

# 1Pa012 ジスルフィド結合ネットワークタンパク質 ヒトインスリンの QCLO 法による全電子計算

(アドバンスソフト(株)/(株)日立製作所 機械研究所) ○稲葉亨  
(東京大学 生産技術研究所) 田原才静  
(株)富士総合研究所) 西川宜孝  
(東京大学 生産技術研究所) 佐藤文俊  
(東京大学 生産技術研究所/アドバンスソフト(株)) 柏木浩

## 【序】

本研究グループはタンパク質の性質を電子レベルで理解することを目標に、タンパク質のための密度汎関数法プログラム ProteinDF<sup>1)</sup>を開発し、その全電子計算の収束過程を強力にサポートする擬カノニカル局在化軌道法 (QCLO; Quasi-Canonical Localized Orbital) 法<sup>2)</sup>を提案している。これらのプログラムにより、経験がほとんどなくともタンパク質の分子軌道 (MO) 計算が達成できるようになる見込みである。

QCLO は局在化軌道 (LO) の一種であり、例えばアミノ酸残基といった任意の範囲に局在化しているが、その中ではカノニカル軌道の性質を持つ軌道である。この QCLO をつなげて、タンパク質の高品位な初期値 MO を作成することで、タンパク質の全電子計算をほぼ確実にこなせ、同時に SCF 計算の繰り返し回数を低減できる。

本手法の有効性を示すため、本研究ではジスルフィド結合が複雑に絡まりあったヘテロダイマータンパク質であるインスリン (図 1) に QCLO 法を適用し、その全電子計算を実行した。

## 【計算モデルの作成】

インスリンの立体構造データは PDB<sup>3)</sup>の pdb1HLU.ent を用いた。局所構造の歪みが大きい場合 AMBER<sup>4)</sup>を用いて、1935 個の水分子中で 20ps の分子動力学計算を行い、安定構造を得た。続いて、水分子を全て取り出し、インスリン中の荷電を持つカチオンや N 末端、アニオンと C 末端のアミノ酸をすべて中性化した。すなわち、アニオンや C 末端は溶媒側に水素を加え、カチオンや N 末端は溶媒側の水素を抜き取った構造をインスリン計算モデルとした。原子数、電子数はそれぞれ 784、3078 であった。

## 【インスリン全電子計算のシナリオ】

インスリンの全電子計算は図 2 に示す 4 段階の計算過程を経て達成された。ここで、黒の楕円は計算したペプチド鎖を模式的に示している。また、インスリンはα鎖、β鎖がそれぞれ 21 残基、30 残基のペプチド鎖で構成されているが、便宜上、β鎖の残基番号はα鎖からの通し番号とした。

まず、STEP1 では、アミノ酸 1 残基ごとの全電子計算を行った。STEP2 では、簡便に STEP1 により得られたアミノ酸残基の近似電子密度を合成して初期値とし、残基番号 1~3、2~4、3~5

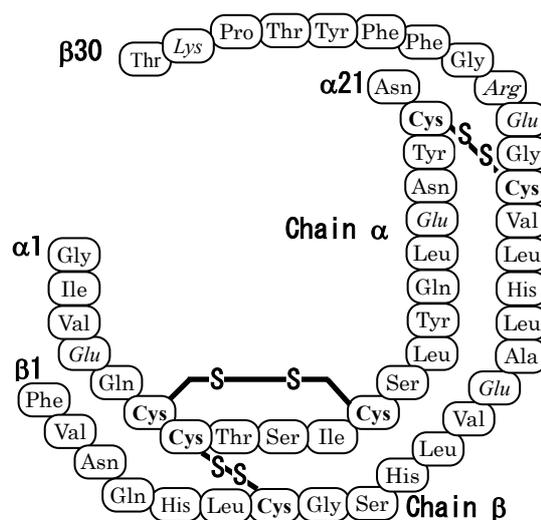


図 1 インスリンのアミノ酸配列図

Chain α																														
Num.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21									
Residue	Gly	Ile	Val	Glu	Gln	Cys	Cys	Thr	Ser	Ile	Cys	Ser	Leu	Tyr	Gln	Leu	Glu	Asn	Tyr	Cys	Asn									
STEP1	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●									
STEP2	●		●		●		●		●		●		●		●		●		●		●									
STEP3	●			●			●			●			●			●			●			●								
STEP4	●					●					●					●					●					●				

Chain β																																			
Num.	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51					
Residue	Phe	Val	Asn	Gln	His	Leu	Cys	Gly	Ser	His	Leu	Val	Glu	Ala	Leu	His	Leu	Val	Cys	Gly	Glu	Arg	Gly	Phe	Phe	Tyr	Thr	Pro	Lys	Thr					
STEP1	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●				
STEP2	●		●		●		●		●		●		●		●		●		●		●		●		●		●		●		●				
STEP3	●			●			●			●			●			●			●			●			●			●			●				
STEP4	●					●					●					●					●					●					●				

図2 インスリンの全電子計算の計算シナリオ

といった重なりを持つ3残基の全電子計算を行った。STEP3以降QCLOを用いた。STEP3残基番号1~9の初期値は、STEP2により得られた1~3の1~2の2残基と、2~4、3~5、4~6、5~7、6~8の真中の1残基、7~9残基の8~9の2残基のQCLOを取り出して合成し、初期値を作成した。同様にSTEP4では、STEP3の1~9、8~15、14~21、22~30、29~37、36~44、43~51のペプチド鎖のQCLOを合成して初期値を作成し、51残基の計算を行った。

このシナリオのポイントは、複雑なトポロジーを生むジスルフィド結合に対応するために、これを構成するCysを持つペプチド鎖の構造は相方のCysもつなげた状態で切り出すことである(例: STEP2のTyr19-Cys20(-Cys40)-Asn21、Val39-(Cys20)-Cys40-Gly41)。これにより、中間STEPのペプチド鎖は全ての共有結合部で必ず互いに1残基の重なりを持つ。

### 【計算結果】

上記シナリオの全電子計算にはMPIにより並列化された密度汎関数法プログラムProteinDFを用いた。このプログラムは、MOだけでなく電子密度と交換相関ポテンシャルもガウス型(補助)基底関数で展開している。本計算ではValence Double型の基底関数セットを使用しており、軌道数、補助基底関数の数はそれぞれ4443、8074となった。最終的なインスリン全体の全電子計算は35回のSCF計算を要し、SCF計算1回にかかった実時間は、900MHzのItanium2プロセッサ32台を用いた場合、平均約1時間であった。

タンパク質はサイズが大きい上に、半導体的な電子状態を持ち占有軌道と非占有軌道との間のエネルギーギャップはわずかしかない。加えて、インスリンは複雑なトポロジー構造を持っている。このような計算系において、本研究ではSTEP3における各ペプチド鎖の長さを指定しただけで、その全電子計算に成功した。現在、タンパク質の表面に関するよりリアルな取り扱いについて検討している。本研究は、文部科学省ITプログラム「戦略的基盤ソフトウェアの開発」において実施された。

### 参考文献:

- 1)Sato, F. et al.: Calculation of all-electron wavefunction of hemoprotein cytochrome c by density functional theory, Chem. Phys., Lett. 341, pp645-651, 2001
- 2)Kashiwagi, H. et al.: Convergence process with quasi-canonical localized orbital in all-electron SCF calculation on proteins, Molecular Physics, 101, pp81-86, 2003
- 3)Protein Data Bank: <http://www.rcsb.org/pdb/>
- 4)AMBER7: <http://www.amber.ucsf.edu/amber/amber.html>