

## Role of Aromatic Amino Acids in Carbohydrate Binding. Laser Photo CIDNP (Chemically Induced Dynamic Nuclear Polarisation) and Molecular Modeling Study of Hevein-domain Containing Lectins

Hans-Christian Siebert [a,b], Claus-Wilhelm von der Lieth\* [c], Robert Kaptein [b], Ukun M.S. Soedjanaatmadja [d], Johannes F.G. Vliegthart [b], Christine S. Wright [e] and Hans-Joachim Gabius [a]

[a] Institut für Physiologische Chemie der Ludwig-Maximilians-Universität, München, Germany;

[b] Bijvoet Center for Biomolecular Research, University of Utrecht, The Netherlands;

[c] German Cancer Research Center, Zentrale Spektroskopie, Im Neuenheimer Feld 280, D - 69120 Heidelberg; Germany (W.vonderlieth@dkfz-heidelberg.de)

[d] Biochemisch Laboratorium and Bioson Research Institute, University of Groningen, The Netherlands;

[e] Department of Biochemistry and Molecular Biophysics, Virginia Commonwealth University, USA

Received: 15 May 1996 / Accepted: 6 August 1996 / Published: 27 September 1996

### Introduction

The recognition of carbohydrate determinants by lectins is involved in mediation of intercellular binding and elicitation of biosignalling processes [1,2]. Structural aspects of this glycobiological interplay can in principle be disclosed by multidimensional nuclear magnetic resonance (NMR) experiments and/or X-ray crystallography. However, these techniques can suffer from a number of inherent restrictions, for example molecular mass in multidimensional NMR studies and in case of X-ray crystallography a lack of information about the dynamic behaviour of molecules in solution. The side chains of tyrosine, tryptophan and histidine often involved in carbohydrate recognition are able to produce CIDNP signals [3] after laser irradiation in the presence of a suitable radical pair-generating dye. Elicitation of such a response in protein simplifies accessibility of the respective groups to the light-absorbing dye. This technique allows in close combination with molecular modeling tools to monitor on an atomic level the effect of ligand binding to a receptor if CIDNP-reactive amino acids are affected.

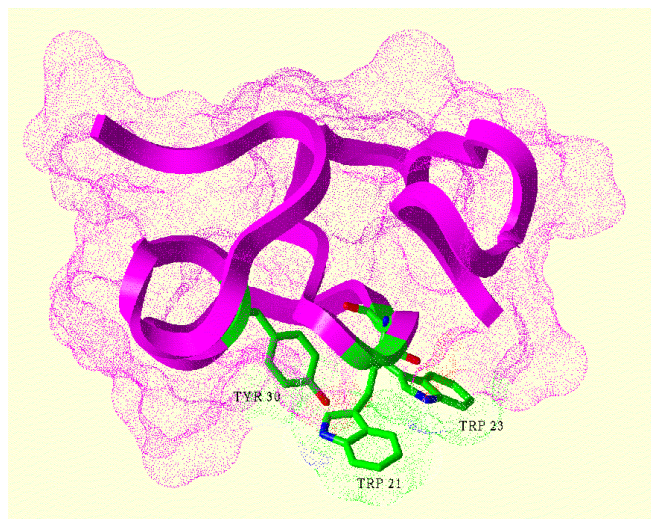
Experimentally, the shape and intensity of CIDNP signals are determined in the absence and in the presence of specific glycoligands. When the carbohydrate ligand is bound,

CIDNP signals of side chain protons of tyrosine, tryptophan or histidine residues are altered, e. g. they can be broadened and of reduced intensity or disappear completely.

To determine the actual value of this method in glycosciences, several well-characterised N-acetyl-glucosamine-binding lectins from different species with increasing molecular complexity are selected as models. In detail, laser photo CIDNP and molecular modeling experiments of hevein pseudo-hevein, *Urtica dioica* agglutinin (UDA), wheat germ agglutinin (WGA) and its B domain (domB) have been carried out in the absence and in the presence of their specific ligands. Hevein (43 amino acid residues, PDB entry 1HEV) and pseudo-hevein (43 amino acid residues, modelled by homology from PDB entry 1HEV) are small lectins of rubber trees (*Hevea brasiliensis*) with ragged C-terminal sequences. The amino acid sequences of the two proteins differ only in six positions from each other. Among these differences the exchange of Trp 21 present in hevein to Tyr 21 in pseudo-hevein is especially notable in the context of CIDNP experiments.

*Urtica dioica* agglutinin (UDA) from stinging nettle rhizomes is a member of the chitin-binding family of plant lectins whose structure consists of two hevein domains. It has a single chain of 89 amino acids, constituting two binding sites with different affinities for oligomers of N-acetylglucosamine.

\* To whom correspondence should be addressed



**Figure 1:** The backbone of hevein (PDB-entry 1HEV) is drawn in magenta. The aromatic amino acids TRP-21, TRP-23 and TYR 30 are displayed in green. The water accessible surface is indicated with a dotted surface.

The isoforms of wheat germ agglutinin (WGA PDB entry 7WGA), which are built up from four hevein domains can also bind N-acetylglucosamine. Their structures with a subunit molecular mass of 16 kDa and aggregation of two chains display a considerably increased complexity in retention to the other lectins of this family. The carbohydrate-specific unit, namely the B-domain, was cloned enabling to logically extend the list of test substances.

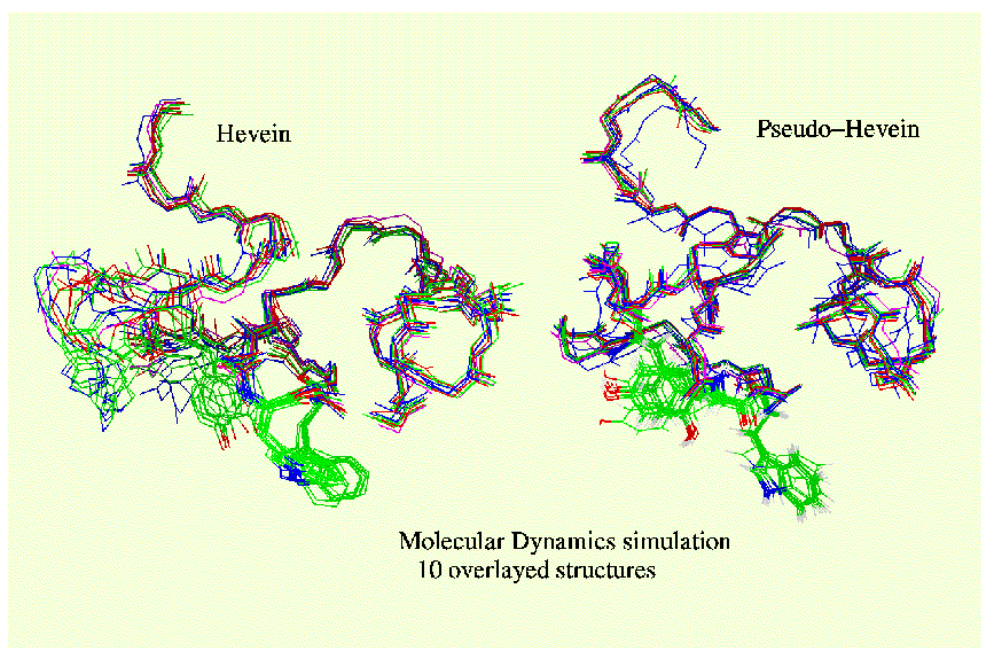
The lectins where no X-ray or NMR-derived structures were available were modelled using standard homology modelling techniques (ProMod). All lectins are relaxed by MD-

simulations in aqueous solution using the CVFF force field. The surface accessibility of the aromatic amino acids is calculated for a certain number of minimised structures taken from MD-simulations. The comparisons of the calculated surface accessibility with the CIDNP signals showed a good agreement.

CIDNP has to be considered as an excellent tool for protein-carbohydrate interaction studies in solution supplementing the commonly used panel of X-ray-crystallography, high-resolution multidimensional NMR and molecular modeling.

### Computational details

The starting structures of the lectins were either taken from the Brookhaven Protein Data Bank or generated using knowledge-based homology modeling technique. The generation of the H-atoms, automatic assignment of partial charges for each atom of the molecule were accomplished using the INSIGHTII software. The center of mass of the lectin was placed in the center of a 30x30x30 Å large water box. A simple point charge model for the water was used allowing the atoms of the water molecules to vibrate. Periodic boundary conditions were applied using a double cutoff of 10 and 12 Å. Each molecular ensemble was submitted to a molecular dynamics simulation using the CVFF force field at a temperature of 300 K, with an equilibration time of 20 ps, a production period of 100 ps, an integration step of 1 fs and a dielectric constant of 1. Each 250 steps of integration a conformation was stored. From the 400 individual conformations of the production time the ten with the lowest potential energy were automatically selected and submitted to a energy minimization of the complete ensemble using the conjugate gradient method. Delineation of possible conforma-



**Figure 2:** 10 overlaid conformation taken from a molecular dynamics simulation (see computational details) of hevein and pseudo hevein. All selected structures were minimized.

tions for domB was achieved using different approaches. I) singling out the domB from the X-ray structure of the complete lectin (PDB-entry 7WGA). II) knowledge-based homology modeling techniques using the ProMod tools [4]

a) based on 1HEV as 3D template and

b) based on 1WGC,2WGC,1WGT and 1HEV as 3D template.

The MD calculation were performed on an IBM-SP2 parallel machine using four processors. The handling of the molecules and the analysis of the MD results were accomplished with various modules of the INSIGHTII software on an Silicon Graphics Indigo<sup>2</sup> workstation.

## References

1. Sharon and Lis, *Trends Biochem. Sci.* **1993**, *6*, 221-226.
2. Gabius and Gabius (eds.) *Glycosciences: Status and Perspectives*, Chapman&Hall, Weinheim, 1996.
3. Kaptein et.al. *Nature* **1978**, *274*, 293-294.
4. Peitch, M.C. *Bio/Technology* **1995**, *13*, 658 - 660.