

蛋白質原子座標の情報処理プログラム・システム : PSPCS

武富, 敬
九州大学大型計算機センター研究開発部

水野, 裕重
九州大学理学部物理学教室

郷, 信広
九州大学理学部物理学教室

<https://doi.org/10.15017/1474881>

出版情報 : 九州大学大型計算機センター広報. 13 (2), pp.171-188, 1980-06-10. 九州大学大型計算機センター
バージョン :
権利関係 :



蛋白質原子座標の情報処理プログラム・システム — PSPCS —

武富 敬^{*}, 水野 裕重^{**}, 郷 信広^{**}

1. はじめに

分子生物学のもたらした主要な成果の一つとして、遺伝物質核酸の塩基配列として担われている遺伝情報が、蛋白質の一次元アミノ酸配列 — 蛋白質の一次構造 — を決定しているということの発見があげられる。この事は、取りも直さず、生物の持つ多様な機能は、蛋白質あるいはその集合体の機能に還元されることを意味する。1960年代の初めに、初めて蛋白質のX線結晶解析に成功し、この分子の三次元立体構造 — 蛋白質の三次構造 — が原子レベルで解明されて以来、X線解析により原子座標データの得られている蛋白質分子の数は、着実に増大しつつある。この間、このX線解析の成果により明らかにされた事のうちに最大のものの一つは、蛋白質の立体構造の中に、蛋白質の持つ機能に関する情報が含まれているということの発見である。今日では、いかなる蛋白質の物理化学的研究にも、立体構造のイメージが源を出し、いわば欠かせないベースになったと言ってもよい。

ところで、一般の蛋白質研究者にとって、原子座標データを加工して自分の研究に必要な情報を取り出すプログラムの作成を考えたとしても、その事自体それほど容易なことではない。そこで、それを支援するツールとしての役割を果たす、簡単に利用できるプログラム群を用意しておくことは意義のあることである。また、もし、そのようなプログラム群が組織的に体系づけられて開発されたとすれば、そこから産み出される情報は、さらに次の段階の基礎データとして使えるような、蓄積可能な学術情報としての価値を持つ性質のものである。

以上のような観点から、1976年度から3年間、筆者の一人、郷を中心として、科研費特定研究「情報システムの形成過程と学術情報の組織化」の中に一つのグループを作成し、蛋白質原子座標データを処理するプログラムの開発に取り組んできた。このグループの作成するプログラム・システムを「Program System for Protein Conformation Study」と呼び、PSPCSと略称する[1]。また、1979年度は、科研費の生物物理学分科における試験研究「蛋白質原子座標の情報処理プログラムのシステム化」において、それまでに蓄積されたプログラム群の見直し、書き替えを行い、プログラム・システムとして整備することを試みた。以下は、このPSPCS作成グループの活動とPSPCSの利用法の紹介である。

2.では、蛋白質原子座標データの入手に関し、Protein Data Bankの活動について述べる。

3.では、PSPCS作成グループの活動について、まずPSPCSへの入力データフォーマットの整備・統一について述べ、次に開発されたプログラム群の内容について、最後にPSPCS作成グループ以外で開発された有用なプログラムの導入について述べる。4.では、PSPCSの利用法について述べる。5.では、筆者の一人、水野によって開発された、蛋白質原子座標データの簡便な検索システムについて紹介する。

2. 蛋白質原子座標データ —Protein Data Bank—

蛋白質分子のX線結晶解析が成功し、その成果が論文誌に発表されても、原子座標データは、かなり膨大な量

* 九州大学大型計算機センター研究開発部

** 九州大学理学部物理学教室

にのぼるので掲載されない。代わりに、現在では、これらの原子座標データの預託・仲介機関として、Protein Data Bank (以下、PDBと略称する) が設立され、アメリカのBNL (Brookhaven National Laboratory) において、データの収集・標準化・領布などの管理業務が執り行われている[2]。日本では、大阪大学蛋白質研究所付属結晶解析研究センターが、BNLと連携して我国の利用者に対する窓口業務を行っている[3]。

```

HEADER      PROTEINASE INHIBITOR (TRYPSIN)                01-NOV-76   3PTI   3PTI   3
COMPND      TRYPSIN INHIBITOR                             3PTI   4
SOURCE      BOVINE (BOS TAURUS) PANCREAS                 3PTI   5
AUTHOR      R.HUBER,D.KUKLA,A.RUEHLHANN,G.EPP,H.FORMANEK, 3PTI   1
            2 J.DEISENHOFER,W.STEIGEMANN                 3PTI   2
REMARK      1 REFERENCE 1                                3PTI   3
REMARK      1 AUTH   J.DEISENHOFER,W.STEIGEMANN           3PTI   4
REMARK      1 TITL   CRYSTALLOGRAPHIC REFINEMENT OF THE STRUCTURE OF 3PTI   6
REMARK      1 TITL 2 BOVINE PANCREATIC TRYPSIN INHIBITOR AT 3PTI   7
REMARK      1 TITL 3 1.5 ANGSTROMS RESOLUTION             3PTI   8
REMARK      1 REF    ACTA CRYSTALLOGR.,SECT.B            U. 31   238 1975 3PTI   9
REMARK      1 REFM   ASTM ACBCAR DK ISSN 0001-5520       107 3PTI  10

REMARK      2 RESOLUTION. 1.5 ANGSTROMS                   3PTI  32
REMARK      3 REFINEMENT. BY REAL-SPACE PROCEDURE OF R. DIAMOND WITH 3PTI  33
REMARK      3 INTERMEDIATE RECALCULATION OF ELECTRON-DENSITY MAP AND 3PTI  34
REMARK      3 CORRECTIONS BY HAND. SEE REFERENCE 1 ABOVE. 3PTI  35
                                                    3PTI  36

SEQRES      1 59 ACE ARG PRO ASP PHE CYS LEU GLU PRO PRO TYR THR GLY 3PTI  41
SEQRES      2 59 PRO CYS LYS ALA ARG ILE ILE ARG TYR PHE TYR ASN ALA 3PTI  42
SEQRES      3 59 LYS ALA GLY LEU CYS GLN THR PHE VAL TYR GLY GLY CYS 3PTI  43
SEQRES      4 59 ARG ALA LYS ARG ASN ASN PHE LYS SER ALA GLU ASP CYS 3PTI  44
SEQRES      5 59 MET ARG THR CYS GLY GLY ALA             3PTI  45
FTNOTE      1 THESE ATOMS WERE NOT FOUND IN THE ELECTRON DENSITY MAP. 3PTI  47
FTNOTE      1 THEIR COORDINATES WERE GENERATED USING STEREOCHEMICAL 3PTI  48
FTNOTE      1 CRITERIA.                                     3PTI  49
FORMUL      2 HOH *47(H2 O)1                               3PTI  61
HELIX       1 H1 SER 47 GLY 56 1                           3PTI  50
SHEET       1 S1 2 ALA 16 ALA 25 0                         3PTI  51
SHEET       2 S1 2 GLY 28 GLY 36 -1                        3PTI  52
SSBOND      1 CYS 5 CYS 55                                  3PTI  62
SSBOND      2 CYS 14 CYS 38                                 3PTI  63
SSBOND      3 CYS 30 CYS 51                                 3PTI  64
CRYST1      43.100 22.900 48.600 90.00 90.00 90.00 P 21 21 21 4 3PTI  53
ORIGX1      1.000000 0.000000 0.000000 0.000000          3PTI  54
ORIGX2      0.000000 1.000000 0.000000 0.000000          3PTI  55
ORIGX3      0.000000 0.000000 1.000000 0.000000          3PTI  56
SCALE1      .023202 0.000000 0.000000 0.000000          3PTI  57
SCALE2      0.000000 .043668 0.000000 0.000000          3PTI  58
SCALE3      0.000000 0.000000 .020576 0.000000          3PTI  59
ATOM        1 C ACE 0 27.468 28.283 -2.381 1.00 0.00 1 3PTI  60
ATOM        2 O ACE 0 27.655 28.721 -1.239 1.00 0.00 1 3PTI  61
ATOM        3 CH3 ACE 0 28.349 28.731 -3.546 1.00 0.00 1 3PTI  62
ATOM        4 N ARG 1 26.522 27.417 -2.687 1.00 0.00 3PTI  63
ATOM        5 CA ARG 1 25.603 26.909 -1.661 1.00 0.00 3PTI  64
ATOM        6 C ARG 1 26.256 26.217 -.466 1.00 0.00 3PTI  65
ATOM        7 O ARG 1 27.268 25.519 -.612 1.00 0.00 3PTI  66
ATOM        8 CB ARG 1 24.610 25.885 -2.248 1.00 0.00 3PTI  67
ATOM        9 CG ARG 1 25.171 24.458 -2.379 1.00 0.00 3PTI  68
ATOM        10 CD ARG 1 24.124 23.541 -2.979 1.00 0.00 3PTI  69
ATOM        11 NE ARG 1 23.644 24.058 -4.269 1.00 0.00 3PTI  70
ATOM        12 CZ ARG 1 24.299 23.863 -5.432 1.00 0.00 3PTI  71
ATOM        13 NH1 ARG 1 23.797 24.379 -6.539 1.00 0.00 3PTI  72
ATOM        14 NH2 ARG 1 25.434 23.152 -5.473 1.00 0.00 3PTI  73
ATOM        15 N PRO 2 25.669 26.420 .698 1.00 0.00 3PTI  74
ATOM        16 CA PRO 2 26.191 25.807 1.926 1.00 0.00 3PTI  75
ATOM        17 C PRO 2 26.248 24.283 1.838 1.00 0.00 3PTI  76
ATOM        18 O PRO 2 25.462 23.661 1.112 1.00 0.00 3PTI  77
ATOM        19 CB PRO 2 25.364 26.408 3.014 1.00 0.00 3PTI  78
ATOM        20 CG PRO 2 24.531 27.554 2.480 1.00 0.00 3PTI  79
ATOM        21 CD PRO 2 24.639 27.414 .964 1.00 0.00 3PTI  80

```

図1. Protein Data Bankのトリプシンインヒビターのデータレコードの抜粋

表1. PDBのデータフォーマットの説明

レコード識別子	レコードの内容
HEADER	この蛋白質の機能, PDBへの登録の日付, データコード名 (3文字で蛋白質を表わし, その前に付けた1桁の数字で, 精密化などの異なるデータであることを示す), データコード名は, 本稿5.でコマンドのオペランドとして使用されるので重要である.
COMPND	蛋白質名
SOURCE	この蛋白質を採取した物質名
AUTHOR	この蛋白質を解析した研究者名
REMARK	参照論文, X線分解能など
SEQRES	この蛋白質のアミノ酸残基数と, アミノ酸配列順序
FTNOTE	特別の残基や原子に関する脚注
HELIX	著者の報告した二次構造 α -ヘリックスの部分
SHEET	著者の報告した二次構造 β -シートの部分
SSBOND	この蛋白質のS-S結合の部位
ATOM	原子の通し番号 (原子一連番号), 原子名, 属するアミノ酸残基の略名とその番号, x, y, zの各座標値, 占有率, 温度因子

九州大学関係では, 筆者の一人, 郷のもとに, 阪大からPDBデータの入った2400 ftの磁気テープ1本が送られてきて, 九大大型計算機センターにクローズド磁気テープとして保管されているので, 一般の利用者も使える形になっている. BNLから発行されるPDBニュースレターの最新号(1980年1月)によると, 現在, t-RNAも含めて126セットの原子座標データが集録されている.

図1に, PDBのデータレコードの抜粋を示す. ある一定のフォーマット(BNLフォーマットと呼ぶ)に従って書かれた, 1レコード80バイトのカード・イメージのデータである. 各レコードは, 最初の6欄のレコード識別子により識別される. 表1に, 主なレコード識別子と, その識別子により区別されるレコードの内容との対応表を掲げたので, 図1の内容は容易にご理解いただけるものと思う.

3. PSPCS作成グループの活動

3.1. 入力データフォーマット — PSPCSフォーマット —

BNLフォーマットによるPDBデータは, よく考案されてはいるものの, いくつかの点で不満足な面も見受けられる. 例えば, アミノ酸残基名や原子名などは, BNLフォーマットでは英字コードだけだが, これらに数値コードを与えておけば, 応用プログラムを作成する側としては, 英字と番号の両方で検索できたり, 特に番号については論理IF文や計算型GOTO文が活用できたり, そのまま配列の添字として使用できたり, また, 番号の組み合わせで短時間に検索できるなど, 応用プログラムの構造を簡単にする上で, そのメリットは大きい.

そこで, 応用プログラムを意識した立場のフォーマットとして, BNLフォーマットの1レコード80バイトの

後ろ60バイトに、上述の数値コード、それに共有結合のデータを付加した、1レコード140バイトの新しいフォーマット（PSPCSフォーマットと呼ぶ）を考案し、これをPSPCSの入力データフォーマットとすることにした。後者の共有結合データは、アミノ酸配列と原子座標の一次データから計算によって出すことのできる二次データではあるが、応用プログラム中にしばしば現われるものであり、プログラムの構造を簡単にする上で、その役割は大きいので、あたかも一次データのごとく付加してある。

```

HEADER  PROTEINASE INHIBITOR (TRYPSIN)          01-NOV-76   3PTI   3 11
COMPND  TRYPSIN INHIBITOR                      3PTI   4 12
SOURCE  BOVINE (BOS TAURUS) PANCREAS           3PTI   5 13
AUTHOR  R. HUBER, D. KUKLA, A. RUEHLMANN, D. EPP, H. FORMANEK,
AUTHOR  2 J. DEISENHOFER, W. STEIGEMANN          3PTIB  1 14
REMARK  1 REFERENCE 1                          3PTIB  2 14
REMARK  1 AUTH J. DEISENHOFER, W. STEIGEMANN     3PTIB  3 14
REMARK  1 TITL CRYSTALLOGRAPHIC REFINEMENT OF THE STRUCTURE OF
REMARK  1 TITL 2 BOVINE PANCREATIC TRYPSIN INHIBITOR AT 3PTIB  4 16
REMARK  1 TITL 3 1.5 ANGSTROMS RESOLUTION        3PTIB  5 16
REMARK  1 REF ACTA CRYSTALLOGR. SECT. B        3PTIB  6 16
REMARK  1 REFN ASTM ACBCAR DK ISSN 0001-5520    3PTIB  7 16
REMARK  107                                     3PTIB  8 16

REMARK  2 RESOLUTION. 1.5 ANGSTROMS             3PTI  32 16
REMARK  3 REFINEMENT. BY REAL-SPACE PROCEDURE OF R, DIAMOND WITH 3PTI  33 16
REMARK  3 INTERMEDIATE RECALCULATION OF ELECTRON-DENSITY MAP AND 3PTI  34 16
REMARK  3 CORRECTIONS BY HAND. SEE REFERENCE 1 ABOVE. 3PTI  35 16

SEGRES  1 59 ACE ARG PRO ASP PHE CYS LEU GLU PRO PRO TYR THR GLY 3PTI  41 21
SEGRES  2 59 PRO CYS LYS ALA ARG ILE ILE ARG TYR PHE TYR ASN ALA 3PTI  42 21
SEGRES  3 59 LYS ALA GLY LEU CYS GLN THR PHE VAL TYR GLY GLY CYS 3PTI  43 21
SEGRES  4 59 ARG ALA LYS ARG ASN ASN PHE LYS SER ALA GLU ASP CYS 3PTI  44 21
SEGRES  5 59 MET ARG THR CYS GLY GLY ALA        3PTI  45 21
FTNOTE  1 THESE ATOMS WERE NOT FOUND IN THE ELECTRON DENSITY MAP. 3PTI  46 22
FTNOTE  1 THEIR COORDINATES WERE GENERATED USING STEREOCHEMICAL 3PTI  47 22
FTNOTE  1 CRITERIA.          3PTI  48 22
FORMUL  2 HOH *47(H2 O1)      3PTI  49 22
HELIX   1 H1 SER 47 GLY 56 1 3PTIB 61 24
SHEET   1 S1 2 ALA 16 ALA 25 0 3PTI  50 31
SHEET   2 S1 2 GLY 28 GLY 36 -1 3PTI  51 32
SSBOND  1 CYS 5 CYS 55        3PTI  52 32
SSBOND  2 CYS 14 CYS 38       3PTIB 62 34
SSBOND  3 CYS 30 CYS 51       3PTIB 63 34
CRYST1  43.100 22.900 48.600 90.00 90.00 90.00 F 21 21 21 4 3PTIB 64 34
ORIGX1  1.000000 0.000000 0.000000 0.000000 3PTI  53 51
ORIGX2  0.000000 1.000000 0.000000 0.000000 3PTI  54 52
ORIGX3  0.000000 0.000000 1.000000 0.000000 3PTI  55 52
SCALE1  .023202 0.000000 0.000000 0.000000 3PTI  56 52
SCALE2  0.000000 .043668 0.000000 0.000000 3PTI  57 53
SCALE3  0.000000 0.000000 .020576 0.000000 3PTI  58 53
ATOM    1 C ACE 0 27.468 28.283 -2.381 1.00 0.00 1 3PTI  60 61 1 0 022 11
ATOM    2 O ACE 0 27.455 28.721 -1.239 1.00 0.00 1 3PTI  61 61 2 0 022 23
ATOM    3 CH3 ACE 0 28.349 28.731 -3.546 1.00 0.00 1 3PTI  62 61 3 0 022 31
ATOM    4 N ARG 1 26.522 27.417 -2.687 1.00 0.00 3PTI  63 61 4 0 115 12
ATOM    5 CA ARG 1 25.603 26.909 -1.661 1.00 0.00 3PTI  64 61 5 0 115 21
ATOM    6 C ARG 1 26.256 26.217 -.466 1.00 0.00 3PTI  65 61 6 0 115 31
ATOM    7 O ARG 1 27.268 25.519 -.612 1.00 0.00 3PTI  66 61 7 0 115 43
ATOM    8 CB ARG 1 24.610 25.885 -2.248 1.00 0.00 3PTI  67 61 8 0 115 51
ATOM    9 CG ARG 1 25.171 24.458 -2.379 1.00 0.00 3PTI  68 61 9 0 115 61
ATOM   10 CD ARG 1 24.124 23.541 -2.979 1.00 0.00 3PTI  69 61 10 0 115 71
ATOM   11 NE ARG 1 23.644 24.058 -4.269 1.00 0.00 3PTI  70 61 11 0 115 82
ATOM   12 CZ ARG 1 24.299 23.863 -5.432 1.00 0.00 3PTI  71 61 12 0 115 91
ATOM   13 NH1 ARG 1 23.797 24.379 -6.539 1.00 0.00 3PTI  72 61 13 0 115 102
ATOM   14 NH2 ARG 1 25.434 23.152 -5.473 1.00 0.00 3PTI  73 61 14 0 115 112
ATOM   15 N PRO 2 25.669 26.420 .698 1.00 0.00 3PTI  74 61 15 0 220 12
ATOM   16 CA PRO 2 26.191 25.807 1.926 1.00 0.00 3PTI  75 61 16 0 220 21
ATOM   17 C PRO 2 26.248 24.283 1.838 1.00 0.00 3PTI  76 61 17 0 220 31
ATOM   18 O PRO 2 25.462 23.661 1.112 1.00 0.00 3PTI  77 61 18 0 220 43
ATOM   19 CB PRO 2 25.364 26.408 3.014 1.00 0.00 3PTI  78 61 19 0 220 51
ATOM   20 CG PRO 2 24.531 27.554 2.480 1.00 0.00 3PTI  79 61 20 0 220 61
ATOM   21 CD PRO 2 24.639 27.414 .964 1.00 0.00 3PTI  80 61 21 0 220 71

```

図2. PSPCSフォーマットによるトリプシンインヒビターのデータレコードの抜粋。

現在は、まだ共有結合のデータを付加する機能が欠けている。

図2に、PSPCSフォーマットによるデータレコードの抜粋を示す。BNLフォーマットからPSPCSフォーマットへのデータ変換プログラムは、後述するように、PSPCS中に一つの応用プログラムとして含まれている。要約すると、PSPCSフォーマットの利点としては、

- 1) 各種の数値コードを付加することにより、検索が容易になり、また、それによって応用プログラムの構造が簡単化できて、開発の手間を軽減することができる。
- 2) BNLフォーマットの1レコードを長くしただけで、それとの一対一対応は保たれているので、BNLフォーマットで書かれたプログラムも何の変更もなしに通することができる。

がある。

入力データフォーマットの些少な変更に対しても、応用プログラムを書き替えなければならないことの苦痛は大きい。この事を無しで済ます、すなわちデータ独立性を保つ最良の方法は、Data Base Management System (DBMS と略称) によりデータベースを構築することである[4]。また、そうすれば、前述のPDBデータのように、シーケンシャルなアクセスだけに依らずにすむので、検索時間も短縮できる。これに関しては、阪大大型計算機センター、磯本征雄氏らによる、DBMS, INQを用いた「蛋白質データベースPROTEIN-DB」の研究開発がある[3, 5]。

3.2. PSPCSのプログラム群

表2に、PSPCS作成グループにより開発されたプログラムの主なものと、その概要を掲げる[1, 6, 7, 8]。

表2. PSPCS中の主なプログラムとその概要

プログラム名が明示してあるものが、現在使用できるものである。長野氏の、アミノ酸配列から二次構造を予測する、著名な二次構造予測プログラムは、原子座標データをなまに使う訳ではないが、PSPCSに含まれている。

分 類	プログラムの内容	作成者	プログラム名
フォーマット変換	BNL座標データのフォーマット変換	九大・理 水野裕重 宮沢三造	CONVRT
分子構造表示	蛋白質分子鎖立体構造のステレオ投影図プロット	九大・理 水野裕重 郷信広 花田光弘	STEREO
	蛋白質分子の透視図	群馬大・教養 中田吉郎	
距離マップ	XYプロットによる距離マップの図示	九大・理 水野裕重	DISMAP
	グラフィックディスプレイによる距離マップの図示	京大・化研 大島玄久	
	PRINTRONIXによる距離マップの印刷	早大・理工 輪湖 博	
(Φ , Ψ)	Φ , Ψ 値のRamachandran プロット	京大・化研 大島玄久 西川健 大井竜夫	
エネルギーマップ	Φ , Ψ 面の等高線のプロット	早大・理工 輪湖 博	
構造の探索・予測	近傍原子の探索	九大・計セ 武富 敬	NASH
	α -ヘリックスの探索	東大・薬 長野晃三	
	蛋白質における二次構造の同定	九大・理 宮沢三造 郷 通子	
	二次構造予測法	東大・薬 長野晃三	
種々の量の計算	BENDER模型用プログラム	九大・理 栄 智英 郷 信広	BENDER
	主鎖のボンド長、ボンド角および二面角の計算	九大・理 郷 通子	
	Accessible Surface Areaの計算	九大・理 水野裕重 郷 信広	ASA
	Voronoi Polyhedronの計算	九大・理 宮沢三造 郷 通子	
	二つの立体構造の重ね合わせ	京大・化研 西川 健 大井竜夫	SUPPOS
	蛋白質分子の水素原子座標の計算	群馬大・教養 中田吉郎	
	蛋白質分子の変形とエネルギー地図の作成	群馬大・教養 中田吉郎	
	Lab Quip モデルの原子座標の修正	東大・薬 宮本秀一 長野晃三	
	multiple moments の計算および電荷分布図のプロット	東大・理 中村春木 和田昭允	
	蛋白質内部の局所的な電場、電位の計算	東大・理 中村春木 和田昭允	

現在、PSPCSのプログラムは、まだ完全に整備されている訳ではなく、表2でプログラム名が明示してあるものだけが使用できる。PSPCSでは、原則としてソースプログラムを公開する特にサブルーチン形式のプログラムでは、蛋白質の立体構造の複雑さに対応できるほどの機能を持って書かれていることは少ないと思われるので、各利用者で自分の研究に必要なように改編して、適宜、自分のプログラムに組み込んで使用するのが望ましいと思われる。

3.3. PSPCS以外の有用なプログラムの導入

PSPCS作成グループ以外にも、蛋白質原子座標データを処理するプログラムを開発している研究者は少ないと思われる。有用なプログラムについては、どしどし導入して、PSPCSを一つのかかなり完備したソフトウェアパッケージにしたいと考えているので、筆者まで連絡をとられることを希望する。その候補の一つは、名大・理学部・物理学教室の別府良孝氏による、分子構造表示プログラム「NAMOD」である[9, 10]。

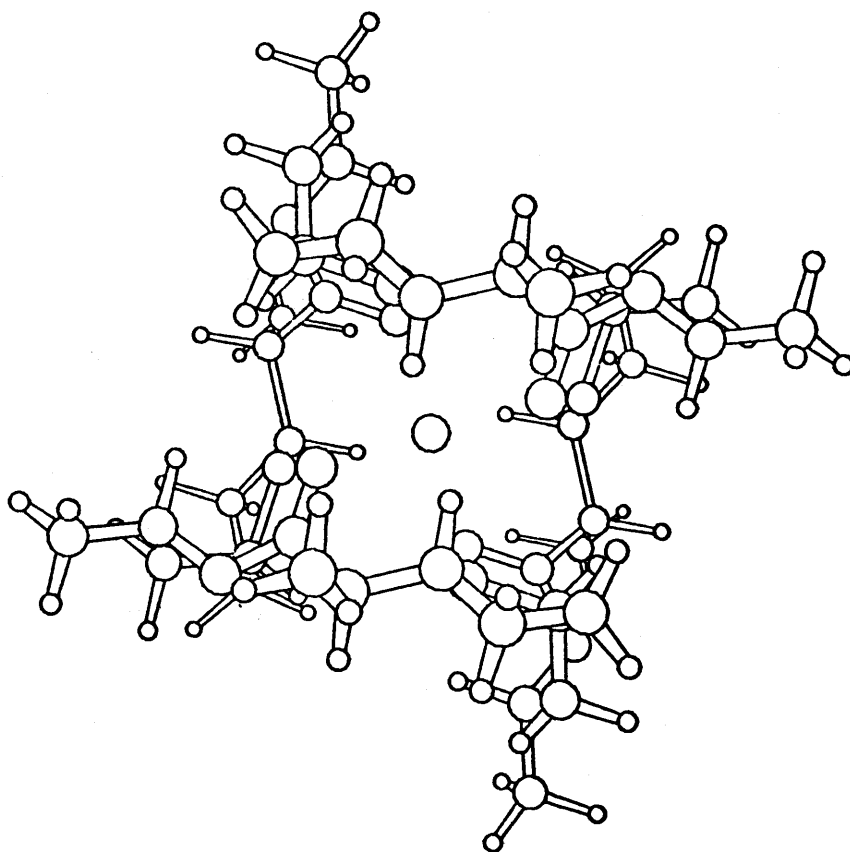


図3. 分子構造表示プログラムNAMODを用いて描いたノナクチンの遠近図、名大 別府氏の御好意により掲載。

NAMODは、図3.に示すように、球と棒模型で分子構造をグラフィック・ディスプレイ上に表示するもので、隠線消去を施しつつ、手前の原子や結合ほど大きく太く描くので、遠近感のある図の表示に適している。また、

α -炭素原子を結ぶ蛋白質の骨格構造をリボン状に描く, A. D. Mac Lachlan の RIBBON プログラムも考慮したいものの一つである.

4. PSPCSの利用法

PSPCSのプログラムは, 種々の応用プログラムの中に頻繁に現われる機能をプログラム化したサブルーチン形式のものと, ある一まとまりの機能を持つコンプリートプログラム形式のものとから成る. 3.2. で述べたように, すべてのプログラムが整備されている訳ではないが, 現在, 使えるものに対しては, 文献[11]にまとめであり, これが利用者用のマニュアルとなる. このマニュアルを希望される方は, 筆者(郷, 内線4149) まで申し込んでいただきたい.

4.1. サブルーチン形式プログラムの利用例

種々の応用プログラム中で, ある特定の原子の近傍に存在する原子を探索しなければならない場合によく出会う. 原子数の増大に伴い, 探索の手間は急激に増加するので, その高速化のためにメッシュ法を採用したプログラムを作成した.

```
// EXEC FORTXCG
//FORT.SYSIN DD *
C** MAIN PROGRAM
    DIMENSION NATH(2000),NORD(2000),NARD(2000)
    DIMENSION CORD(3,2000),INW(3,2000)
    DIMENSION IZ(200),DIST(200)
    READ(5,100) IN, ICUT, IJK, KKK
    READ(5,200) M, WDM, R
    READ(5,300) NAS, NAE
    CALL NASIN(IN,IJK,M,NATH,NARD,NORD,CORD,N)
    CALL MESH(WDM,N,M,CORD,INW)
    DO 20 IC = NAS, NAE
        CALL NASUB(IC,N,M,NORD,ISA,NRA)
        IEA = ISA + NRA - 1
        DO 10 II = ISA, IEA
            CALL NASH(II,KKK,WDM,R,N,M,CORD,INW,MI,IZ,DIST)
            CALL NASOUT(II,ICUT,M,NATH,NARD,NORD,MI,IZ,DIST)
        10 CONTINUE
    20 CONTINUE
    100 FORMAT(4I2)
    200 FORMAT(I5,2F5.1)
    300 FORMAT(2I5)
    STOP
    END
//GO.SYSIN DD *
1 6 0 1
1000 2.5 5.0
25 30
//GO.FT01F001 DD DSN=F9999.PCSDT(PCS3PTI),DISP=SHR
//
```

図4. プログラムNASH (5つのサブルーチンプログラムから成る) を用いた, 残基番号による近傍原子の探索の例. 整合配列を用いて作成してある.

図4.に、ある蛋白質分子の、残基番号NASからNAEまでの原子について、近傍原子を探索する場合の例を示す。下線部分が、PSPCSに含まれるサブルーチンプログラムである。

4.2. コンプリート形式プログラムの利用例

1) カタログドプロシジャPROCS^{*}の説明

プロシジャ名	記号パラメータ	プロシジャステップ名
PROCS	PROG=プログラム名 $\left[\begin{array}{c} , \text{SYSOUT}=\left\{ \begin{array}{c} \Delta \\ K \\ R \\ S \end{array} \right\} \end{array} \right]$	GO

機能：PSPCSのコンプリート形式プログラムを実行する

記号パラメータ：

PROG：表2.のプログラム名を指定する

SYSOUT：出力クラスを指定する

2) BNL座標データのフォーマット変換

図5.に、3.1.で述べた、BNLフォーマットからPSPCSフォーマットのデータに変換する場合の例を示す。

```
//          EXEC  PROCS,PROG=CONVRT
//GO.SYSIN DD  *
//          0    1    2
//GO.FT01F001 DD DSN=F9999.PCSDT(BNL3PTI),DISP=SHR
//GO.FT02F001 DD DSN=F9999.PCS3PTI,DISP=(NEW,CATLG),UNIT=PUB,
//              SPACE=(TRK,(10,10),RLSE),
//              DCB=(RECFM=FB,LRECL=140,BLKSIZE=6300)
//
```

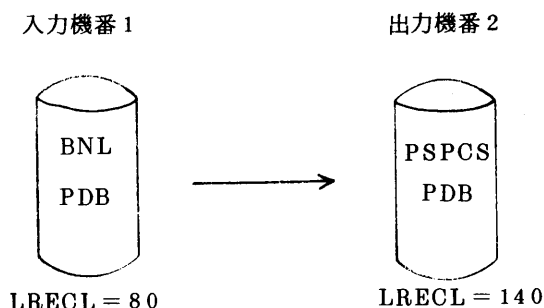


図5. プログラムCONVRTによる、BNLフォーマットからPSPCSフォーマットへのデータ変換の例。

* PROtein Conformation Studyから採った

3) XYプロットによる距離マップの図示

これは、蛋白質分子の三次元空間における折れたたみ構造を、二次元の紙の上に表示する一つの方法である。図6は、カルボキシペプチダーゼAの α -炭素原子間の距離の分布を、三角形地図として残基番号の組に対してプロットしたものである。これにより、 α -ヘリックス、 β -シートなどの二次構造や長距離の相互作用の特徴 [12, 13] などが抽出できる。

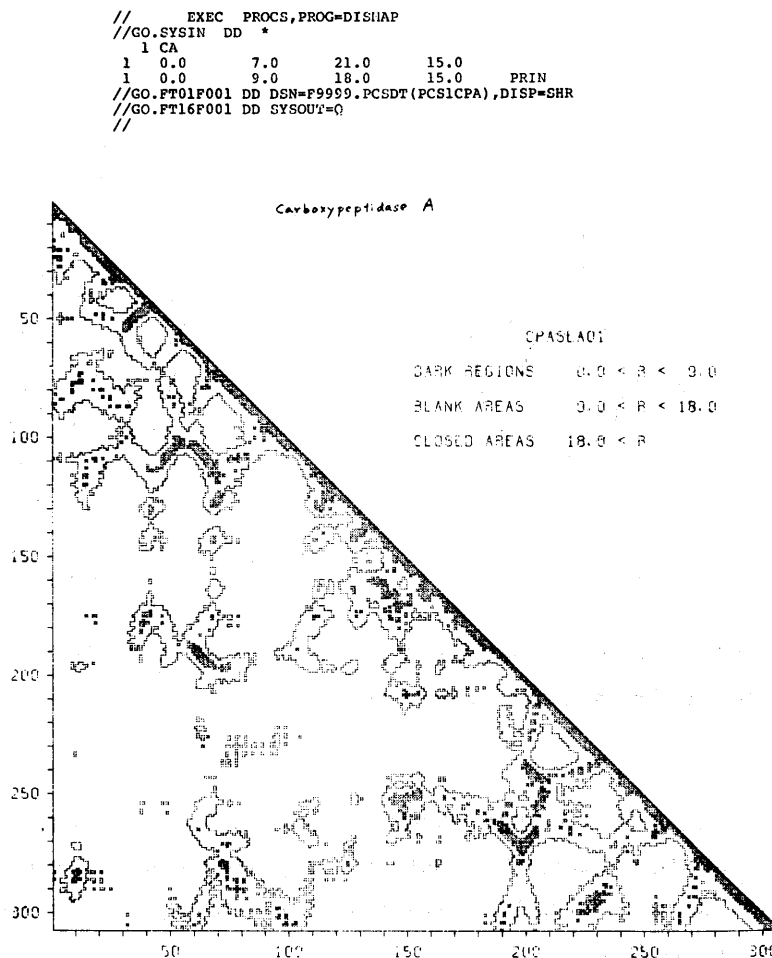


図6. プログラムDISMAPで描いたカルボキシペプチダーゼAの距離マップ。

黒い部分が、9.0 Å以内の残基の組を示している。

4) 蛋白質分子鎖立体構造のステレオ投影図プロット

蛋白質分子の立体構造を、二次元の紙の上に表示して立体的に見るために、ステレオ投影図を描く。図7は、トリプシンインヒビターの α -炭素原子を結ぶ骨格構造と側鎖の配置をステレオ投影図で描いたものである。左右の図の間に白い紙を立て、右眼で右図を、左眼で左図を見ると、少し訓練すれば立体的に見えるようになる。

```

//      EXEC  PROCS,PROG=STEREO
//GO.SYSIN  DD  *
      8.0      8.0      5
      10.0     10.0     5
//GO.FT01F001 DD  DSN=F9999.PCSDT(PCS3PTI),DISP=SHR
//GO.FT16F001 DD  SYSOUT=Q
//

```

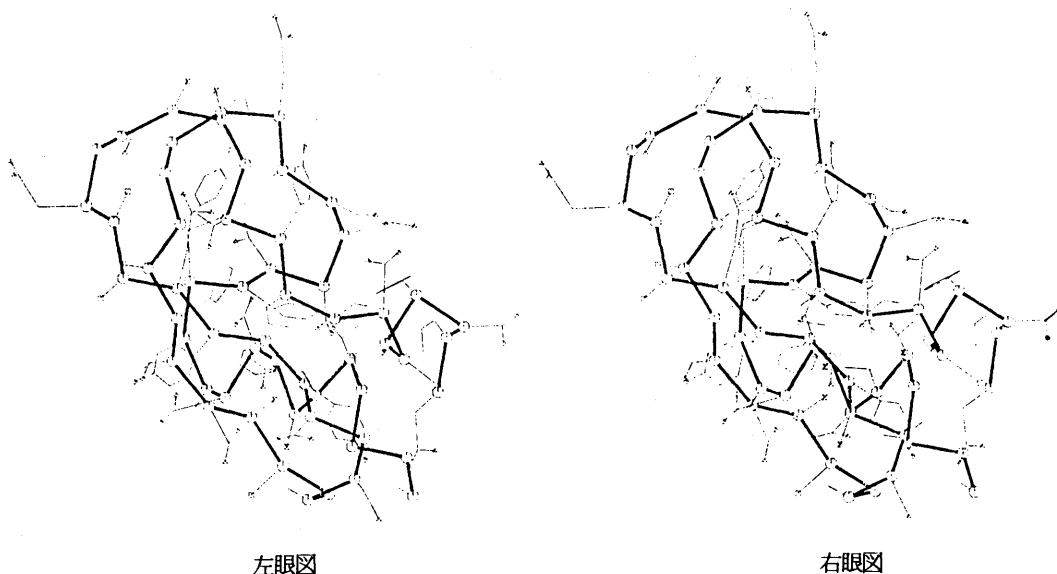


図7. プログラムSTEREOで描いたトリプシンインヒビターのステレオ投影図。
側鎖の配置まで観察できるのが特徴である。

5) その他のプログラム

その他に使用できるものとして、針金模型で蛋白質の骨格構造を製作するために必要な構造パラメータを計算するプログラムBENDER, 相同蛋白質などの類似した二つの立体構造を最適に重ね合わせて構造比較などを行うSUPPOS, 疎水相互作用の定量化などの目的で、蛋白質分子中の各原子の水に対する接触表面積を計算するプログラムASAがある。

5. 蛋白質原子座標検索システム

PSPCSのプログラム群はすべてプロティン・データ・バンク (PDB) のデータに簡単な変換を施した P SPCSデータを入力データとしているが、これを用いて実際に計算を行なう場合には、予めデータについて計算に必要な情報を知っておかねばならない。例えば、 α -ヘリックス部分の溶媒に対する表面積を計算しようとするような場合、目的とする蛋白質分子の α -ヘリックス部分の残基番号と残基名を予め知っておく必要がある。このような場合、従来までは数十ページからなるPDBデータのリストを取って調べるとか、磁気テープに格納されているPDBデータをディスクにコピーしてTSSで調べるとかの方法を取っていた。しかし、これでは作

業に費す労力が大きく、計算費もその分だけ無駄になった。そこで今回、T S Sで上述のことを手軽に調べることができるよう簡単な検索システムを開発したので以下に紹介する。

5.1. システムの概要

このシステムは、現在のところデータベースと6つのコマンドプロシジャから構成されている。データベースはリマークデータセットとデータステイタスデータセットから成る。前者は、PDBデータからATOM, HETATM, CONNECTレコード以外のリマーク部分を抜き出したもので、後者には、データコード名、分子名、PDBに登録された日付、登録者名、 α 炭素だけのデータであるとか主鎖だけのデータであるとかを表わすデータステイタス、一般利用者のためにPDBデータが格納されているセンタークロズド磁気テープのボリューム名、データセット名、及び関連するニュースレター名などの情報が収納されている。用意されている6つのコマンドでデータコード名やレコード名などを指定することによって、上述のデータベースから必要な情報を取り出す仕組みになっている。

5.2. システムの機能

このシステムは以下のような機能を持っている。なお、以下の説明では出力は特に断わらない限り端末に出力することを意味する。ここで用意されているコマンドは、コマンドライブラリとして登録されているので、普通のコマンドと同様に入力すればよい。

- 1) PDBデータの一覧、及びステイタス別のデータの一覧（データコード名、分子名、データ登録者名、格納磁気テープボリューム名、データセット名、関連ニュースレター番号、その他）を得ることができる。
- 2) レコード検索によって、目的とするPDBデータに関して1)で述べた情報以外に、採取資料名、X線解析に関する各種の変換マトリックス、アミノ酸配列、 α -ヘリックス・ β シートなどの二次構造、結合部位、ジスルフィド結合、ヘムなどのヘテロな残基、関連文献名などの情報を得ることができる。
- 3) PDBデータのレコード構成や、ディスク上にコピーする場合のスペースを知ることができる。
- 4) PDBデータのリマーク部分のリスト（ラインプリンタ出力）を得ることができる。
- 5) 6つのコマンドの入力形式・機能を教えるコマンドが用意されている。

5.3. コマンドの説明

(1) REFERコマンド

1) 入力形式

コマンド	オ ペ ラ ン ド	
REFER	データコード名	NUM SNUM PAUSE NOPAUSE

2) 機能

指定したPDBリマークデータから入力促進に対して応答して指定したレコードを出力する。

3) オペランドの説明

a. データコード名

PDBで定められているデータコード名を指定する。（表1参照）

b. NUM/SNUM

NUMを指定すると行番号フィールドまで出力し、SNUMを指定すると行番号フィールドを出力しない。

NUMもSNUMも省略した場合、NUMが指定されたものとみなす。

c. PAUSE/NOPAUSE

PAUSEを指定した場合一画面分（23行）出力した後一旦ポーズし、次の処理を問い合わせてくる。

NOPAUSEを指定した場合はポーズしない。PAUSEもNOPAUSEも省略した場合、NOPAUSEが指定されたものとみなす。

4) 入力促進に対する応答

レコード識別子、またはその並び（最大10個）、または'ALL'を指定する。これらは任意省略形（先頭の文字を含む任意長の部分文字列）で指定してよい。'ALL'を指定した場合、リマーク部分のすべてのレコードを出力する。

5) READYモードへの復帰

入力促進をやめさせREADYモードへ戻したい場合は、入力促進に対して直ちに空行を返すか、応答の並びの最後に'END'を指定する。

(2) DSCANコマンド

1) 入力形式

コマンド	オペランド
DSCAN	データコード名

2) 機能

指定したPDBデータの種別レコード数、ディスクにコピーする場合に必要なスペース量を出力する。

3) オペランドの説明

a. データコード名

REFERコマンドの場合と同じである。

(3) DSTATUSコマンド

1) 入力形式

コマンド	オペランド
DSTATUS	[PAUSE] [NOPAUSE]

2) 機能

入力促進に対する応答として指定したPDBデータまたはプログラムに関して、或は、指定したステータスを持つPDBデータに関して、①エントリー番号、②データ又はプログラムコード名、③分子名又はプログラムの機能、④登録者名、⑤登録日、⑥データステータス又はプログラムサポートの有無、⑦格納されているセンタークロズド磁気テープのボリューム名、⑧データセット名、⑨更新コード、⑩関連ニュースレター番号、⑪更新日、⑫コメントを出力する。

3) オペランドの説明

a. PAUSE/NOPAUSE

REFERコマンドの場合と同じである。

4) 入力促進に対する応答

a. コード名リスト

データまたはプログラムコード名、または▼*AAA▼、またはこれらの並び（最大10個）を指定する。
▼*AAA▼を指定した場合、データコードの後ろ3文字がAAAのデータに関するステータス情報を出力する。コード名リスト指定の場合は、2)の①～⑫の情報全てを出力する。

b. INFORMATION

システム管理者、PDBデータ磁気テープボリューム名と属性、ステータスコード名、更新コード名、最新ニュースレター番号などを出力する。

c. NEW [ALL]

最新のニュースレターで新規登録されたデータ及びプログラムに関する情報を出力する。

d. CORRECTED [ALL]

最新のニュースレターで更新されたデータ及びプログラムに関する情報を出力する。

e. STANDARD [ALL]

標準的データに関する情報を出力する。

f. ALPHA [ALL]

α炭素だけのデータに関する情報を出力する。

g. BACKBONE [ALL]

主鎖原子だけのデータに関する情報を出力する。

h. PROMISED [ALL]

更新が予定されているデータに関する情報を出力する。

i. APPROVAL [ALL]

新規登録が予定されているデータに関する情報を出力する。

j. PREPARE [ALL]

登録準備中のデータに関する情報を出力する。

k. REPLACE [ALL]

古いパラメータセットに代って新しいパラメータセットでリファインし直されたデータに関する情報を出力する。

l. PROGRAM [ALL]

すべてのプログラムエントリーに関する情報を出力する。

m. DATA [ALL]

すべてのデータエントリーに関する情報を出力する。

n. ALL [ALL]

すべてのプログラムエントリーとデータエントリーについての情報を出力する。

- ・ a～nの応答は同時に指定することはできない。もし誤まって指定すると、並びの最初の応答でa～nのうちのどの応答かを判断する。

- ・ b～n の第1 応答は他の応答と区別できる任意省略形で指定することができる。
- ・ c～n の第2 応答で ALL を省略すると2) で述べた出力情報のうち①～③及び⑥～⑧を出力し、指定すると①～⑫の情報全てを出力する。

5) READYモードへの復帰

REFERコマンドの場合と同じである。

(4) PRINTRコマンド

1) 入力形式

コマンド	オ ペ ラ ン ド
PRINTR	CODE (データコード名リスト) NUM SNUM [COLUMN (整数)] [MAXLINE (整数)] [DSTATUS]

2) 機能

指定したデータのリマークをラインプリンタへ出力する。

3) オペランドの説明

a. CODE (データコード名リスト)

データコードまたはデータコードの*表現またはそれらの並びを指定する。並びを指定する場合は、全体を一对の引用符でくくる。

b. NUM/SNUM

REFERコマンドの場合と同じである。

c. COLUMN (整数)

出力位置を指定する。ラインプリンタ用紙の左から数えて整数で指定したバイト目から出力される。省略時はCOLUMN (21) がとられる。

d. MAXLINE (整数)

1 ページに出力する行数を整数で指定する。省略時はMAXLINE (64) がとられる。

e. DSTATUS

現在のPDBの座標データエントリー及びプログラムエントリーのステイタスの一覧をラインプリンタに出力する。DSTATUSを指定した場合、CODE指定をしていても無視される。

(5) RECORDコマンド

1) 入力形式

コマンド	オ ペ ラ ン ド
RECORD	な し

2) 機能

入力促進に対する応答によって、PDBデータのレコード識別子の一覧を出力したり、指定したレコード

識別子を持つレコードに記載されている事項の意味を出力する。

3) 入力促進に対する応答

以下の a～c は任意省略形で指定することができる。

a. NAMES

レコード識別子の一覧を出力する。

b. レコード識別子またはその並び (最大 10 個)

指定されたレコード識別子を持つレコードに記載されている事項の意味を出力する。

c. ALL

全てのレコード識別子とそれらのレコードに記載されている事項の意味を出力する。ALL を指定すると出力後 READY モードに戻る。

(6) TELL コマンド

1) 入力形式

コマンド	オペランド
TELL	NAMES ALL コマンド名

2) 機能

用意されているコマンドの一覧や、指定されたコマンドの入力形式、機能、オペランドの説明を出力する。

3) オペランドの説明

a. NAMES

コマンドの一覧を出力する。

b. ALL

全てのコマンドに関する 2) で述べた情報を出力する。

c. コマンド名

指定したコマンドに関する 2) で述べた情報を出力する。

5.4. 簡単な使用例

以下に REFER コマンドと DSTATUS コマンドの使用例を示す。下線の部分が利用者の入力する部分である。(左端の番号を参照のこと。) REFER コマンドの例は、蛋白質リゾチームに関する情報検索の例である。

```

READY
① REFER ILYZ
② +RECORD ID> HELIX SHEET SSBOND
<RECORDS 'HELIX' FOR ILYZ>
HELIX 1 A ARG 5 HIS 15 1 ILYZ 218
HELIX 2 B LEU 25 GLU 35 1 ILYZ 219
HELIX 3 C CYS 80 LEU 84 5 ILYZ 220
HELIX 4 D THR 89 LYS 96 1 ILYZ 221
<RECORDS 'SHEET' FOR ILYZ>
SHEET 1 S1 2 LYS 1 PHE 3 0 ILYZ 222
SHEET 2 S1 2 PHE 30 THR 40 -1 N THR 40 O LYS 1 ILYZ 223
SHEET 1 S2 3 ALA 42 ASN 46 0 ILYZ 224
SHEET 2 S2 3 SER 50 GLY 54 -1 N ASN 46 O SER 50 ILYZ 225
SHEET 3 S2 3 GLN 57 SER 60 -1 N TYR 53 O ILE 58 ILYZ 226
<RECORDS 'SSBOND' FOR ILYZ>
SSBOND 1 CYS 6 CYS 127 ILYZB 92
SSBOND 2 CYS 30 CYS 115 ILYZB 93
SSBOND 3 CYS 64 CYS 80 ILYZB 94
SSBOND 4 CYS 76 CYS 94 ILYZB 95
③ +RECORD ID> HEADER COMPND SOURCE AUTHOR
<RECORDS 'HEADER' FOR ILYZ>
HEADER HYDROLASE (O-GLYCOSYL) 01-FEB-75 ILYZ ILYZ 3
<RECORDS 'COMPND' FOR ILYZ>
COMPND LYSOZYME (E.C.3.2.1.17) ILYZB 1
<RECORDS 'SOURCE' FOR ILYZ>
SOURCE HEN (GALLUS GALLUS) EGG WHITE ILYZB 2
<RECORDS 'AUTHOR' FOR ILYZ>
AUTHOR R.DIAMOND,D.C.PHILLIPS,C.C.F.BLAKE,A.C.T.NORTH ILYZB 3
④ +RECORD ID> FORMUL HET JRNL
<RECORDS 'FORMUL' FOR ILYZ>
FORMUL 2 HOH *97(H2 O1) ILYZB 91
<RECORDS 'JRNL' FOR ILYZ>
JRNL AUTH R.DIAMOND ILYZB 4
JRNL TITL REAL-SPACE REFINEMENT OF THE STRUCTURE OF HEN ILYZB 5
JRNL TITL 2 EGG-WHITE LYSOZYME ILYZB 6
JRNL REF J.MOL.BIOL. V. 82 371 1974 ILYZB 7
JRNL REFN ASTM JMOBOK UK ISSN 0022-2836 070 ILYZB 8
MSG: NOT FOUND REQUEST(S)- 'HET'
⑤ +RECORD ID> MASTER END
<RECORDS 'MASTER' FOR ILYZ>
MASTER 106 145 0 4 5 11 0 6 1102 1 8 10 ILYZB 96
READY
⑥ DSTATUS
⑦ +REQUEST(S)> ILYZ *CHA 3PTI
80 <ILYZ> LYSOZYME (HEN EGG-WHITE, SET W2) /ST/ /MT/F02213 /DS/DS077
/DEPOSITOR/R.DIAMOND,D.C.PHILLIPS /DATE/ 2.75
/CORRECT/ /NEWS/ /COMMENT/
26 <2CHA> ALPHA-CHYMOTRYPSIN A (TOSYLATED) /ST/R /MT/F02213 /DS/DS029
/DEPOSITOR/J.J.BIRKTOFT,D.M.BLOW /DATE/ 1.75
/CORRECT/ /NEWS/(4) /COMMENT/
27 <3CHA> ALPHA-CHYMOTRYPSIN A /ST/ /MT/F02213 /DS/DS030
/DEPOSITOR/A.TULINSKY,R.L.VANDLEN /DATE/ 8.76
/CORRECT/ /NEWS/(4) /COMMENT/
134 <3PTI> TRYPSIN INHIBITOR (BOVINE PANCREAS) /ST/R /MT/F02213 /DS/DS129
/DEPOSITOR/R.HUBER,J.DEISENHOFER /DATE/11.76
/CORRECT/ /NEWS/(4) /COMMENT/
⑧ +REQUEST(S)> ALPHA
<ENTRIES WITH ALPHA CARBONS ONLY>
8 <1WGA> AGGLUTININ (WHEAT GERM) /ST/A /MT/F02213 /DS/DS015
14 <2BCL> BACTERIOCHLOROPHYLL-A-PROTEIN /ST/RA /MT/F02213 /DS/DS021
22 <1CPB> CARBOXYPEPTIDASE B (BOVINE) /ST/A /MT/F02213 /DS/DS028
53 <1HRB> HEMERYTHRIN B /ST/A /MT/F02213 /DS/DS055
54 <1HMN> HEMERYTHRIN (MET,AQUO) /ST/A /MT/F02213 /DS/DS056
70 <1MCG> IMMUNOGLOBULIN B-J INTACT MCG /ST/A /MT/F02213 /DS/DS068
72 <1RHE> IMMUNOGLOBULIN B-J FRAGMENT(MNMR) RHE /ST/A /MT/F02213 /DS/DS070
73 <1KGA> KDPG ALDOLASE /ST/A /MT/F02213 /DS/DS071
89 <1MDH> MALATE DEHYDROGENASE (CYTOPLASMIC) /ST/A /MT/F02213 /DS/DS085
95 <1MHR> MYOHEMERYTHRIN /ST/A /MT/F02213 /DS/DS091
103 <1PEP> PEPSIN (PORCINE) /ST/A /MT/F02213 /DS/DS099
104 <1PGK> PHOSPHOGLYCERATE KINASE (YEAST) /ST/A /MT/F02213 /DS/DS100
106 <1PGM> PHOSPHOGLYCERATE MUTASE (YEAST) /ST/A /MT/F02213 /DS/DS102

```

```

108 <1PYK> PYRUVATE KINASE (CAT) /ST/AN /MT/ /DS/
120 <1SSI> SUBTILISIN INHIBITOR (STREPTOMYCES) /ST/A /MT/F02213 /DS/DS115
123 <1SOD> SUPEROXIDE DISMUTASE /ST/A /MT/F02213 /DS/DS118
126 <1SRX> THIOREDUXIN (E.COLI, OXIDISED) /ST/A /MT/F02213 /DS/DS121
⑨ +REQUEST(S)> BACKBONE ALL
<ENTRIES WITH BACKBONE ATOMS ONLY>
105 <2PGK> PHOSPHOGLYCERATE KINASE (HORSE) /ST/B /MT/F02213 /DS/DS101
/DEPOSITOR/P.EVANS,C.BLAKE /DATE/ 9.76
/CORRECT/ /NEWS/ /COMMENT/
136 <1SBV> VIRUS COAR PROTEIN(SOUTHERN BEAN MOSA /ST/BN /MT/ /DS/
/DEPOSITOR/M.ROSSMANN /DATE/12.79
/CORRECT/ /NEWS/ /COMMENT/AWAIT NXT DISTRBTN
⑩ +REQUEST(S)> CORRECTED
<CORRECTED ENTRIES>
57 <2DHB> HEMOGLOBIN (HORSE,DEOXY) /ST/ /MT/F02213 /DS/DS058
77 <1LDX> LACTATE DEHYDROGENASE (HOUSE TESTES) /ST/ /MT/F02213 /DS/DS074
78 <1HBL> LEGHEMOGLOBIN (ACETATE,NET) /ST/ /MT/F02213 /DS/DS075
84 <5LYZ> LYSOZYME (HEN EGG-WHITE, SET RS12A) /ST/ /MT/F02213 /DS/DS081
120 <1SSI> SUBTILISIN INHIBITOR (STREPTOMYCES) /ST/A /MT/F02213 /DS/DS115
135 <3PTP> TRYPSIN (DIP INHIBITED) /ST/R /MT/F02213 /DS/DS130
⑪ +REQUEST(S)> PROMISED ALL END
MSG: NOT FOUND REQUEST(S)- 'PROMISED'
READY

```

(使用例の説明)

- ① データコード名 1 LYZ のリマークのレコード検索を開始する。
- ② レコード識別子 HELIX, SHEET, SS BOND を持つレコードを出力させる。
- ③ レコード識別子 HEADER, COMPND, SOURCE, AUTHOR を持つレコードを出力させる。
- ④ レコード識別子 FORMUL, HET, JRN L を持つレコードを出力させる。▼ MSG : ▼ として, HET レコードが 1 LYZ のデータの中になくいることを知らせてきている。
- ⑤ レコード識別子 MASTER を持つレコードを出力した後 READY モードに戻る。
- ⑥ データステータス検索を開始する。
- ⑦ データコード 1 LYZ, コード名の後ろ 3 文字が CHA のデータ, 3 PTI のデータに関する情報を出力させる。
- ⑧ ALPHA 炭素だけのデータを出力させる。ALL は省略する。
- ⑨ 主鎖原子だけのデータを出力させる。ALL を指定する。
- ⑩ 今回のニュースレターで更新されたデータを出力させる。
- ⑪ 登録が予定されているデータに関する情報を ALL 指定で出力させた後 READY モードに戻る。
▼ MSG : ▼ で今のところ登録予定データがないことを示してきている。

6. おわりに

最初に述べたように、今日では、いかなる蛋白質の物理化学的研究も、立体構造のイメージを基礎に据えて議論しなければならない場合が多くなった。一般の蛋白質研究者も、原子座標データを処理して自分の研究に役立てるような場合が、これからますます増えることは間違いない。しかし、日頃計算機にあまり馴染のない研究者にとっては、この間のギャップはかなり大きいと思われる。

PSPCS は、そのような潜在的な利用者にも使いやすいように開発したつもりである。また、ある程度、熟達した利用者には、PSPCS のプログラムを自分の研究に最適に改編して使用したり、PSPCS を使って得た成果を基礎データとして、さらに一段階進んだ研究に使用するなど、その応用できる範囲は広いと思われる。利用

者の方々のいろいろなご意見・ご批判をお願いしたい。

なお、PDBを使用した研究成果を発表する時には、文献[2]を引用することが望まれる。また、PSPCSを利用した場合には、各プログラム名と作成者・改訂者名を明記する必要があることを付加しておく[11]。

PSPCSは、分子科学研究所の計算機センターにも、すでに登録されていて利用できる。また、名大、京大、阪大の各大型計算機センターでも使えるよう現在作業中である。

最後に、いろいろご協力いただいたPSPCS作成グループのメンバーの方々（表2に氏名が出ている）、試験研究グループの方々、日頃いろいろご教示いただいている名大・理学部・化学教室の佐々木教祐氏、阪大大型計算機センターの磯本征雄氏、阪大蛋白質研究所の安岡則武氏、図3の原因を送っていただいた名大・理学部・物理学教室の別府良孝氏に感謝する。

参考文献

1. 郷 信広, 水野裕重, 郷 通子, 宮沢三造, 武富 敬 蛋白質原子座標データとその情報処理, 蛋白質・核酸・酵素, Vol. 23, No. 13, pp. 1336 - 1344, 1978.
2. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. The Protein Data Bank: A Computer-based Archival File for Macromolecular Structures, J. Mol. Biol., Vol. 112, pp. 535 - 542, 1977.
3. 角戸正夫, 安岡則武 蛋白質データベース, 日本結晶学会誌, Vol. 19, pp. 294 - 296, 1977.
4. 武富 敬 データベース, 九大大型計算機センター広報, Vol. 10, No. 4, pp. 257 - 271, 1977.
5. 磯本征雄, 安岡則武, 田中信夫, 松浦良樹, 角戸正夫 学術研究用たんぱく質データベース PROTEIN-DB, 情報処理, Vol. 21, No. 1, pp. 15 - 22, 1980.
6. タンパク質原子座標の情報処理プログラム・システム[I], 科研費報告集, 1977.
7. タンパク質原子座標の情報処理プログラム・システム[II], 科研費報告集, 1978.
8. タンパク質原子座標の情報処理プログラム・システム[III], 科研費報告集, 1979.
9. 別府良孝 電子計算機による分子構造の表示, 生化学, Vol. 51, No. 1, pp. 24 - 28, 1979.
10. 別府良孝 分子構造表示プログラムNAMODについて, 名大大型計算機センターニュース, Vol. 9, No. 2, pp. 123 - 134, 1978.
11. タンパク質原子座標の情報処理プログラム・システム[IV], 科研費報告集, 1980.
これが, PSPCS利用者マニュアルとなる。
12. Taketomi, H., Ueda, Y. & Go, N. Studies on Protein Folding, Unfolding and Fluctuations by Computer Simulation. I., Int. J. Peptide Protein Res., Vol. 7, pp. 445 - 459, 1975.
13. Go, N. & Taketomi, H. Respective Roles of Short- and Long-range Interactions in Protein Folding, Proc. Natl. Acad. Sci., USA, Vol. 75, pp. 559 - 563, 1978.