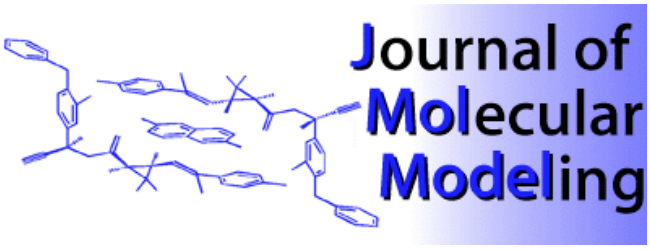
J. Mol. Model. **2000**, *6*, 328 – 340



FULL PAPER

Molecular Graphics - Trends and Perspectives

Jürgen Brickmann, Thomas E. Exner, Matthias Keil, and Richard J. Marhöfer

Physikalische Chemie I and Darmstädter Zentrum für Wissenschaftliches Rechnen (DZWR), Technische Universität Darmstadt, Petersenstraße 20, D-64287 Darmstadt, Germany. Tel: +49-6151-162198; Fax: +49-6151-164298. E-mail: brick@pc.chemie.tu-darmstadt.de

Received: 2 November 1999/ Accepted: 15 December 1999/ Published: 28 February 2000

Abstract The principles of molecular graphics are discussed in context with the optimization of man-machine communication in molecular sciences. The state of the art in this field is demonstrated with several examples. The paper is focussed on the discussion of a strategic basis of the information transfer between human activities and computational processes. It is demonstrated that enhancement of interactivity in the visualization process may lead to the generation of new insight on one side and the development of new computational algorithms based on the human visual pattern recognition strategy.

Keywords Molecular graphics, Man-machine communication, Information transfer, Graphical excellence

Introduction

The term *molecular graphics* first occurred in the seventies when molecular scientists adapted a new method from researchers in the field of structural investigation [1-3]. In the latter field computer graphical representations of electron densities (obtained from x-ray diffraction measurements) as isosurfaces have lead to an enormous step forward in the determination of protein structures: molecular structures (represented as "chicken wire" line graphs [4]) could be fitted interactively to the experimentally determined electron densities represented as nets of lines. The computer graphically supplied technique replaced a standard method based on complicated models made out of sticks and modeling clay. Computer graphical tools in the early stage of molecular graphics were dominated by vector graphical representations on special graphics computer hardware based on calligraphic technology [5,6]: Only lines and dots could be represented and almost all manipulation of the molecular scenarios had to be performed on a main frame computer and then submitted to the graphics hardware. The aims of research in these early days of molecular graphics were not so far away from the aims of chemists in this field today: One was interested in the quantitative treatment of intra- and intermolecular arrangements controlled by distance criteria which could be applied interactively. Distance vectors between selected atomic increments could be recorded and displayed during interactive manipulations of the molecular scenario and a possible overlap of molecular surfaces (represented as a collection of dots (dotted surfaces) based on a hard sphere model - the so-called CPK-model) could hardly be recognized. Almost all of early molecular graphics applications came from the field of pharmaceutical chemistry and many of the applicants were from chemical companies. The new field grew fast. The Molecular Graphics Society (today Molecular Graphics and Modelling Society)

Correspondence to: J. Brickmann

Dedicated to Professor Paul von Ragué Schleyer on the occasion of his 70th birthday

was founded in 1983 and the Journal of Molecular Graphics was set up shortly thereafter. Many of the papers in this journal and others were related to new graphic techniques (see for example Brickmann et al. [7-9]), in particular also as a consequence of the dramatic increase of raster graphics technologies pioneered by the new hardware devices of the Silicon Graphics company and others. While in the beginning of the eighties nearly nobody believed that this technology could ever be fast enough to allow interactivity on the basis of space filling molecular models, the new technique replaced the old one within a few years. New hardware developments and the generation of graphical software standards made it easy to generate images of very high complexity with interactive refresh rate without extensive graphics programming [10-15]. Today the question in the field is no longer "how to represent (technically) a certain molecular scenario" but "what should be represented" [16-19] in order to obtain a maximum of information from the underlying data and to get an optimal insight from the image.

This paper is not focussed on the description of new algorithms or new programs in the field of molecular graphics. We are dealing here with more general aspects, namely the strategic basis of molecular graphics for the optimization of the information transfer between human activity and computational processes and *vice versa*. If not indicated otherwise, we use the term *molecular graphics representation* (MGR) synonymously for the representation of molecular scenarios as screen images, as slides or hardcopies, as video sequences, and also as virtual scenarios which can be inspected and manipulated interactively (locally or over the net).

The paper is organized as follows. In section II some general principles are summarized concerning some basic rules for an optimal presentation termed "graphical excellence" by E. Tufte in his famous book on the visual display of quantitative information [20]. The following section III is related to the human ability of analysis of visually represented information. In particular, the treatment of similarity and complementarity of two objects are discussed. Section IV deals with the transformation of molecular information to a scenario which can be recognized with human pattern recognition abilities. In section V some possibilities for the interactive reduction of complexity of MGM's are outlined while section VI deals with the introduction of graphical languages and the inclusion of these languages in algorithmic descriptions of molecular scenarios. In the final section some conclusions are drawn and perspectives are given.

Graphical excellence as a guideline for the development of molecular graphics representations

Edward Tufte wrote in his book "Envisioning Information" [21] in the first chapter: "*Even though we navigate daily through a perceptual world of three spatial dimensions and reason occasionally about higher dimensional arenas with mathematical ease, the world portrayed on our information displays is caught up in the two-dimensionality of the endless flatlands*[1] *of paper and video screen. All communication between the readers of an image and the makers of an image must now take place on a two-dimensional surface. Escaping this flatland is the essential task of envisioning information - for all the interesting worlds (physical, biological, imaginary, human) that we seek to understand are inevitably and happily multivariate in nature. Not flatlands.*" Molecular objects are essentially three-dimensional, molecular information may be adequately represented even in a space which has much more dimensions. Molecular graphics representations (MGRs) must take this multi-dimensionality into account. Since molecules and molecular data are not objects of our everyday life, there is much room for the creation of MGRs. There are technical possibilities to generate quasi three dimensional representations of model scenarios by placing the human inspector in a large cubic box – the CAVE [22] – wherein images of a scenario are projected to the walls and an interactive device (data glove etc.) can be used in order to interfere with the 3D-world but in this paper this new approach is not in the focus of our consideration. Independently from this new technical possibility, a representation should fulfill some general requirements in order to optimize the information flow between a graphical representation (in a very general sense) and a human inspector. These requirements have been outlined in Tufte's first book [20] "The visual display of quantitative information" under a somewhat different context as *graphical excellence*:

Excellence in statistical graphics consists of complex ideas communicated with clarity, precision, and efficiency. Graphical displays should

(i) show the data

(ii) induce the viewer to think about the substance rather than about methodology, graphic design, the technology of graphic production, or something else

(iii) avoid distorting what the data have to say

(iv) present many numbers in a small space

(v) make large data sets coherent

(vi) encourage the eye to compare different pieces of data

(vii) reveal the data at several levels of detail, from a broad overview to the fine structure

(viii) serve a reasonably clear purpose: description, exploration, tabulation, or decoration

(ix) be closely integrated with the statistical and verbal descriptions of a data set.

The author summarizes the principles of graphical excellence as follows: "Graphical excellence is the well-designed presentation of interesting data, a matter of *substance*, of *statistics*, and of *design*, it consists of complex ideas communicated with clarity, precision, and efficiency. Moreover, graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space and graphical excellence is nearly always multivariate".

Although these principles were formulated mainly for applications of graphical representations in printed media, slides or static two-dimensional screen images, they represent a very good basis for all types of molecular graphics representations. This should be kept in mind while reading

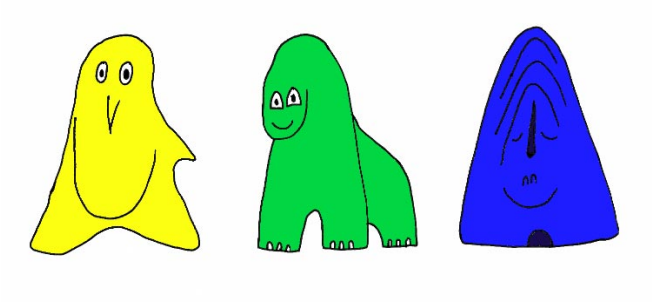


Figure 1 *Size similarity of two-dimensional objects*

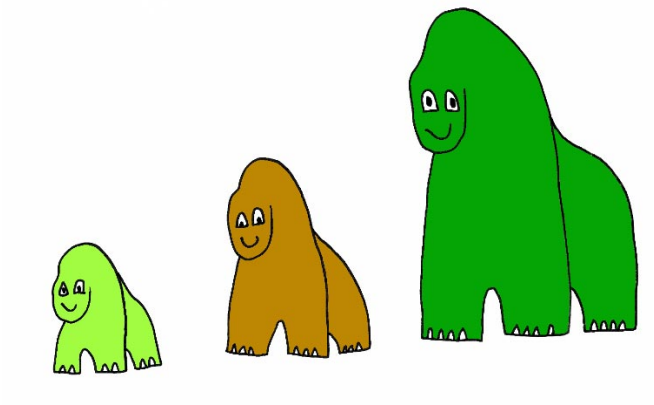


Figure 2 *Global shape similarity of two-dimensional objects*

the following sections. In the next subsection we will give a particular example for a consequent realization.

Pattern recognition abilities of human searchers

The specific recognition of a molecule within a molecular scenario plays an important role in many chemical processes. For instance, it forms the basis for highly specific reactions in biochemistry and catalysis. A large variety of different factors (energetic, entropic, and kinetic, etc.) come into play in a conceptional model approach when we attempt to describe such recognition in a precise way [23]. From a thermodynamic point of view, the specificity of a receptor can be measured by a subgroup A of molecules it recognizes (at a given level of affinity defined by a certain ΔG value) from among a larger ensemble B of molecules which in principle have to be considered. This type of recognition may often be described in terms of the key and lock image first introduced by Emil Fischer [24] in 1894. In order to define the set A and the reference set B a rational approach is clearly needed. In principle, the tools for such an approach are available. ΔG values or relative ΔG values for the binding of a substrate to a given receptor can be calculated on the basis of thermodynamic quantities, and molecular simulations. Such calculations are still very expensive as far as computational effort is concerned and they are only justified for such molecular scenarios for which a preselection process of molecular partners and their relative arrangements has been made on the basis of a simplified model scenario: One has to answer the question what molecular part of molecule A may fit to what molecular part of molecule B. The answer is strongly related to the question of molecular similarity (if different molecules A should be tested as possible partner for an unknown receptor, knowing that one or more molecules from the set are active) or molecular complementarity in the region of a receptor B (if this is known). In order to answer the open questions we thus have to look at the molecules from the point of view of a "molecular inspector" and trying to discover which may belong to a certain class of possible "keys" to fit some given "lock". The search becomes even more complicated when

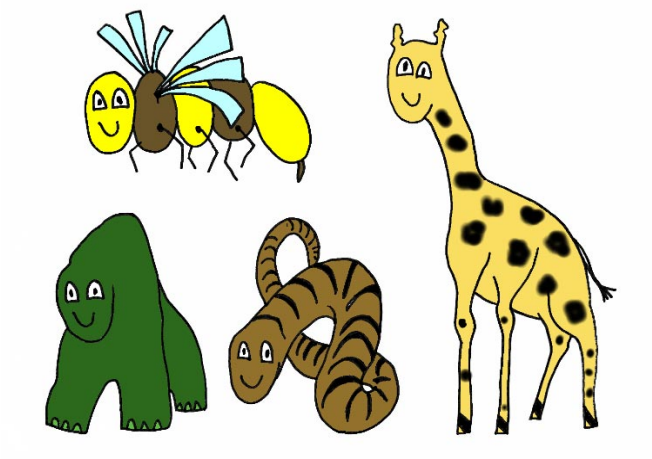


Figure 3 *Partial similarity: All animals are pairwise different but they have a similar face in common*

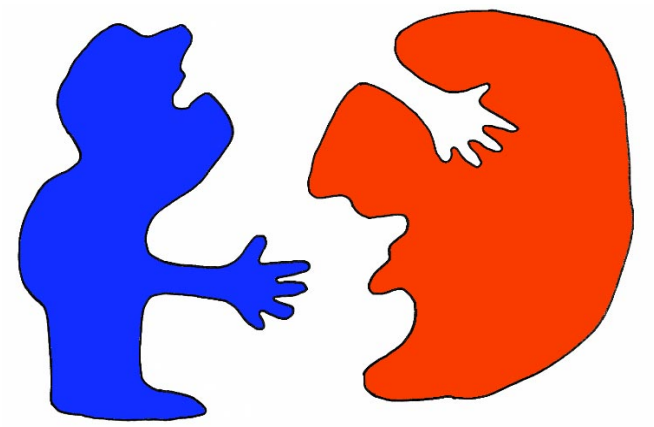


Figure 4 *Shape complementarities of two two-dimensional objects*

the "lock" cannot be specified. In this case one is looking for complementarity between arbitrary regions of one molecule and regions of a second molecule without any knowledge of the search patterns.

There is no doubt that the most effective search procedure - for those instances in which it can be applied - is still the "eyeball" technique used by human "searchers" [25]. One can easily compare different objects and analyze their similarity or dissimilarity without having explicitly defined criteria in hand. This is demonstrated with the cartoons in Figures 1-3. The creatures in Figure 1 obviously show *size similarity*.

For this case, it is relatively simple to transfer criteria for the analysis of size similarity to a molecular scenario calculating diameters, molecular masses, volumes etc. Size similarity is definitely relevant for the analysis of molecular sieves (zeolites and others). The animals in Figure 2 are different in size but similar in shape.

There are several papers published on the analysis of global *shape similarity* of molecules in the last few years [26-30]. It turns out that an algorithmic solution of this problem is not an easy task while the eyeball technique can be still applied very effectively. This becomes even more obvious if one is interested in the analysis of partial similarity. This is demonstrated with the creature in the cartoon shown in Figure 3.

For a human inspector it is relatively easy, to answer the question "What is similar among these animals?" just by applying the eyeball technique. The remarkable fact in this case is that one does not have to define the criteria for the search in advance. It seems to be obvious that the brain determines these criteria interactively in a hierarchical manner: After short inspection one will find that the animals have similarly shaped faces. In what way is this example relevant for an application in molecular sciences? A simple example may demonstrate this relevance: The molecule benzamidine and the protein PTI are both inhibitors for the activity of trypsine. No global similarity search algorithm has any chance of finding this result simply as a consequence of the fact that these molecules only show partial similarity. However, the application of the eyeball technique is also not trivial in this case. The molecular scenarios have to be transferred to MGRs which are acceptable for the human "search engine". The same is true for complementarity searches. The two objects shown in Figure 4 have two areas of complementarity.

The first one (a hand with fingers) obviously fits to a glove type cavity, a human profile type shape is possibly recognized only on a second view. In any case, it is relatively easy to see by inspection that a regularly shaped object (the key) "probably fits" into a rigid surface of complementary shape (the lock) if pattern elements for this analysis are stored in the brain. This should be kept in mind while developing MGR tools. An optimally tailored graphical interface has to give the inspector the possibility for an interactive modification (as a whole or in detail) of the representation in such a way that his own abilities for pattern recognition are supported in the best way. In the next section some of the possibilities for the generation of MGRs are presented in order to show the bandwidth of the present technology.

Transformation of molecular information to a representation which is convenient for human interactions

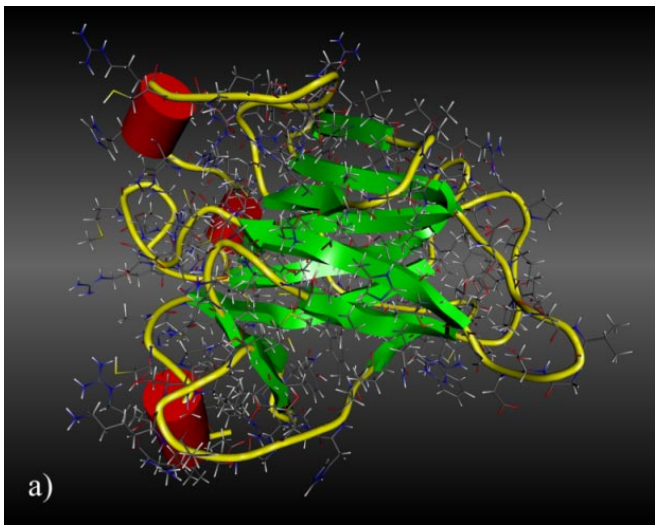
Chemists have a long tradition in inventing and applying model scenarios for the rational analysis and communication of molecular information. The standard models are formed out of wood, metal or polymer material. They can help a chemist estimate whether a certain arrangement of molecular parts necessary for a particular reaction can be reached. However, the use of materialized models can only give a rough insight into the elementary interactions of molecules, i.e. these models give only a poor representation of the way molecules "see each other". The use of present computer graphics techniques enables the chemist to extend his model world significantly. MGRs have to follow the traditional rules of model building in chemistry in order to be widely accepted by chemists. On the other hand, with the new technology chemists have the great opportunity to generalize old ideas and to establish model scenarios which take into account the building laws of the microscopic world. A prerequisite for this is an effective man-machine communication, i.e. the "molecular point of view" has to be transformed into pictures which can be easily analyzed with the human recognition capacity. It is not the aim of this paper to review all the different techniques presently used in the field of molecular visualization. A collection of images (see Figure 5) generated in the lab of the authors may serve as an overview.

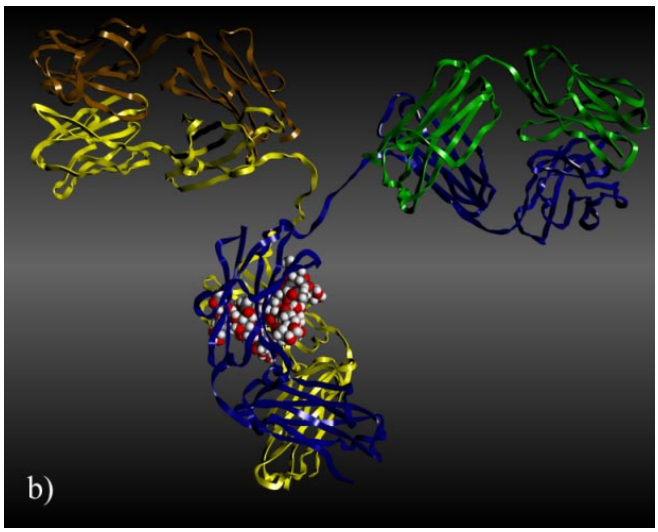
More details can be found in a recent review from our group [31]. In this section we focus on two concepts. The first is the concept of molecular surfaces, which can be adequately used in order to transform information into a visual representation of molecular properties which may serve as a first trial for an analysis with the aid of the eyeball technique. The second representation is related to the visualization of volumetric data.

While for the molecular surfaces the brain may find structure elements from the real world (like knobs, ridges, holes, canyons, saddles etc. as building blocks for shape recognition, there are only a few such elements available in the case of 3D volumetric information. An example for the latter may be fog of different density. Missing a separating surface between different space areas obviously leads to confusion in the recognition process.

Molecular surfaces

It is clear that atoms and molecules do not have a surface like macroscopic objects of our everyday life. Nevertheless, as a consequence of the statement given above it is very convenient to define such a molecular surface as a separation for "inside" and "outside" and as a screen for the representation of properties reflecting the interaction with other molecules.





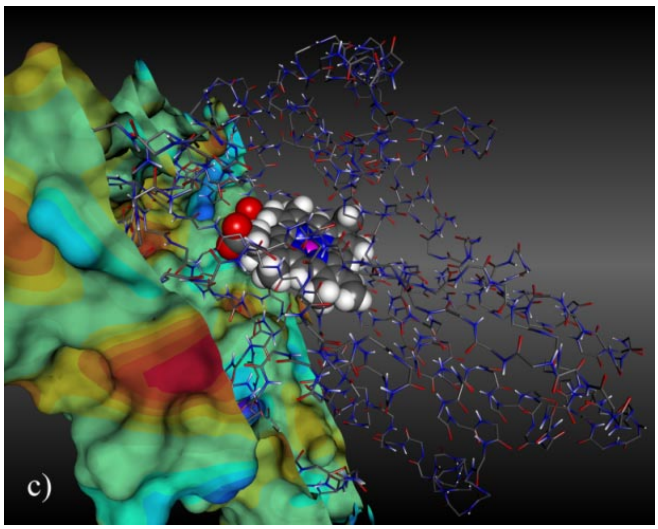
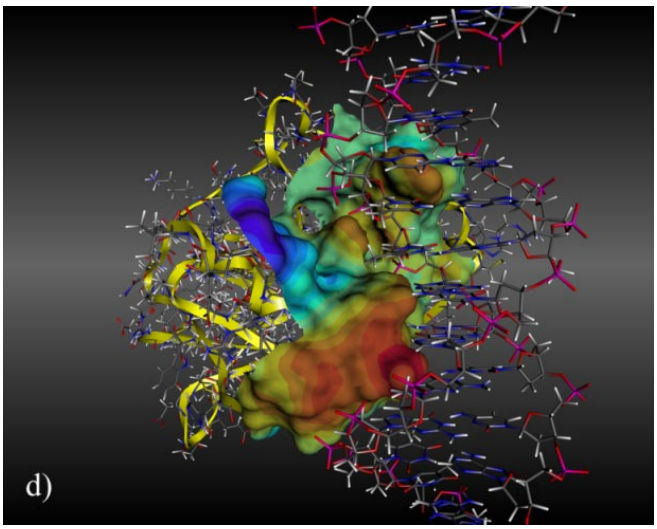
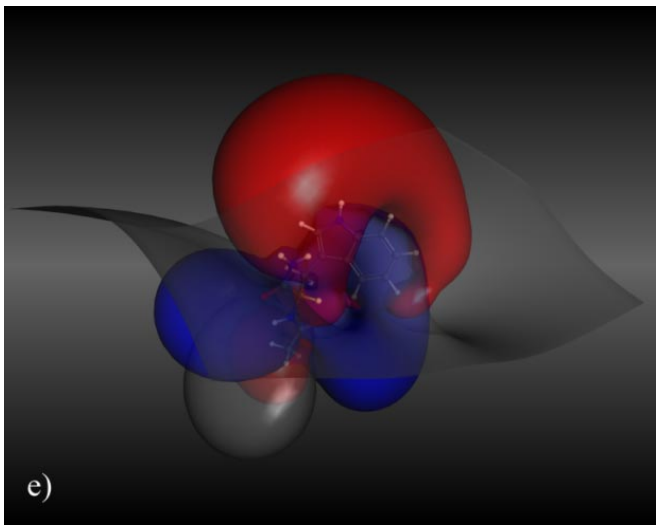
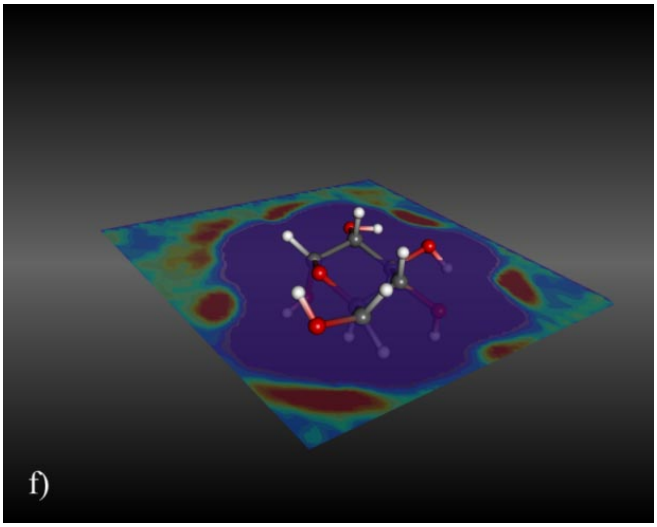
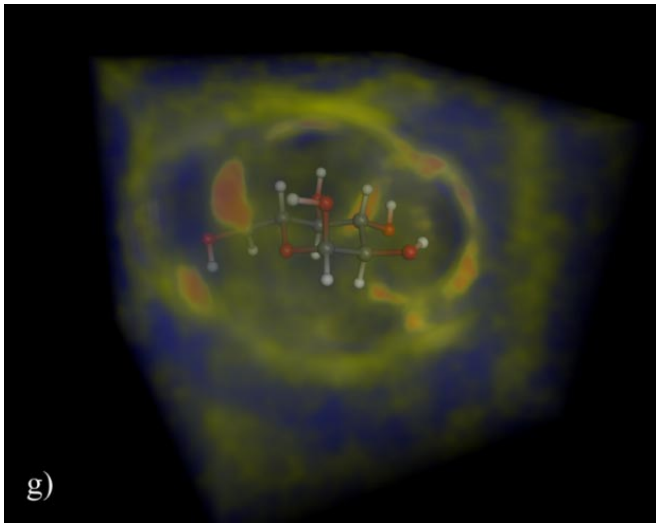


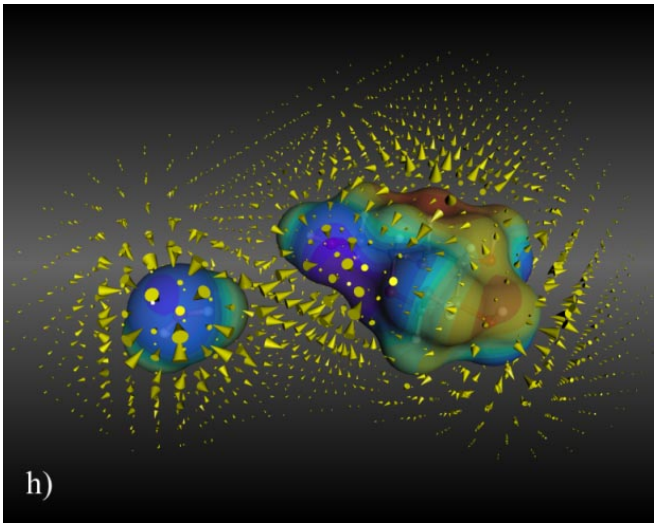
Figure 5 (continues next page) *Molecular graphics representations (MGRs) for different molecular scenarios and different applications. (***a***) MGR of the protein p53: The full structure is shown in stick representation. In addition the secondary structure is shown in a simplified manner: red drums represent alpha helixes, green ribbons beta sheets and the yellow tubes are the loops; (***b***) Ribbon representation of the light and heavy chains of an antibody protein (IgG). This MGR is particularly useful for the submission of a rapid overview on the secondary structure. Oligosaccaride ligands are highlighted through a CPK-representation; (***c***) The protein cytochrome P450: parts of the molecular surface are blended out in order to make the active center visible. The color coding on the molecular surface represents local hydrophobicity. The heme group in the active center of the protein is represented in CPK representation while the backbone of the protein is shown as a stick model; (***d***) Capped stick representation of a p53-DNA complex: The molecular surface is only shown in the binding area. The electrostatic potential of the protein is mapped onto this surface by texture mapping technology (see text): red indicates a positive blue a negative potential value. The backbone of the protein is shown as a yellow ribbon; (***e***) Isosurface representation (see text) of the electrostatic potential around a tryptophane derivative molecule. Red and blue surfaces represent a positive and negative potential value of identical absolute value, respectively, gray is the isosurface for neutral potential value. The individual surfaces are shown with transparency in order to submit a better insight; (***f***) Water density (as obtained from molecular dynamics simulations) around a glucose molecule. The density is shown as a contour map with texture mapping technology on a plane intersecting the molecular area. The plane can be moved interactively in order to obtain an overview over the complete density function (red: high density, blue: low density); (***g***) Water density around a glucose molecule (as in fig. 5f) represented by 3D texture mapping. The fog type representation allows an immediate recognition of ”hot spots” (values of high density in a small volume area); (***h***) Electrostatic interaction between a glucose and a water molecule. The electrostatic potential values of both molecules are represented by texture mapping technology on the molecular surface. The yellow cones are representations of the electric field between the molecules. The top of the cones point in field direction, their size is proportional to the field strength; (***i***) Separating surface between a glucose and a water molecule. This surface is generated from the condition that every surface point has to have an identical closest distance to both of the molecular partners. The separating surface representation is very effectively applicable for interactive docking procedures.*

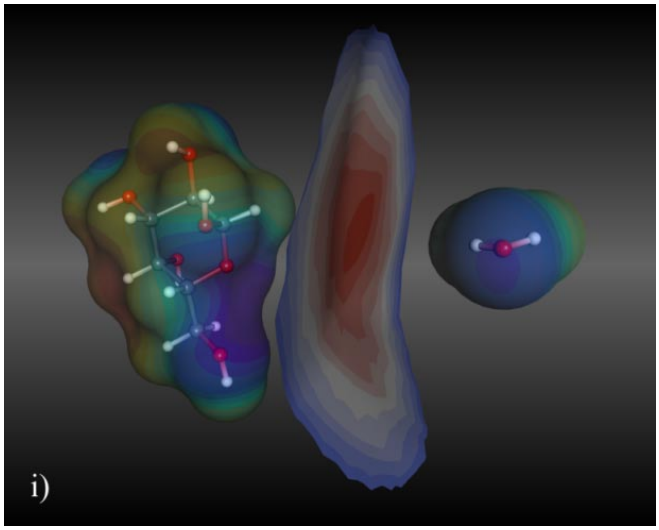

d)


e)


f)


g)


h)


i)

A smooth molecular surface can be generated by rolling around another hard sphere model particle on the CPK surface. This model, which was first introduced by Richards, [32-35] forms some reference standard for molecular surface generations in many molecular modeling packages. The contact surface representation gives the chemist some insight into the molecular shape as it would be seen from a particle of given size. Surfaces generated with the same test particle (e.g. a water molecule with an effective sphere radius of r = 1.4 Å) can be compared qualitatively and quantitatively. Moreover, the contact surface, generated with a water probe, is well suited in order to discuss shape fitting (for example of two proteins [17]). Formal molecular surfaces have become important tools for the interpretation of molecular properties, interactions and processes [10,11,36-38]. A detailed review has been given by Mezey [35].

The molecular surface concept is not only useful for a representation of the bulkiness and the shape of molecules. These surfaces can be used as screens for the visualization of arbitrary properties using color coding techniques. Color coding is a popular means of displaying scalar information on a surface [39]. In interactive molecular graphics, high contrast color code variation can be realized by using texture mapping techniques which are available on graphical workstations and high end PC's to represent a color ramp as a 1D texture (see Figures 5c, 5d, 5h, 5i). Texture mapping is a technique that applies an image to an object's surface as if the image were a decal or cellophane shrink-wrap. The image exists in a parametric coordinate space called the texture space [15,40-42].

Mapping the calculated property into texture space instead of color space ensures that the coloring evaluated at every pixel is taken from information lying in-between the values of the relevant vertices. High contrast variation in the color code is then possible, even on sparsely tessellated surfaces. It is important to note that, although the texture is one-dimensional, it is possible to tackle a three-dimensional problem, because the dimensionality of the texture space does not affect the object space.

The independence of texture and object coordinate space is well suited to accommodate immediate changes to the meaning of the color, i.e. by applying simple 3D transformations in texture space. Translation allows readjustment of the zero line of the color code, while scaling of the texture changes the range of the mapping. Such modifications may be performed in real-time.

Similar to the 1D texture used as a color code on a molecular surface, the texture space may be extended to 2D or even 3D (see Figure 5), incorporating additional information with each additional dimension, such that a maximum of three independent properties can be simultaneously be visualized. Special care must be taken not to overload the surface with too much information However, texture mapping technology is not only valuable for the visualization of complex information, it can also be used for the reduction of complexity by filtering out interactively information from the graphical representation. Filtering property information on a molecular surface is able to generate more insight in two different ways:

(1) The filter allows the scientist to distinguish between important and irrelevant information.

(2) The filter puts an otherwise qualitative property into a quantitative context.

In both cases the information can be filtered using a function mapping which suppresses all information not exceeding a specific threshold, or a continuous filter may be used to allow for a more fine grained quantification.

Filtering may be implemented analogously to the color coding technique presented above, if one uses 2D or 3D texture maps. A useful application results from the filtering of properties such as the electrostatic potential and the local hydrophobicity (see Figure 6, Video 1).

The enormous progress in the application of texture mapping in MGR technology is not simply related to the new way of image manipulation per se but to the possibility that both the object space and the texture space can be manipulated interactively, i.e. with this technique the inspector has a relatively large freedom for generating an image which is optimal for his own reception. This is demonstrated with the scenarios presented in Video 1 (Figure 6). In Figure 6 the electrostatic potential (calculated on the basis of MO calculation) and the molecular free energy surface density, MolFESD [43-44] are mapped on the molecular surface with the aid of 2D-texture map. While the first quantity is represented by a color code the second is used as a filtering property which can be handled interactively. Taking into account that positive MolFESD values correspond to hydrophobic, negative to hydrophilic areas, one can interactively relate these two quantities with the aim to generate, for example, a new receptor model while comparing different molecules with the same strategy. This has recently been demonstrated for a set of sweeteners [44].

Three dimensional data fields

Visualization and quantification of data fields in 3D space becomes increasingly important in molecular science. There are many possibilities for such a visualization. Some examples are shown in Figure 5. For all of them interactivity play an even more prominent role than for the investigations on the basis of molecular surfaces.

As has been mentioned above, there are only a few binding blocks in the human brain which can be used for the analysis and recognition of 3D bulky information. One of the possibilities for a transformation of a volumetric data field to a convenient visualization for the application of the eyeball technique is the isosurface representation.

Here only this representation should be shortly reviewed. Isosurfaces are the two-dimensional analogs to iso-lines in hydrographic or topographic maps. Among a variety of isosurface generation algorithms [10,45-48], the marching cube [45] has become one of the most popular ones because of its speed and multiple applicability. With this algorithm isosurfaces can be generated interactively on workstations as

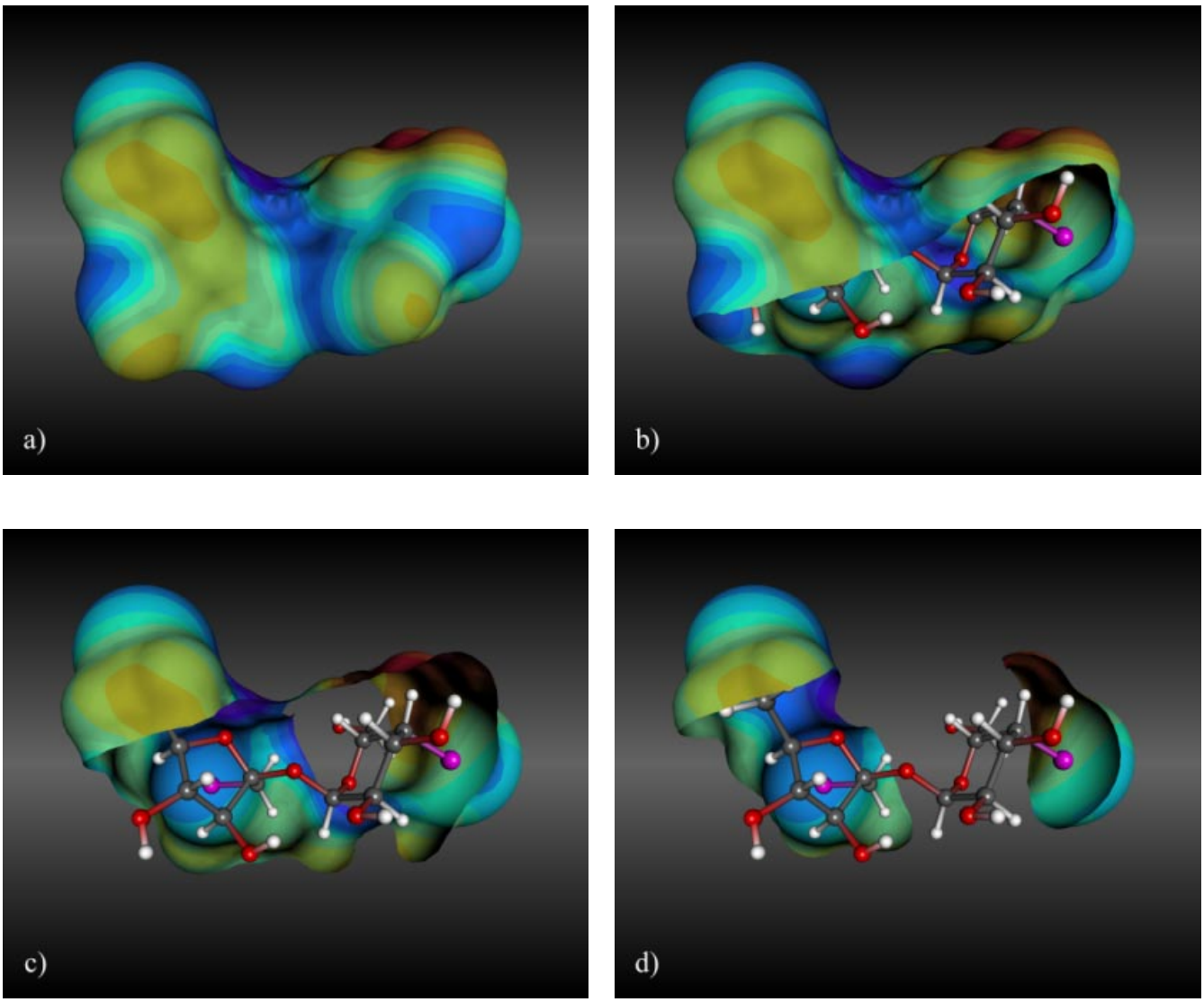


Figure 6 (video 1) *Molecular surface of a sucralose molecule. The surface is color coded according to the electrostatic potential. 2D-texture mapping technology is applied in order to blend out certain hydrophobicity values of the surface (according to the molecular free energy concept MolFESD [43,44]). Four different filtering situations are shown in 6a – 6d. In video 1 a continuous change of the filtering process is shown as a time sequence*

well as on high end PC's. In contrast to iso-lines in two dimensional maps, a collection of isosurfaces (for example related to equidistant iso-values in property space) cannot be inspected simultaneously simply as a consequence of the fact that in almost all cases most of these surfaces are hidden behind others. One can occasionally circumvent this problem by using transparent surfaces (see Figure 5e), but this technique leads to a drastic increase of complexity of the image and so to a decrease of graphical excellence. A solution is the introduction of a "time sequence" of such surfaces.

An illuminating example for the application is the comparison of HOMO-(highest occupied molecular orbital) and LUMO (lowest unoccupied molecular orbital) as is demonstrated in Video 2 (Figure 7). There is no particular iso-value which can be optimally chosen for the clear recognition of the topology of these orbitals and the characteristic difference between the HOMO and LUMO state. The choice depends on the personal reception ability of the human inspector as well as on the physical quantity the inspector is interested in: for the understanding of spectroscopic properties (transition moments etc.) one certain isosurface may become highly informative while for the analysis of chemical reactivity another value may be more expressive. The simultaneous change of isosurfaces and molecular orientation allows the inspector to find out the best representation. Moreover, it may happen that even the continuous change of the surfaces under interactive control is the key property for graphical excellence (in Tufte's definition, see section 2) in this case.

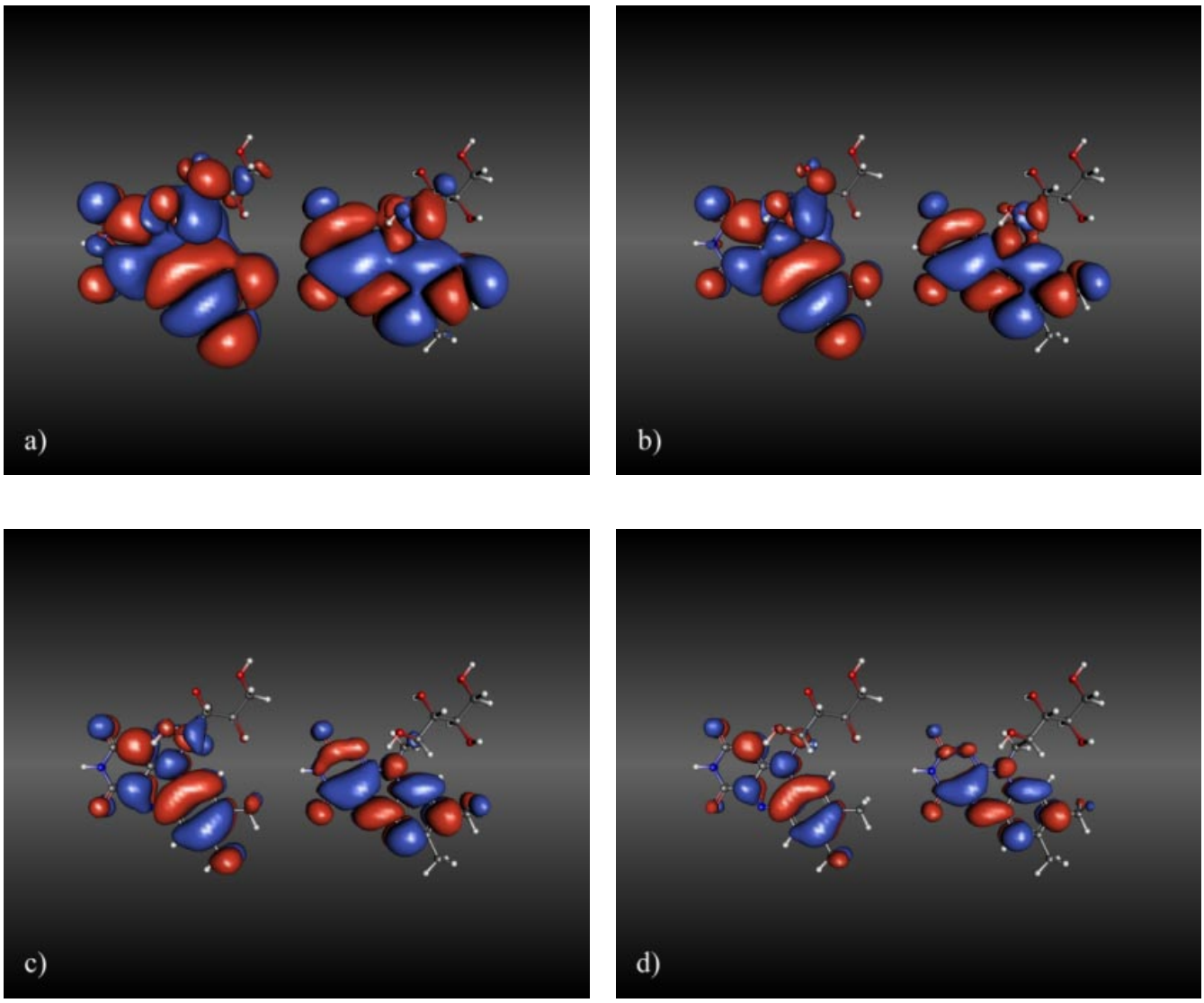


Figure 7 (video 2) *HOMO- (left) and LUMO- (right) wave functions for the vitamin B2 molecule. a-d: Isosurfaces for the negative (blue) and positive (red) amplitude for decreasing iso-values. In the video sequence the interactive change of the iso-value as well as the orientation of the molecules is shown as the result of an interactive treatment*

The reduction of complexity for the enhancement of insight

One of the key questions in the field of computer aided molecular modeling is related to the molecular docking process. The qualitative and quantitative treatment are both controlled by an effective handling of the question of molecular complementarity.

The problem to be solved can be formulated easily as follows: Find structures for complexes AB built from two molecular components A and B in solution for which the free energy $\Delta G_{ass}^{AB} = G_{AB}-G_A-G_B$ takes a (relative or absolute) minimum, and for a given receptor (say A) find those molecules B' out of a large set B for which $\Delta G_{ass}^{AB'} << \Delta G_{ass}^{AB}$ holds. The solution of the general problem has two components, the *computational component* (related to the *computation* of free energy differences $\Delta G_{ass}^{AB'}$) and the *classification component* (the definition of the set B'). In this work, we only stress the second one.

The classification problem deals with the question of how to define for a given molecule A and a reference set B the set B' of *possible* docking partners. This problem is strongly related to the question of molecular similarity or molecular complementarity of the molecule A and those from the set B' in the region of a receptor (if this is known). We are thus looking at the molecules from the point of view of a "molecular inspector" and trying to discover which may belong to a certain class of possible "keys" to fit some given "lock". The search becomes even more complicated when the "lock"

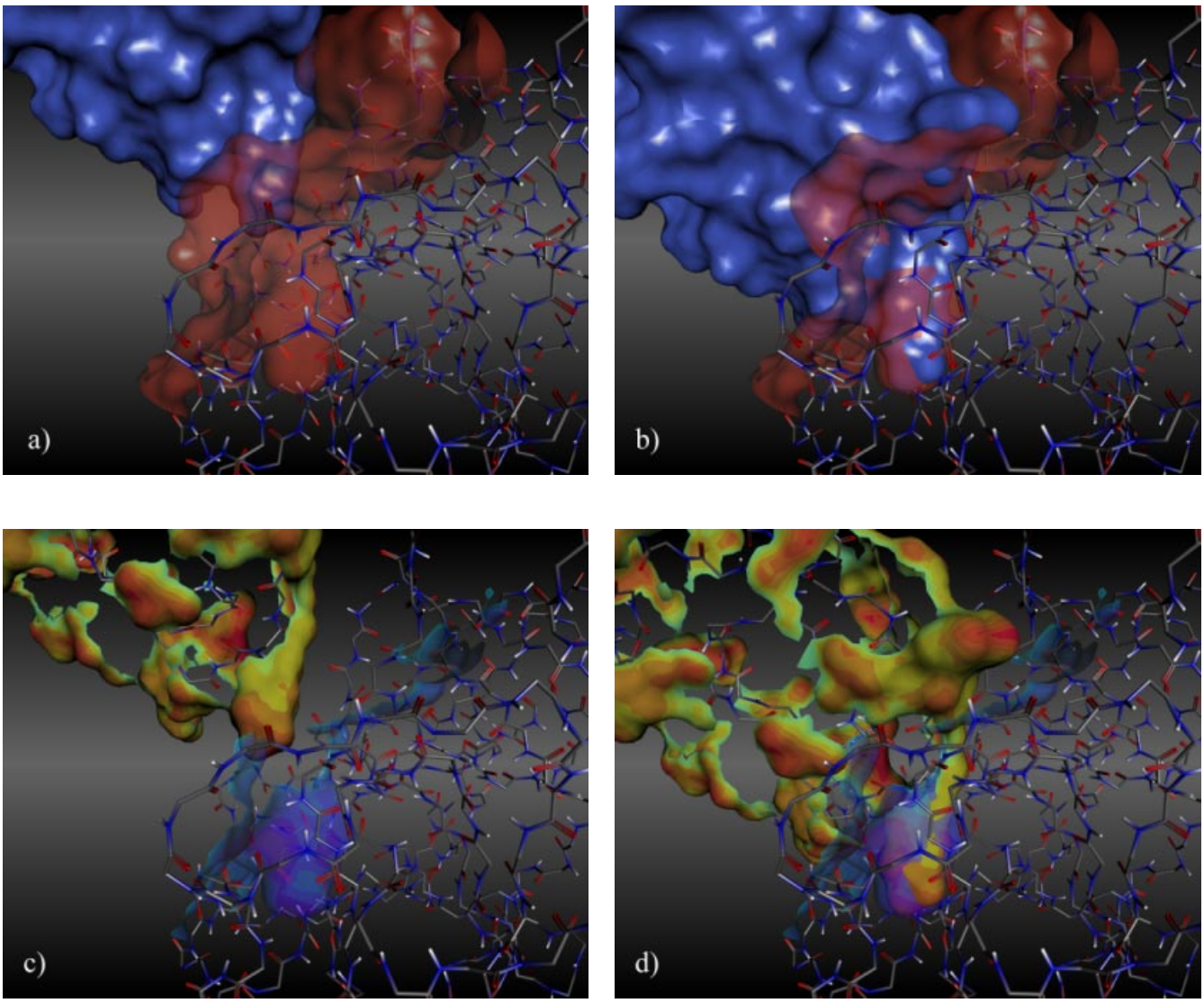


Figure 8 (video 3) *Interactive docking of the proteins trypsine and pancreatic trypsine inhibitor (PTI) using texture mapping and filtering according to the surface topography index (STI [17]). From figs. 8a and 8b it is seen that there is only a minor chance for a success of the interactive docking because of the hidden surface problems. In figs. 8c and 8d this problem is reduced by just showing those areas in both molecules which are complementary in shape according to the STI classification. Video 3 shows the interactivity of the whole process*

cannot be specified. In this case one is looking for complementarity between arbitrary regions of one molecule and regions of a second molecule without any knowledge of the search patterns. For simplicity we restrict the following discussion to the comparison of molecular shapes.

It has been demonstrated in section III that the pattern recognition abilities can be well applied for the investigation of possible shape complementarity of 2D objects. How can this be transferred to 3D objects like proteins? The surfaces of these molecules “look” like the surfaces of highly irregular potatoes. Moreover, for an interactive shape matching there are additional problems which are again (like in the isosurface analysis) related to the hidden surface problem: in order to analyze whether two surface elements are close to each other one has to inspect their interference area. This can only be done if irrelevant parts of the scenario are blended out, i.e if the human searcher has visual access to the relevant surface patches. Such an outblending can again be performed with the aid of interactive texture mapping technology. This is demonstrated in Video 3 (Figure 8). Therein, different areas of the molecular surface can be removed from the graphical representation on the basis of the value of the surface topography index (STI) which varies as $0 \leq \mathrm{STI} \leq 4$ and is calculated from the local canonical curvatures of the molecular surface. The STI values are related to the topography as [17]

STI = 0 → bag STI = 1 → cleft STI = 2 → saddle

STI = 3 → ridge STI = 4 → knob STI = –1 → plateau

i.e. one can select surface areas of a given topography (a knob for example) of one molecule interactively and that of topographical complementarity (a hole in this case) and show the corresponding surface elements. With this reduction of the surface complexity, the eyeball technique can be easily applied. It is important to say that the information filtering can be performed on an interactive time scale, i.e. one can change the surface complementarity criteria during the search.

The transfer of visual information to an algorithm

The example in the last section demonstrates that the reduction of information can help to optimize the use of the eyeball technique for the recognition of shape complementarity. Nevertheless, the "eyeball" technique has a variety of limitations. These are significant in all those cases when there is no way of transforming the scenario into a representation where the human senses are able to recognize data or features. Another limitation is related to the large numbers of objects within a search. If one has to check all molecules stored in a structural database (10^4-10^5 molecular structures) in order to find those molecules which, in principle, can be considered as possible "keys" for a given receptor (set B', see above), the "eyeball" technique will no longer be applicable, simply for pragmatic reasons. Such a search can be done only using the increasing power of modern computational technology.

How can strategies based on human recognition be used for the development of algorithms which can be applied in molecular recognition processes, at least in a preselective manners. This question is somewhat related to a paradigm change in the application of MGRs in the man-machine interaction in the field of molecular science. Where in the traditional use MGRs play an important role in the information transfer process from a simulated world (based on well defined algorithms) to the human brain, the situation is now completely reversed: the abilities of the human senses should be transferred to an activity of a computational process. This is by no means an easy task, because the strategies of the human searchers are not known in all details.

One trial towards a partial solution of the problem has recently been published by the group of the authors [49]. In this approach the concepts of fuzzy logic and, in particular, the introduction of linguistic variables has been applied. Fuzzy set theory has been successfully applied in different areas of pattern recognition and at different stages of the recognition process (for references see Exner et al. [49]). In our formalism the molecular surface are subdivided in surface patches, which can be characterized linguistically according to the STI. The values of the classification variables are knob, ridge, saddle, cleft, bag, and plateau (Figure 9). The fuzzy logic strategy for this subdivision as well as the algorithms for the establishment of surface complementarity were controlled by visual experience of a human searcher. It is clear that the transformation of topographical information to a small set of linguistic variables is a dramatic reduction of complexity. Moreover the reduction method can only be vaguely defined (as a matter of principle). Nevertheless it could be demonstrated [49] with three examples (trypsine-PTI, HLE-ovomucoid inhibitor, α-chymotrypsine – ovomucoid inhibitor) that the proposed method can be very effectively used for the prediction of initial guesses for biomolecular cases particularly in those cases where the binding sites are not known.

Conclusions and perspectives

It has been demonstrated above that molecular graphics representation (MGRs) can be used effectively in many cases in man-machine communication, i.e. the communication between a computational process handling a molecular scenario and the human brain. It also has been demonstrated with a few examples that a straightforward application of computer graphics tools is helpful in any case in the field of molecular graphics. From a computer graphics point of view, there are nearly no limitations for the complexity of a visual representation of a 3D scenario on a screen or even in a three dimensional environment like the CAVE. Today, hardware and software tools can be handled easily without extensive programming effort. One can say that almost all the problems from the side of computer graphics technology are solved. Is this also true for molecular graphics applications? A naive conclusion could be that there is no need for a further develop-

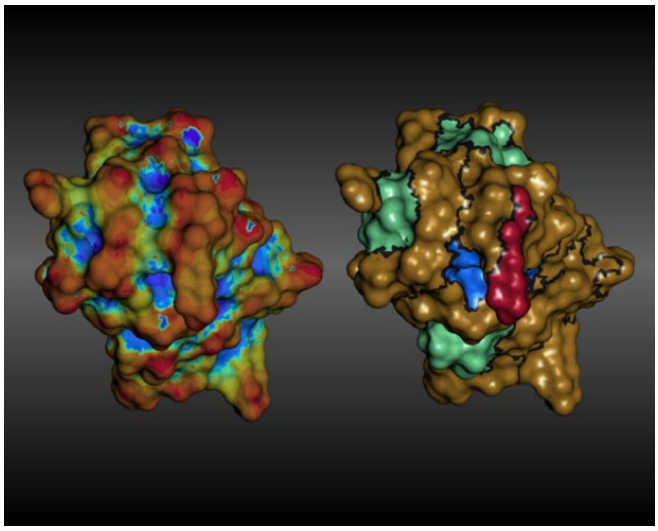


Figure 9 *Molecular surface of the pancreatic trypsine inhibitor (PTI) color coded according to the surface topography index (left) and segmented into surface patches which can be classified with linguistic variables (right). red: knob; brown: ridge; green: saddle; blue: cleft*

ment in this field culminating in an advice to chemists who want to communicate to a computer simulated molecular scenario with the aims to obtain new insights and (or) to obtain quantitative information from the visual inspection (or interaction) process as: Take a standard package for scientific visualization, feed in your molecular information and start to work. This type of applying MGRs may be helpful for all types of standard applications but it does not automatically fulfill the postulates of E. Tufte (see section II), i.e. there is no guarantee for "graphical excellence". The optimization of the man-machine communication in molecular sciences along the criteria outlined in section II requires an active participation of the human searcher in the actual creation process of a MGR. One can see the information behind a flood of data, if the representation is tailored according to the individual needs. This requires a maximum of freedom for the selection of those parts of a scenario which may be optimal for the answer of actual questions and an optimal representation for the best individual reception. At the moment there is no general development line to be seen. Some possible extensions of the standard use of graphical tools and the integration of computational processes have been described in this paper, but this can only be a beginning. There are two aspects which may be used as a guideline for further developments

(1) the introduction of a common "molecular graphical language" and

(2) the introduction of a meta language in order to handle graphical representation and computational processes simultaneously.

The first aspect can be seen as an analog to the introduction of 2D molecular graphs (as representations of 3D structures) which has lead to a substantial increase of the development of chemistry at the beginning of this century – all chemist understand this representation in more or less the same way - and the second as a generalization and quantification of reaction trees.

Acknowledgment This work was supported by the Fonds der Chemischen Industrie, Frankfurt

Supplementary material available Three videos showing different molecular scenarios are available in MPEG-format as supplementary materials. These MPEG-videos can be viewed with different video playback programs. On Windows systems the Microsoft Windows Media Player (available at http://www.microsoft.com/windows/mediaplayer/) and on Unix systems the program mpeg_play (available at http://brmc.berkeley.edu/frame/research/mpeg/mpeg_play.html) can be used.

References

1. Levinthal, W. *Sci. Am.* **1966**, *214*, 42-52.
2. Palmer, T. C.; Hausheer, F. H. *J. Mol. Graphics* **1993**, *6*, 149-154.
3. Jones, T. A. *J. Appl. Chystallogr.* **1978**, *11*, 268-272.
4. Dreiding, A. S. *Helv. Chim. Acta* **1959**, *42*, 1339-1344.
5. Warme, P. K. *Comput. Biomed. Res.* **1977**, *10*, 75-82.
6. Smith, G. M.; Gund, P. *J. Chem. Inf. Comp. Sci.* **1978**, *18*, 207-210.
7. Brickmann, J. *J. Mol. Graphics* **1983**, *1*, 62-67.
8. Brickmann, J. *J. Mol. Graphics* **1983**, *1*, 78-82.
9. Brickmann, J. *Int. J. Quantum Chem.; Quant. Chem. Symp.* **1984**, *18*, 647-659.
10. Heiden, W.; Schlenkrich, M.; Brickmann, J. *J.Comput.-Aided Mol. Des.* **1990**, *4*, 255-269.
11. Heiden, W.; Schlenkrich, M.; Zachmann, C.-D.; Brickmann, J. In *Software Development in Chemistry*; Gasteiger, J., Ed.; Springer: Berlin, Heidelberg, 1989; pp 115-121.
12. Waldherr-Teschner, M.; Goetze, T.; Heiden, W.; Knoblauch, M.; Vollhardt, H.; Brickmann, J. In *Advances in Scientific Visualization*; Post, F. H.; Hin, A. J. S., Eds.; Springer: Berlin, 1992; pp 58-67.
13. Heiden, W.; Goetze, T.; Brickmann, J. *J. Comput. Chem.* **1993**, *14*, 246-250.
14. Brickmann, J.; Goetze, T.; Heiden, W.; Moeckel, G.; Reiling, S.; Vollhardt, H.; Zachmann, C.-D. In *Data Visualization in Molecular Science - Tools for Insight and Innovation*; Bowie, J. E., Ed.; Addison-Wesley: Reading, 1995; pp 83-97.
15. Waldherr-Teschner, M.; Henn, C.; Vollhardt, H.; Reiling, S.; Brickmann, J. *J. Mol. Graphics* **1994**, *12*, 98-105.
16. Heiden, W.; Moeckel, G.; Brickmann, J. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 503-514.
17. Heiden, W.; Brickmann, J. *J. Mol. Graphics* **1994**, *12*, 106-115.
18. Zachmann, C.-D.; Kast, S. M.; Brickmann, J. *J. Mol. Graphics* **1995**, *13*, 89-97.
19. Brickmann, J.; Heiden, W.; Vollhardt, H.; Zachmann, C.-D. In *Proceedings of the 28th Annual Hawaii International Conference on System Sciences*; Hunter, L.; Shriver, B. D., Eds.; IEEE Computer Soc. press: Los Alamos, 1995; pp 273-282.
20. Tufte, E. R. *The Visual Display of Quantitative Information*; Graphics Press: Cheshire, USA, 1983.
21. Tufte, E. R. *Envisioning Information*; Graphics Press: Cheshire, USA, 1990.
22. Cruz-Neira, C.; Sandin, D. J.; DeFanti, T. A.; Kenyon, R. V.; Hart, J. C. *Communications of the ACM* **1992**, *35*, 65-72.
23. *Molecular Recognition Mechanisms*; Delaage, M., Ed.; VCH: Weinheim/New York, 1991.
24. Fischer, E. *Chem. Ber.* **1894**, *27*, 2985-2993.
25. Zimmermann, H. J. *Fuzzy Set Theory and Its Applications*; Kluwer Academic: Dordrecht/Boston, 1991.
26. Blaney, F.; Edge, C.; Phippen, R.; Burt, C. In *Data Visualization in Molecular Science - Tools for Insight and Innovation*; Bowie, J. E., Ed.; Addison-Wesley: Reading, 1995; pp 99-129.
27. Mezey, P. G. *Molecular Surfaces*; VCH: Weinheim, 1993.
28. Perkins, T. D. J.; Dean, P. M. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 155-172.

29. Perkins, T. D. J.; Mills, J. E.; Dean, P. M. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 479-490.
30. Meyer, P. G.; Richards, W. G. *J. Comput.-Aided Mol. Des.* **1991**, *5*, 427-439.
31. Brickmann, J.; Exner, T.; Keil, M.; Marhöfer, R.; Moeckel, G. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R.; Allinger, N. L; Clark, T.; Gasteiger, J.; Kollmann, P. A.; Schaefer III, H. F.; Schreiner, R. P., Eds.; Wiley & Sons: Chichester, 1998; pp 1679-1693.
32. Richards, F. M. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151-176.
33. Connolly, M. *J. Appl. Cryst.* **1983**, *16*, 548-558.
34. Connolly, M. *Science* **1983**, *211*, 709-713.
35. Mezey, P. G. In *Reviews in Computational Chemistry*; Lipkowitz, K.; Boyd, D., Eds.; VCH: Weinheim, 1990; pp 265-294.
36. Langridge, R.; Ferrin, T. E.; Kunz, J. D.; Connolly, M. L. *Science* **1981**, *211*, 661-666.
37. Pang, L.; Lucken, E.; Weber, J.; Bernardelli, G. *J. Comput.-Aided Mol. Des.* **1991**, *5*, 285-291.
38. Karfunkel, H. R.; Eyraud, V. *J. Comput. Chem.* **1989**, *10*, 628-634.
39. Dill, J. C. *Comput. Graphics* **1981**, *15*, 153-161.
40. Catmull, E. Ph.D. thesis, University of Utah, 1974.
41. Crow, F. C. *Comput. Graphics* **1984**, *18*, 207-212.
42. Williams, L. *Comput. Graphics* **1983**, *17*, 1-11.
43. Pixner, P.; Heiden, W.; Merx, H.; Moeckel, G.; Moeller, A.; Brickmann, J. *J. Chem. Inf. Comp. Sci.* **1994**, *34*, 1309-1319.
44. Schmidt, F.; Schilling, B.; Jäger, R.; Brickmann, J. *J. Comput.-Aided Mol. Des.* **1999**, submitted.
45. Lorensen, W.; Cline, H. *Comput. Graphics* **1987**, *21*, 163-169.
46. Max, N. *J. Mol. Graphics* **1984**, *2*, 8-13.
47. Connolly, M. *J. Appl. Cryst.* **1985**, *18*, 499-505.
48. Henn, C.; Teschner, M.; Engel, A.; Aebi, U. *J. Struct. Bio.* **1996**, *116*, 92-96.
49. Exner, T.; Brickmann, J. *Commun. Math. Comput. Chem. (Match)* **1998**, *38*, 53-80.

[1] The idea of "flatland" is based on the classic by A. Square [Edwin A. Abbott] Flatland: A Romance of Many Dimensions (London, 1884)