

Brookhaven National Laboratory
Protein Data Bank
Newsletter

The Protein Data Bank has grown appreciably in the six months since the last issue of this newsletter as evidenced by the following table.

<u>Date</u>	<u>No. Sets Held</u>	<u>Atomic Coordinates</u>	
		<u>No. Sets Distributed</u>	<u>No. Recipients</u>
1973	15	106	14
1974	17	99	14
1975	40	513	31
1976 (Oct. 15th)	72	1101	38

These distribution statistics represent activity only at Brookhaven. Duplicate copies of the file are maintained at the Crystallographic Data Centre in Cambridge (England) and at the University of Tokyo (Japan). These sites distribute data independently.

All of the new structures obtained since this past spring have been entered in the new 80-column format and work is proceeding on conversion of the older parameter sets to the new format. After approval by depositors, new-style entries for old parameter sets replace the original ones. Thus far, four such sets have been replaced. In addition, six sets are out for approval and 10 in preparation. A detailed list of our holdings of current data sets is enclosed. In addition to the listed entries we have structure-factor data and outdated coordinate sets for several structures--inquiries regarding these are invited.

As depositors and users who have examined data in the new format will be aware, this new file structure can accommodate much additional information not previously included in the Bank. The new file structure is significantly improved by the integration of text, specification of secondary structures, coordinate transformations, non-standard connectivity and other useful information with the coordinates. Full details of the file structure would take too much space here but a complete description is available by filling out and returning the request form.

Effective November 1, 1976 a charge of \$34.30 per magnetic tape will be made to recipients of data from the Bank (at Brookhaven). Currently all atomic coordinate data in the Bank will fit on one 2400' reel of tape at

800 bpi with 20 records per block, so that a complete copy of the data can be obtained for a single charge of \$34.30 (user-supplied tape). Persons requesting data from the Bank should include a purchase order with their request. It is most convenient for us to send all available coordinate sets at one time, but if specific entries are requested we ask that these be referred to by the code associated with each entry (see list of holdings).

The last newsletter implied the eventual availability of computer programs from the Data Bank. Unfortunately the task of updating the database has fully occupied us and so no appropriate programs have been written here. It is, however, our intention in due course to provide programs for connectivity/chirality checking, ϕ/ψ plot generation, diagonal plot generation, and display list generation. In addition we are prepared to accept and distribute computer code useful to people involved in either the provision of macromolecule data or its interpretation, and people wishing to have us handle this distribution for them are invited to deposit code. The rules for this are simple--i) The code, complete with full documentation, must be in 80 col. machine-readable form. ii) We will not accept responsibility for the completeness, correctness or usefulness of the code, nor will we (usually) undertake to distribute "fixes" to users unless these are requested. iii) We will act as a passive clearing house for user complaints and will correct the code on instructions from the depositor.

From time to time errors are found in the data files. Corrections for these are either obtained from or referred to the depositor for verification before any data files are changed. Recipients of erroneous data have not usually been notified of these changes, however. In order to improve matters we intend to maintain a file of corrections which will become part of this newsletter. This service will begin with the next edition (spring 1977).

We are aware that the Bank lacks data for many protein structures, in spite of our recent rapid rate of growth. We will continue to solicit data as we become aware of new structures, and we urge investigators to deposit atomic coordinates, structure factors and phases when these become available. All new data received are made available to R. J. Feldmann at NIH for use in generation of the Atlas of Macromolecular Structures on Microfiche (AMSOM).

Mail, depositions or other correspondence to the Protein Data Bank may be addressed to any one of the three people listed below. Our institutional address is

Department of Chemistry
Brookhaven National Laboratory
Upton, New York 11973

Personnel

Ms. F. C. Bernstein	Tel (516) 345-4382
Dr. T. F. Koetzle	(516) 345-4384
Dr. G. J. B. Williams	(516) 345-4383

PROTEIN DATA BANK HOLDINGS - October 15, 1976

Code	Name	Depositor	Status	Code	Name	Depositor	Status
IADK	Adenylate Kinase	G. Schulz	SCA	LDH02	Lactate dehydrogenase	M. Rossmann	LCTR
IADH	Alcohol Dehydrogenase (ADP-Rib)	C.-I. Brändén	SCP	TERNRY01	Lactate dehydrogenase/NAD/Pyruvate	M. Rossmann	LCT
2ADH	Alcohol Dehydrogenase (Orthophen)	C.-I. Brändén	SCP	LYRS5D01	Lysozyme Hen egg white set RS5D	R. Diamond	LCT
3CHA	Alpha-chymotrypsin	A. Tulinsky	SC	LYRS6A01	Lysozyme Hen egg white set RS6A	R. Diamond	LCT
CHYM02	Alpha-chymotrypsin (tosyl)	D. Blow	LCR	LYRS9A01	Lysozyme Hen egg white set RS9A	R. Diamond	LCT
1FAB	Antigen binding fragment (New)	R. Poljak	SC	LYRS12A1	Lysozyme Hen egg white set RS12A	R. Diamond	LCT
IREI	Bence-Jones Immunoglobulin REI	O. Epp, R. Huber	SC	LYRS1601	Lysozyme Hen egg white set RS16	R. Diamond	LCT
CMPP6A01	Calcium-binding parvalbumin Set 6A	R. Kretsinger	LC	LYS0W201	Lysozyme Hen egg white set W2	R. Diamond	LCT
CMPP6H01	Calcium-binding parvalbumin Set 6H	R. Kretsinger	LC	LYR0H001	Lysozyme Hen egg white water molecules	R. Diamond	LC
CMPP6I01	Calcium-binding parvalbumin Set 6I	R. Kretsinger	LC	IMDH	Malate dehydrogenase	L. Banaszak	SCA
1CAB	Carbonic anhydrase B	K. Kannan	SC	IMBN	Myoglobin (sperm whale)	H. Watson	SC
1CAC	Carbonic anhydrase C	K. Kannan	SC	2MBN	Myoglobin (sperm whale, met)	T. Takano	SC
1CPA	Carboxypeptidase A	W. Lipscomb	SC	3MBN	Myoglobin (sperm whale, deoxy)	T. Takano	SC
1CHG	Chymotrypsinogen	J. Kraut	SCN	IPAP	Papain, native	J. Drenth	SCN
CONCNA02	Concanavalin A	G. Reeke, G. Edelman	LCT	2PAP	Papain (Ace-ala-ala-phe-ala, Cys-25)	J. Drenth	SCP
3CNA	Concanavalin A	K. Hardman	SCR	3PAP	Papain (Cys deriv of Cys-25)	J. Drenth	SCP
CYT501	Cytochrome B5	F. S. Mathews	LC	4PAP	Papain (oxidized Cys-25)	J. Drenth	SCP
CYT502	Cytochrome B5	F. S. Mathews	LT	5PAP	Papain (Tos-lys, Cys-25)	J. Drenth	SCP
1CYT	Cytochrome c (Albacore, Oxidized)	R. Dickerson	SC	6PAP	Papain (Bzoxo-gly-phe-gly, Cys-25)	J. Drenth	SCP
2CYT	Cytochrome c (Albacore, Reduced)	R. Dickerson	SC	7PAP	Papain (Bzoxo-phe-ala, Cys-25)	J. Drenth	SCP
1CYC	Cytochrome c (Bonito, heart)	M. Kakudo	SC	2PTI	Pancreatic trypsin inhibitor	R. Huber	SCRN
CYT0C201	Cytochrome c ₂	J. Kraut	LC	1PGK	Phosphoglycerate kinase (yeast)	H. Watson	SCA
155C	Cytochrome c ₅₅₀	R. Timkovich	SC	2PGK	Phosphoglycerate kinase (horse)	P. Evans, D. Phillips	SCB
1EST	Elastase	H. Watson	SC	1PAB	Prealbumin (human, plasma)	S. Oatley, D. Phillips	SC
1FDX	Ferredoxin	L. Jensen	SC	RNASES01	Ribonuclease S	H. Wyckoff	LC
1FXN	Flavodoxin (clostridium MP)	M. Ludwig	SC	RUDXN02	Rubredoxin	L. Jensen	LCR
LOBGPD01	Glyceraldehyde-3-P-dehydrogenase Lobster (green)	M. Rossmann	LCT	STAPHN01	Staphylococcal nuclease	F. A. Cotton, E. Hazen	LC
LOBGPD02	Glyceraldehyde-3-P-dehydrogenase Lobster (red)	M. Rossmann	LCT	1SGB	Streptomyces Griseus Proteinase B	M. James	SCA
1MHB	Hemoglobin (horse, met)	M. Perutz, G. Fermi	SCN	DISMUT01	Superoxide dismutase	J. & D. Richardson	LCA
1DHB	Hemoglobin (horse, deoxy)	M. Perutz, G. Fermi	SCN	SUBTLE1	Subtilisin BPN'	J. Kraut	LC
1HHB	Hemoglobin (human, deoxy)	M. Perutz, G. Fermi	SC	2SBT	Subtilisin novo	J. Drenth	SC
1FDH	Hemoglobin (human, fetal, deoxy)	J. Frier	SC	1SRX	Thioredoxin	B.-O. Soderberg	SCA
LAMP1	Hemoglobin (lamprey)	W. Hendrickson	LCT	THERM002	Thermolysin (refined)	B. Matthews	LC
1YHX	Hexokinase (yeast) BIII	T. Steitz	SCB	THERM001	Thermolysin (unrefined)	B. Matthews	LC
1HIP	High Potential Iron Protein	J. Kraut	SCN	1TIM	Triose phosphate isomerase	I. Wilson, D. Phillips	SC
				1TNA	Transfer RNA (yeast, phe)	J. Sussman, S.-H. Kim	SCP
				2TNA	Transfer RNA (yeast, phe)	M. Sundaralingam	SCP

Status Codes

L 140 column "Diamond style" format
 S 80 column Protein Data Bank format
 C atomic coordinates held
 T torsion angles held
 R replaces an out-of-date set

P in preparation
 N new format entry with contributor for approval
 A alpha carbon atoms only
 B backbone only

Please note any address correction above.

Fold 1

Ms. F. C. Bernstein
Chemistry Department
Brookhaven National Laboratory
Upton, New York 11973

Fold 2