ELSEVIER

Application of promolecular ASA densities to graphical representation of density functions of macromolecular systems

X. Gironés,* R. Carbó-Dorca,* and P.G. Mezey†

**Institute of Computational Chemistry, University of Girona, Campus Montilivi, 17071 Girona, Catalonia, Spain*
†Department of Chemistry and Department of Mathematics and Statistics, University of Saskatchewan, 110 Science Place, Saskatoon, S7N 5C9 Canada

In this article we report the application of the Promolecular Atomic Shell Approximation (Promolecular ASA) to the graphical representation of the density function (DF) of large macromolecular systems. Promolecular ASA DF, constructed from previously computed and fitted atomic densities, provides a fast and practical representation of Molecular IsoDensity Contours (MIDCOs). These representations can be extended to macromolecular systems composed by >1000 atoms easily and with low computational costs, allowing the visualization of protein DF. The method is at first presented with a small molecule (2,4,6-trinitrophenol), comparing the resulting ASA MIDCOs with direct ab initio *contours. For macromolecular tests the Promolecular ASA densities are also applied to the generation of macromolecular density surfaces of two proteins: myoglobin (2541 atoms) and gene V protein (1362 atoms).* © *2001 by Elsevier Science Inc.*

Keywords: ASA, promolecular densities, density function, MIDCOs, graphical representation, macromolecular systems

INTRODUCTION

The representation and visualization of molecular shape constitute a powerful tool in understanding a molecular system and its chemical behavior or biological response. As molecular shape may be a useful concept to many scientists, several models have been introduced to describe molecular shape.

Concerning local shapes, the earliest quantum chemical approaches are those of the orbital representations of Mulliken.[1–5] Quantum chemistry provides, in general, the most reliable shape representations,[6] strongly connected to molecular similarity concepts based on quantum mechanics.[7,8] Novel macromolecular Quantum chemistry approaches provide new linear time-dependent computational techniques and *ab initio* quality results for truly large systems, including proteins.[9–13] The special aspects of local shape vs. global shape are reflected in The Holographic Electron-Density Theorem.[14]

Shape analysis on a simple level also has a rich history. A first crude approximation to molecular shape is given by the wire-frame model, where the molecule is simply represented using vectors corresponding to the bonds. This model can be improved using the ball-and-stick model, which constructs the molecular structure by means of small balls, representing the existing atoms in the molecule, linked by cylinders, accounting for the bonds (simple, double, triple) between the constituent atoms. In this way ball-and-stick models provide a discrete vision of molecular shape using atoms and bonds as primary building blocks. An alternative to this last model is ribbon representation of the molecular structure, which is mainly used in biochemistry to visualize proteins.[15] These models, however, become inappropriate when the molecular surface is to be explored, since they are unable to create a valid representation of the effective occupied space in the molecular environment.

A common and simple procedure to obtain a molecular surface is the space-filling model, which assigns to each atom a sphere of its Van der Waals radius. This empirical method provides a better understanding of the molecular shape, as it accounts for some individual properties of the corresponding atoms. Within this procedure the molecular shape is repre-

Corresponding author: R. Carbó-Dorca, Institute of Computational Chemistry, University of Girona, Campus Montilivi, 17071 Girona, Catalonia, Spain. Tel.: +34 972418367; fax: +34 972418356.

E-mail address: director@iqc.udg.es

Journal of Molecular Graphics and Modelling 19, 343–348, 2001

655 Avenue of the Americas, New York, NY 10010

1093-3263/01/$–see front matter
PII S1093-3263(00)00083-8

sented by a set of fused spheres. Other empirical approaches to molecular surface are the solvent accessible[16] and Connolly surfaces.[17] The solvent accessible surface corresponds to the molecular envelope of the surface of each atom that is accessible to the center of a probe sphere of a given radius. The Connolly surface is composed of two kinds of surface patches: the part of the Van der Waals surface of each atom that is accessible to a probe sphere of a given radius, generally 1.4 Å (sphere including most of a water molecule), and the inward-facing part of the probe sphere when it is simultaneously in contact with more than one atom (the reentrant surface).

However, these methods lack realism when representing the molecular shape as they truncate the existing interactions of the molecule. A practical way to introduce better molecular shape representations is the use of the system's electronic density function (DF). The DF describes the actual molecular electronic distribution and reflects its real, fuzzy characteristics, hence DF provides a suitable molecular shape representation.[6] Among the several strategies to obtain DF, the most used are the *ab initio* procedures, which are implemented in available quantum chemistry software packages, such as Gaussian 98[18] or Gamess,[19] with an assorted collection of basis sets. Even though this methodology provides precise DF, insofar as the basis used is accurate enough, the procedure requires a large amount of computer time, which implies that large systems may not be accurately described within reasonable time limits. So far, direct *ab initio* calculations are presently restricted to molecules containing no more than 150 atoms, with the exception of periodical systems.

To overcome this problem, some procedures have been described to obtain accurate DF in computationally inexpensive ways. Successful linear, time-dependent *ab initio* quality approaches include the MEDLA[20,21] and ADMA methods.[13] Other techniques use data from experimental sources or from previous theoretical calculations to simplify the computation. Alternative semi-empirical methodologies, even though they produce valuable results, fail when they deal with molecular features that were not previously parameterized. A related procedure, developed in our laboratory, is the Promolecular Atomic Shell Approximation (Promolecular ASA),[22] which will be explained below. The major use of Promolecular ASA DF has been devoted to Molecular Quantum Similarity Measures,[23] however, the successful application of Promolecular ASA densities to molecular shape representations was demonstrated in a previous work.[24] It was shown that Promolecular ASA DF were capable of generating isodensity contours comparable to those obtained by direct *ab initio* methods, while only using a fraction of the time and computational resources. This method opened a door to the possible use of Promolecular ASA densities to the simple generation of Molecular IsoDensity COntours (MIDCOs). The intention of this article, as a continuation of the previous study, is to prove the applicability of Promolecular densities as an easy, inexpensive method for representing ASA MIDCOs of large macromolecular systems.

The article is organized as follows. First, the main characteristics of the Promolecular ASA approach are detailed and illustrated with a graphical example. Next, MIDCOs for two proteins, myoglobin and gene V, obtained with the exposed methodology, are presented and commented upon. Finally, the conclusions of this work, as well as proposals for future research are given.

PROMOLECULAR ATOMIC SHELL APPROXIMATION (PROMOLECULAR ASA)

The theoretical basis of Promolecular densities[22] overlays the ideas of the ASA approach.[25] ASA is a model in which the molecular electron DF is fitted with spherical 1*S*-type Gaussian functions to the original *ab initio* DF. This methodology allows the simplification of the mathematical expression of the DF, thus reducing the computational cost of the subsequent computations, with a reasonable error margin.

As the fitting procedure to obtain the ASA DF was computationally tedious and, in addition, a previous *ab initio* calculation was needed, this methodology was hard to extend to large molecular systems. To expand the applicability of ASA, another simplification was applied, resulting in the Promolecular ASA approach. The Promolecular approach considers the whole molecular DF as a linear combination of discrete atomic contributions, which in turn are expressed as linear combinations of spherical 1*S*-like Gaussian functions. The Promolecular ASA DF can be mathematically formulated as:

$$\rho_A^{PASA}(\mathrm{r}) = \sum_{a \in A} P_a \rho_a^{ASA}(\mathrm{r} - \mathrm{R}_a), \tag{1}$$

where P_a is the atomic number of the a-th atom. In this way, the molecular density is a simple addition of atomic densities, ρ_a^{ASA}, centered at the atomic positions R_a. These atomic densities can be expressed as:

$$\rho_a^{ASA}(\mathrm{r} - \mathrm{R}_a) = \sum_{i \in a} w_i |S_i(\mathrm{r} - \mathrm{R}_a)|^2, \tag{2}$$

where S_i are the normalized spherical 1*S* Gaussian functions and w_i are the coefficients of the linear expansion, which must fulfill the following conditions:

$$w_i \in \mathrm{R} \ \wedge \ w_i > 0, \ \forall \, i \in a, \tag{3}$$

to guarantee that the DF is positive definite in its whole range, and

$$\sum_{i \in a} w_i = P_a, \tag{4}$$

to assure that the integral of ρ_a^{PASA} over all space is the number of electrons of molecule *A*, thus ensuring the proper definition of charge distribution function.

As seen in Equation 1, Promolecular ASA densities require atomic ASA DF. For this reason and for each atom, *ab initio* calculations and fitting computations have been carried out to parameterize the optimal exponents of the S_i functions and coefficients w_i. These results are stored in a database for later calculations.[26] So, the construction of Promolecular ASA DF is done simply by means of these previously computed ASA densities and atomic coordinates. Molecular coordinates can be extracted from many sources, such as crystallographic data or theoretical calculations.

A Graphical Example

To illustrate the validity of Promolecular densities when representing MIDCOs, a graphical example is provided. The chosen molecule is 2,4,6-trinitrophenol, which presents interesting

features such as aromaticity, due to the presence of the benzene ring and an internal hydrogen bond. The DF of the selected molecule has been constructed in two different ways. The first involves an *ab initio* method computed with Gaussian 98,[18] in this case the calculation chosen was a restricted Hartree-Fock calculation using the 6-31G basis set. The second one uses the Promolecular ASA approach, employing previously computed functions present in the exponent and coefficient database.[26]These parameters were obtained from fitting atomic calculations using the 6-31G basis set to spherical 1*S* Gaussian functions. In both cases, the employed molecular geometry is the same and has been optimized using the AM1 Hamiltonian with the Ampac 6.01 software package.[27]

From the computed DF, the corresponding MIDCOs that will be used to analyze the differences between the two methodologies are created. MIDCOs are created by evaluating the DF in a grid. The molecule is placed on a grid of points large enough to adequately envelop the molecule. At each point of the grid, the corresponding DF is evaluated and stored. Then a curvature analysis is performed over the previously computed points. The curvature analysis is used to determine features of the molecular shape, like convexity and concavity in the molecular surfaces, and to observe how these features change by modifying the isodensity level. The different representations have been graphically plotted with AVS software[28] and are presented in Table 1.

As seen in the graphical representations in Table 1, there are no significant differences between the Promolecular ASA MIDCOs and the respective *ab initio* ones. At first sight, high-density levels (>0.300) reflect the atomic positions in space, intermediate density levels ($0.150-0.300$) define the bonding pattern, and the low-density levels outline the volume and steric features of the molecular shape for both methods. However, a closer inspection of the figures reveals that the carbon–carbon bonds are manifested earlier in the *ab initio* representation at a density value of about 0.300, whereas there is no sign of them in the Promolecular approach at the same density level. In the Promolecular density, hydrogens appear at a lower density value, 0.209, whereas they are already defined in the *ab initio* density. These facts are due to the nature of the Promolecular ASA method. As DF is collapsed at the nuclear positions with 1*S* Gaussian functions, the bond pattern appears at a later stage, compared with the *ab initio* procedure, which includes higher-quantum number functions and consequently offers a better description. At the same time, the description of hydrogen atoms within the Promolecular framework is not accurate enough, when compared with the *ab initio* description.

However, these differences may not seem important if the reduced time needed to obtain the representations is taken into account. While the Promolecular density grid was calculated in seconds, the *ab initio* procedure lasted about 20 minutes using the same computer (a DEC Alpha 3000 with 64 MB RAM). In addition, the Promolecular ASA DF was constructed nearly instantly with the coefficients and exponents of the database,[26] whereas the *ab initio* calculation needed about 35 minutes to be completed. Furthermore, MIDCOs obtained for high-density levels are similar enough to perform qualitative comparisons. For low-density levels, figures from both methodologies become almost identical.

Besides the MIDCOs comparison, curvature analysis confirms the applicability of Promolecular densities to represent molecular shape. The regions of the surfaces close to the

Table 1. Graphical comparison and curvature analysis of 2,4,6-trinitrophenol, computed with Promolecular ASA and *ab initio* levels

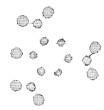
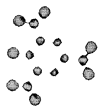
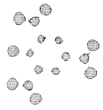
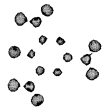
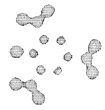
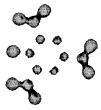
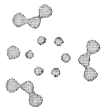
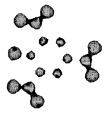
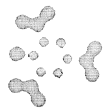
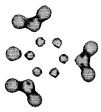
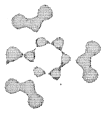
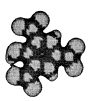
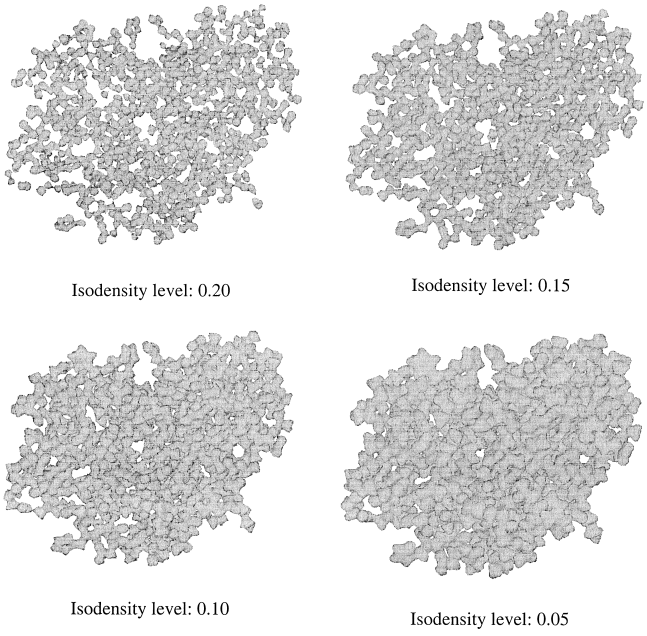
	Promolecular ASA		*Ab initio*	
Density Value	Density function	Curvature analysis	Density function	Curvature analysis
0.500				
0.400				
0.300				
0.209				
0.170				
0.150				
0.100				
0.050				
0.045				
0.037				
0.020				

Table 2. Representation of myoglobin MIDCOs



atomic locations present a convex shape, whereas the intermediate vicinities evidence a locally saddle-type appearance (D_1 domains in the terminology of Mezey[6]). As the density level decreases, the convex regions tend to grow until a completely convex surface would be obtained for very low-density levels. The curvature tendencies for both methods are entirely equivalent, evidencing the validity of Promolecular ASA DF to provide realistic and inexpensive MIDCOs.

Construction of Macromolecular MIDCOs

In this section, we intend to show the applicability of Promolecular ASA DF as a powerful tool for creating MIDCOs of macromolecular systems. As examples, two proteins will be presented: myoglobin and gene V protein. The calculation involves the same steps as the previous example:

- obtaining the atomic coordinates
- constructing of the Promolecular ASA DF
- creating a grid large enough to completely surround the selected molecule
- calculating the DF value at each grid point
- representing the obtained MIDCOs with the AVS program.[28]

The atomic coordinates have been extracted from the Protein Data Bank (PDB),[29] the PDB Id of the proteins are 1MBN for myoglobin and 1VQB for gene V protein. The original files contained several solvating water molecules, which were removed before the calculations were carried out. The construction of the Promolecular DF has been done using the parameters present in our database,[26] but in this case the fitting parameters used correspond to the 3-21G basis set, even though this basis set offers a poorer description. It was also demonstrated in a previous work[24] that acceptable ASA MIDCOs could be derived from such fitting. In this way, the number of functions used is small, which translates to a reduction of computational time while still preserving a good visualization of the main, important features of the molecular shape, as will be seen in the examples.

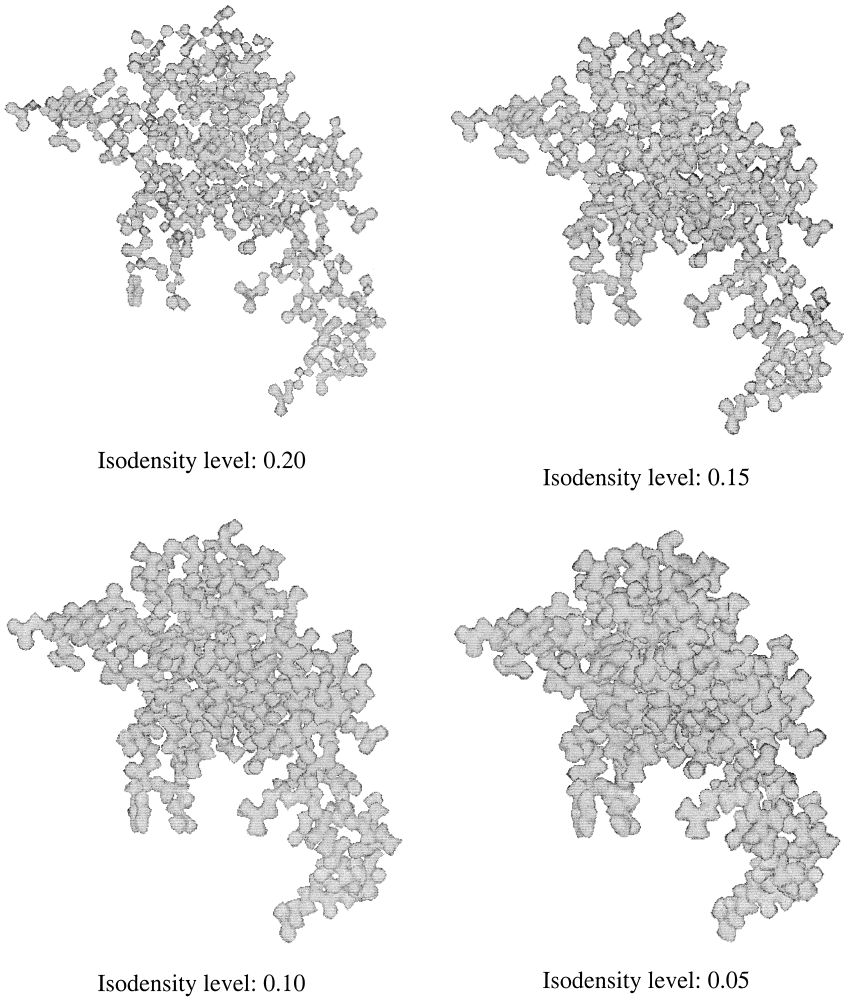
The first example of this is myoglobin, a protein that stores molecular oxygen. Kendrew and Watson,[30] first described this protein by extracting it from whale sperm, as myoglobin is present in larger quantities in marine mammals due to their extra oxygen requirements when doing deep dives under sea. The group responsible for this capacity is an iron atom encased in a heme group; this protein is also called ferric iron metmyoglobin. This molecule was chosen because it was the first protein whose atomic structure could be determined, and thus it laid the foundation for an era of biological understanding. This protein is formed by 2,541 atoms and 5,065 1*S* Gaussian functions are needed to construct de Promolecular ASA DF. The dimensions of the grid of points were $91\times83\times85$ atomic units (au) and the spacing between the points (resolution) was chosen to be 1 au. The MIDCOs obtained for myoglobin are presented in Table 2.

As seen from the MIDCOs presented in Table 2, the representation of the molecular shape shows the same tendencies as the previous 2,4,6-trinitrophenol. At high-density values the atomic positions are visualized, as are high-order bonds. Intermediate values provide the bonding pattern and low-density levels reflect the bulkiness of the molecular system. These graphical representations can provide valuable information about reaction centers or pathways due to the volume of information they provide, accounting for the realistic sterical features of MIDCOs. In addition, these molecular representations are easily obtained with low computational costs. In this case, the evaluation of the Promolecular DF at all grid points required about 9 hours to complete in the test machine, whereas a direct *ab initio* calculation and DF evaluation for this system would last years in the most modern computers.

The second example presented consists of the gene V, a DNA binding protein. The structure of this protein was determined by Terwilliger et al.,[31] and was extracted from Bacteriophage F1. This molecular structure was chosen because its MIDCOs were previously studied by Mezey et al. within the MEDLA approach.[20,21] hence these methodologies can be compared. Briefly, MEDLA can construct in linear time *ab initio* quality MIDCOs of large molecules from previously computed fuzzy molecular fragments, which when adequately translated and rotated reproduce the entire molecular DF. The molecular fragments are chosen to posses environment similar to that of the final molecule, assuming that the fuzzy DF is almost transferable. This procedure requires calculating many small molecular systems, regarded as building blocks, at *ab initio* levels, which requires low computational costs. In addition, as the fragments resulting from these calculations are stored in a database, they can be used in the construction of other MIDCOs. In contrast, in the ASA method, the gene V protein is formed by 1,362 atoms and 2,709 1*S* Gaussian functions are needed to fit the DF at the 3-21G basis set. In this case, the dimensions of the grid box were $95\times107\times83$ au and the resolution, as in the previous example, has been set to 1 au. The obtained ASA MIDCOs for gene V protein are presented in Table 3.

As seen in Table 3, the ASA MIDCOs obtained follow the same model as the previous examples. As the DF level decreases, atomic locations, bonding patterns, and volumetric information are described. When the obtained ASA MIDCOs are compared

Table 3. Representation of gene V MIDCOs



with the previously studied ones,[21] no relevant differences arise at first. However, because the MEDLA approach [20,21] is based on fuzzy density fragments, it is better able to reflect the interatomic interactions. Promolecular ASA, as found in discrete atomic densities, results in a poorer description of these features. However, as only these precomputed atomic contributions are required in Promolecular ASA, there is no need to recompute them and thus the fitted DF obtained can be used throughout regardless of the molecular environment.

CONCLUSIONS

We have demonstrated the practical application of Promolecular ASA DF to the graphical representation of MIDCOs. Using this simplified model to obtain DF has been proved to achieve valuable representations of the molecular shape and opens a new door for the potential use of fitted densities in the scientific areas where accurate descriptions of molecular surfaces are needed. Future applications of Promolecular ASA densities, besides Quantum Molecular Similarity calculations, are intended to be developed. These new capabilities will be focused on obtaining various molecular properties such as electrostatic potentials and interaction energy surfaces. Considering the low computational requirements of the Promolecular Approach, these new features will also be easily applied to macromolecular systems.

ACKNOWLEDGMENTS

This research was completed while X. Gironés stayed at the University of Saskatchewan, and has been supported by a European Commission Project No. ENV4-CT97-0508. X.G. benefited from a predoctoral fellowship from the University of Girona. We also acknowledge financial support from the Fundació Maria Francisca de Roviralta and from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1 Mulliken, R.S. Electronic population analysis on LCAO molecular wave functions. I. *J. Chem. Phys.* 1955, **23**, 1833–1840
2 Mulliken, R.S. Electronic population analysis on LCAO molecular wave functions. II. Overlap populations, bond

orders, and covalent bond energies. *J. Chem. Phys.* 1955, **23**, 1841–1846
3 Mulliken, R.S. Electronic population analysis on LCAO molecular wave functions. III. Effects of hybridization on overlap and gross AO populations. *J. Chem. Phys.* 1955, **23**, 2338–2342
4 Mulliken, R.S. Electronic population analysis on LCAO molecular wave functions. IV. Bonding and antibonding in LCAO and valence-bond theories. *J. Chem. Phys.* 1955, **23**, 2343–2346
5 Mulliken, R.S. Criteria for the construction of good self-consistent-field molecular orbital wave functions, and the significance of LCAO-MO population analysis. *J. Chem. Phys.* 1962, **36**, 3428–3439
6 Mezey, P.G. *Shape in chemistry: an introduction to molecular shape and topology*. VCH Publishers, New York, 1991
7 Carbo, R. (Ed.) *Molecular similarity and reactivity: from quantum chemical to phenomenological approaches*. Kluwer Acad. Publ., Dordrecht, The Netherlands, 1995
8 Johnson, J.A., and Maggiora, G.M. (Eds.) *Concepts and applications of molecular similarity*. Wiley, New York, 1990
9 Mezey, P.G. Quantum chemistry of macromolecular shape. *Int. Rev. Phys. Chem.* 1997, **16**, 361–388
10 Mezey, P.G. Shape analysis. In: *Encyclopedia of computational chemistry*. Schleyer, P.v.R., Allinger, N.L., Clark, T., Gasteiger, J., Kollman, P.A., Schaefer III, H.F., and Schreiner P.R., Eds., John Wiley & Sons, Chichester, UK, 1998, Vol. 4, 2582–2589
11 Mezey, P.G. Computational microscopy: pictures of proteins. *Pharmaceutical News* 1997, **4**, 29–34
12 Mezey, P.G. Functional groups in chemistry. *Adv. Quant. Chem.* 1996, **27**, 163–222
13 Mezey, P.G. Quantum similarity measures and Löwdin's transform for approximate density matrices and macromolecular forces. *Int. J. Quantum Chem.* 1997, **63**, 39–48
14 Mezey, P.G. The holographic density theorem and quantum similarity measures. *Mol. Phys.* 1999, **96**, 169–178
15 Richardson, J.S. The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* 1981, **34**, 167–339
16 Lee, B., and Richards, F.M. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* 1971, **V55**, 379–400
17 Agishtein, M. E. Fuzzy molecular surfaces. *J. Biomolec. Struct. Dynamics* 1992, **9(4)**, 759–768
18 Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Zakrzewski, V.G., Montgomery, Jr., J.A., Stratmann, R.E., Burant, J.C., Dapprich, S., Millam, J.M., Daniels, D., Kudin, K.N., Strain, M.C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G.A., Ayala, P.Y., Cui, Q., Morokuma, K., Malick, D.K., Rabuck, A.D., Raghavachari, K., Foresman, J.B., Cioslowski, J., Ortiz, J.V., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R.L., Fox, D.J., Keith, M. T., Al-Laham, A., Peng, C.Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P.M.W., Johnson, B., Chen, W., Wong, M.W., Andres, J.L., Gonzalez, C., Head-Gordon, M., Replogle, E.S., and Pople, J.A. Gaussian 98, Revision A.6, Gaussian, Inc., Pittsburgh PA, 1998
19 Schmidt, M.W., Baldridge, K.K., Boatz, J.A., Elbert, S.T., Gordon, M.S., Jensen, J.H., Koseki, S., Matsunaga, N., Nguyen, K.A., Su, S., Windus, T.L., Dupuis, M., and Montgomery, J.A. General atomic and molecular electronic structure system. *J. Comput. Chem.* 1993, **14**, 1347–1363
20 Walker, P.D., and Mezey, P.G. Molecular electron density lego approach to molecule building. *J. Am. Chem. Soc.* 1993, **115**, 12423–12430
21 Walker, P.D., and Mezey, P.G. Ab initio quality electron densities for proteins: A MEDLA approach. *J. Am. Chem. Soc.* 1994, **116**, 12022–12032
22 Amat, L., and Carbó-Dorca, R. Quantum similarity measures under atomic shell approximation: First order density fitting program using elementary Jacobi rotations. *J. Comput. Chem.* 1997, **18**, 2023–2039
23 Carbó, R., Leyda, L., and Arnau, M. How similar is a molecule to another? An electron density measure of similarity between two molecular structures. *Int. J. Quantum Chem.* 1980, **17**, 1185–1189
24 Gironés, X., Amat, L., and Carbó-Dorca, R. A comparative study of isodensity surfaces using ab initio and ASA density functions. *J. Mol. Graph. Model.* 1998, **16**, 190–196
25 Constants, P., and Carbó, R. Atomic shell approximation. Electron density program algorithm restricting coefficients to positive values. *J. Chem. Inf. Comput. Sci.* 1995, **35**, 1046–1053
26 Promolecular ASA coefficients and exponents for an assorted set of atoms and for different basis sets can be consulted and downloaded in our web site, http://iqc.udg.es/cat/similarity/ASA/funcset.html
27 Ampac, version 6.0, 1994, Semichem, 7128 Summit, Schawnee, KS 66216DA
28 AVS, rev. H, 1994, Advanced Visual Systems Inc., 300 Fifth Avenue, Waltham, MA 02154, USA
29 The PDB can be consulted through its website: http://www.rcsb.org/pdb/index.html
30 Watson, H.C., The stereochemistry of the protein myoglobin. *Prog. Sterochem.* 1969, **4**, 299–303
31 Skinner, M. M., Zhang, H., Leschnitzer, D.H., Guan, Y., Bellamy, H., Sweet, R.M., Gray, C.W., Konings, R.N., Wang, A.H., and Terwilliger, T.C. Structure of the gene V protein of Bacteriophage F1 determined by multiwavelength X-ray diffraction on the selenomethionyl protein. *Proc. Nat. Acad. Sci. USA.* 1994, **91**, 2071–2075.