

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Molecular Graphics and Modelling 25 (2006) 380–388

*Journal of
Molecular
Graphics and
Modelling*

www.elsevier.com/locate/JMGM

Techniques for visualization of carbohydrate molecules

Michelle Kuttel^{*}, James Gain, Anton Burger, Ian Eborn*Collaborative Visual Computing Laboratory, Computer Science Department, University of Cape Town,
Private Bag X3, Rondebosch 7701, South Africa*

Received 17 October 2005; received in revised form 21 February 2006; accepted 21 February 2006

Abstract

Standard molecular visualizations, such as the classic ball-and-stick model, are not suitable for large, complex molecules because the overall molecular structure is obscured by the atomic detail. For proteins, the more abstract ribbon and cartoon representations are instead used to reveal large scale molecular conformation and connectivity. However, there is currently no accepted convention for simplifying oligo- and polysaccharide structures.

We introduce two novel visualization algorithms for carbohydrates, incorporated into a visualization package, *CarboHydra*. Both algorithms highlight the sugar rings and backbone conformation of the carbohydrate chain, ignoring ring substituents. The first algorithm, termed *PaperChain*, emphasizes the type and conformation of the carbohydrate rings. The second, *Twister*, emphasizes the relative orientation of the rings. We further include two rendering enhancements to augment these visualizations: silhouettes edges and a translucent overlay of the ball-and-stick atomic representation.

To demonstrate their utility, the algorithms and visualization enhancements are here applied to a variety of carbohydrate molecules. User evaluations indicate that they present a more useful view of carbohydrate structure than the standard ball-and-stick representation. The algorithms were found to be complementary, with *PaperChain* particularly effective for smaller carbohydrates and *Twister* useful at larger scales for highlighting the backbone twist of polysaccharides.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Carbohydrate; Polysaccharide; Visualization; Silhouette; Transparency

1. Introduction

The principle purpose of molecular visualization is to emphasize the important structural features of a molecule, from which chemical properties and function can be inferred. Existing molecular visualization algorithms have their origin in physical constructions and can be broadly categorised into “ball-and-stick” and “space-filling” representations. The classic ball-and-stick model represents atoms as colour-coded spheres and bonds as cylinders to emphasise molecular connectivity, while “space-filling” representations, such as a CPK (Corey–Pauling–Koltun) surface [1], highlight the overall shape and contours of a molecule. However, both types of visualization suffer from overcrowding, in that the overall molecular structure is swamped by the atomic detail. This is particularly true for larger molecules.

More abstract representations that dispense with explicit representation of atoms have been developed to highlight the overall conformation and connectivity of large protein molecules. Current abstract visualizations focus on proteins [2,3]. Ribbon diagrams, invented by Richardson [4] and computerized by Carson and Bugg [5,6], display a narrow curved surface fitted to the backbone of the protein molecule. Three-dimensional cartoon diagrams of proteins [4] use cylinders to represent the α -helices and flat arrows to represent the β -strands. Connecting regions are shown as a thin tube tracing the backbone of the protein chain. Both of these methods have proved highly effective and popular for protein visualization.

Molecular structures of many other classes of molecule, such as carbohydrates or lipids, are still typically visualized using a ball-and-stick representation. In the case of carbohydrates, both the conformation of the glycosidic linkages between the rings and the pucker of the characteristic hemiacetal rings are of interest. The conformation of the glycosidic linkages between the rings determines the overall

^{*} Corresponding author. Tel.: +27 21 650 5107; fax: +27 21 689 9465.

E-mail addresses: mkuttel@cs.uct.ac.za (M. Kuttel),
jgain@cs.uct.ac.za (J. Gain).

conformation of the carbohydrate backbone. In addition, each carbohydrate ring has a variety of possible pucker conformations which are difficult to distinguish in conventional visualizations. For example, the six-atom glucopyranose ring has six distinct boat forms, six twist-boat forms and two chair forms. In general, the two chair conformations of the glucopyranose ring (4C_1 and 1C_4) are favoured, because the boat and skew forms cause crowding of the atoms in the ring and are thus significantly higher in energy. In particular, the 4C_1 conformation of α -D-glucose, with all ring substituents in the equatorial position, is preferred over the 1C_4 conformation in which all substituents are in the axial orientation.

In this paper, we present two algorithms that overcome current shortcomings in carbohydrate visualization by conveying the location, orientation and conformation of ring structures. The *PaperChain* visualization depicts the conformation of individual rings by fitting a polyhedron through the participating atoms and the ring centroid. The polyhedron is then coloured according to ring type and pucker. This algorithm thus emphasises information useful in understanding smaller molecule chains or subsections of larger chains. In contrast, the *Twister* visualization highlights the relative orientation of rings. Each ring is depicted as a disc, and linked to the next by a curved ribbon with rectangular cross-section. *Twister* thus serves the same purpose for carbohydrates as ribbon diagrams do for proteins, conveying the global structure of larger molecular chains. In addition to the above algorithms, we investigate the utility of two non-standard rendering enhancements applied to the novel visualizations. These are the use of silhouette edges and a ghosted ball-and-stick overlay. Finally, the efficacy of the two algorithms and rendering enhancements for carbohydrate visualization is demonstrated by application of the techniques to a variety of carbohydrate molecules as well as evaluation of the *CarboHydra* package [7] by a panel of research chemists.

The remainder of this paper is structured as follows. Section 2 covers the *PaperChain* algorithm, the *Twister* algorithm and rendering enhancements in detail. Section 3 presents illustrations of the algorithms and visualization enhancements applied to a sample set of molecules and results of the algorithm evaluation. The paper then concludes with a summary and recommendations for future work (Section 4).

2. Methods

Conventionally, a visualization is divided into two stages: filtering (in which meaningful information is extracted from data) and mapping (in which the information is visually formatted) [8]. Here the filtering phase involves identification of the carbohydrate rings and the linkages between them. Side-chains are ignored. In the mapping phase, the ring structures are presented with either the *Twister* or the *PaperChain* algorithm. The mapping can be further enhanced by incorporating the available rendering options.

The two algorithms, *PaperChain* and *Twister*, and additional rendering improvements have been implemented using the C++ programming language, the OpenGL graphics library and the

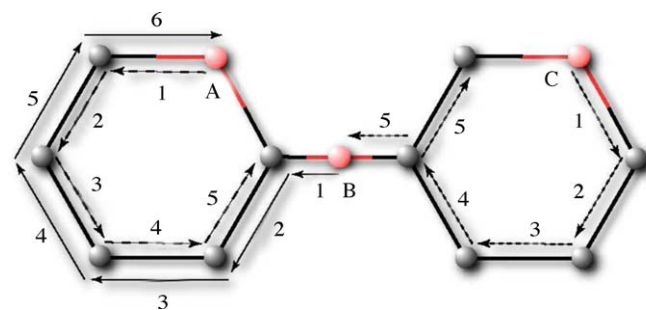


Fig. 1. Breadth-first search for three potential ring structures. Searches originate from one oxygen-bonded carbon, proceed in one direction around potential rings, and terminate successfully on encountering the second oxygen-bonded carbon. The A, B and C searches are marked with dashed, full and dotted arrows, respectively, with each numbered according to search depth.

FOX Windowing Toolkit [22] in *CarboHydra*, an Open Source customised molecular visualization package for the Windows environment.¹

2.1. Identifying rings and linkages

To begin the filtering phase, a molecular structure file in the Protein Data Bank (PDB) file format [9] is loaded. The atomic connectivity of the carbohydrate molecule, if not already provided, is inferred by creating bonds between atoms whose separation is less than the sum of their van der Waals radii. Atomic rings and inter-ring linkages are then identified by a breadth-first search strategy. This guarantees that the first ring or linkage found is the smallest ring. As the carbohydrate rings are restricted to hemiacetyl rings, the list of atoms is traversed and a ring search is seeded whenever an oxygen with two bonded carbons is encountered. These searches are carried out breadth-first with one carbon as the source atom and the other as the destination. Fig. 1 illustrates this procedure for three potential rings (side-chains are ignored in the interests of clarity). In this illustration, identification of ring A is a straightforward search without any branching. The search for ring C first explores the oxygen in the linkage between rings (on step 5) before the terminating carbon is found. The search for ring B correctly fails to find a ring on the sixth step, since its originating oxygen is part of a linkage. Once a cycle has been identified, its list of ring atoms is recorded and reversed, if necessary, to make traversal consistent with the IUPAC definition of ring orientation [10]. This numbering allows for the definition of a “top” and “bottom” for each ring and hence a consistent way of defining a ring’s orientation.

While a breadth-first search guarantees that the first cycle found is also the shortest, it tends to explore an expanded search space in achieving this. Fortunately, in the case of carbohydrates, there are several pruning factors: rings are typically six atoms or less, carbon atoms generate at most three outgoing search branches, and non-carbon atoms act as branch terminators (since apart from the starting oxygen, rings contain only carbon atoms).

¹ Available from the authors on request.

Though the work here is focussed on representations of carbohydrates, a more general definition of a ring as containing any cycle of atoms can be useful for creating novel visualizations of molecules, such as DNA, which contain other cyclic structures. Therefore, an option to define rings in this general manner was retained in the ring identification algorithm implemented in *CarboHydra*.

After all rings have been found, it is necessary to identify the atoms comprising linkages between rings. To do this, a highly constrained breadth-first search is spawned at each ring atom. A given search is forbidden from exploring other atoms in the originating ring and completes successfully once an atom from another ring is encountered. Ring pairs that contain common atoms are ignored.

2.2. PaperChain algorithm

The *PaperChain* algorithm focuses on portraying the conformations of individual hemiacetal rings. This is achieved by fitting a non-planar polyhedron through the ring atoms. Links between rings do not follow the atomic bonds, but are instead abstracted simply as narrow cylinders directly from one ring to the next.

To form the ring, adjacent ring atoms are connected to form an N -sided polygon. This polygon is invariably non-planar, and so requires triangulation before rendering. In order to do this, a central vertex is required to connect the atoms. The *PaperChain* algorithm uses the collective centroid of the atoms for the central vertex.

For each ring identified by the filtering phase, the algorithm proceeds as follows:

- (1) *Calculate the ring centroid and normal.* The averaging plane of the ring is encoded by a point (the centroid) and vector (Newell's normal). Newell's method [11] provides the components of an averaging normal:

$$n_x = \sum_{i=0}^{N-1} (y_i - y_{\text{succ}(i)})(z_i + z_{\text{succ}(i)}) \quad (1)$$

$$n_y = \sum_{i=0}^{N-1} (z_i - z_{\text{succ}(i)})(x_i + x_{\text{succ}(i)}) \quad (2)$$

$$n_z = \sum_{i=0}^{N-1} (x_i - x_{\text{succ}(i)})(y_i + y_{\text{succ}(i)}) \quad (3)$$

The function `succ()` supplies the index of the next vertex in the ring, wrapping as required. Thus, `succ(i) = i + 1`; $i = 0, \dots, N - 2$ and `succ(N - 1) = 0`. It is based on the observation that the components (n_x, n_y, n_z) of a polygon's normal vector are proportional to the area of the polygon's orthographic projection onto the YZ, XZ and XY planes, respectively. This still provides a well-behaved approximation when the polygon is non-planar.

- (2) *Create the ring polyhedron.* To begin, the centroid is moved slightly in the direction of its normal vector. The ring's upper surface is a triangle fan created by joining this offset centroid to the centers of the ring atoms. The ring's

underside is similar, except that the centroid is offset in the opposite direction. A platter-like polyhedron, which faithfully preserves the original atom locations, is then formed by fusing the upper- and undersides.

- (3) *Mark the ring oxygen atom.* The location of the oxygen atom in the carbohydrate ring structure is of particular significance as it indicates the orientation of the ring. So, while other atoms (from the ball-and-stick model) are removed, the red oxygen is always retained.
- (4) *Colour the ring polyhedron according to ring type and pucker.* The ring identified polyhedra are then automatically coloured according to the number of atoms in the ring and the type and degree of pucker of the ring. Carbohydrates contain predominately furanose (five atoms) or pyranose (six atoms) rings, the colouring scheme is focussed on these rings. For each N -membered ring, $N - 3$ ring puckering parameters are calculated according to the method detailed by Cremer and Pople [12]. The puckering of pentose rings is described by two puckering parameters: a phase angle, ϕ and an amplitude, q . For six-membered rings, Cremer and Pople defined a convenient coordinate system (θ, ϕ, Q) which maps all types of puckering onto the surface of a sphere. Here the radial coordinate, Q , is the total puckering amplitude (the degree of puckering) and the angles θ, ϕ the type of puckering. The polar positions correspond to chair conformations, and the equatorial positions to various twist-boat forms. Boat conformations occur along the equator at $0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ$ and 300° . Twist-boat conformations occur at $30^\circ, 90^\circ, 150^\circ, 210^\circ, 270^\circ$ and 330° . The colouring scheme aims to clearly differentiate boat from chair conformations and twist-boat from boat conformations. Thus, the colour of a pentose ring is defined according to a red, green, blue (RGB) colour scheme as

$$R = |\sin(\theta)| \cdot Q \quad (4)$$

$$G = |\cos(\theta)| \cdot Q \quad (5)$$

$$B = |\sin(3\phi) \sin(\theta)| \cdot Q \quad (6)$$

where R is the red component, G the green component and B is the blue component. This effectively maps all points in the sphere onto a single quadrant. According to this scheme, typical chair conformations are green, the higher energy boat conformations are red and twist-boat conformations purple (the blue component is only appreciable for the equatorial regions). In addition, the more extreme the pucker, the more intense is the colour. For contrast and easy identification, the colour of pentose rings is defined as blue, with intensity determined by the degree of pucker, i.e.,

$$R = 0 \quad (7)$$

$$G = 0 \quad (8)$$

$$B = Q \quad (9)$$

Thus, in our scheme, truly planar rings are black, which conforms to a typical colouring of carbon. Non-pentose or hexose rings are assigned a default colour, which is set by the user.

Finally, for each linkage identified between rings, a narrow connecting cylinder is placed, starting and ending at the linkage atoms on the connected rings.

2.3. Twister algorithm

The *Twister* algorithm highlights the relative orientation of adjacent carbohydrate rings in a polysaccharide. Each ring is depicted as a disc, and linked to the next by a curved ribbon with rectangular cross-section. These ribbons twist along their length to align with the discs at either end. Relative orientation of the rings in a carbohydrate is further clarified by colouring top and bottom surfaces differently. Because of its reliance on a definition of a “top” and “bottom” for each ring, the *Twister* algorithm is only applied to carbohydrate rings.

For each ring identified in the filtering phase, the *Twister* algorithm performs the following steps:

- (1) *Calculate ring centroid and normal.* As with the *PaperChain* algorithm, calculating the ring centroid (c) and Newell normal (n) is a necessary precursor to constructing a ring representation.
- (2) *Create a disc.* Each ring is represented by a disc (effectively a flattened cylinder) centered at c and with major axis oriented along n . The disc radius (r) is derived by projecting ring atoms (a_i) onto the disc plane and finding that which is furthest from the centroid.

For each linkage identified between rings, the *Twister* algorithm performs the following steps:

- (1) *Fit a Hermite curve.* The backbone of each linking ribbon is a Hermite curve ($Q(t)$) [13] whose coefficients are derived from the position (P_0, P_1) and tangents (T_0, T_1) at the curve endpoints. In this case, the positions are the starting and terminating atoms of the link and the tangents are aligned with the vector from the corresponding endpoint to its ring centroid, reversed as necessary. The magnitude of each tangent is calculated using the chord-length parametrisation of Rogers and Adams [14], which ensures that equal steps of the spline parameter t will move approximately equal distances along the curve.
- (2) *Create a ribbon.* A ribbon in this context is a generalized cylinder with a rectangular cross-section that follows the underlying Hermite curve. The cross-sections are positioned and oriented to match local reference frames uniformly sampled along the curve. The challenge is to derive frames that naturally follow the underlying twists and turns of the curve. The Frenet frame admits analytic computation but behaves poorly for degenerate curvature, such as at inflection points and along straight sections. Instead we employ Bloomenthal’s rotation minimizing frame [15], which propagates a starting frame along the curve using local incremental adjustments. Unfortunately, this means that the final frame does not necessarily align with the destination disc, so in a backpropagation step

frames are additionally rotated about the tangent in proportion to their distance along the curve.

The *Twister* algorithm is flexible: by removing the discs and starting the ribbons at ring centroids, a plain ribbon diagram of a carbohydrate can be generated. This visualization is similar to the conventional ribbon representation for proteins.

2.4. Rendering enhancements

There are a variety of rendering techniques, suggested by psychology and visualization research, that can serve as cognitive aids to understanding when applied to the visualizations. Here we consider two: silhouettes and a transparent overlay of the ball-and-stick model.

2.4.1. Silhouettes

The efficacy of feature lines, such as contours and silhouettes, in understanding complex shapes is well recognized [16]. Research in Cognitive Psychology suggests that silhouettes are a particularly important determinant of shape. We exploit this by emphasizing object silhouettes with darker edges. Essentially, there are three choices in implementing silhouette rendering [17]: object space, where calculations are performed directly on the 3D model, typically producing an analytic description of silhouette edges; image space, where calculations use image buffer projections of the model, often by exploiting graphics hardware; or a hybrid, potentially combining the best features of both. In the interests of a simple, robust, real-time solution we choose a modification of Raskar and Cohen’s [18] image precision method. Back-facing polygons are rendered in wireframe as thick black lines. These are occluded in the z -buffer by front-facing polygons except around the silhouette at the join between front and back faces where they are only half covered. An outline-only style reminiscent of pen-and-ink illustrations can be achieved by flat shading front faces in the background colour.

2.4.2. Ball-and-stick overlay

Schneiderman’s visual information-seeking mantra [19]: “overview first, zoom and filter, then details on demand”, describes a hierarchy of visualization tasks with wide applicability. In *CarboHydra* the *Twister* algorithm provides an overview, while the *PaperChain* algorithm can be used when the focus is on substructures. Both algorithms filter extraneous data. However, the side-chain detail should still be available on demand. To achieve this, a ball-and-stick model is overlaid on the principle visualizations with controllable transparency, so that a user can adjust the focus between the levels of primary (backbone), secondary (ring conformation) and tertiary (side-chain) structure. This transparency involves only alpha blending. True transparency, which requires object sorting, was not implemented as the depth complexity in the overlay is not very great and the overlay is required for identification of side-chains and ring orientation. In addition, the option to view the overlay in either colour or greyscale is included in the main menu in *CarboHydra* (Fig. 2).

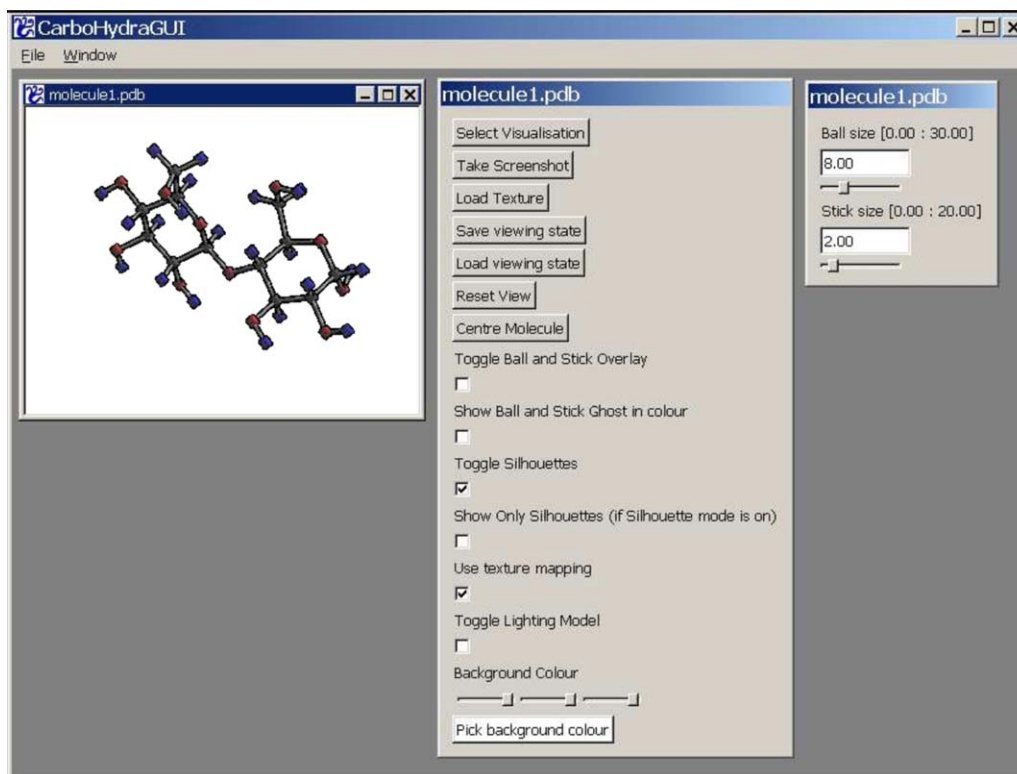


Fig. 2. View of the *CarboHydra* graphical user interface for a Windows platform, implemented using the FOX Windowing Toolkit [22].

2.5. Evaluation

The system was tested by research chemists according to the “expert evaluation” paradigm, in which users conversant with the problem domain are asked to evaluate and provide feedback on a system. The objective was a qualitative estimate of whether further work on carbohydrate visualization is justified. Four subjects were given a demonstration of the system, and then encouraged to interact with it, filling in a questionnaire as they did so. The questionnaire asked them to rate the visualization algorithms and the rendering options for each of a number of sample molecules.

3. Results and discussion

The *CarboHydra* package (Fig. 2) was positively received by the evaluation panel: all subjects favoured the novel visualizations over the ball-and-stick model. Example visualizations using the two algorithms and the rendering enhancements are shown and discussed below.

3.1. PaperChain visualization

The *PaperChain* algorithm highlights the conformation of the sugar rings in a polysaccharide, with both the fold of the polyhedra and their colour. For example, a 4C_1 chair and a twist-boat form of the α -D-glucopyranose ring are shown in Fig. 3. The conformation of the carbohydrate rings is not immediately apparent in the ball-and-stick visualizations. However, a clear contrast between the low-energy 4C_1 chair and higher energy

twist-boat forms of the α -D-glucopyranose ring can be seen in the *PaperChain* visualizations (Fig. 3(b) and (d)). The chair conformation is coloured green and the twist-boat purple according to their puckering parameters.

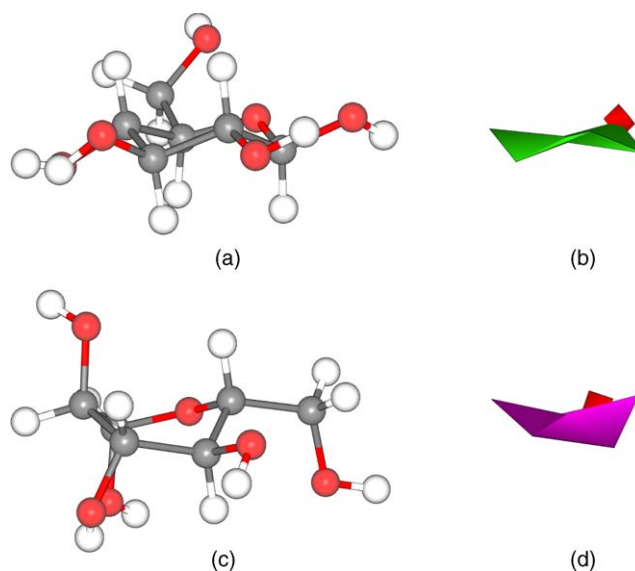


Fig. 3. Comparative visualizations of α -D-glucose in a chair and twist-boat conformation: (a) 4C_1 chair conformation viewed with a standard ball-and-stick algorithm. (b) 4C_1 chair conformation viewed with the *PaperChain* algorithm. Chair conformations are coloured green. (c) Twist-boat conformation viewed with a ball-and-stick algorithm. (d) Twist-boat conformation viewed with the *PaperChain* algorithm. Twist-boat conformations are coloured purple. Orientation of the rings in the *PaperChain* visualizations is indicated by the location of the red oxygen atom on the ring polyhedron.

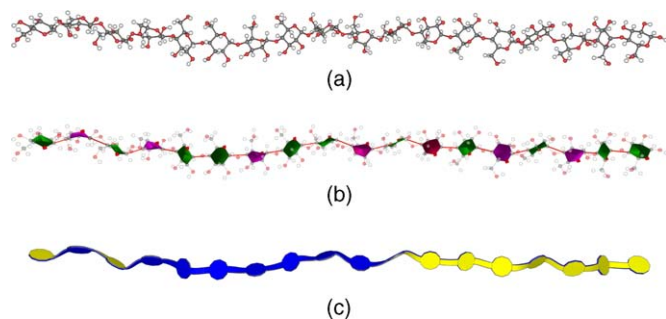


Fig. 4. Visualizations of an 18-unit α -1,4-linked glucan molecule with gluco-pyranose rings in both 4C_1 chair and twist-boat conformations. (a) Ball-and-stick visualization. (b) *PaperChain* visualization using silhouettes and a transparent overlay of the ball-and-stick model. This visualization highlights the ring conformations: chair conformations are coloured green, twist-boat conformations purple. (c) *Twister* visualization highlighting the relative orientation of the rings. The “top” and “bottom” of the glucan rings (defined according to IUPAC convention) are coloured in blue and yellow, respectively.

Fig. 4 shows the conformation of an 18-unit amylose oligosaccharide (α -1,4-D-glucan) that has been subjected to a stretching force. The force has converted some of the rings from their low-energy chair conformation into a higher-energy twist-boat conformation. In the ball-and-stick visualization (Fig. 4(a)), the conformation of the rings is not immediately apparent. However, in Fig. 4(b), the *PaperChain* visualization allows one to identify at a glance sugar rings that are not in their preferred low-energy conformations. Further, Fig. 6(b) shows a *PaperChain* visualization of the same chain in its typical unstrain helical conformation. Here all the rings are in clearly in the chair conformation.

The *PaperChain* algorithm is simply dependent on the identification of atomic cycles and thus can also be used to provide novel visualizations of molecules which contain heterocycles and other cyclic structures. This requires the alternate form of the ring identification algorithm mentioned in Section 2.1. An example of a B-DNA fragment visualized with the *PaperChain* algorithm is shown in Fig. 5(b). The location and orientation of the DNA bases is made apparent in the *PaperChain* visualization. The furanose rings are coloured blue, with intensity adjusted to the puckering amplitude: flatter rings appear blacker.

3.2. Twister visualization

The principal advantage of the *Twister* visualization for polysaccharides is that it highlights the relative orientation of successive carbohydrate rings by the twist of the connecting ribbon. This is further emphasized by using contrasting colours for the “top” and “bottom” of the carbohydrate rings. Comparisons of the *Twister* and *PaperChain* algorithms are shown in Figs. 4 and 6. In these figures, the twist of the carbohydrate backbone is not apparent in the ball-and-stick or *PaperChain* visualizations, but is clearly seen in the *Twister* visualizations. A further example is shown in Fig. 7, which depicts a *Twister* visualization of the crystal structure conformation of the 14-unit cyclic cyclotetradecaamylose [20]. This molecular has two characteristic twists in the chain,

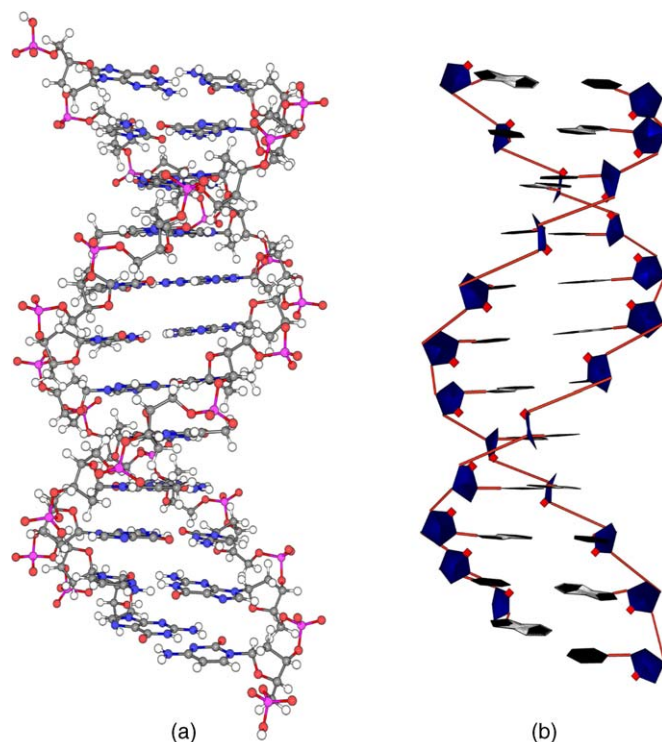


Fig. 5. Two alternate visualizations of a B-DNA molecule. (a) Ball-and-stick visualization; (b) *PaperChain* visualization, using silhouette enhancements. The furanose rings are coloured blue.

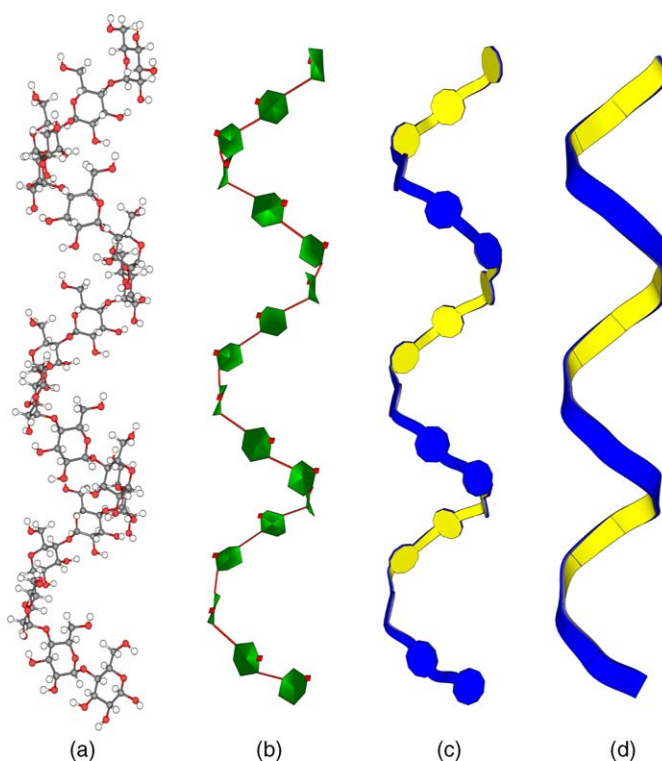


Fig. 6. Visualizations of an 18-unit amylose (α -1,4-D-glucan) oligosaccharide. (a) Ball-and-stick visualization; (b) *PaperChain* visualization; (c) *Twister* visualization; (d) ribbon variant of *Twister* visualization. The segmentation of the ribbon follows the separate carbohydrate rings.

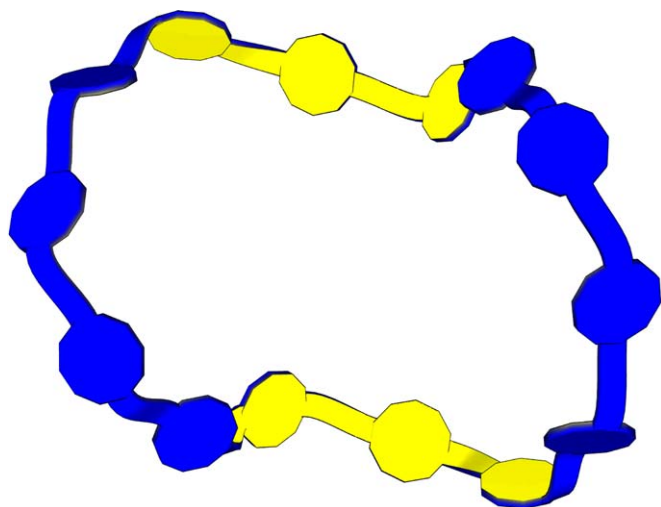


Fig. 7. A *Twister* visualization of cyclotetradecaamylose in its crystal structure conformation [20]. The “top” and “bottom” of the glucan rings (defined according to IUPAC convention) are coloured in blue and yellow, respectively.

termed “band-flips”, which are clearly shown by the *Twister* visualization.

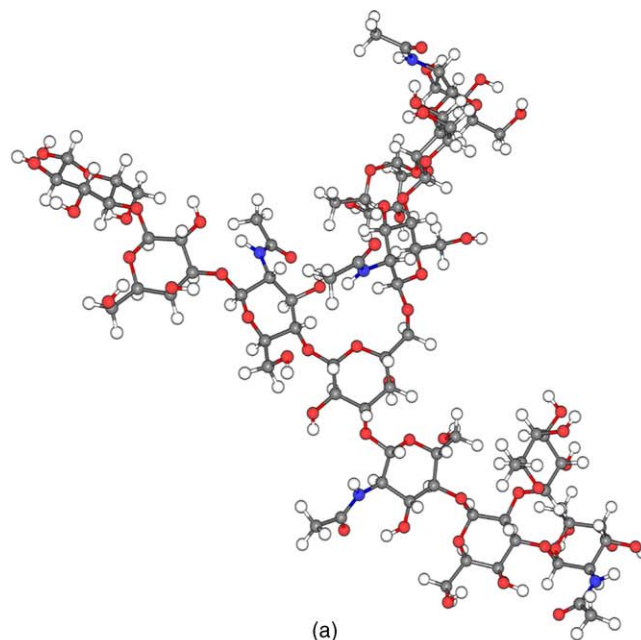
More complex examples are the blood antigen A molecule depicted in Fig. 8 and the N-linked carbohydrate in Fig. 10 (both were built using the built using the SweetII facility [21]). While the relative orientations of the pyranosering in the molecules are not clear in the ball-and-stick visualizations (Figs. 8(a) and 10(a)), they are clearly visible in the *Twister* visualizations (Figs. 8(b) and 10(b)).

As implemented in *CarboHydra*, the *Twister* algorithm provides flexibility in the sizes and widths of the rings and the ribbons. Fig. 9 shows the Options Menu for the *Twister* visualization in the *CarboHydra* package. This allows the user to adjust the size, width and resolution of both the ribbons and the rings. If rings are made invisible and the ribbon endpoint stretched into the ring centroid, a pure ribbon diagram can be created. An example of this is shown in Fig. 6(d). Evaluation by the panel indicates that this representation is particularly appealing to chemists familiar with protein representations.

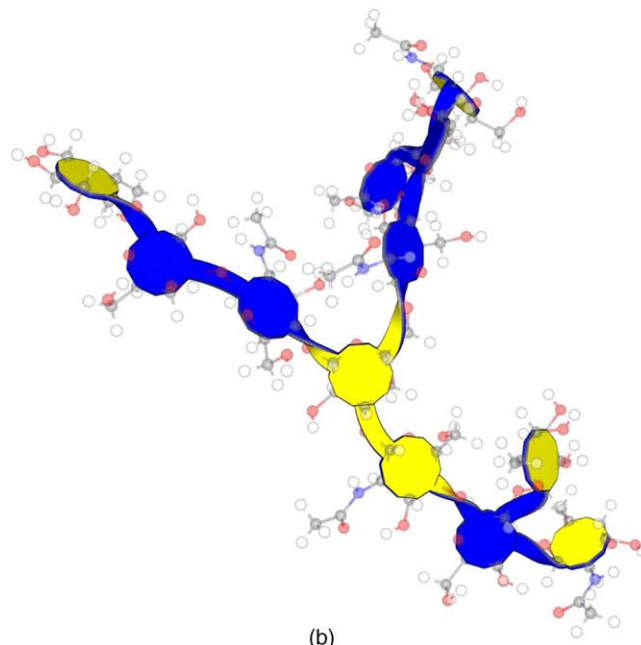
Comments from the evaluation panel show that *Twister* is particularly suited to visualizing the tertiary structure of large carbohydrate molecules. However, one subject indicated that the *Twister* algorithm provided insufficient information about the carbohydrate rings—specifically, that it was impossible to identify the bond points of the linkages between rings.

3.3. Rendering enhancements

The evaluation panel’s comments on the usefulness of silhouette edges corroborates previous research on the subject [16]: silhouettes are effective at highlighting molecular contour and shape. However, one member of the evaluation panel thought that rendering silhouette edges alone in *Twister* was insufficient to depict the twist in the ribbons, suggesting that some internal edges should also be rendered. The silhouette-only depictions are clear black-and-white representations of the



(a)



(b)

Fig. 8. Example structure of a blood antigen A carbohydrate molecule generated by the SweetII modelling package [21]. Shown are (a) ball-and-stick and (b) *Twister* visualizations. The *Twister* visualization incorporates transparent overlay of the ball-and-stick visualization and silhouette enhancements. The “top” and “bottom” of the glucan rings (defined according to IUPAC convention) are coloured in blue and yellow, respectively.

carbohydrate. This rendering option was deemed useful for producing images for publication, as it is similar to many sketched schematic drawings of molecules used in text books. This abstraction of the shape of the molecule is particularly useful when applied to large molecules, an example of which is shown in Fig. 10(b).

The system of transparent overlay of the atomic ball-and-stick structure was found to be a useful tool for providing context when viewing a molecule using either of the novel algorithms (Figs.

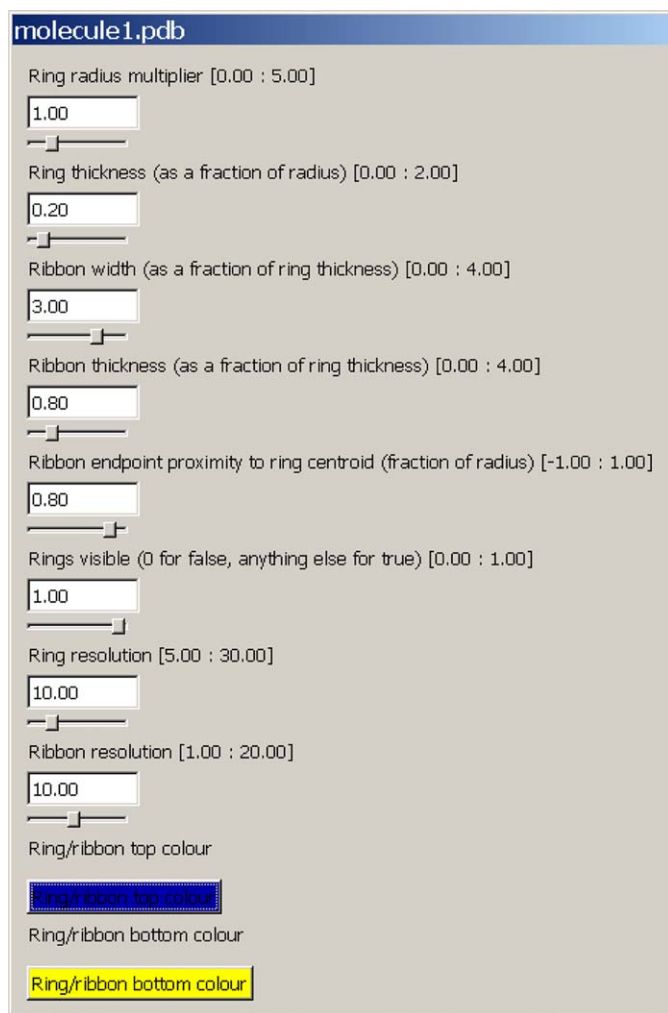


Fig. 9. *Twister* options menu in the *CarboHydra* package. The user can adjust both the ring radius and thickness and the ribbon radius and thickness. In addition, the ring endpoint proximity to the ring centroid can be altered. The rings can be rendered invisible and the resolution of both the rings and the ribbons adjusted.

4(b), 8(b) and 10(a)). In particular, combination of the *PaperChain* algorithm with the ball-and-stick overlay emphasizes ring location and conformation, while retaining atomic detail in the side-chains.

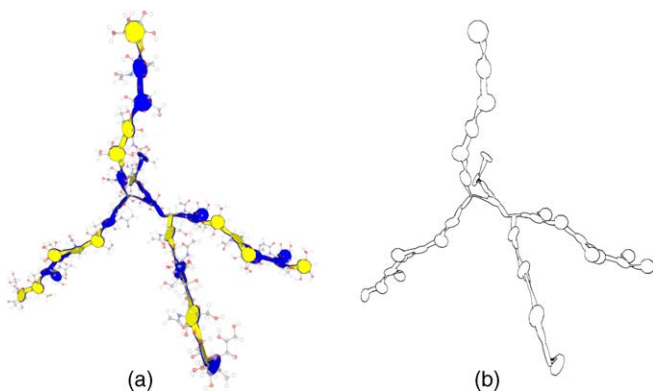


Fig. 10. *Twister* visualizations of a complex N-linked carbohydrate built using the *SweetII* facility [21]: (a) standard *Twister* visualization with ball-and-stick overlay; (b) silhouette-only *Twister* visualization.

4. Conclusions

The *PaperChain* and *Twister* algorithms presented here produce novel visualizations of carbohydrates that highlight aspects of molecular structure obscured by traditional ball-and-stick visualizations. The two schemes are complementary, providing solutions at fine- and coarse viewing scales respectively. The *PaperChain* algorithm emphasizes secondary structure: the connectivity and conformation of individual carbohydrate rings. The *Twister* algorithm's ability to show clearly the relative orientation of connected rings makes it a useful tool for depicting tertiary structure.

The addition of silhouettes and a ghosted ball-and-stick overlay were found to enhance the *PaperChain* and *Twister* visualizations. Emphasizing object silhouettes with darker edges highlights the visualization geometry and can be used to create simple black-and-white depictions of carbohydrates for schematic diagrams. The use of a regular checkerboard texture assists in perception of twist in the ribbons generated by the *Twister* algorithm. An optional ghosted overlay of the atomic ball-and-stick representation compensates for lack of contextual detail in the two new visualizations.

The *PaperChain* algorithm can also be used with suitable visualization enhancements for depicting non-carbohydrate rings and thus provide novel visualizations of large complex molecules such as DNA and lignin.

Having established the merit of the novel algorithms, future work will encompass rigorous analysis of the relative strengths and weaknesses of the search techniques in order to optimize execution times and reduce space requirements. In addition, improvements in the generation of the geometry should accelerate rendering. Finally, it is anticipated that the incorporation of these algorithms as a module within a standard molecular graphics package will provide novel and useful visualizations of glycoproteins.

References

- [1] R.B. Corey, L. Pauling, Molecular models of amino acids, peptides and proteins, *Rev. Sci. Instrum.* 24 (8) (1953) 621–627.
- [2] W. Humphrey, A. Dalke, K. Schulten, Visual molecular dynamics, *J. Mol. Graphics* 14 (1) (1996) 33–38.
- [3] J.A. Lopera, J.N. Sturgis, J.-P. Duneau, Ptuba: a tool for the visualization of helix surfaces in proteins, *J. Mol. Graph. Modell.* 23 (4) (2005) 305–315.
- [4] J.S. Richardson, The anatomy and taxonomy of protein structure, *Adv. Protein Chem.* 34 (1981) 167–218.
- [5] M. Carson, C. Bugg, Algorithm for ribbon models of proteins, *J. Mol. Graphics* 4 (2) (1986) 121–122.
- [6] M. Carson, Ribbon models of macromolecules, *J. Mol. Graphics* 5 (2) (1987) 103–106.
- [7] I. Eborn, A. Burger, M. Kuttel, J. Gian, Carbohydrate: rendering carbohydrate cartoons, Tech. Rep. CS04-18-00, Department of Computer Science, University of Cape Town, 2004.
- [8] K. Brodlie, *Animation and Scientific Visualization: Tools and Applications*, Academic Press, 1993.
- [9] F.C. Bernstein, T.F. Koetzle, G.J. Williams, E.E. Meyer, M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, The protein data bank: a computer-based archival file for macromolecular structures, *J. Mol. Biol.* 112 (3) (1977) 535–542.
- [10] A.D. McNaught, Nomenclature of carbohydrates (recommendations 1996), *Adv. Carbohydr. Chem. Biochem.* 52 (1997) 43–177.

- [11] I.E. Sutherland, R.F. Sproull, R.A. Schumaker, A characterization of ten hidden-surface algorithms, *ACM Comput. Surv.* 6 (1) (1974) 1–55.
- [12] D. Cremer, J.A. Pople, A general definition of ring puckering coordinates, *J. Am. Chem. Soc.* 97 (6) (1975) 1354–1358.
- [13] G. Farin, *Curves and Surfaces for Computer-Aided Geometric Design: A Practical Guide*, 4th ed., Academic Press, 1997.
- [14] D.F. Rogers, J.A. Adams, *Mathematical Elements for Computer Graphics*, McGraw-Hill, 1976.
- [15] J. Bloomenthal, Calculation of reference frames along a space curve, in: *Graphics Gems*, Academic Press, 1990, pp. 567–571.
- [16] I. Biederman, G. Ju, Surface versus edge-based determinants of visual recognition, *Cognitive Psychol.* 20 (1) (1988) 38–64.
- [17] T. Isenberg, B. Freudenberg, N. Halper, S. Schlechtweg, T. Strothotte, A developer's guide to silhouette algorithms for polygonal models, *IEEE Comput. Graph. Appl.* 23 (4) (2003) 28–37.
- [18] R. Raskar, M. Cohen, Image precision silhouette edges, in: *SI3D'99, Proceedings of the 1999 Symposium on Interactive 3D Graphics*, ACM Press, 1999, pp. 135–140.
- [19] B. Schneiderman, The eyes have it: a task by data type taxonomy for information visualizations, in: *Proceedings of the 1996 IEEE Conference on Visual Language*, IEEE Computer Society Press, 1996, pp. 336–343.
- [20] J. Jacob, K. Geler, D. Hoffmann, H. Sanbe, K. Koizumi, S.M. Smith, T. Takaha, W. Saenger, Band-flip and kink as novel structural motifs in α -(1 \rightarrow 4)-D-glucose oligosaccharides. Crystal structures of cyclodeca- and cyclotetradecaamylose, *Carbohydr. Res.* 322 (1999) 228–246.
- [21] A. Bohne, E. Lang, C. von der Lieth, Sweet—www-based rapid 3D construction of oligo- and polysaccharides, *Bioinformatics* 15 (1999) 767–768.
- [22] J. van der Zijp, FOX Toolkit, <http://www.fox-toolkit.org/fox.html>, 2005.