------------------------|---|---|
| | Ε | ISLIM成果報告会2012 |
| <u>ور</u> | | |
| (なり) (なり) (なり) (なり) (なり) (なり) (なり) (| > モデル統合環境PLATOの構築 サブシステムレベルでモデル構築・並列化を行い、共有フォーマットにてIOレベルで結合 することで大規模モデル構築を可能にした > 助びシステムモデルレベルでは、眼球運部にて微小サッカード生成のメカニズムを > サブシステムモデルレベルでは、眼球運部にて微小サッカード生成のメカニズムを Computational Levelで明らかにした > 税細胞モデルならびに双極細胞 - 神経節細胞層における信号処理(細胞応答)を情報量を 指標として評価中 本プロジェクトにて、上記成果を学術論文5本、国際会議13件、国内会議11件で報告した 今後の課題 > Platonicイン分フェースにおける速度低下の改善 > 創本モデルの速度向上 > 離体モザイクで実施した情報量評価の地領域への適用 > 確休モザイクで実施した情報量評価のもないの部品 | G.g. 彼小昭环代里期による7次34月77月上なC い計画で注め/る 13 |

「京」による

無脊椎動物嗅覚系シミュレーション -背景と現状と将来への展開-

神崎 売平 次世代計算科学研究開発プログラム 脳神経系研究開発チーム



発表者紹介

- 1986年3月 筑波大学大学院生物科学研究科生物物理化学専攻博士課程修了
- 1987年1月 アリゾナ大学神経生物学部門博士研究員
- 1991年7月 筑波大学生物科学系助手、講師、助教授
- 2003 年 4 月 筑波大学生物科学系教授
- 2004 年 4 月 東京大学大学院情報理工学系研究科知能機械情報学専攻教授
- 2006 年 4 月 東京大学 先端科学技術研究センター生命知能システム分野教授

その間

1996 年 9月-1999 年 8月 科学技術振興事業団さきがけ研究 21 研究員(兼任) 2003 年 4月-現在 アリゾナ大学神経科学部 Adjunct Professor

研究分野

神経行動学、感覚生理学、生物-ロボット融合学





ISLiM ソフトウェア研究開発報告会(2013.1.10-11)





ISLiM ソフトウェア研究開発報告会 (2013.1.10-11)



生命基盤ソフトウェア開発・高度化チーム成果報告

ISLiM ソフトウェアの高度化

泰地 真弘人

次世代計算科学研究開発プログラム 生命体基盤ソフトウェア開発・高度化チーム チームリーダー



発表者紹介

1992 年 3 月 東京大学理学系研究科物理学専攻博士課程修了

1992年4月 東京大学教養学部助手

1997年1月 統計数理研究所助教授

- 2002 年 4 月 理化学研究所ゲノム科学総合研究センターチームリーダー
- 2008 年 4 月 理化学研究所基幹研究所グループディレクター
- 2011 年 4 月 理化学研究所生命システム研究センター生命モデリングコアコア長

研究分野

高性能計算・計算機アーキテクチャ

ISLIM ソフトウェアの高度化

泰地 真弘人

次世代計算科学研究開発プログラム 生命体基盤ソフトウェア開発・高度化チーム チームリーダー

1. 目的

京コンピュータは、80,000 プロセッサ、640, 000 コア以上を有する大規模な超並列計算機です。 このような規模の計算機を使いこなすためには、 並列処理、特に京コンピュータに通暁した研究者 による支援体制が必要です。そのために、分子動 力学計算コアソフトウェアの開発などの独自開 発を通じた人員育成を行うと同時に、他チームで 開発されたソフトウェアの性能評価・開発支援を 行っています。さらに、開発容易化のためのミド ルウェア開発・可視化ソフトウェア開発を行い、 産業利用に向けた大規模仮想化合物ライブラリ の開発を進めています。

2. 現時点での成果

2.1 分子動力学コアソフトウェア

「京」用アプリケーションの最適化ノウハウの蓄 積・提供および最適化の参考コードの提供のため、 分子動力学コアプログラムの開発・最適化を行な いました。主要な計算部分であるカットオフ法カ ーネルでは 60%以上の演算効率を達成しました。 カーネル部分の SIMD 化率は全命令の 57%、浮 動小数点演算命令に関しては 99%以上となり、

「京」への最適化が十分行なわれています。最適 化阻害要因(特に C++コンパイラの場合)に関する 知見は、他チームの支援業務にも活かされていま す。ノードあたり 6000 原子程度の計算規模に対 し、全ノード並列で演算速度 4Pflops・効率 40% を達成しており、ノードあたり 1000 原子以上の 計算規模であれば「京」全ノードまでの並列性が あることを確認しました。

長距離相互作用計算においては、現在主流の PME法はFFTが大域通信を必要とするため「京」 の全ノード規模での並列性は得られないと考え られるので、高速多重極展開法(FMM)を実装中で す。単純な実装での評価ではありますが、「京」 全ノード規模での並列性を確認できています。

開発したコードは、「京」での最適化・分子動力

学計算の参照コードとして自由に再利用できる ように公開する予定です。

2.2 ソフトウェア高度化

ソフトウェアの性能評価・チューニング支援を行っています。第一走者アプリケーションを中心に、 性能評価報告を行いました。チューニング支援については、特に神戸拠点における京に向けての開発支援を中心に、開発実施本部との窓口・京に向けたチューニング外注の窓口機能を担っています。

2.3 可視化ソフトウェア

大規模データを扱える並列可視化システム ISL-LSVを開発しました。

2.4 並列処理ミドルウェア

ソフトウェア開発容易化のためのミドルウェア Sphere を開発しました。連続体シミュレーショ ンを中心に、ソルバ群の組み合わせにより容易に 並列化されたソフトウェア開発を可能にします。 本ソフトウェアは臓器全身スケール研究チーム で開発されている ZZ-EFSI/ZZ-HIFU、細胞スケ ール研究チームで開発されている RICS で用いら れています。

2.5 大規模仮想化合物ライブラリ

10 億化合物規模の合成可能性をもつ仮想化合物 ライブラリを、東大工学部船津研究室を中心に開 発しています。

2. プロジェクト終了時の達成目標

プロジェクト終了時に、より多くのアプリケーションが京コンピュータ上で十分な性能を発揮していることが我々の目標です。また、創薬応用に向けた 10 億化合物規模の大規模化合物ライブラリを完成させます。

以上





ISLiM ソフトウェア研究開発報告会 (2013.1.10-11)



MD コアプログラムの最適化と性能評価

大野 洋介

次世代計算科学研究開発プログラム 次世代生命体統合シミュレーション研究推進グループ 生命体基盤ソフトウェア開発・高度化チーム



発表者紹介

1996年3月 東京大学大学院理学系研究科天文学専攻博士課程修了

1996年4月 理化学研究所 基礎科学特別研究員

1999 年 4 月 理化学研究所 計算科学研究室 協力研究員

2001 年 4 月 理化学研究所 ものつくり情報技術統合化研究プログラム 研究員

2004 年 4 月 理化学研究所 ゲノム科学総合研究センター(GSC) 研究員

2008 年 4 月 理化学研究所 次世代計算科学研究開発プログラム 上級研究員 (現兼務)

2012 年 4 月 理化学研究所 生命システム研究センター(QBiC) 上級研究員

研究分野

計算科学、専用計算機

| œ. | N | C. | | 4 |
|----|--|-------------|--|----------------------|
| 目的 | 「京」での最適化ノウハウを他チームへ提供する。 コンパイラ・ハードウェアの能力を把握する。 キロパイラ・ハードウェアの能力を把握する。 特に、最適化の知見を描のチームへ提供する。 目本的な最適化の知見を他のチームへ提供する。 目本的田可能なコードとして提供する。 日本利田可能なコードとして提供する。 最適化事例 最適化事例 最適化事がおりまし」 | 現在までの研究開発成果 | Cutoffカーネルの演算効率60%以上 SIMD, Software-pipelining データ形式の変更 データ形式の変更 単純配列 "IP文のマスク演算化 "IP文のマスク演算化 "IPOFUネットワークに適した通信 10FUネットワークに適した通信 10FUネットプークに適した通信 10FUネットプークに通信が少ない PMEで使われるFFTIC比べ大域通信が少ない | ISLIMンフトウェア研究開発報告会 |
| | | | | |
| | <section-header><section-header><section-header><text><text></text></text></section-header></section-header></section-header> | 炎 想・アプローチ | 研究開発コードの概要 基本的な分子動力学計算 麦本的な分子動力学計算 デンサンブル:NVE,NVT,NPT デンサンブル:NVE,NVT,NPT 時間積分: velocity Vertet 時間積分: velocity Vertet シーロン相互作用: cutoff, PME, FMM ・ 一のン相互作用: cutoff, PME, FMM ・ 並列化: 直方体空間分割, cell index method ・ アプロー子 ・ シブロー子 ・ シブロー子 ・ アプロー子 ・ アプロー子 ・ アプロー子 ・ アプロー子 ・ アプロー子 ・ アプロー子 ・ シブロー子 ・ ション ・ シブロー子 ・ ション ・ シブロー子 ・ シブロー子 ・ ション ・ ション<td>ISLIMソフトウェア研究開発報告会 3</td> | ISLIMソフトウェア研究開発報告会 3 |

| C. MEMA | | 9 | C. Hand | 82,944 | 48x48x36* | 242,044,272 112.085 | 91.641 | 15.067 | 5.376 | 4.031 | 38.0 | |
|---------------|--|--------------------|----------------------------------|-----------------|----------------|----------------------------------|-------------------|--------------|---------|--------------------|------------|---|
| | | | | 79,872 | 48x52x32 | 322,340,330 112.414 | 91.262 | 15.820 | 5.332 | 3.871 | 37.9 | 4 step |
| 。最適化 | 変換 | 報告会 | лg | 32,768 | 32x32x32 | z14,377,964 111.672 | 92.329 | 14.012 | 5.332 | 1.599 | 38.1 | energy every 2 |
| カーネル | な た や か や 端 に か や い 端 に や い か い い い い や い か い か い か い か か い 一 か む い か い い の い い い い い い い い い い い い い | フトウェア研究開発 | ak Scali | 4,096 | 16x16x16 | 20,/9/,248 | 91.601 | 14.011 | 5.574 | 0.201 | 38.3 | de, calculate ogy 48x54x3; |
| utoff J | 设 まで い し し 他 の 一 た の 一 に と の に の し が し い し い し の し が し い し い し い し い し の し の し い し い し の し の | A WINSI | We | 4 512 | 4 8x8x8 | / 3,349,000 8 110.535 | 2 91.528 | 4 13.694 | 2 5.313 | 3 0.025 | 0 38.5 | 42 atom/noo ysical topolo |
| | イコ 必 参 照 brefe 参 照 brefe ーネトッシー ーキトッシー ーキトッシー ーキー 小し た の オ フ デ デ ー オ し ボ の の 子 し の の の の の の の 一 の の の 一 の の 一 の 一 の 一 | | | ė | 4X4X | 418,70 | 91.62 | n 12.12 | 5.31 | 0.00 | 39. | off 28Å, 6,5 fferent to ph |
| | ・ キョー・ S ー ー や間・座・ M 単・カ・ ジ接ゴ標か G 純 個 小 o | | | mber of nodes | ology | niber of atoms te (ms / step) | Force | Communicatio | Other | formance :LOPS) | ciency (%) | * Crt |
| ISI YU | | | ISI YW | Nur | Top | | | | | Per (PF | Effic | |
| Sutoff カーネル性能 | Number of atoms per CPU6,542Cutoff radius28 Å + 2 Å marginCutoff radius28 Å + 2 Å marginNumber of pairs per atom8,835(+ 2,031 in margin)Kernel FLOPS per CPU80, 516 MFLOPSEfficiency62.9 %SIMD ratio57.2 % *SIMD ratio57.2 % *9 % of floating-point calculations are SIMD | ISLIMソフトウェア研究開発報告会 | Cutoff法の並列性能 | 26,797,248 atom | 7,536,726 atom | 3,349,656 atom 418.707 atom | , 6.542 atom/node | | 3 | 818 atom/node | - mons at | Cutoff 28Å, calculate energy every 4 step |



創薬プラットフォーム -大規模バーチャルライブラリの開発-

船津公人

VLSVL 合成可能な化学構造および反応スキームからなる 大規模仮想ライブラリ 発表者



発表者紹介

- 1983年3月 九州大学大学院理学研究科化学専攻博士課程修了
- 1984年3月 豊橋技術科学大学工学部物質工学系 助手
- 1988年3月 豊橋技術科学大学工学部知識情報工学系 助手
- 1992 年 4 月 豊橋技術科学大学工学部知識情報工学系 助教授
- 2004年4月 東京大学大学院工学系研究科 教授

研究分野

ケモインフォマティックス(化学情報学) 関連する研究分野としてプロセスシステム工学、有機合成化学、分析化学、触媒化学、 創薬化学など







ISLiM ソフトウェア研究開発報告会(2013.1.10-11)


| キ や や キ キ や や ・ ・ ・ ・ ・ ・ ・ ・ | チ令を群を補ノ」 6位DBJUは出したTransformを注田1 と | PTIPをます。 ので、 なたののの、 ので、 ので、 ので、 ので、 ので、 ので、 ので、 | | 創生した化合物は、人力構造群の特徴を引き継ぎ、 華アだいころ水会物を挿アオスケ 華アだいころ構造を出ナオス値向を確認した | 米门やいしゅうゴジウゴー、ター、米口やいしゅすりのコン、タマシウ番号へい | ■新規性: | 市販化合物ライブラリとの対比で 重複0.004%、 新規性の確保を確認済み | ■分布: (Lipinski's ruleの指標によるチェック) | 新規な化合物構造からなる出力構造群は、 1 本博学部の公本去芝士社語ナス43、 桂仲左に叫「アハス | 人が構造群のガヤで有工地版するが、特性で文映している | ■規模: 。 本 // 竹井 // 井井 / オーン | 人刀化子権宣辞に対して、順合成方向にそれぞれ3段階、 | 各方向で5億件の化学構造による 総10 <mark>億化合物からなる</mark> バーチャルライブラリを構築し、ライブラリの特性評価を進める | ISLIMソフトウェア研究開発報告会 18 |
|---|------------------------------------|--|--------------|---|--------------------------------------|-------------|---------------------------------------|-----------------------------------|--|----------------------------|-------------------------------|----------------------------|--|-----------------------|
| | | 入力化合物群 B | input output | 98.5% 88.1% | 99.9% 99.9% | 9.9% 99.6% | 89.3% 87.0% | 99.5% 95.0% | 99.4% 95.1% | 88.3% 79.3% | 98.9% 91.0% | 99.9% 99.6% | れる可能性が増す | 17 |
| 分布の影響 | | 入力化合物群 A | input output | 85.6% 75.3% | 100.0% 99.9% | 99.9% 99.7% | 79.9% 79.9% | 98.2% 93.0% | 99.0% 96.4% | 70.7% (64.8%) | 97.3% 89.9% | 99.9% 99.7% | (の影響を受ける :出力構造群が得ら | び研究開発報告会 |
| 群の特性: | | rule of 5 | | <500 | 1 2 √ | ≤10 | 1 √2 | ≦10 | ≦140 | ī値 | 結合数 | 《素結合受容体数 | 、力構造群の特徴 月すれば、適切な | ISLiMソフトウェア |
| _ | | | | | 数 | 本数 | 直(cLogP) | 数 | Ų) | 配係数計算 | /S. 回転可能 | -体数 VS. 水 | 分布は、入 造群を活用 | |

ポスター一覧



Orange: REMD Red, Blue and Green: MD

Conformation

MD{T4}

MD{T3

MD{T2

MD{T1

Replica-exchange Molecular Dynamics

Accept!

eject

K. Hukushima, and K. Nemoto, J. Phys. Soc. Jpn. 65, 1604, (1996)
 Sugita, and Y. Okamoto, Chem. Phys. Lett. 314, 141–151, (1999)
 Sugita, A. Kitao, and Y. Okamoto, J. Chem. Phys., 113, 6042-6051,(2000)

Conformation

Accept!

10000

Free

Replica 4

Replica 3

Replica 2

Replica 1



<u>Naoyuki Miyashita^{1,2,3},</u> Suyong Re⁴, and Yuji Sugita^{1,3,4}



Introduction Replica-exchange Molecular Dynamics (REMD)

Replica-exchange molecular dynamics (REMD) simulation [2] is one of popular enhanced conformational sampling methods, and it has been widely used to predict the protein conformation, configuration and so on. In the original REMD simulation, the dynamics of noninteracting replicas of the system, each at a specified temperature, are performed by molecular dynamics (MD) simulation. Pairs of replicas that have nearest neighbor temperatures are exchanged every few steps using the Metropolis criterion to ensure detailed balance.

Multi-dimensional replica-exchange method (MREM) [3] is one of the extentions. It exchanges not only nearest neighbor temperatures but also the nearest neighbor parameter sets that relate to a potential energy. Exchange probability can be written:

 $\min[1, \exp(-\Delta)], \qquad \Delta = \beta_m(U_{\lambda m}(q_j) - U_{\lambda m}(q_i)) - \beta_n(U_{\lambda n}(q_j) - U_{\lambda n}(q_i))$

where, $\beta_m = 1/(k_B T_m)$, T_m : temperature, k_B : Boltzmann constant, and , $U_{\lambda m}(q_i)$ is the potential energy of m-th state at i-th replica.

MREM needs huge computer resources, and is promising method for the future enhanced sampling. We have developed Replica-exchange Interface (REIN) program package to perform MREM simulation using conventional MD program packages, such as NAMD2, MARBLE and so on.





→ NAMD2 and MARBLE are available. MARBLE Fortran95/03, MPI 2 **Binary format** NAMD2 Control RFIN **Conventional MD packages** MPI_COMM_SPAWN etc... Feature 2) Multi-dimensional REMD (MREM) →It can exchange not only nearest neighbor temperatures, but also nearest neighbor distances etc... The parameters depend on the function of "constraint" in the used MD package program. Feature 3) Three supportive software are included →Input builder, Batch builder, REMD converter Supportive software Now available Area http://www.islim.org/islim-dl_j.htm, input builde REMD converter batch builde Twitter: @rein_devel E-mail: yukimya+rein@gmail.com REIN Input 🗼 Submit job Output One of the applications Structure prediction of N-glycans in solution We performed REMD simulation to predict the structures of N-glycans with and without bisecting GlcNAc using REIN [4]. These simulations show that the N-glycan modification like the bisecting GlcNAc can select a certain "key" in the framework of "bunch of keys" mechanism. **REIN/NAMD2 on RICC** 260 hours/512 core ~10,000 atoms, 300K-500K 64 replica x 52 ns = 3.3 micro sec. [4] S. Re, N. Miyashita, Y. Yamaguchi, Y. Sugita, *Biophys. J.* 101, L44-L46, (2011) [5]Wataru Nishima, <u>Naoyuki Miyashita</u>, Yoshiki Yamaguchi, Yuji Sugita, and Suyong Re, J. Phys. Chem. B, 116 (29), pp 8504–8512, (2012)

















謝辞

本発表の結果の一部は、理化学研究所が実 施している京速コンピュータ"京"の試験利用 によるものです。

P.

赤:細

青:核

緑:ミ

トコン

ドリア

KX= 16 />
<VoxelWidth
wx="1.000000"
wy="1.000000"
wz="1.000000" />

</Elem>

</Elem> </Elem>

<Elem name="M2' id="2">

M2. svx" no" /> 11:E1″

"sph" E1.sph" no" /> Data>



Fig. 5 それぞれの物質の産生量について、トロンビンはほぼ同量、 フィブリンはCase1が高濃度となった。

4000 6000 8000 10000 12000 14000 step 0.00E+00

6000 8000 10000 12000 1400









謝辞

本研究は「次世代計算科学研究開発プログラム」の助成を受けて行なわれた 本研究の計算結果の一部は、RIKEN Integrated Cluster of Clusters (RICC) システムを利用して得られた.

参考文献

Sugiyama, K., Ii, S., Takeuchi, S., Takagi, S. and Matsumoto, Y. (2010) Full Eulerian simulations of biconcave neo-Hookean particles in a Poiseuille flow, Comput. Mech., 46: 147-157. Nagano, N., Sugiyama, K., Takeuchi, S., Ii, S., Takagi, S. and Matsumoto, Y. (2010) Full-Eulerian finite-difference simulation of fluid flow in hyperelastic wavy channel, J. Fluid Sci. Tech., 5·475-490

Sugiyama, K., Ii, S., Takeuchi, S., Takagi, S. and Matsumoto, Y. (2011) A full Eulerian finite difference approach for solving fluid-structure coupling problems, J. Comput. Phys., 230: 596-627.

Ii, S., Sugiyama, K., Takeuchi, S., Takagi, S. and Matsumoto, Y. (2011) An implicit full Eulerian method for the fluid-structure interaction problem, Int. J. Numer. Meth. Fluids, 65: 150-165.

Ii, S., Gong, X., Sugiyama, K., Wu, J., Huang, H., and Takagi, S. (2012) A full Eulerian fluid-membrane coupling method with a smoothed volume-of-fluid approach, Comm. Comput. Phys., 12:544-576

Takagi, S., Sugiyama, K., Ii. S., and Matsumoto, Y. (2012) A review of full Eulerian methods for fluid-structure interaction problems, J. Appl. Mech., 79: 010911.

Ii, S., Sugiyama, K., Takagi, S. and Matsumoto, Y. (2012) A computational blood flow analysis in a capillary vessel including multiple red blood cells and platelets, J. Biomech. Sci. Engrg., 7: 72-83.









現在取り組んでいる課題:肥大型心筋症のメカニズム解明

0-7 血管内の血流シミュレーションソフトウェア開発・検証研究



Introductions

• Many computational simulations of blood flow have been done by using a body-fitted mesh and a fixed mesh to understand correlation between hemodynamics and the development of vascular diseases.

Aim

• To introduce a supporting tool for hemodynamics analysis on voxel system of Eulerian frame to reduce costs(especially man-based time) and to produce user-friendly tools for non-experts (i.e. medical doctors).

•Long-term purpose is to study the blood flow in a very complicated vessel, a compliant vessel and a deformable organ without the body-fitted mesh and a re-mesh process.

Numerical Models

• Simplified s-shaped geometry for femoral artery (Model I)

· Realistic geometry for carotid artery (Model II & III)



Numerical Methodology



 Volume integral over a control volumes in Navier-Stokes equations is calculated on staggered grids in several spatial finite difference discretizations and time discretization with FAVOR (fraction area/volume obstacle representation) method and HSMAC (Highly Simplified MAC) Method.

•FAVOR method on fixed grid and voxel size play a role in the definition of contour of vascular surface as well as results in computational simulations in Figure 4.

• Boundary conditions (Parabolic steady inlet flow (Remean = 500), no-slip wall and free velocity gradient outlet condition)

Results

• The comparisons of velocity of combination method of both FVM and FDM in this study (Method I) and spectral/hp element method (Method II)[1].



Spatial Convergence test with the double bend model Method I. Estimation of accuracy order by substituting the equations as below.

Method II. Grid Convergence Index (GCI) method is based on estimated fractional error.

a design of the second s

Method III. Estimation of accuracy order with an assumption that the results on the finest voxel is very close to exact solution.

| $1 \leq 1 \leq m \leq n \leq n$ |
|---|
| and that the endance is a significant $r = \log \log \frac{1}{2} e^{-2}$, here everywhere |
| where the HTM particular is a specific mean of the mean of the barrier dependence $\lambda_{\rm eff}$ |
| serve line in |
| |
| |

ورد المحمد الم

. .

Conclusions

Our findings confirm the voxel based simulation software can be applied to blood flow simulations for biomedical analysis without a surface reconstruction procedure and a complicate mesh generation. **Reference**

Kelerenc

[1] K. E. Lee, K. H. Parker, C. G. Caro, S. J. Sherwin, The spectral/hp element modelling of steady flow in nonplanar double bends, International Journal for Numerical Methods in Fluids 57 (2008) 519-529

Acknowledgements

This research was supported by Research and Development of the Next-Generation Integrated Simulation of Living Matter, a part of the Development and Use of the Next-Generation Supercomputer Project of the Ministry of Education, Culture, Sports, Science and Technology.



MVC





R

三澤計治¹,長谷川亜樹¹,角田達彦^{1,2} (理化学研究所 次世代計算科学研究開発プログラム¹,理化学研究所 ゲノム医科学研究センター²)

研究目的

- 近年のDNA配列決定技術の進歩により、大量のゲノム配列が解析されるようになりました。ゲノム配列は、個人の間に違いがあります。これを利用し、数万人に対して、ゲノム全体に渡り、患者群とコントロール群との遺伝子頻度を統計的に解析することによって、疾患関連遺伝子を発見する手法がゲノムワイド関連解析(GWAS)です。
- GWASでは多重検定によるtype I errorが問題になります。多重検定の問題に対しては、Bonferroni correctionが広く行われていますが、SNP間でlinkage disequilibriumがあるため、疾患関連遺伝子を見落 とす恐れが出てきます。
- 我々は、多項分布を用いて、SNP間にlinkがある場合でも正確にtype I errorを計算する式を導きました。 それに基づき、genome wideのP値を適切なthresholdを設定する方法を開発しました(Misawa et al. 2008)。 また、高速計算のため、マルコフ連鎖モンテカルロ法(MCMC法)でtype I errorを計算する方法を開発しま した。これらの方法の他にpermutationに基づいたRAT法も開発されています(Kimmel and Shamir 2006)。 これらの方法は時間がかかるため、ゲノムワイドの解析を行う際には、高速に計算する必要があります。

ParaHaploの開発

- 我々は京速コンピュータ「京」(右図)の上で、ゲノムワイド関連解析の ためのプログラムParaHaploを開発しました。「京」は8万ノード・64万コ アの計算ユニットからなる非常に速いスーパーコンピュータです。
- ParaHaploでは、ゲノムワイドのhaplotype多型データを分割し、データ を分散入力させた上で、ユニット間・コア間のハイブリッド並列により、 Misawa et al. (2008)の方法とRAT法の両方に関して、高速計算を実 現しました (Misawa and Kamatani 2009, 2010, 2011)。
- ParaHaploのソースコードは以下のサイトで公開されています。 http://sourceforge.jp/projects/parallelgwas/releases/51104



実際のデータをParaHaploで解析した例

Wellcome Trust Case-Control Consortium (WTCCC)から、2型糖尿病の患者1,924人とコントロール2,938人の遺伝子型データの提供を受けました。右図は22番染色体に関してparaHaploを用いたハプロタイプ関連解析を行った結果です。赤い点がparaHaploによって初めて2型糖尿病と関連があると推測された領域になります。



謝辞

- 京での計算に関しては京速コンピュータ「京」の試験利用での結果です。
- また、PCクラスタでの性能計測に関しては理化学研究所情報基盤センターのRICCを使用しています。

参考文献

Kimmel, G. and Shamir, R. (2006) A fast method for computing high-significance disease association in large population-based studies. Am J Hum Genet, **79:** 481-92.

 Misawa K, Fujii S, Yamazaki T, Takahashi A, Takasaki J, Yanagisawa M, Ohnishi Y, Nakamura Y, Kamatani N (2008) New correction algorithms for multiple comparisons in case-control multilocus association studies based on haplotypes and diplotype configurations. J Hum Genet 53:789-801
 Misawa K, Kamatani N (2009) ParaHaplo: A program package for haplotype-based whole-genome association study using parallel computing. Source

Code Biol Med **4**:7 Misawa K, Kamatani N (2010) ParaHaplo 2.0: a program package for haplotype-estimation and haplotype-based whole-genome association study using parallel computing. Source Code Biol Med **5**:5

Misawa K, Kamatani N (2011) ParaHaplo 3.0: A program package for imputation and a haplotype-based whole-genome association study using hybrid parallel computing. Source Code Biol Med 6:10



参考文献

1. * <u>Fujimoto A</u>, Nakagawa H, Hosono N, Nakano K, Abe T, Boroevich KA, Nagasaki M, Yamaguchi R, Shibuya T, Kubo M, Miyano S, Nakamura Y, Tsunoda T (2010) Whole-genome sequencing and comprehensive variant analysis of a Japanese individual using massively parallel sequencing. **Nat Genet** 42: 931–6

2. The International Cancer Genome Consortium (2010) International network of cancer genome projects. Nature 464, 993-8.

3. Fujimoto A, Totoki Y (equally contributed) et al. (2012) Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. **Nature Genetics** 44: 760-764

为 D-3: ExRAT 拡張RAT法による2SNP組合せの全ゲノム関連解析ソフトウェア

開発責任者: 角田達彦 (理化学研究所ゲノム医科学研究センター)

概要:

遺伝子間相互作用が発症リスクを変化させる疾患関連遺伝子の組合せ を全ゲノムで探索する。2SNP間の全組合せを超並列に行う方法と、SNP 間の連鎖不平衡(相関)も考慮した、より精密な方法の2種類を実装。前 者で全組合せをスクリーニングし、後者で経験的p値を求める手順を想定。





単一SNPの解析で相互作用の解析は困難 Logistic回帰と対数尤度比検定

目的と方法のまとめ

- ・目的:2つの遺伝子(SNP)が同時に作用して疾患 の原因となるものを探索する
- ・データ:全染色体上の50万SNP × 数千人 → 500000 x 500000 / 2 = 10¹¹組合せ
- 解析方法1:単純組合せのロジスティック解析 (all comb)

解析方法2: Permutation test に基づく経験分布 から計算し、より厳密なp値を出す (RAT)

2つのプログラムの適用方法

- ステップ1: 単純組合せロジスティック回帰 all_comb - 全体の計算量は比較的軽め
- 概念がわかりやすいので、多くの人が使うはず
- データ並列に適している
- 最初のスクリーニングに使う

J. ステップ2: permutation test による RAT

- ゲノム上の性質(連鎖不平衡)も考慮した計算
- 計算量が重いので、スクリーニング後の精査へ

n(m-1)/2個のSNP組み合わせ β個の組それぞれ 可能な分割表を求め 真のデータの統計量より 大きな統計量を与える パーミュテーション数hを計算 160

「京」上のall_comb

パーミュテーションテスト

ゲノムワイド関連解析の基盤構築 → 解析対象SNPは50万 ボンフェローニ補正は保守的すぎる

Permutation Test 統計量に関する事前の仮定を必要とせず現実的な補正

しかし通常のランダムパーミュテーションテスト(SPT)は 極めて計算負荷が高い

高速化 RAT (rapid association test) Kimmel and Shamir (2006)

Minine and Shanin (2006) Importance Samplingによる近似 分割表作成 + X²検定 → 尤, → 尤度比検定に拡張



赤に含まれるある1個のパーミュテーションの生起確率

サンプリング数を十分にとって足しあわせる → 赤の比率

RAT法の拡張

各種線形回帰モデルと尤度比検定への拡張

ŧ

2SNPの相互作用付きロジスティック回帰モデル

 $logit(P) = \mu + \alpha x_A + \beta x_B + \gamma x_A x_B$ t t ■ SNP B SNP間の の効果 相互作用

ExRAT:並列化

パーミュテーション メインルーチン

olis-Hasting法による分割表抽出

・ 抽出された分割表から ペーミュテーションd を抽出

d が得られる確率rをhから算出 ▼ rの平均 = p値の推定値

→ MPIによる並列実行 (x SNP組合せ数)

5

. SNP A の効果

より複雑な解析に対応するために

ケースとなる確率P

SNP数: m

↓ プリプロセス

からB個を選択

1

■:β~1000 -:MH法反復数~1000

- シノンスラ Co, I32/一トの 业列1生能を実証. データ:患者708人, 一般集団3397人, 8314SNP, 14100回 permutationで評価. ハイブリッド並列化対応: Preprocess処理とpermutation処理に OpenMPでのスレッド化の実装

- PCクラスタで8,192ノードの並列性能を実証.

Preprocess処理の見直しによるさらなる高速化:分割表探索処 理での計算処理の見直しにより、約40%程度高速化計算処理を 探索中に積み上げていけるように処理を変更し、全体の計算量 の削減に成功した).

遺伝子どうしが相互作用を起

こして疾患への発症リスクを

上昇する現象とそれらの遺伝

子を新たに疾患関連遺伝子と

して発見することができる

- ExRAT ver. 1.0- r440, phase III-0, 100% - 富士通コンパイラ(RICC, FX1,「京」,の全て)対応済, 実行成功, buildも可能.

P.

- MPIプロセスの増加によるメモリ使用量増加への対応:非同期通信用のためのバッファ確保が原因だったため、通信処理を見直し、処理完了待ちのバッファの大きさを最小とすることで対処した.
- WTCCCの実データを解析し、評価と結果を出した.

| Dataset | WTCCC type 2 diabetes |
|-----------------------------|------------------------------|
| Sample size | 1,924 cases + 2,938 controls |
| #SNPs (MAF>0.05, after QC) | 360,297 |
| #combinations | 6.49x10 ¹⁰ |
| #runs | 25 |
| Cores used / run | 24,576 |
| Disk space (stage in) / run | 1.5TB |
| Memory used / run | 123GB |
| Time / run (average) | 41 min. 17 sec. |
| Total time | ~17h. (24,576 cores used) |

MPIIによる 並列化







druggable gene networks from genome-wide DNA expression profiles, Pacific Symposium on Biocomputing (PSB 2006), 11, 559-571. tt, S. et al. (2004). Finding optimal models for small gene 0

Angiogenesis 12 (3), 221-229.

associa e34247.



東京大学

D-6: SiGN-SSM



開発責任者: 宮野 悟 (東京大学 医科学研究所 ヒトゲノム解析センター)



Yamaguchi et al. Finding module-based gene networks with state-space models - Mining high-dimensional and short time-course gene expression data. IEEE

systems by state space models. Genome Informatics, 21: 101-113, (2008).

203

Tamada et al. SiGN-SSM: open source parallel software for estimating gene networks with state space models.





SBiP を使用した京での SiGN の実行

実行中

Lange Contractor

Inches in the local division of the local di

E -

3ootstrap) ● SiGN-BN (NNSR) ● SiGN-SSM こ上で実行するためのコンポーネントを実装し

実行後

パイプライン上で京コ: ボーネントと既存コン ボーネントを結合する。 とで、京での出力デー をシームレスに処理で るようにした。

-

開発責任者: 宮野悟 (東京大学 医科学研究所 ヒトゲノム解析センター)

SBiP 開発メンバー: 関谷 弥生, 斉藤 あゆむ, 山口 類, 井元 清哉, 島村 徹平, 新井田 厚, 玉田 嘉紀, 宮野 悟

SiGN-BN (HC+Bootstrap)

SIGN-BN (NNSR)

概要

本研究は、データ解析融合チームで開発した技術とプログラム を、相互に連携させるソフトウェアプラットフォームとしての SBiP を構築することを目標とすしている。平成 24 年度は、平成 23 年度の開発で達成した SiGN を東京大学ヒトゲノム解析セン ターをテスト環境として実装した経験を踏まえ、京のゲートウェイ ノードから SiGN などの解析コンポーネントが実行できるように 改良を行い、ソフトウェアとしての完成度を高めた。ユーザは SBiP に用意されているさまざまな解析コンポーネントを組み合 わせながら、カスタマイズした解析フローを実行可能である。例 えば SiGN-SSM、SiGN-L1、SiGN-BN などの遺伝子の発現 情報からネットワークを推定するコンポーネント群を京の上で実 行し、それらの結果をユーザに整理し視覚化して表示することが 可能である。オンラインドキュメントの整備と改良は継続して行っ ている。また、平成 22 年度までに取得した肺がん及び乳がん 関連データについて、SBiP 上の解析パイプラインで解析処理を 行う上で必要となった解析コンポーネントの改良を行った。




















[※] H-1: cppmd 大規模並列用MDコアプログラム

開発責任者: 大野 洋介(理化学研究所生命体基盤ソフトウェア開発・高度化チーム)

Molecular Dynamics Simulation

The cell is constructed by molecules such as water, protein etc. Every molecules are constructed by atoms. These atoms influence each other and move complexly. Molecular dynamics (MD) simulation calculates the interaction and motion of atoms and reveal the structure and functions of the bio-molecules.

Because coulomb force that is force between charged atoms is affect to long distance, we must calculate force between all pair of atom. It takes very long time when the number of atom is large. Using parallel computer, it requires many

communication. Therefore we use many techniques that reduce calculation and communication considering coulomb force is small at long distance.



Figure1 Example of Protein (hide water)

Parallel Performance

We used spatial decomposition by rectangular for parallelization. One CPU calculate atoms in one rectangular box. When region that one CPU has is small, the communication time is larger than that of calculation. Total simulation time is decrease proportional to reciprocal of the number of CPU (N_{CPU}) when N_{CPU} is small, but this decrease is limited at large N_{CPU} . These decrease of simulation time against to the number of CPU for simulation with fixed size is "**strong scaling**". We obtained the result that the limit was 100 atoms per CPU (Table 2 and Figure 2). Full K computer will perform simulation with over 10 million atoms.

The performance increase when the size of simulation per CPU is fixed is "weak scaling". In the case of the number of atom per CPU is 6500, the performance was completely scaled (Table 3).

Table 3 Weak Scaling

| Time consumption for 6,542 atom/GPU | | | | | | | |
|-------------------------------------|---------------|---------|-----------|------------|-------------|-------------|-------------|
| Number of Node | | 64 | 512 | 4,096 | 32,768 | 79,872 | 82,944 |
| Number of Atom | | 418,707 | 3,349,656 | 26,797,248 | 214,377,984 | 522,546,336 | 542,644,272 |
| Time (ms / step) | Total | 109.058 | 110.535 | 111.186 | 111.672 | 112.414 | 112.085 |
| | Force | 91.622 | 91.528 | 91.601 | 92.329 | 91.262 | 91.641 |
| | Communication | 12.124 | 13.694 | 14.011 | 14.012 | 15.820 | 15.067 |
| | Other | 5.312 | 5.313 | 5.574 | 5.332 | 5.332 | 5.376 |
| Performance (PFLOPS) | | 0.003 | 0.025 | 0.201 | 1.599 | 3.871 | 4.031 |
| Efficiency (%) | | 39.0 | 38.5 | 38.3 | 38.1 | 37.9 | 38.0 |

Molecular Dynamics Simulation on the K computer

Using the K computer, we achieved 4.39 PFLOPS, 42.9% efficiency with 522 million atoms (Table 1).

Table1 Performance of Massive MD Simulations

| Number of CPU | 79,872 | 82,944 | |
|------------------------|---------------|------------------|--|
| Number of atom | 522 million | 543 million | |
| Number of atom per CPU | 6542 | 6542 | |
| Time (for 1000 step) | 116.4 sec | 112.1 sec | |
| Performance | 4.39 PFLOPS | 4.03 PFLOPS | |
| Efficiency | 42.9 % | 38.0 % | |
| | Reduce CPI to | Topology was not | |

fit simulation hox

Table 2. Strong Scaling time consumptions for 3,349,656 atoms

| Numb | er of CPU | 64 | 512 | 4,096 | 32,768 |
|--------------------------|--------------------|--------|--------|-------|--------|
| Numb | er of atom per CPU | 52,338 | 6,542 | 817 | 102 |
| Amount of communication* | | 2,232 | 936 | 504 | 342 |
| Target of communication | | 26 | 26 | 124 | 342 |
| Time (ms / step) | Total | 823.07 | 110.54 | 25.90 | 10.77 |
| | Force | 731.66 | 91.53 | 12.50 | 2.41 |
| | Communication | 59.28 | 13.69 | 11.48 | 7.38 |
| | Other | 32.13 | 5.31 | 1.92 | 0.98 |

* volume of data swap area

Table 4 Force calculation at one CPU 6,532 atom

| Time consumption | 77.45 sec | | |
|------------------|---------------|--|--|
| Performance | 80.949 GFLOPS | | |
| Efficiency | 63.2 % | | |
| SIMD ratio | 57.2 % | | |

SIMD ration include all instruction

Single CPU Performance

Atoms assigned to a CPU are distributed to 8 cores. When the number of atom per CPU is larger than several hundred, the load unbalance is small and 8 cores well perform.

SIMD instruction is applied to force calculation because the number of atom affect to an atom is large even if ignore far atoms.

212

Figure 2 Strong Scaling

adjusted





