PS22.03.09 π -TURNS IN PROTEINS AND PEPTIDES; CLASSIFICATION, OCCURRENCE, SEQUENCE AND HYDRATION: A PROTEIN DATA BANK ANALYSIS S. Ramakumar, K. R. Rajashankar, Department of Physics, Indian Institute of Science, Bangalore-560012, India

This paper describes the conformational features of 6→1 hydrogen bonded turns (π-turns) observed in proteins and peptides. The polypeptide conformations under the constraint of 6→1 hydrogen bond were selected from Brookhaven Protein Data Bank (PDB) using a FORTRAN program developed in our laboratory. The data set consisted of the co-ordinates of 228 non-homologous protein chains, determined by X-ray crystallography to better than 2.5 Å resolution. Totally 486 π-turns were located. 367 π-Turns were found to have 5th residue in left handed α-helical(α_L) region of Ramachandrans map and were classified as $\pi_{\alpha L}$-turns. These turns are generally observed at the C-terminal end of helices, a result in harmony with previous observations. 286 $\pi_{\alpha L}$-Turns define 'Schellman motif' (6→1 & 5→2 hydrogen bond). 111 π-Turns had 5th residue in α_R region, a novel finding and were referred by the name $\pi_{\alpha R}$-turns. $\pi_{\alpha R}$-Turns generally occur in α-helices as a distortion. Four π-turns were seen to be mirror images of $\pi_{\alpha L}$-turns and hence were termed as $\pi_{\alpha L}$'-turns. The 5th residue in rest of the four π-turns were seen to adopt ß-conformation, hence these four were named as π_{β}-turns. The former two classes of π-Turns ($\pi_{\alpha L}$ & $\pi_{\alpha R}$) form major classes. Nine π-turns observed so far in oligopeptides share the features of $\pi_{\alpha L}$-turns. Sequence analysis shows that hydrophilic residues are preferred at position 2,3 and 4 of $\pi_{\alpha L}$-turns while position 1 and 6 prefer hydrophobic residues. Residue 5 (α_L) is mainly Gly and less often Asn. A high preference for Pro after the C-termini is observed in both the major classes. Poor α-helix formers like His, Tyr and Asn were found to be preferred for $\pi_{\alpha R}$ -turns, where as good α-helix former Ala is not preferred. The discussion will include the details of classification, context of occurrence, sequence and hydration of π-turns in proteins and peptides.

(K.R.R. thanks CSIR(India) for a fellowship)

PS22.03.10 AUTODEP - FACILITATING DEPOSITION TO THE PROTEIN DATA BANK THROUGH THE NEW WEB-BASED SUBMISSION FORM. D.R.Stampf, E.E.Abola, N.O.Manning, D.Xue, and J.L.Sussman, Brookhaven National Laboratory, P.O. Box 5000, Upton, NY, 11973-5000 USA, and Bioinformatics Unit, Weizmann Institute of Science, 76100 Rehovot, Israel.

The Protein Data Bank (PDB) is an archive of experimentally-determined three-dimensional structures of biological macromolecules, serving a global community of researchers, educators, and students. The database contains atomic coordinates, bibliographic citations, primary and secondary structure information, as well as crystallographic structure factors and NMR experimental data.

In order to streamline the deposition procedure, both to reduce the depositors' time spent in filling out the deposition form and to increase the accuracy and information content of a PDB entry, a World Wide Web-based graphical user interface (AutoDep) has been designed.

The AutoDep form may be populated with data from any released PDB entry, a previously submitted but not yet released entry, or it may be filled out from scratch. Additionally, PDB-formatted output from refinement and rebuilding programs can be submitted in lieu of filling out the form.

Preliminary verification of each field in the form is done during the interactive session. The depositor is able to view the resultant bibliographic portion of the PDB entry at any time in order to check his or her progress. A simple script is provided that can be copied to a window on the home computer which will FTP the coordinates and related files to the PDB.

The AutoDep procedure is based on CIF and was built using CIF-based data dictionaries. Any changes to the deposition procedures or information content of PDB entries are easily accommodated by editing the dictionary files.

The PDB is supported by a combination of Federal Government Agency funds and user fees. Support is provided by the U.S. National Science Foundation, the U.S. Public Health Service, National Institutes of Health, National Center for Research Resources, National Institutes of General Medical Sciences, National Library of Medicine, and the U.S. Department of Energy under contract DE-AC02-76CH00016.

PS22.03.11 PDB VALIDATION PROCEDURES - PAST, PRESENT AND FUTURE J.P. Rose, S. Swaminathan, E. Abola, N.O. Manning, J.L. Sussman, Chemistry Department, Brookhaven National Laboratory, P.O. Box 5000, Upton, NY, 11973-5000 USA.

The Protein Data Bank (PDB) is an archive of experimentally-determined three-dimensional structures of biological macromolecules, serving a global community of researchers, educators, and students. As the PDB has evolved over the past 25 years, the validation procedures used to evaluate the data have also evolved from simple bond length and angle calculations to the extensive suite of validation checks including PROCHECK, WhatCheck, PDBINFO, PKB and local programs used today.

The PDB is currently developing procedures to allow automated deposition with minimal staff intervention. For these procedures to work, better methods of identifying and reporting possible errors and/or outliers are required. PDB is therefore working with the community to develop guidance on the validation of protein crystallographic data, coordinates, and the relation between them.

The PDB is also working closely with Axel Brunger and Jian-Sheng Jiang of Yale University to incorporate some of the procedures available through X-PLOR into PDB's validation suite. A special version of X-PLOR will be available from BNL to all depositors for validation purposes only.

The PDB realizes that it must build consensus among its depositors regarding the diagnosis and reporting of outliers. It must also give clear guidance to users on how to interpret and make use of these annotations.

The PDB is supported by a combination of Federal Government Agency funds. Support is provided by the U.S. National Science Foundation, the U.S. Public Health Service, National Institutes of Health, National Center for Research Resources, National Institutes of General Medical Sciences, National Library of Medicine, and the U.S. Department of Energy under contract DE-AC02-76CH00016.

PS22.03.12 PDB SUBMISSION WITH X-PLOR. J.-S. Jiang, N. O. Manning*, E. E. Abola*, A. T. Brünger, J. L. Sussman*, The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA, *Protein Data Bank, Brookhaven National Laboratory, Upton, NY 11973, USA

Submission tools have been developed for the X-PLOR program that output PDB records compliant with the coordinate file format of the Protein Data Bank at Brookhaven. The implementation of the submission tools is accomplished by using a set of easy-to-use "macro" files. These macros will allow easy and convenient distribution of changes of the deposition format without the need for changing the underlying X-PLOR code.