2D molecular graphics: a flattened world of chemistry and biology

Peng Zhou and Zhicai Shang

Submitted: 31st December 2008; Received (in revised form): 17th February 2009

Abstract

Molecular graphics provides an intuitive way for representation, modeling and analysis of complex chemical and biological systems. It is now widely used in the theoretical chemistry, structural biology, molecular modeling and drug design communities. Traditional molecular graphics techniques mainly dedicate to showing molecular architectures at three-dimensional (3D) level. However, in some occasions the two-dimensional (2D) representation of molecular configurations, profiles, behaviors and interactions may be more readily acceptable for audiences, especially when we need to describe abstract information in a straightforward way or to present numerous data in schematic diagrams. In recent years, 2D representation methods/tools have been developed rapidly for various purposes, ranging from the aesthetic depiction of atomic arrangement for small organic molecules to schematic layout of complicated nonbonding network across the biomolecular binding interfaces, and have received considerable interest in the fields of chemistry, biology and medicine. In this article we first propose the term of 2D molecular graphics to cover the spectrum of 2D representing chemical and biological systems, we also give a comprehensive review on the methods, tools and applications of 2D molecular graphics.

Keywords: 2D molecular graphics; schematic diagram; chemical and biological system; bioinformatics

INTRODUCTION

Graphic language, symbolic language and written language represent the human communication ways at different abstract levels, in which the graphic language presents information in an intuitive way and, apparently, is the most understandable manner for audiences—‘a picture is worth a thousand words’ [1]. In the fields of chemistry, biology and medicine, molecular graphics is widely used for study and dissemination of molecular structures and functions. It is assumed that microscopic world could be visualized on computer screen using a dozen of metaphors, such as the use of spheres to represent atoms or ribbons to represent protein chains, because we are generating synthetic images of objects that are far smaller than the wavelength of light [2]. Traditionally, the subject of ‘molecular graphics’ is always associated with the stereo representation of molecular architectures in three-dimensional (3D) space by using various models such as CPK, BALL & STICK, and SURFACE, etc. [3]. In recent years, however, such a viewpoint has been changed with the emergence of a number of methods/tools which dedicate to schematically depicting chemical and biological information on two-dimensional (2D) diagrams. These simplified diagrams are aesthetic and straightforward for the exhibition of molecular configurations, profiles, behaviors and interactions. Therefore, this area is considered as a branch of molecular graphics, and here we call it 2D molecular graphics. In fact, 2D and 3D visualization methods are complementary. They are used for different purposes and/or different occasions. For example, a rotatable molecular stereograph provides a faithful representation of available data but requires considerable operator time and practice to thoroughly perceive and is often an inconvenient way to communicate information among peers. A flattened...
image can be designed to be very easily understood even at a glance and can be readily shared as a printed hardcopy.

**OVERVIEW OF 2D MOLECULAR GRAPHICS**

Computer is a powerful tool to implement graphical operations and its development has had a dramatic impact upon 2D molecular graphics. It should always be remembered, however, that there is much more to 2D molecular graphics than the naissance of computer. The history of 2D molecular graphics could be traced back to the proposal of benzene’s structure by Kekule [4]. After that, the chemical structure became the most frequently used concept in chemistry and the related fields. In recent decades, 2D molecular graphics was further developed and applied in molecular biology. For example, protein topology diagram and RNA secondary structure diagram give a clear expression of the arrangement of secondary structure elements and functionally structural domains; the macromolecular surface map can intuitively reflect the distribution situation of residues on a protein surface. Nowadays, 2D molecular graphics is also a useful tool in computer-assisted drug design; pharmacologists can easily identity and conveniently analyze pharmacophore and nonbinding interaction of receptor–ligand complexes by using diagrammatic analysis applications. These works suggest the 2D molecular graphics is a promising branch of molecular graphics and will be further developed rapidly in future. Here we categorize 2D molecular graphics into three aspects in terms of the research object and purpose: (i) Generation and drawing of chemical structures. (ii) Representation and annotation of biomolecules (proteins and nucleic acids). (iii) Schematic layout of molecular interactions.

**CLASSIFIED DISCUSSION OF 2D MOLECULAR GRAPHICS**

**Generation and drawing of chemical structures**

Generating 2D chemical structures from molecular formulas, connection tables or spatial coordinates belongs to the earliest application of computer in chemistry. Nowadays, graphical drawing of chemical structures is a relatively mature technique, readily available in many commercial and academic software such as ChemDraw [5], Symyx/Draw (formerly named ISIS/Draw) [6], OGHAM [7], CACTVS [8], etc. The demand for depiction of molecular structures is great, because the stylistic conventions of such pictograms constitute a ‘natural language’ for chemists which can be rapidly comprehended, and is thus highly suitable for communication. Fortunately, chemical structures constitute a limited subset of all possible graphs. If one considers only organic chemistry, there can be formulated a relatively small set of rules which cover most of the common environments in which an atom type might occur. Once these rules are stated, the structure can be pieced together and an overall depiction obtained [9].

Correct, aesthetic and understandable are the basic requirement for generating chemical structures. In the early stage, researchers mainly adopted some simple constraints to generate chemical structures. First attempt by Zimmerman and Feldman [10, 11] used a library of basic substructures and combined them according to a set of rules. Although this method is particularly effective for drawing edge-fused polycyclic ring systems such as steroids, it is limited in its generality by the size and nature of both the template set and the rule set, with the most difficult cases being bridged polycyclic systems. In the same time or later, several methods were proposed for the similar purpose. However these methods are all incapable of handling complex molecules and thus do not to be applied widely [12–14]. After 1980, intelligent algorithms were introduced into this field, obtaining a good effect on both computational efficiency and scale. Shelley [15] presented a heuristic approach to avoid overlapping atoms and bonds and to facilitate similar representations of the structural diagram for identical or similar structures. Hibbert [16] employed genetic algorithm to generate and display chemical structures, in which the atomic coordinates or bond angles are encoded into chromosome and the distances between all pairs of bonded and non-bonded atoms are served as the fitness function. The DEPICT program written by Weininger [17] was designed to convert SMILES format, the linear notation of molecular graph, into a depiction of molecular structure. The output display allows all aspects of SMILES representation of structure to be verified easily, including aromaticity, formal charge, bond order assignment and hydrogen attachment. A comprehensive review on the research status
of 1980s and 1990s for structure diagram generation can be found in Ref. [18].

After 2000, a number of methods and programs were proposed in elaborate consideration to generate, depict and edit 2D molecular structures. Their application scope covers not only the common organic compounds but also some special cases such as organometallic complexes, borides, carbon nanotubes, cage-shaped carbon molecules, etc. (Figure 1). Boissonnat et al. [19] firstly defined the concept of molecular family that was a set of molecules sharing a connected supertree. Using the correspondences between atoms provided by the supertree, the drawing performing of molecular graphs can be accurately coordinated by a breadth-first traversal approach. Fricker et al. [20] further proposed an algorithm that was based on the classical scheme of a drawing queue placing the molecular fragments in a sequential way. This method extended the traditional prefabricated units developed for complex ring systems to automatically created drawing units for chains and rings which will then be assembled in a sequential manner. Recently, Clark and co-workers [9] have stated the two goals of a successful structure depiction method: (i) to maximize the fraction of molecular connection graphs that are readily discernible to the observer and (ii) to maximize the fraction of graphs which are aesthetically ideal and functionally equivalent to the efforts of a skilled artist. In addition, two works are deserved to be mentioned: (i) one is the chiral depiction algorithm by Maehr [21] who used the targeted redeployment of established stereobonds to serve as the stereodescriptors that can differentiate between enantiopure and racemic compounds and those that are enantiopure but whose chirality sense is unknown; (ii) another is the small-molecule topology generator PRODRG [22] which takes input from given coordinates or various 2D formats and automatically generates molecular topologies suitable for X-ray refinement of protein–ligand complexes. On the other hand, two java-based programs, JChemPaint [23] and MCDL editor [24], were published for generating and editing chemical structures. They are different from previous structure-generating programs in that their source codes are shared and collaboratively developed through the Internet, and these characteristics lead to that the open source programs can be rapidly spread and widely applied in the science community.

### Representation and annotation of proteins

#### Protein topology representation

Protein topology was originally investigated systematically by Crippen [25, 26] and gradually becoming a common concept of protein structures. 2D representation of protein topology structures is a useful tool in protein searching [27] and fold comparison [28, 29]. However, the main application of protein topology representation is, of course, to straightforwardly show the structural pattern and spatial arrangement of protein architectures. Protein topology representations can be classified into two styles: cartoon diagram and topology diagram.

Cartoon diagram utilizes various symbols such as circles, squares, triangles and lines to represent protein’s substructures (Figure 2a), which was first adopted by Schulz et al. [30] and Rossmann et al. [31]. Levitt and Chothia [32] used a similar way to illustrate the arrangement of α-helices and β-sheets in 31 globular proteins with known conformations. Based upon the representation results, they arranged protein structures into four classes that are well known today, i.e. all-α proteins, all-β proteins, α + β proteins and α/β proteins. Cartoon diagrams were later drawn manually in some publications to elucidate the arrangement fashion of protein.

---

**Figure 1:** Some examples of 2D chemical structures generated by different algorithms. (a) Organometallic complex (ChemDraw [5]). (b) Boride (CACTVS [8]). (c) Carbon nano-tube (Clark’s method [9]). (d) Cage-shaped carbon molecule (OGHAM [7]).
structural elements [33, 34]. Flores et al. [35] have described a general algorithm for automatically generating protein cartoon diagrams, which was further improved by Westhead et al. [36, 37] and integrated into the TOPS database [38]. Recently, May et al. [39] proposed a modified version of protein cartoon diagram called PTGL in which the secondary structure units were notated as small circles and quadrates and the spatial neighbourhoods were represented as arcs between these units.

Topology diagram consists of parallel/anti-parallel arrows (for strands) and cylinders (for helices) (Figure 2b). There were considerable publications that adopted topology diagrams to illustrate protein structural configurations [40–42], and several software tools were also available which take atomic coordinate data input and produce protein topology diagrams in different styles [43, 44]. The key notation describing topology diagrams were first introduced by Richardson [45] for β-sheet topologies and further developed by Koch, Flower and Grigoriev et al. [46–48]. The topology diagrams are not only competent for clearly expressing simple protein motifs, such as the Greek-key and the Jelly-roll barrel motifs [49], but also performing well for showing complex protein global structures.

A normal style for protein structure annotations is the linear representation which unfolds the structured proteins into a linear state, and then uses different patterns and colors to annotate structural implications. One of the mostly used aspects of linear representation is to show the arrangement of secondary structures in protein configuration. This is frequently adopted by database pages and publications [52]. Another important aspect of linear representation is to display structural features from sequence alignment [53, 54]. In addition, linear representation has many other applications in information presentation of protein structures at sequence level, such as protein hydrophilicity plot [55], flexibility plot [56] and structure validation plot [57] (GCG [58] and PSAweb server [59] are two useful tools to draw these plots based on protein sequences and structures). Here we present three interesting applications of linear representation: (i) DOMPLOT [60] is a standalone program used to generate line–box diagrams of protein chains in terms of their structural domains. This program is written in C and thus can be implemented efficiently. (ii) Sipos et al. [61] developed a rapid visual approach, pairwise repeat homology diagram, to visualize the protein sequence repeats detected by a profile HMM in a linear pair of sequences and to highlight their homology relations inferred by a phylogenetic tree. A Perl-script t2prhd was also provided by them for drawing the diagrams. (iii) Kolodny and Honig [62] proposed a method called VISTAL that mapped the Cα–spatial distances between two superposed proteins into a 2D sequence alignment which described the superposed structures as a series of matched secondary structure elements, colored according to the distance of their Cα atoms.
Besides the linear representation, protein structure annotation can also be performed using various fashions for different purposes. Wako and Scheraga [63] applied a distance constraint approach to 2D modeling of proteins in order to visualize the nature of protein folding and to examine the relative roles of different ranges of interaction. Campagne et al. [64–66] developed a web-based server RbDe that greatly simplifies the construction of schematic diagrams of proteins; the yielded residue-based diagrams displayed the sequence of a given protein in the context of its secondary and tertiary structure and highlighted the amino acid residues in different colors according to the physicochemical properties. Kannan et al. [67] proposed a logo-based strategy to express the information about neighbor composition densities for residues with different secondary structures, and this method was used to ascertain the reason why a particular sequence with a given amino-acid composition should be fold into a specific structural class. Furthermore, the solvent accessibility of amino acids in structured proteins is a critical property of proteins [68] and can be calculated by many approaches such as MSMS [69], NACCESS [70] and SAVOL3 [71]. However, there was only one tool called ASAView [72, 73] that can provide the schematic representation of accessible surface area for the independent amino acid residues in buried state, by using spiral plot or histogram.

Another important application of protein structure annotation is the sequence logo which was originally described by Schneider and Stephens [74]. This method graphically represents the information present in a multiple sequence alignment. Each position in the logo is represented by a stack consisting of letters occurring at the corresponding position in the alignment. The height of each character is proportional to its frequency, where the most common character is placed at the top of the stack. The height of the entire stack indicates the information content available at that position. Currently, several tools for creating protein sequence logos are available via the web interfaces [75–77]. Sequence logo can also be used for DNA and RNA.

**Intra-protein nonbonding/bonding plot**
The global structure of proteins is stabilized by hydrophobic packing occurring in the cores, whereas the conformational specificity is accurately adjusted only by the forming or breaking of the bonding (e.g. disulphide bridges) and nonbinding (e.g. hydrogen bonds) interactions between intra-protein elements. It is conceivable that the complicated interaction network of protein structures is difficult to be understood in the stereoviews. Therefore, several graph theory-based methods were proposed to visualize the disulphide bridges and hydrogen bond networks of proteins in the 2D pages.

2D visualization of hydrogen bond networks. HERA [78] is the first program for automatically plotting hydrogen bonding diagrams of protein structures. It can also be used to calculate the connectivity of β-strands and to extract simple structural motifs such as hairpins or Greek keys. Another similar tool is HBNG [79] which gives an enumeration of favorable topologies of hydrogen bond networks in protein structures and determines the effect of cooperativity and anticooperativity on protein stability and folding.

2D visualization of disulphide bridges. Protein disulphide bridge topology has been investigated intensively [80–83] and is usually schematically represented by topology pattern plot. The topological stereoisomerism of multiple-disulfide polypeptide chain in stable folded proteins was first characterized from graph theoretic analysis of its covalent structure by Mao [84]. Based on that, Benham and Jafri [85] made a systematical study on disulfide bonding patterns and defined a simple linear graph to illustrate the linking patterns of disulfide bridges in protein sequences. The similar way was also used to analyze and visualize the topological stereochemistry of protein hydrogen bonds [86].

**Protein surface map**
The original work of projecting biological surface onto a 2D plane was made by Rossmann and Palmenberg [87] who used the ‘roadmaps’ to depict the residue distribution on the capsids of human rhinovirus and mengovirus. Chapman [88] extended the application scope of ‘roadmaps’ to the external surfaces of proteins and gave a clear definition of the boundaries between labeled residues in the maps. Recently, researcher’s interest has been transferred from protein global surfaces to some concerned local regions. Stahl et al. [89] proposed a computer-based strategy for automatically mapping protein surface cavities. Using this approach they successfully examined 176 metalloproteinases containing zinc cations in the protein active sites. Lee and Tzou [90]
described an earth-map representation of the directionality of protein surfaces in the DNA-binding regions. Gabdoulline et al. [91] developed a tool MolSurfer for analyzing, characterizing and visualizing the physicochemical properties of protein–protein-binding interfaces in 2D pages.

**Transmembrane protein display**

Transmembrane proteins play many important and diverse functional roles in cells and biological processes [92]. A number of methods and tools were reported for producing 2D transmembrane diagrams from protein primary sequences or advanced structures, such as the above mentioned RbDe [93] which can also be used to generate residue-based schematic diagram of transmembrane proteins. Moreover, the G protein-coupled receptor database (GPCRDB) [94] provides a 2D visualization tool Viseur [95] to display snake like diagram for the queried GPCRs. VHMPT [96] is a transmembrane topology plotting program that allows users to modify the layout of the generated topology, to label specific amino acid or amino-acid groups, and to annotate with arrows and texts. TMRPres2D [97] automates the generation of uniform, high analysis 2D graphical images of α/β-helical or α/β-barrel transmembrane proteins. TOPO2 [98] is a simple graphics program using a sequence along with input provided by the user to produce an image of a transmembrane protein. TEXtopo [99] creates the transmembrane topologies from PHD predictions [100] or SwissProt database files [101], and the output diagrams use shades and rich decorations to emphasize conserved residues or functional properties of the residue side-chains.

**Representation and annotation of nucleic acids**

**2D representation of DNA**

The mostly used 2D representation of DNA is, just like proteins, the linear representation of which the main applications are concentrated on sequence characteristics display and alignment plotting, these functions are available in many software suites such as BioViews [102], viewGene [103], SeqVISTA [104], etc. DNA sequence logo is also a popular fashion for annotating nucleotide’s conservation or other properties [105]. For example, the SequenceWalker [106] is a special tool adopting logo style to display how binding proteins and other macromolecules interact with individual bases of a given nucleotide sequences. Another important aspect of DNA 2D representation is the random walk plot in which the four nucleotides A/T and C/G are served as movers indicating respectively left/right and up/down moves, a DNA sequence can thus ‘walk’ on a 2D plane, from 5’ to 3’, to form a polyline plot [107–109]. However, the main uses of DNA walk plot are in sequence characterization [110] and analysis [111], rather than the 2D visualization. Moreover, there has been a simple but interesting instance of visualizing complete genomes, called DNA rainbow [112]. This method is a beautiful mix of science and art, which assigns a color to every of the four bases in a chromosome and renders it as a multicoloured picture.

In a previous review, Roy et al. [113] have given a detail description of the DNA graphical representation.

**2D representation of RNA**

The focus of RNA 2D representation is how to generate non-overlapping arrangement and layout of RNA secondary structure elements under an aesthetically pleasing consideration. In a mathematical sense, the secondary structure of RNA is a topological structure rather than a geometric structure. It remains unchanged even when under distortion, as long as the connectivity relation between nucleotides is not changed. Han et al. [114] have classified the RNA 2D representation methods into conventional and unconventional representations. The published works on 2D representation of RNA are listed in Table S1 in Supplementary Data.

**Conventional representation of RNA.** The conventional representation illustrates a polygonal display of RNA global structure in which loops are drawn as regular polygons, thereby generating clear and compact representation, especially for the long RNAs. In the early stage, Lapalme et al. [115] presented a simple algorithm for short RNA segments. Shapiro et al. [116] later developed a method to circumvent the problem of overlapping portions of the RNA. Bruccoleri and Heinrich [117] described an improved algorithm that is particularly suitable for drawing large RNA molecules with very limited overlap of strands. Gautheret et al. [118] gave a brief review on the early studies of RNA display. In the 1990s, RNA 2D representation obtained a rapid development in both methodology and software. These works laid emphasis on different aspects ranging from the generation of overlap-free polygonal displays [119]
and richly annotated diagrams [120] to development of general graphic libraries [121] and efficient computer algorithms [122]. After 2000, RNA conventional representation has gradually grown to be mature. Recently, it trends to developing combinatorial application suites that assemble a number of modules with diverse functions to automatically implement the complete procedure of RNA representation, including sequence alignment, energy optimization, secondary prediction, and so on.

Unconventional representation of RNA. The unconventional representation pays attention to local characteristics of RNA structures (e.g., loops, domes, pseudoknots, base pairs, etc.) or uses special fashions to define RNA global structures. Compared to conventional representation, methods and tools for unconventional representation are relatively few but developing rapidly in recent years. Martinez [123] early proposed a multiple approach for identifying and depicting local structures in the form of RNA hairpins. After that, Han et al. [124] presented an algorithm drawing H-type pseudoknots based on RNA graph theory, and this method was further improved in several aspects including pseudoknot types, display quality and computational efficiency [125]. Another important tool in the RNA unconventional representation area is the RNAView [126], which creates a schematically diagram for all base pairs in a RNA molecule using previously developed theory and symbols by the same group [127, 128]. RNAView was further modified to Rna2DViewer and integrated into the S2S package [129]. Like the RNAView, MC-Annotate [130] can also be used to annotate information about RNA local structures. The output of this program is composed of a 2D structural graph that contains the annotations, and a series of HTML documents, one for each nucleotide conformation and base–base interaction present in the input 3D structure. Moreover, the tool of jViz.Rna developed by Wiese et al. [131] is a Java program that represents RNA global structures using various strange patterns, such as linked graph, circle graph, dot plot, etc. In this literature they also gave a brief description of some useful RNA visualization tools.

Schematic layout of molecular interactions

Protein–small molecule interaction
Protein–small molecule interactions are accompanied with the binding and association processes of enzyme–substrate, receptor–ligand, target–drug, etc. Knowledge of these interactions is important to understanding how proteins function in biological systems and to developing drugs with specific selectivity. 2D representation of protein–small molecule interactions dedicates to straightforwardly demonstrating the hidden active sites and the complicated nonbonding networks across the binding interfaces. However, only three tools are available that can be used to automatically generate schematic 2D diagrams of protein–small molecule interactions.

LIGPLOT is a well-known program written by Wallace et al. [132]. It employs HBPLUS [133] and NACCESS [70] to identify and compute hydrogen bonds and solvent accessible areas around the ligand molecule, and subsequently arranging these information into a neat page containing a complete description of the nonbinding effects on the given complex’s interface (Figure 3a). LIGPLOT adopts a series of algorithms to avoid overlapping and to flatten ring groups, and provides diverse output styles that can be easily changed by users.

In the MOE platform, Clark and Labute [134] employed scientific vector language (SVL) to fulfill the procedure of automatic depiction of protein–ligand complexes. The output diagrams are annotated with a substitution contour, solvent exposure, cofactors, covalently bound linkages and a series of identified nonbonding interactions and conserved residues (Figure 3b). This method can also be applied to aligned sets which contain multiple ligands, or multiple members of a protein family, in which case the ligand orientations and protein residue placement will show consistent trends throughout the series.

Stierand et al. [135, 136] described an algorithm that is based on a combinatorial optimization strategy which can solve parts of the layout overlapping problem nonheuristically. The depicted molecules are represented as structure diagrams according to their chemical nomenclatures and arranged to maximize aesthetic ideals (Figure 3c). This method is called PoseView and available as a web server by which the complete protocol of layout procedure can be implemented efficiently and conveniently.

Protein–nucleic acid interaction
Today, only a single software tool for the automatic creation of protein–nucleic acid complex diagrams called NUCPLOT [138] is available. It is a standalone program that can automatically identify...
interfacial residues and interactions between them from the input PDB coordinates of the complex, and generates a plot in which all interaction information are intuitively shown in a schematic manner. The program creates the output files consisting of nucleic acid chains, interfacial elements, hydrogen bonds, van der Waal contacts and covalent bonds between the protein and the nucleic acid. The resulting diagram is both clear and simple and allows immediate identification of important interactions within the structure. The program works well for both double-stranded and single-stranded nucleic acids binding with proteins, so it can be used for both cases of protein–DNA complexes and partial protein–RNA complexes.

**Protein–protein interaction**

The ready availability of structural data of protein complexes, both from experimental determination, such as by X-ray crystallography, and by theoretical modeling, such as protein docking, has made it necessary to find ways to easily interpret the results at 2D level. The simplest way used for 2D representation of protein–protein interactions is the linking plot in which the protein sequences are represented as two lines, each from one protein subunit of the complex, and the residue-pairs in interactions are linked linearly. Several methods like MONSTER [139] and iPfam [140] can be used to automatically generate linking plots from protein complex structures. The main disadvantage of linking plot is that the presenting information is too cursory to give a comprehensive insight into the protein-binding interfaces. For that, QContacts [141] adopts a different strategy called contacting map to exhibit nonbonding interactions in the binding interfaces. This map can be viewed as a matrix; residues in turn from one subunit sequence define the columns, and those from another define the rows, the matrix elements thus represent the corresponding residue-pairs and, if in interaction, labeled by different symbols indicating different interaction types. A more aesthetically pleasing, informative diagram showing protein–protein interactions can be generated by using DIMPLOT [132], which is a sister program of above-mentioned LIGPLOT and thus its function and output style are same to that of LIGPLOT.

Recently, our group developed a graphical user interface (GUI)-based package 2D-GraLab [142] for

![Figure 3: 2D representation of the cyclin-dependent kinase 2 (CDK2)-oxindole inhibitor complex (PDB entry: 1fvt) (below). These diagrams were generated using (a) LIGPLOT [132], (b) MOE [134] and (c) PoseView [135]. The stereoview of the complex (above) was produced using CHIMERA [137].](image-url)
automatically generating comprehensive representation of diverse nonbonding (and covalent) interactions across the protein-binding interfaces, including hydrogen bonds, salt bridges, van der Waals interactions, hydrophobic contacts, π–π stackings, disulfide bonds, desolvation effect and loss of conformational entropy. This program takes standard PDB format as the input file and gives two types of PostScript outputs, referring to individual schematic diagram (for each nonbinding type) and summarized schematic diagram (for all nonbinding types). In addition, 2D-GraLab provides a complete protocol to define and calculate geometric and energetics parameters for different nonbonding types, such as length, angle, free energy, stability, etc., and subsequently presents these parameters with interfacial structure elements in an elaborately designed page. To ensure these interactions are determined accurately and reliably, methods and standalone programs employed in 2D-GraLab are all widely used in the chemistry and biology communities.

CONCLUSIONS

2D molecular graphics is the subject that uses an intuitive, understandable and accurate manner to present configuration and information of chemical and biological systems in planar pages. It should be always noted that the two points: (i) The data and parameters used by 2D molecular graphics are all taken from 3D structures and models, and further mapped into 2D representations by a consensus algorithm. Therefore, the diagrams of 2D molecular graphics are different from those of inexact drafted pictures that are used just for schematic demonstration. (ii) 2D molecular graphics lays emphasis on graphic representation, rather than being severed as calculation tool. Therefore, those studies using 2D graph and geometry just for the purposes of mathematical tool or modeling approach, such as chemical graph theory and DNA topological theory, do not belong to the spectrum of 2D molecular graphics.

Acknowledgements

The authors would like to express their gratitude to the editors and the anonymous reviewers for their professional, intensive and useful comments. They also thank the people for providing access of pictures, programs, servers and databases used in this article.

SUPPLEMENTARY DATA

Supplementary data are available online at http://bib.oxfordjournals.org/.

References


58. Genetics Computer Group, Science Drive, Madison, WI, USA.


70. Hubbard SJ, Thornton, JM. NACCESS. Department of Biochemistry and Molecular Biology, University College London, UK.

71. Pearlman RS, Skell JM. SAVOL3. College of Pharmacy, University of Texas, Austin, TX, USA.


