BIOINFORMATICS APPLICATIONS NOTE Vol. 15 no. 9 1999
Pages 767–768

SWEET — WWW-based rapid 3D construction of oligo- and polysaccharides

A. Bohne[1], E. Lang[2] and C. W. von der Lieth[1]

[1]Deutsches Krebsforschungszentrum, Zentrale Spektroskopie, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany and [2]Universität Hildesheim, Institut für angewandte Sprachwissenschaft, Marienburger Platz 22, D-31141 Hildesheim, Germany

Received on January 22, 1999; revised on April 30, 1999; accepted on May 4, 1999

Abstract

Summary: *SWEET is a WWW-based tool which rapidly converts the commonly used carbohydrate sequence information directly into a preliminary but reliable 3D model which can be visualised and written to files in several ways.*
Availability: *SWEET is accessible via the Internet at http://www.dkfz-heidelberg.de/spec/.*
Contact: *a.bohne@dkfz-heidelberg.de or w.vonderlieth@dkfz-heidelberg.de*
Supplementary information*: The current version of SWEET generates only one conformation out of a manifold. Several authors have analysed possible conformations of high-mannose N-linked glycans using a combination of NMR methods and computational approaches showing that such molecules are rather flexible populating normally several conformations for each glycosidic linkage. The displayed model exhibits for all glycosidic linkages a conformation which is in accordance with the reported variations of Φ, ψ and ω values for specific linkage (see http://www.dkfz-heidelberg.de/spec/sweet2/doc/input/sba_example.html).*

Discovery notes

The most common form of protein or lipid modification is glycosylation (Dwek, 1996; Gabius and Gabius, 1997). Mono-, oligo- or polysaccharides (collectively termed glycans) with highly diverse structures are the most versatile co- or post-translational modifications of peptides and proteins (Sharon, 1998). Glycolipids are a structurally heterogeneous group of membrane components found in all species ranging from bacteria to man. The biological functions of glycans include structural effects such as the stabilisation of certain conformations, modulation of the functional activity of a protein, and/or presentation of ligands for specific binding interactions which mediate cell–cell interactions, for example. The diversity of carbohydrate structures found in glycoconjugates is enormous. This is due to the fact that in glycans, in contrast to proteins and nucleic acids, the individual residues can be linked by many different linkage types. Thus, not only linear polymers but also highly branched structures can be formed. It has been proposed that these factors contribute to the exquisite potential of oligosaccharides to establish a complex coding system for biological information (Laine, 1997).

In general, it is difficult to crystallise complex carbohydrates; therefore, only a few X-ray crystal structures are accessible for the many oligosaccharides of biological interest. Various NMR techniques in combination with computational approaches are often needed to elucidate in detail the conformational properties of oligosaccharides in solution (Siebert *et al*., 1997; von der Lieth *et al*., 1997).

Unlike the other classes of biopolymers, oligosaccharides cannot yet be characterised in terms of their secondary structural or three-dimensional motifs. Consequently no knowledge-based approach such as the analysis of homologous sequences can be applied. Compared to other biological macromolecules, glycans with biological interest have relatively short sequences, normally ranging from two to fifteen residues. Therefore, systematic search approaches for exploring the conformational space of glycans via rotations about the glycosidic bonds (which constitute the most flexible parts of such molecules) as well as rule-based approaches can be applied to generate 3D structures for glycans.

The purpose of SWEET is to provide a software tool which rapidly converts the commonly used sequence information of a complex carbohydrate directly into a preliminary but reliable 3D model. The combination of several well-known approaches fulfils this requirement in an efficient way. First, the oligomer is constructed using a library of monosaccharides which are linked according to the linkage information provided in the sequence definition. This procedure is similar to the POLYS approach (Engelsen *et al*., 1996). Second, a simple, fast and well-established procedure (Imberty *et al*., 1990, 1991; von der Lieth *et al*., 1997) for exploring the conformational space of each glycosidic linkage is used to generate pre-optimised conformations (see table in Bohne *et al*., 1998) to judge the results which can be achieved by this procedure for disaccharides). Third, these preliminary conformations are optimised using a complete molecular

mechanics force field. The MM3 force field (Allinger *et al*., 1989) has been proven to be well suited for complex carbohydrates, and the software package TINKER (Kong and Ponder, 1997) has been interfaced to SWEET for carrying out this optimisation.

The user interface of W3-SWEET provides an input spreadsheet consisting of a grid of symbols for the monosaccharide residues and cell defining the linkage type, or a facility to enter the sequence in alphanumeric form or to paste the sequence information from another source. The nomenclature of the monosaccharide residues and their linkages follows commonly used definitions for carbohydrates. To avoid ambiguities all structural specifications such as d or l (isomers) and p or f (pyranose or furanose rings), which are often omitted in the literature, must be explicitly given in the sequence. Thus, the complete saccharide sequence information can be easily created by anyone who is familiar with the carbohydrate nomenclature.

Several methods for visualising and downloading the generated structures and related information are implemented. The coordinates of the resulting optimised macromolecule structure are stored in PDB format and handled using a wide variety of commercial molecular modelling programs. An ASCII file containing a protocol of the session and simple representations of the Φ and ψ conformational maps used to find the conformation with minimal energy are provided. Errors and failures for each step of the procedure are also reported here. The generated oligosaccharide can be visualised in W3-SWEET with a molecular display program such as RASMOL/CHIME (Sayle and Milner-White, 1994) or WEBMOL (Walther, 1997). WEBMOL also allows one to measure directly internal coordinates such as distances, bond angles and dihedral angles of the generated structure.

Currently, SWEET is able to deal with all sorts of linear and branched oligo- and polysaccharides. Further developments of SWEET will also allow the construction of macrocyclic structures (cyclodexdrins), glycolipids, lipopolysaccharides as well as glycosyl-phosphatidylinositol membrane anchors. Since carbohydrates are rather flexible molecules which normally populate more than one conformation at room temperature (Woods, 1995; Kozár and von der Lieth, 1997), we will include efficient algorithms which enable quick access to all major conformations.

References

Allinger,N.L., Yuh,Y.H. and Lii,J.H. (1989) Molecular mechanics. I. The MM3 force field for hydrocarbons. *J. Am. Chem. Soc*., **111,** 8551–8566.

Bohne,A., Lang,E. and von der Lieth,C.-W. (1998) W3-SWEET: carbohydrate modeling by Internet. *J. Mol. Model*., **4,** 33–43.

Dwek,R.A. (1996) Glycobiology: toward understanding the function of sugars. *Chem. Rev.*, **96,** 683–720.

Engelsen,S.B., Cros,S., Mackie,W. and Pérez,S. (1996) A molecular builder for carbohydrates: application to polysaccharides and complex carbohydrates. *Biopolymers*, **39,** 417–433.

Gabius,H.-J. and Gabius,S. (1997) *Glycosciences – Status and Perspectives*. Chapman & Hall, Weinheim.

Imberty,A., Gerber,S., Tran,V. and Pérez,S. (1990) Data bank of three-dimensional structures of disaccharides. A tool to build 3 D structures of oligosaccharides. Part I. Oligo-mannose type N-glycans. *Glycoconjugate J*., **7,** 37–54.

Imberty,A., Delage,M.M., Bourne,Y., Cambillau,C. and Pérez,S. (1991) Data bank of three-dimensional structures of disaccharides: Part II, N-acetyllactosaminic type N-glycans. Comparison with the crystal structure of a biantennary octasaccharide. *Glycoconjugate J*., **8,** 456–483.

Kong,Y. and Ponder,J.W. (1997) Calculation of the reaction field due to off-center point multipols. *J. Chem. Phys*., **107**, 481–492.

Kozár,T. and von der Lieth,C.-W. (1997) Efficient modeling protocol for oligosaccharides: from vacuum to solvent. *Glycoconjugate J*., **14**, 925–933.

Laine,R.A. (1997) The information-storing potential of the sugar code. In Gabius,H.-J. and Gabius,S. (eds), *Glycosciences – Status and Perspectives*. Chapman & Hall, Weinheim, pp. 1–14.

Sayle,R.A. and Milner-White,E.J. (1994) RASMOL: biomolecular graphics for all. *Trends Biochem. Sci*., **20**, 374–376.

Sharon,N. (1998) Glycoproteins now and then: a personal account. *Acta Anat*., **161,** 7–17.

Siebert,H.-C., von der Lieth,C.-W., Gilleron,M., Reuter,G., Wittmann,J., Vliegenthart,J.F.G. and Gabius,H.-J. (1997) Carbohydrate-protein interactions. In Gabius,H.-J. and Gabius,S. (eds), *Glycosciences – Status and Perspectives.* Chapman & Hall, Weinheim, pp. 291–310.

von der Lieth,C.-W., Kozár,T. and Hull,W.E. (1997) A (critical) survey of modeling protocols used to explore the conformational space of oligosaccharides. *J. Mol. Struct. (Theochem.)*, **395–396,** 225–244.

Walther,D. (1997) WebMol – a Java-based PDB viewer. *Trends Biochem. Sci*., **22,** 274–275.

Woods,R.J. (1995) Three-dimensional structures of oligosaccharides. *Curr. Opin. Struct. Biol*., **5,** 591–598.