## Chapter 9

## Molecular Graphics



MolArch ${ }^{+}$- Molecular Architecture Modelling Program
S. Immel,
http://caramel.oc.chemie.tu-darmstadt.de/immel/.

## MolArch $^{+}$- Molecular Architecture Modelling Program

S. Immel, http://caramel.oc.chemie.tu-darmstadt.de/immel/.

Many of the experimental and molecular modelling studies described in the preceding Chapters of this work required highly specific, computer-assisted tools for processing and visualizing molecular data. For this purpose, the preliminary Molecular Architecture Modelling Program MolArch ${ }^{+[1]}$ has been developed for analyzing geometries of molecular structures and generating high-quality color graphics and animations. Main emphasis has been put on the possibility to automatically process large numbers of structure files and molecular configurations without the need to manually intervene at each step. The basic idea was to provide tools for the efficient and fast visualization, particularly including the animation of molecular ensembles from pre-computed data for the purpose of scientific and educational presentations.

Besides basic graphical representations, MolArch ${ }^{+}$was designed to visualize at low-cost advanced objects such as atomic ellipsoids, non-protein ribbon models, molecular surfaces, and 3D-grid based properties with modes that are not available with other programs, or which have only become available latter with expensive commercial applications. In an effort to not copy computational procedures such as molecular mechanics (MM) and molecular dynamics (MD), MolArch ${ }^{+}$provides interfaces or shared file formats for data exchange with numerous applications used for molecular modelling (structure generation and optimization) and crystallography. Although numerous programs for the computation of molecular properties are available for UNIX and LINUX operating systems, there is still a need for visualization software on these computers. Of particular importance was the use of highest-quality color graphics provided by external ray-tracing programs without restrictions on the resolution of the final graphics. All display parameters and settings used by MolArch ${ }^{+}$ had to be user-definable in order to provide a highly versatile visualization tool.

This Chapter briefly describes the options offered by MolArch ${ }^{+}$software package (current version v7.50 24.Nov.2003), and provides some examples of various types of molecular graphics that can be generated. ${ }^{[1]}$

## Technical Notes

The MolArch ${ }^{+}$program is written almost completely in FORTRAN with some minor parts in C, containing about 125,000 lines of source code. The program was designed for LINUX computer systems and was tested on RedHat and SuSe distributions of this operating system (releases $7 . x$ or latter), but may also be used with other LINUX distributions; it has also been exported to SGI (IRIX) and IBM (AIX) machines. MolArch ${ }^{+}$only features a very simple graphics display based on the

VOGLE ("Very Ordinary Graphics Library") interface without graphical menus, but a powerful command line. All commands and options may be entered in abbreviated form via this command line, or may be specified directly on the controlling shell console. The advantages of this design are full batch processing capabilities and the possibility to automatically carry out recurring tasks on large ensembles of molecular structure files.

## Supported File Types and External Modules

Molecular structures, and related objects may be imported or exported by MolArch ${ }^{+}$through a large variety of different file formats. The following list provides an overview on the most important file types supported (file name extensions in parenthesis):

- Protein Data Bank ${ }^{[2]}$ (PDB) files as the most commonly used file type for molecular structures.
- Files used by the force-field programs $P^{\prime} M^{[3]}$ (EIN, OPT), MacroModel ${ }^{[4]}$ (OUT, MAC), CHARMm ${ }^{[5]}$ (RST, DCD), and HyperChem ${ }^{[6]}$ (HIN).
- Crystallographic Information Files ${ }^{[7]}$ (CIF) and ShelIX ${ }^{[8]}$ (INS) data files of solid-state structures of organic and inorganic compounds.
- Crystal structure data files exported from the data base maintained by the Cambridge Crystallographic Data Centre ${ }^{[9]}$ (CCDF, FDAT file type); file type used by the crystallographic programs Conquest ${ }^{[10]}$ and Mercury. ${ }^{[11]}$
- Spartan ${ }^{[12]}$ and Gaussian ${ }^{[13]}$ output files (including atomic charges); for Gaussian 3D-grid and density properties (electron densities, molecular orbitals, potential data, etc.) can be imported from the cube grid files (CUB).
- Various MolArch ${ }^{+}$specific file formats for molecular structures, surfaces, related objects, and script (batch) processing files.

MolArch+ also includes interfaces to the following external programs and modules:

- Interface to the force-field program $P I M M^{[3]}$ for the energy optimization of geometries of organic structures.
- MolCont ${ }^{+[14]}$ program for 2D- and 3D-contouring for orthogonal and polar coordinate data sets.
- MolSurf ${ }^{+[15]}$ script for the generation of Hirshfeld- (crystal packing) ${ }^{[16]}$ and Connolly-type (contact and solvent-accessible $)^{[17]}$ molecular surfaces and surface mapped properties; molecular surfaces can be triangulated by modified procedures based on algorithms described by Brickmann et al. ${ }^{[18]}$
- Mo/Grid ${ }^{+[19]}$ program for the generation adiabatic energy potential surfaces by the method of "prudent ascent ${ }^{[120]}$ using the force-field program PIMM.
- MolSymm ${ }^{+[21]}$ program and external module ${ }^{[22]}$ for the analysis of molecular symmetry, symmetry elements, and point groups.
- XFarbe ${ }^{[23]}$ program for the visualization of 2D-data sets; external module for finding maxima, minima, and saddle points on energy potential surfaces; external modules used by MolArch ${ }^{+}$for generating molecular geometries along reaction coordinates.
- GnuPlot ${ }^{[24]}$ program for scientific plotting of data sets.

The programs MolCont ${ }^{+}$and MolSymm ${ }^{+}$are directly interfaced to MolArch ${ }^{+}$; the other programs mentioned may be used separately to process data.

## Geometry Analysis Options

The MolArch ${ }^{+}$program may be used for the analysis of a wide range of geometry parameters of molecular structures which are listed below:

- Standard interatomic distances and bond lengths, angles, torsion angles, as well as virtual angles and torsions between vectors.
- Puckering amplitudes of rings relative to least-squares (LSQ) best-fit mean ring planes and out-of-plane deviations.
- Tilt angles ${ }^{[25]}$ between rings and LSQ-planes; flip angles ${ }^{[26]}$ for out-of-plane deviations.
- Cremer-Pople parameters ${ }^{[27]}$ for five- ${ }^{[28]}$ and six-membered ${ }^{[29]}$ rings (including the automatic detection of cyclopentane, cyclohexane, furanoid, and pyranoid ring conformations).
- $\mathcal{J}_{\mathrm{H}-\mathrm{C}-\mathrm{C}-\mathrm{H}}{ }^{[30]}$ and $\mathcal{J}^{3} \mathrm{CO-C-H}{ }^{[31]}$ NMR coupling constants.
- Automatic detection of Cahn-Ingold-Prelog $(R / S)$ stereo descriptors for tetrahedral centers, bond-orders (single, double, triple, and quadruple bonds), aromaticity (for rings), and of hydrogen-bonds and their geometry parameters (including symmetry descriptors for crystal structures).
- Molecular shape parameters such as mean-radius of rings, radius of gyration, circularity and asphericity ${ }^{[16]}$ (for both molecular structures and surfaces).
- Molecular connectivity indices ${ }^{[32]}(\chi)$ of various order.
- Least-squares 3D-fitting ${ }^{[33]}$ and superimposition of series of molecules or molecular fragments (sub-structures) with different weighting schemes.

Most geometry parameters can be manipulated and reset to user defined values (distances, angles, torsions, tilt and flip angles, puckering, etc.) in the structure editing section of MolArch ${ }^{+}$; atom types and names can be changed. All methods related to molecular geometries (analysis and manipulation) can be applied to single molecules or series of molecular configurations with the option of automated averaging of data sets. Most notably, MolArch ${ }^{+}$offers the possibility to simultaneously manipulate or analyze and average geometry parameters of similar or identical (symmetry related) molecular fragments in automated procedures.

## Visualization and Molecular Graphics

All molecular graphics contained within this work have been generated using MolArch ${ }^{+}$. Although MolArch ${ }^{+}$features a simple graphics display only, the program is fully interfaced to produce highest-quality color models with the POVRAY ("Persistence of Vision") ${ }^{[34]}$ ray-tracing program or any VRML ("Virtual Reality Modelling Language" for 3D molecular models) ${ }^{[35]}$ viewing program. The latter option is of particular use in the process of presenting 3D-structures on the internet with the possibility of rotating and modifying these options on the client side. A large number of different viewing modes is available, and almost all display parameters may be modified via user specific settings. Figure 1 provides a overview on the various modes of visualization that are available with MolArch ${ }^{+}$; many more examples can be found in the original publications in the preceding Chapters of this work.


Figure 1 (continued on the next pages). Various examples for original molecular graphics generated using MolArch ${ }^{+}$; all models were finally rendered using the POVRAY ray-tracing program. - (a) Basic modes for the representation of structures (wire, capped-stick, ball-and-stick, and CPK-type models of pyridine); - (b) Labeling of selected geometry parameters in molecular structures (distances, angles, and virtual torsion angles of a $\beta$-turn mimeticum); - (c) Symmetry elements (center of inversion $i$, mirror planes, $C_{n}$ rotary and $S_{n}$ rotary reflection axes) for molecular geometries $\left(\left[\mathrm{Fe}(\mathrm{Pyr})_{6}\right)\right]^{2+}$, point group $\left.T_{\mathrm{h}}\right)$; - (d) Transition state modelling with semi-transparent bonds (Diels-Alder reaction of cyclopentadiene with acetylene dicarboxylic acid dimethyl ester); - (e) Molecular geometry and coordination polyhedra of organometallic structures (trigonal-bipyramids and pentagonal-pyramids in the structure of $\left.\mathrm{PtRh}_{4}(\mathrm{CO})_{4}\left(\mathrm{Cp}^{*}\right)_{4}\right)$; - (f) Inorganic solid-state structures and coordination polyhedra (tetrahedrons and octahedrons for $\mathrm{CulnOPO}_{4}$ ); - (g) Anisotropic thermal ellipsoids of crystal structures (X-ray structure of per-2,3-anhydro- $\alpha$-cyclomannin methanol hydrate); - (h) Crystal structures including transparent Hirshfeld surfaces visualizing the mode of molecular packing (two unit cells for the lowtemperature solid-state structure of acetylene); - (i) Molecular surfaces with color-coded mapped properties (molecular electrostatic potential of benzene); - (j) Clipped surfaces (benzene, surface clipped at iso-contour values of the electrostatic potential); - (k) Solid iso-electron-density models of molecular orbitals (HOMO of salicylic acid computed using Gaussian); - (I) Transparent models of molecular orbitals (LUMO of salicylic acid, the transparency was adjusted according to the computed electron density); - $(m+n$ ) Ribbon models (structure of T3R3 human insulin and bacterial ribonuclease P RNA with solid-ring representation of the base pairs); - (o) Visualization of 3D-grid based densities and properties of molecules using transparent iso-density contours (molecular dynamics derived relative water densities around $\alpha$-cyclodextrin providing a view on the hydration shell of this compound in aqueous solution); - (p) Display of grid properties and potentials with excluded volumes defined by molecular surfaces (electrostatic potential of salicylic acid); - For additional examples see all Figures in the preceding Chapters of this work or the MolArch ${ }^{+}$website. ${ }^{[1]}$



Figure 1 (continued from previous page).

Among the possibilities offered by MolArch ${ }^{+}$are the following modes of visualization:

- Organic and inorganic molecular structures with different modes of display (wire-, capped-, ball-and-stick-, and CPK-type molecular models), including the visualization of bond orders (single to quadruple bonds) and aromaticity (ring type visualization instead of successions of double bonds; Fig. 1a).
- Selected bond distances, angles, torsions, planes, and other geometry descriptors (Fig. 1b).
- Symmetry elements of molecules (center of inversion $i$, mirror planes $\sigma, C_{n}$ rotary and $S_{n}$ rotary reflection axes; Fig. 1c).
- Transition state geometries with transparent bonds (Fig. 1d).
- Coordination polyhedra including tetrahedrons, octahedrons, cubes, dodecahedrons, icosahedrons, truncated icosahedrons, prisms, pyramids (square, pentagonal, and hexagonal), and trigonal bi-pyramids (Fig. 1e-f).
- Atom types (labels) and names, as well as atom mapped (color-coded) properties.
- ORTEP ${ }^{[36]}$ style thermal isotropic and anisotropic ellipsoids (Fig. 1g).
- Crystal structures including unit cells (Fig. 1h).
- Molecular surfaces (Connolly-, solvent-accessible-, and Hirshfeld-type surfaces) in dotted and solid form with various modes of color-coded display of surface related (mapped) properties, surfaces can be colored, sliced, and clipped independently, including the possibility to display transparent or opaque surface models (Fig. 1h-j).
- Molecular and atomic (schematic) orbitals (Molden ${ }^{[37]}$ style display; Fig. 1k-l).
- Ribbon models of arbitrary (including non-protein) structures (Fig. 1m-n).
- 3D-grid and density data (Fig. 1o-p).

Options are included to generate series of molecular images for large sets of molecular configurations along reaction coordinates or molecular dynamics simulations, offering the possibility to generate animations on chemical reactions, molecular dynamics, and molecular recognition processes.

The command line interface and the batch-processing capabilities make MolArch ${ }^{+}$ a highly versatile tool for the visualization, animation, and automated geometry analysis of a wide range of molecular scenarios. The MolArch ${ }^{+}$website includes additional archives of molecular graphics generated in the course of the studies described in this work. In addition, a variety of educational courses with MolArch ${ }^{+}$ generated graphics and animations on this site are aimed at basic and advanced students of organic chemistry. Examples include the structure, symmetry, and chirality of organic molecules; animations on various basic reaction mechanisms of organic chemistry are provided (Cope and Claisen-rearrangement and degenerate versions thereof, Diels-Alder reactions, sigmatropic ring-opening and ring-closure mechanisms, $S_{N} 1$ and $S_{N} 2$ reaction, bromination of alkenes, and various conformational rearrangements; see also Fig. 2).

In total, MolArch ${ }^{+}$is a highly useful program for the analysis and visualization of molecular structures for scientific and educational purposes. ${ }^{[1]}$


Figure 2. Examples for MolArch ${ }^{+}$generated graphics on the MolArch ${ }^{+}$website. ${ }^{[1]}$ The top entry presents supporting graphics and information on current research projects (molecular geometries of cyclodextrin-derived crown acetals, for details see Chapter 6 of this work); on the bottom a webpage providing additional information on lectures for students is shown (MolArch ${ }^{+}$animations for the mechanism of the racemization of [8]helicene).

## References

[1] S. Immel, MolArch ${ }^{+}$- MOLecular ARCHitecture Modeling Program v7.50, Technical University of Darmstadt, Germany, 2003; additional MolArch ${ }^{+}$generated graphics can be found on the web site at http://caramel.oc.chemie.tu-darmstadt.de/immel/.
[2] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, Nucl. Acids. Res. 2000, 28, 235-242; PDB - Protein Data Bank, Research Collaboratory for Structural Bioinformatics (RCSB); http://www.rcsb.org/pdb/.
[3] A. E. Smith and H. J. Lindner, Journal of Computer-Aided Molecular Design 1991, 5, 235-262; H. J. Lindner, M. Kroeker, PIMM - Closed Shell PI-SCF-LCAO-MO Molecular Mechanics Program, Technical University of Darmstadt, Germany, 1995; http://www.oc.chemie.tudarmstadt.de/alte_hp/lin/.
[4] F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, and W. C. Still, Journal of Computational Chemistry 1990, 11, 440-467; MacroModel - Molecular Modelling Program, Schrödinger, Inc., Portland, OR, USA; http://www.schrodinger.com/Products/macromodel.html.
[5] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. 1983, 4, 187-217; A. D. MacKerell, Jr., D. Bashford, R. L. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. JosephMcCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, J. Phys. Chem. B 1998, 102, 35863616; CHARMm - Chemistry at HARvard Macromolecular Mechanics, Accelrys Inc., Burlington, MA, USA, 2001-2003; http://www.accelrys.com/insight/charmm.html.
[6] HyperChem - Molecular Modelling Program, Hypercube, Inc., Gainesville, FL, USA; http://www.hyper.com/.
[7] CIF - Crystallographic Information File, International Union of Crystallography (IUCr); http://www.iucr.org/iucr-top/cif/.
[8] G. M. Sheldrick, SHELX - Crystal Structure Determination from Single-crystal Diffraction Data, University of Göttingen, Germany; http://shelx.uni-ac.gwdg.de/.
[9] F. H. Allen, J. E. Davies, J. J. Galloy, O. Johnson, O. Kennard, C. F. Macrae, E. M. Mitchell, G. F. Mitchell, J. M. Smith, and D. G. Watson, J. Chem. Inf. Comput. Sci. 1991, 31, 187-204.
[10] ConQuest - Program for Searching and Retrieving Information from the Cambridge Structural Database (CSD), Cambridge Crystallographic Data Centre (CCDC), Cambridge, UK; http://www.ccdc.cam.ac.uk/.
[11] Mercury - Crystal Structure Visualisation and Exploration, Cambridge Crystallographic Data Centre (CCDC), Cambridge, UK; http://www.ccdc.cam.ac.uk/.
[12] Spartan - Molecular Mechanics and Quantum Chemical Calculations, Wavefunction, Inc., Irvine, CA, USA; http://www.wavefun.com/.
[13] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, J. A. Pople, Gaussian - Electronic Structure Program, Gaussian, Inc., Pittsburgh, PA, USA; http://www.gaussian.com/.
[14] S. Immel, MolCont ${ }^{+}$- 2D- and 3D-Contouring Program for Orthogonal and Polar Coordinate Data Sets v7.35, Technical University of Darmstadt, Germany, 2003; MolCont ${ }^{+}$is part of the MolArch ${ }^{+}$program package.
[15] S. Immel, MolSurf ${ }^{+}$- Program for the Generation of Triangulated Molecular Surfaces v7.35, Technical University of Darmstadt, Germany, 2003; MolSurf ${ }^{+}$is part of the MolArch ${ }^{+}$program package.
[16] J. J. McKinnon, A. S. Mitchell, and M. A. Spackman, Chem. Eur. J. 1998, 4, 2136-2141.
[17] B. Lee and F. M. Richards, J. Mol. Biol. 1971, 55, 379-400; F. M. Richards, Ann. Rev. Biophys. Bioeng. 1977, 6, 151-176; M. L. Connolly, J. Appl. Crystallogr. 1983, 16, 548-558; M. L. Connolly, Science 1983, 221, 709-713.
[18] W. Heiden, M. Schlenkrich, and J. Brickmann, J. Comp.-Aided Mol. Design 1990, 4, 255-269; W. Heiden, T. Goetze, and J. Brickmann, J. Comput. Chem. 1993, 14, 246-250.
[19] S. Immel, MolGrid ${ }^{+}$- Program for the Generation of Adiabatic Energy Potential Surfaces v7.35, Technical University of Darmstadt, Germany, 2003; MolGrid ${ }^{+}$is part of the MolArch ${ }^{+}$program package.
[20] R. W. W. Hooft, J. A. Kanters, and J. Kroon, J. Comput. Chem. 1991, 12, 943-947.
[21] S. Immel, MolSymm ${ }^{+}$- Program for the Analysis of Molecular Symmetry and Point Groups v7.35, Technical University of Darmstadt, Germany, 2003; MolSymm ${ }^{+}$is part of the MolArch ${ }^{+}$ program package.
[22] S. Pachkovsky, symmetry.c - Brute Force Symmetry Analyzer, 1996; http://bioinformatics.org/cgi-bin/cvsweb.cgi/ghemical/small-utilities/.
[23] A. Preusser, XFarbe - Filled Contours with Bicubics - Visualization of 2D-Arrays and Finding Maxima, Minima, and Saddle Points, Fritz-Haber-Institut of the Max-Planck Gesellschaft, Berlin, Germany; http://www.fhi-berlin.mpg.de/~grz/pub/xfarbe.html.
[24] GnuPlot - Scientific Plotting; http://www.gnuplot.info/.
[25] F. W. Lichtenthaler and S. Immel, Liebigs Ann. 1996, 27-37.
[26] S. Immel, G. E. Schmitt, and F. W. Lichtenthaler, Carbohydr. Res. 1998, 313, 91-105.
[27] D. Cremer and J. A. Pople, J. Am. Chem. Soc. 1975, 97, 1354-1358.
[28] G. A. Jeffrey and R. Taylor, J. Comput. Chem. 1980, 1, 99-109.
[29] G. A. Jeffrey and J. H. Yates, Carbohydr. Res. 1979, 74, 319-322.
$[30] \quad$ C. A. G. Haasnoot, F. A. A. M. D. Leeuw, and C. Altona, Tetrahedron 1980, 36, 2783-2792.
[31] I. Tvaroska, Carbohydr. Res. 1989, 189, 359-362.
[32] L. H. Hall and L. B. Kier, Topol. Indices Relat. Descriptors QSAR QSPR 1999, 307-360; L. H. Hall and L. B. Kier, J. Mol. Graphics \& Modelling 2001, 20, 4-18; L. B. Kier and L. H. Hall, J. Mol. Graphics \& Modelling 2001, 20, 76-83.
[33] D. J. Heisterberg, QTRFIT - Rigid Body Rotation and Fitting Program, The Ohio Supercomputer Center, Columbus, Ohio, USA, 1991; http://www.ccl.net/cca/software/SOURCES/FORTRAN/fitest/fitest.shtml.
[34] POV-Ray - Persistence of Vision Raytracer v3.5, Persistence of Vision Development Team; http://www.povray.org/.
[35] H. Vollhardt, C. Henn, M. Teschner, and J. Brickmann, J. Mol. Graphics 1995, 13, 368-372.
[36] M. N. Burnett, C. K. Johnson, ORTEP-III - Oak Ridge Thermal Ellipsoid Plot Program for Crystal Structure Illustrations, Oak Ridge National Laboratory, Oak Ridge, TN, USA; http://www.ornl.gov/ortep/.
[37] G. Schaftenaar and J. H. Noordik, J. Comput.-Aided Mol. Des. 2000, 14, 123-134; G. Schaftenaar, MOLDEN - A Pre- and Post-processing Program of Molecular and Electronic Structure, Centre for Molecular and Biomolecular Informatics, University of Nijmegen, Nijmegen, Netherlands; http://www.cmbi.kun.nl/~schaft/molden/molden.html.

