PERMOL: Restraint based Protein Homology Modeling using DYANA or CNS

Andreas Möglich, Daniel Weinfurtner, Wolfram Gronwald, Till Maurer and Hans Robert Kalbitzer*

Institut für Biophysik und physikalische Biochemie, Universität Regensburg, D-93040 Regensburg, Germany.

Abstract
PERMOL is a new restraint based program for homology modeling of proteins. Restraints are generated from the information contained in structures of homologous template proteins. Employing the restraints generated by PERMOL three-dimensional structures are obtained using MD programs such as DYANA or CNS. In contrast to other programs PERMOL is mainly based on the use of dihedral angle information which is optimally suited to preserve the local secondary structure. The global arrangement of these elements is then facilitated by a small number of distance restraints. Using PERMOL homology models of high quality are obtained. A key advantage of the proposed method is its flexibility allowing the inclusion of data from other sources such as experimental data and the use of modern molecular dynamics programs to calculate structures.

Availability
The software and a detailed manual are available free of charge (http://www.biologie.uni-regensburg.de/Biophysik/Kalbitzer/permol/permol.html).

Contact
hans-robert.kalbitzer@biologie.uni-regensburg.de

Program Overview
Homology modeling with PERMOL is accomplished in several stages. Initial sequence alignment between the target sequence and the homologous template proteins is carried out with CLUSTAL_X (Thompson et al., 1997). In several modules the PERMOL software evaluates the three-dimensional structures of the template proteins and derives different types of spatial restraints such as averaged dihedral angle restraints for use with the MD programs DYANA (Güntert et al., 1997) and CNS (Brünger et al., 1998), both of which are routinely used for structure determination by NMR spectroscopy. Simulated annealing runs yield an ensemble of homology structures. PERMOL is written in Perl/Tk (http://www.perl.org) and has been tested on Unix (IRIX), Suse Linux and various Windows systems (98, 2000 and XP).

Program Description
As input PERMOL takes one or several protein structures in the PDB format and a text file containing the sequence of the target protein in one-letter code. Using CLUSTAL_X the sequences of the homologous template proteins are aligned to the target sequence. The resulting alignment is imported by PERMOL. Within three separate program modules different kinds of restraints are generated from the spatial information contained in the template structures: Dihedral angles, inter-atomic distances and hydrogen bonds are automatically translated into restraints for MD runs. A graphical user interface shows the aligned sequences with the individual residues colored depending on their degree of sequence conservation. Restraints are only generated for residues which have been selected before. Amino acids can be selected or deselected either individually or as a class based on their degree of sequence conservation. After appropriate selection the template structures are evaluated and strategies based on homology modeling seem particularly rewarding.

Here, we present the program PERMOL which can be employed to generate protein homology models using restraint molecular dynamics (MD) simulation programs.
spatial properties, e. g. dihedral angles, are computed. Finally, from these data restraints are generated. Several options are available which control how the spatial information is translated into restraints. For example, the restraint for a specific dihedral angle can be defined as the mean value observed in the template structures for this angle plus/minus the corresponding standard deviation. The individual restraints are weighted relative to each other according to the degree of homology of the amino acid residues involved. Output files can be created for use with either DYANA or CNS. Standard simulated annealing protocols in torsion angle space are used to calculate an ensemble of structures. From these the best in terms of the MD target function are selected as the homology model of the target protein.

**Discussion**

As an example we used PERMOL along with DYANA to calculate a homology structure for the HPr protein from *E. coli*. Comparison with the experimentally determined X-ray structure (Jia et al., 1993) yielded a value of 0.15 nm for the root mean square deviation of the backbone atom positions in the two structures. PERMOL was also successfully tested on the the ligand binding domain of human nuclear receptor PPARγ (Uppenberg et al., 1998). A related approach for homology modeling was published by Sali and Blundell (1993). In their program MODELLER spatial restraints are expressed as so-called probability density functions which are derived from structural features observed in homologous proteins. In some respects the PERMOL approach is also related to the program for fold prediction of helical proteins recently published by Zhang et al. (2002). Here, however, restraints are not derived from template proteins but are predicted based on primary and secondary structure of the target protein. Other previously published programs utilizing restrained molecular dynamics simulations include for example the programs by Kolinski et al. (2001) and Brocklehurst et al. (1993).

The high quality of the homology models we generated demonstrates that MD programs which are used for NMR structure determination can also be employed for the purpose of homology modeling. Using the graphical user interface of PERMOL full control over which restraints enter the modeling process is retained. As the created restraint files are text files, they can be easily edited and joined with data from other sources. Thus we believe that PERMOL could also be particularly useful in the context of NMR structure determination. Homology models could help in the process of resonance and NOE cross-peak assignment as well as in the analysis of residual dipolar couplings. The modeled structures can be validated by comparison to experimental data. Further, incomplete experimental data could be combined with spatial restraints derived from template proteins to yield reasonable structures. Recently, PERMOL has been successfully employed for the determination of the solution structure of a mutant form of HPr from *S. carnosus* (Möglich et al., 2004, in press). A modified version of PERMOL has also been incorporated into the NMR program suite AUREMOL (Gronwald and Kalbitzer, 2004).

**Acknowledgements**

The authors thank Dr. R. Döker and Dr. W. Kremer for helpful discussions. Financial support by the European Union is gratefully acknowledged.

**References**


