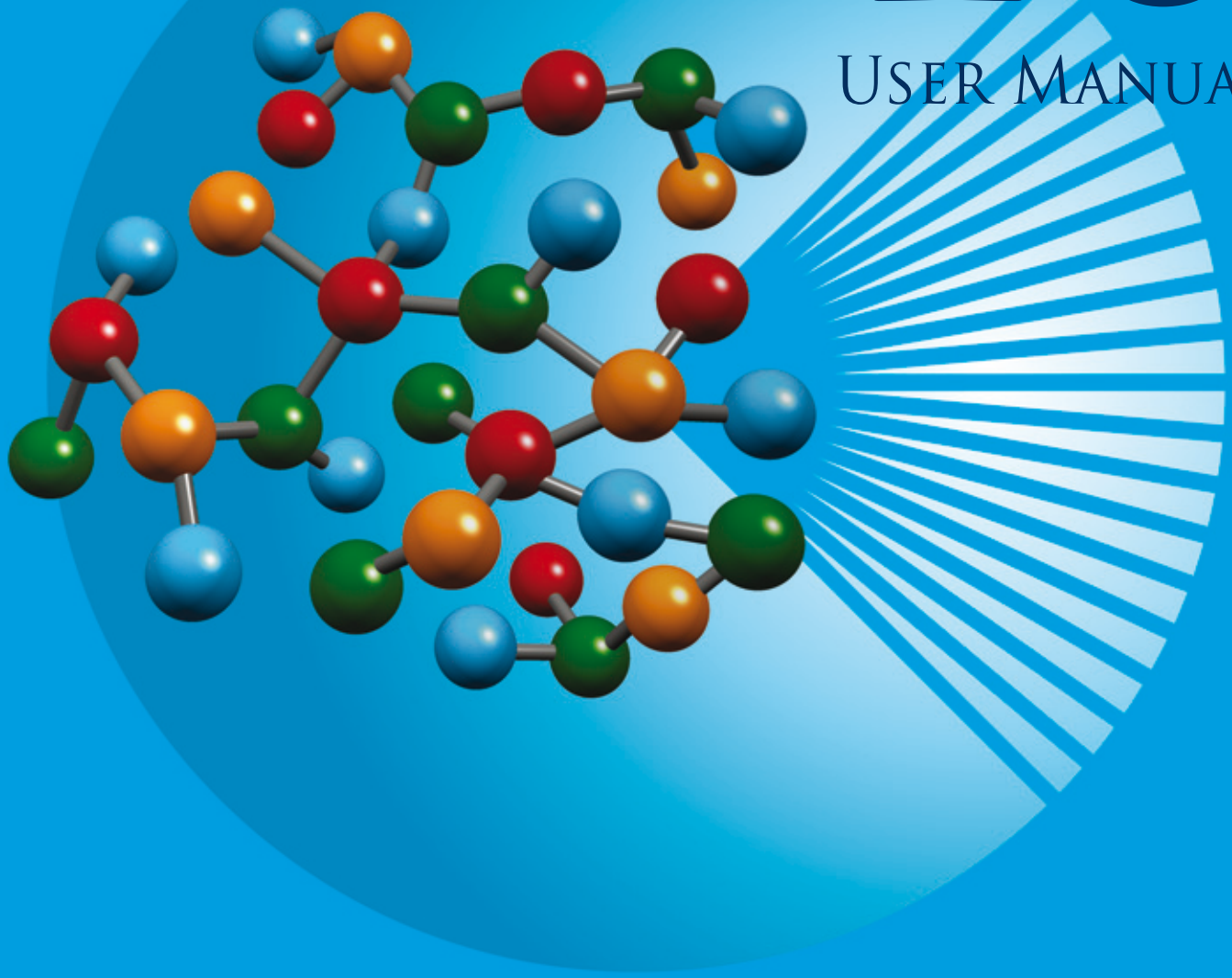
CEPOS INSILICO
SCIENCE FOR SCIENTISTS

FREE ACADEMIC VERSION 10

USER MANUAL



PARASURF™

Impressum

The Old Vicarage
132 Bedford Road
Kempston
BEDFORD, MK42 8BQ
United Kingdom
www.ceposinsilico.com

Manual Timothy Clark

Layout **www.eh-bitartist.de**

TABLE OF CONTENTS

PROGRAM HISTORY

Release Date	Version	Platforms
1st July 2005	ParaSurf´05™ initial release (Revision A1)	32-bit Windows 32-bit Linux Irix
1st January 2006	ParaSurf´05™ Revision B1 (customer-feedback release)	
1st July 2006	ParaSurf´06™ Revision A1	32-bit Windows 32-bit Linux 64-bit Linux Irix
1st July 2007	ParaSurf´07™ Revision A1	
1st July 2008	ParaSurf´08™ Revision A1	32-bit Windows 64-bit Windows 32-bit Linux 64-bit Linux
22nd August 2008	ParaSurf´08™ Revision A2 (minor bug fix release)	
16th December 2008	ParaSurf´08™ Revision A3 (minor bug fix release)	
1st July 2009	ParaSurf´09™ Revision A1	
1st September 2009	New Vhamil.par file including PM6 and first-row transition metals in AM1*	
1st February 2010	ParaSurf´09™ Revision B1 (additional atom-centred descriptors)	
1st July 2010	ParaSurf´10™ Revision A1	
1st July 2011	ParaSurf´10™ Academic Version	

1 INTRODUCTION

ParaSurf™ is a program to generate isodensity or solvent-excluded surfaces from the results of semiempirical molecular orbital calculations, either from VAMP [1] or a public-domain version of MOPAC modified and made available by Cepos InSilico. [2] The surface may be generated by shrink-wrap [3] or marching-cube [4] algorithms and the former may be fit to a spherical harmonic series. [5] The principles of these two techniques are explained below, but for comparison Figure 1 shows default isodensity surfaces calculated by ParaSurf™ for a tetracycline derivative. The surfaces are color-coded according to the electrostatic potential at the surface.

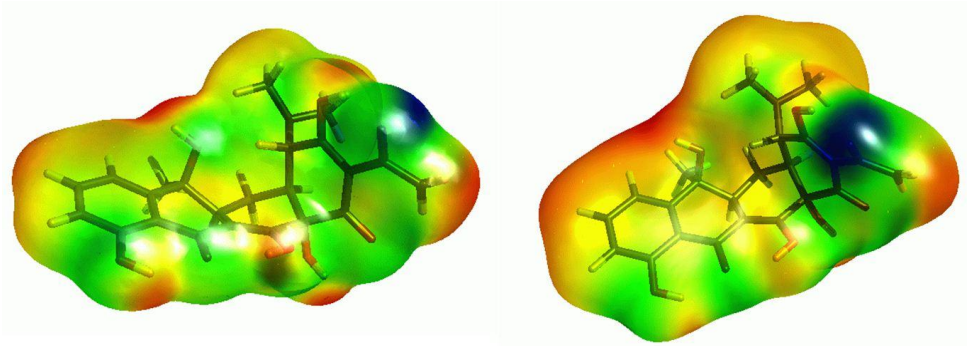


Figure 1 Marching-cube (left) and shrink-wrap (right, fitted to a spherical-harmonic approximation) isodensity surfaces calculated with ParaSurf™ using the default settings

Four local properties, the molecular electrostatic potential (MEP), [6] the local ionization energy (IE_L), [7] the local electron affinity (EA_L), [8] and the local polarisability (α_L) [8] are calculated at the points on the surface. Two further properties, the local hardness (η_L), [8] and the local electronegativity (χ_L) [8] can be derived from IE_L and EA_L.

The local properties can be used to generate a standard set of 81 descriptors [9] appropriate for quantitative structure-property relationships (QSPRs) for determining physical properties.

ParaSurf™ can also generate local enthalpies and free energies of solvation [10] and integrate them over the entire molecular surface to give the enthalpy or free energy of solvation. ParaSurf™ can read so-called *Surface-Integral Model* (SIM) files that allow it to calculate properties such as, for instance, the enthalpy and free energy of hydration and the free energies of solvation in *n*-octanol and chloroform. The surface-integral models are expressed as summations of local solvation energies over the molecular surface. These local solvation energies can be written to the ParaSurf™ surface file.

ParaSurf™ is the first program to emerge from the ParaShift collaboration between researchers at the Universities of Erlangen, Portsmouth, Southampton, Oxford and Aberdeen. It is intended to provide the molecular surfaces for small molecules (i.e. non-proteins) for subsequent quantitative structure-activity relationship (QSAR), QSPR, high-throughput virtual screening (HTVS), docking and scoring, pattern-recognition and simulation software that will be developed in the ParaShift project.

1.1 Changes relative to ParaSurf´09™

ParaSurf'10™ has been enhanced relative to its predecessor in order to provide improved flexibility and a more comprehensive range of descriptors and features. The changes are outlined below:

1.1.1 Local electron affinity for AM1 and other Hamiltonians with d-orbitals as polarisation functions*

Calculating the local electron affinity with AM1* led to spurious results with ParaSurf'09™ because the *d*-polarisation functions dominated the summation. A new technique [11] has been introduced to fix this problem in ParaSurf'10™. A new command-line option requests that the local electron affinity be calculated exactly as in ParaSurf'09™ to ensure continuity.

1.1.2 The PM6 Hamiltonian

ParaSurf'10™ can be used with PM6. [12] This was also the case with parSurf'09 if the post-release Vhamil.par file were used. PM6 is available for 70 elements.

*1.1.3 Second generation surface-integral models; local hydrophobicity**
*1.1.4 Molecular fragments**
*1.1.5 Atom-centred descriptors**

1.2 Isodensity surfaces

Isodensity surfaces [13] are defined as the surfaces around a molecule at which the electron density has a constant value. Usually this value is chosen to approximate the van der Waals' shape of the molecule. ParaSurf™ allows values of the isodensity level down to 0.00001 $e^- Å^{-3}$. Lower values than this may result in failures of the surface algorithms for very diffuse surfaces.

1.3 Solvent-excluded surfaces

The solvent-excluded surface is obtained by rolling a spherical solvent molecule of radius r_{solv} over the surface of the molecule as shown in Figure 2. The surface of the solvent molecule defines the molecular surface, so that the yellow volume in Figure 2 becomes part of the molecule.

* Only available in the full version.

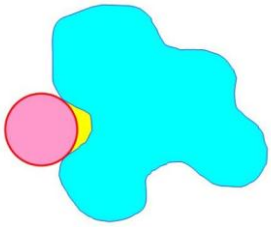


Figure 2 2D-representation of a solvent-excluded surface

1.4 Solvent-accessible surfaces

Solvent-accessible surfaces are obtained in the same way as solvent-excluded surfaces but the outer surface of the solvent sphere is used to define the molecular surface, as shown in Figure 3.

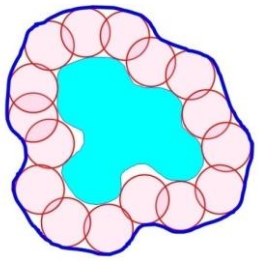


Figure 3 The solvent-accessible surface is obtained by rolling a spherical "solvent molecule"

1.5 Shrink-wrap surface algorithm

Shrink-wrap surface algorithms [3] are used to determine single-valued molecular surfaces. Single-valued in this case means that for any given radial vector from the centre of the molecule the surface is only crossed once (vectors **A** and **B** in Figure 4) and not multiply (vectors **C** and **D** in Figure 4):

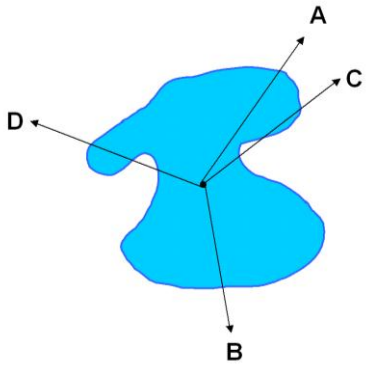


Figure 4 2D-representation of a molecular surface with single-valued (A and B) and multiply valued (C and D) radial vectors from the centre

Single-valued surfaces are necessary for spherical-harmonic fitting (see 1.4). Thus, spherical-harmonic fitting is only available for shrink-wrap surfaces in ParaSurf™. The shrink-wrap algorithm works by starting outside the molecule (point **a** in Figure 5) and moving inwards along the radial vector until it finds the surface (in our case defined by the predefined level of the electron density, point **b** in Figure 5). Thus, the shrink-wrapped surface may contain areas (marked by dashed lines in Figure 5) for which the surface deviates from the true isodensity surface.

These areas of the surface, however, often have little consequence as they are situated above indentations in the molecule that are poorly accessible to solvents or other molecules. The shrink-wrapped surfaces generated by ParaSurf™ should normally be fitted to a spherical-harmonic series for use in HTVS, similarity, pattern-recognition or high-throughput docking applications. The default molecular centre in ParaSurf™ is the centre of gravity (CoG). In special cases in which the CoG lies outside the molecule, another centre may be chosen.

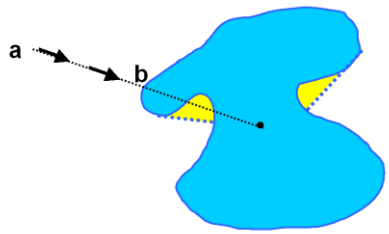


Figure 5 2D-representation of the shrink-wrap algorithm. The algorithms scans along the vector from point a towards the centre of the molecule until the electron density reaches the preset value (point b). The algorithm results in enclosures (marked yellow) for multi-valued radial vectors

Figure 6 shows a spherical-harmonically fitted shrink-wrap surface for a difficult molecule. The areas shown schematically in Figure 5 are clearly visible.

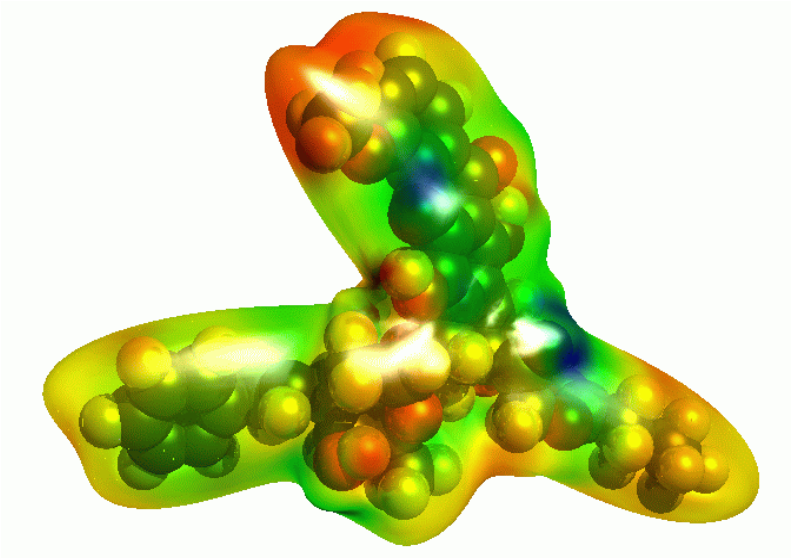


Figure 6 Spherical-harmonic approximation of a shrink-wrap isodensity surface. Note the areas where the surface does not follow the indentations of the molecule

1.6 Marching-cube algorithm

The marching-cube algorithm [4] implemented in ParaSurf™ does not have the disadvantage of being single-valued like the shrink-wrap surface. It cannot, therefore, be fitted to a spherical harmonic series and is used as a purely numerical surface primarily for QSPR applications or surface-integral models. [10] The algorithm works by testing the electron density at the corners of cubes on a cubic lattice laid out through the molecular volume. The corners are divided into those “inside” the molecule (i.e. with a higher electron density than the preset value) and those “outside”. The surface triangulation is then generated for each surface cube and the positions of the surface points corrected to the preset electron density.

1.7 Spherical-harmonic fitting

Complex surfaces can be fitted to spherical harmonic series to give analytical approximations of the surface. [5] The surfaces are fit to a series of distances $r_{\alpha,\beta}$ from the centre along the radial vector defined by the angles α and β as:

$$r_{\alpha,\beta} = \sum_{l=0}^{N} \sum_{m=-l}^{l} c_l^m Y_l^m \tag{1}$$

Where the distances $r_{\alpha,\beta}$ are linear combinations of spherical harmonics Y_l^m defined as:

$$Y_l^m(\alpha,\beta) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(\cos\alpha) e^{im\beta} \tag{2}$$

where $P_l^m(\cos\alpha)$ are associated Legendre functions and l and m are integers such that $-l<=m<=l$. In the above form, spherical harmonics are complex functions. Duncan and Olson [14] have used the real functions

$$Y_l^m(\alpha,\beta) = N_{lm} P_l^m(\cos\alpha)\cos|m|\beta \tag{3}$$

where N_{lm} are normalization factors, to describe molecular surfaces using spherical harmonics.

ParaSurf™ not only fits the surface itself (i.e. the radial distances) to spherical harmonic expansions, but also the four local properties (see 1.8). In this way, a completely analytical description of the shape of the molecule and its intermolecular binding properties is obtained. [15] This description can be truncated at different orders l depending on the application and the precision needed. Thus, a simple description of the molecular properties (shape, MEP, IE_L, EA_L and α_L) to order 2 consists of only five sets of nine coefficients each, or 45 coefficients. These coefficients can be rotated, overlaps calculated etc. [5] to give fast scanning of large numbers of compounds.

Note that, because of the approximate nature of the spherical-harmonic fits, the default isodensity level for the shrink-wrapped surface (0.00002 $e^{-}Å^{-3}$) is lower than that (0.0003 $e^{-}Å^{-3}$) appropriate for an approximately van der Waals' surface using the marching-cube algorithm. The lower value avoids the surface coming too close to atoms. Note also that the fits are incremental, which means that the order chosen for a given application can be obtained by ignoring coefficients of higher order in the spherical-harmonic series.

In some cases, the default resolution of the molecular surface does not allow fitting the spherical-harmonic expansion to very high orders without introducing noise ("ripples") on the fitted surface. In this case, the calculated RMSD becomes larger at higher orders of the spherical-harmonic expansion. ParaSurf'10™ recognizes this condition and truncates the fitting procedure at the optimum value. This can be recognized in the output because the RMSD for later cycles remains constant and the coefficients of the higher order spherical harmonics are all zero. This guarantees the optimum fit in each case and is important for applications that use either the spherical-harmonic coefficients themselves or the hybridization coefficients.

The choice of centre for fitting to a spherical-harmonic expansion is critical. ParaSurf'10™ therefore goes through a multi-step procedure in order to find a suitable centre. This procedure is retained for all molecules for which the ParaSurf'08™ found a suitable centre. However, if the algorithms implemented in ParaSurf'08™ fail to find a suitable centre, the additional technique implemented in ParaSurf'10™ will probably work.

The problem with many molecules is that, for instance, the centre of mass does not lie within the molecular volume. This can easily be the case for, for instance, U- or L-shaped molecules. The procedure implemented in ParaSurf'10™ works as follows:

1. The program first calculates the centre of mass and tests whether it lies within the volume of the molecule. If it does, it is used as the molecular centre. If not, the program moves on to the next step.
2. ParaSurf™ calculates the principal moments of inertia of the molecule and derives a centre from them by assuming that the molecule is U- or V-shaped. The procedure tries to place the centre at the base centre of the molecule. This procedure was implemented in ParaSurf'08™ as a fallback if the centre of mass proved unsuitable. If it also fails to find a suitable centre, ParaSurf'10™ moves on to a third option, which finds a centre for all but the most difficult molecules.
3. The new procedure first searches for the largest plane in the molecule (i.e. the one that contains the most atoms). This search has some leeway, so that the atoms must not all lie exactly in the plane. As a second step, the second largest plane is sought. The molecular centre is then placed in the hinge area between the two planes, as illustrated in Figure 7:

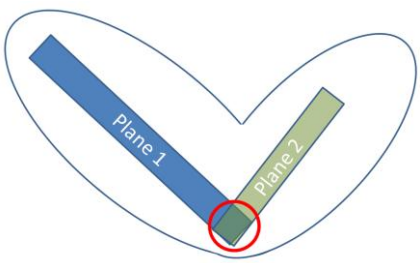


Figure 7 Schematic representation of the planes and hinge area used to determine the centre for spherical-harmonic expansions

1.8 Local properties

The local properties calculated by ParaSurf™ are those related to intermolecular interactions. Local properties, sometimes inaccurately called fields in QSAR work, are properties that vary in space around the molecule and therefore have a distribution of values at the molecular surface. The best known and most important local property in this context is the molecular electrostatic potential, which governs Coulomb interactions, but the MEP only describes a part of the intermolecular interaction energy, so that further local properties are needed.

1.8.1 Molecular electrostatic potential

The MEP is defined in ParaSurf™ as the energy of interaction of a single positive electronic charge at the position ***r*** with the molecule. Within quantum mechanical (semiempirical or *ab initio* molecular orbital (MO) theory, density functional theory (DFT)) the MEP (V(***r***)) is described [**6**] as:

$$MEP(\mathbf{r}) = \sum_{i=1}^{n} \frac{Z_i}{|\mathbf{R_i} - \mathbf{r}|} - \int \frac{\rho(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r}' - \mathbf{r}|} \quad (4)$$

where n is the number of atoms in the molecule, Z_i is the nuclear charge of atom i located at $\mathbf{R_i}$ and $\rho(\mathbf{r})$ is the electron-density function of the molecule. This expression, however, involves integrating the electron density, a time-consuming calculation. ParaSurf™ therefore uses two different approximate models for calculating the MEP.

1.8.1.1 The natural atomic orbital/PC (NAO-PC) model

The NAO-PC model [**16, 17**] uses a total of nine point charges, one positive charge at the nucleus and eight negative ones distributed around it, to describe the electrostatics of a non-hydrogen atom with a valence-only *s*- and *p*-basis set for the semiempirical Hamiltonians MNDO, [**18**] AM1 [**19**] and PM3. [**20**] The negative charges are located at the charge centres of each lobe of the natural atomic orbitals, which are obtained by diagonalising the one-atom blocks of the density matrix. [**16**] The NAO-PC charges are calculated by VAMP and output in the .sdf file for use in ParaSurf™. The NAO-PC model is therefore only available when using ParaSurf™ with VAMP .sdf input. NAO-PC charges are also not available for semiempirical Hamiltonians such as MNDO/d [**21**] or AM1* [**22**] that use *d*-orbitals in the basis set.

1.8.1.2 The multipole model

The integrals needed to evaluate **Eq. 4** in MNDO-type methods use a multipole approximation [**18, 21**] that extends to quadrupoles. We can therefore also use this approximation to calculate atom-centred monopoles, dipoles and quadrupoles for each atom in the molecule. [**23**] This multipole model is applicable to all methods, including those with *d*-orbitals, and can be used with MOPAC output files as input to ParaSurf™.

1.8.2 Local ionization energy, electron affinity, hardness and electronegativity

The local ionization energy $IE_L(\mathbf{r})$ is defined [7] as a density-weighted Koopmans' ionization potential at a point $\mathbf{r}$ near the molecule:

$$IE_L(\mathbf{r}) = \frac{-\sum_{i=1}^{HOMO} \rho_i(\mathbf{r})\varepsilon_i}{\sum_{i=1}^{HOMO} \rho_i(\mathbf{r})} \tag{5}$$

where $HOMO$ is the number of the highest occupied MO, $\rho_i(\mathbf{r})$ is the electron density at point $\mathbf{r}$ due to MO i and ε_i is its Eigenvalue. The local ionization energy describes the tendency of the molecule to interact with electron acceptors (Lewis acids) in a given region in space. [7, 8]

The definition of the local electron affinity is a simple extension of **Eq. 5** to the virtual MOs: [8]

$$EA_L(\mathbf{r}) = \frac{-\sum_{i=LUMO}^{norbs} \rho_i(\mathbf{r})\varepsilon_i}{\sum_{i=LUMO}^{norbs} \rho_i(\mathbf{r})} \tag{6}$$

The local electron affinity is the equivalent of the local ionization energy for interactions with electron donors (Lewis bases). [8]

Eq. 6 requires that the occupied and virtual orbitals be approximately equivalent to each other. This is not the case for semiempirical Hamiltonians (such as AM1*) that include *d*-orbitals as polarisation functions or for extensive basis sets in Hartree-Fock *ab initio* or in Density-Functional theory (DFT) calculations. A new technique has therefore been defined [11] to exclude pure polarisation functions from the summation in **Eq. 6**. This technique is now default in ParaSurf'10™ and gives reliable results. For continuity, a new command-line option (EAL09) has been introduced to request that the calculation of the local electron affinity be performed exactly as in ParaSurf'09™ and earlier versions.

Two further, less fundamental local properties have been defined. [8] These are the local hardness, η_L:

$$\eta_L = \frac{\left(IE_L - EA_L\right)}{2} \tag{7}$$

and the local electronegativity, χ_L:

$$\chi_L = \frac{\left(IE_L + EA_L\right)}{2} \tag{8}$$

1.8.3 Local polarisability

Within the NDDO, the molecular electronic polarisability is easily accessible using the parameterized version [24] of the variational technique introduced by Rivail, [25] which can also be partitioned into an additive polarisability scheme. [26] This allows us to define the local polarisability, α_L, at a point near the molecule as

$$\alpha_L(\mathbf{r}) = \frac{\sum_{j=1}^{norbs} \rho_j^1(\mathbf{r}) q_j \bar{\alpha}_j}{\sum_{j=1}^{norbs} \rho_j^1(\mathbf{r}) q_j} \tag{9}$$

where q_j is the Coulson occupation and $\bar{\alpha}_j$ the isotropic polarisability attributed to atomic orbital *j*. The density ρ_j^1 is defined as the electron density at the point in question due to an exactly singly occupied atomic orbital *j*. The sum is now over atomic orbitals, rather than MOs as for the other local properties. Thus, the local polarisability is a simple occupation-weighted sum of the orbital polarisabilities in which the contribution of each AO is determined by the density of the individual AO at the point being considered.

1.8.4 Field normal to the surface

The electrostatic field (the first derivative of the potential) normal to the molecular surface is closely related to the electrostatic solvation energy in implicit solvation models. [27, 28] This field also has the advantage that it is largely independent of the total molecular charge, so that charged molecules can be compared with neutral ones. If the molecular electrostatic potential is used for this purpose, the charge of ions leads a shift in the potential descriptors, so that molecules and ions with different charges cannot be compared directly. The direction of the normal field (inwards or outwards) also defines, for instance hydrogen-bond donors and acceptors specifically.

1.9 Descriptors

A set of 81 molecular descriptors derived from the MEP, local ionization energy, IE_L, electron affinity, EA_L, electronegativity, χ_L, hardness, η_L, and polarisability, α_L has been defined for QSPR-studies. [9] These and several related descriptors calculated and output by ParaSurf™ are defined in the following table.

Table 1: The descriptors calculated by ParaSurf™

Descriptor	Description	Formula/ Reference	Symbol in CSV file
μ	Dipole moment		**dipole**
μ_D	Dipolar density	[21]	**dipden**
α	Molecular electronic polarisabilty	[29]	**polarizability**
MW	Molecular weight		**MWt**
G	Globularity	[30]	**globularity**
A	Molecular surface area		**totalarea**
VOL	Molecular volume		**volume**
V_{max}	Maximum (most positive) MEP	[31]	**MEPmax**
V_{min}	Minimum (most negative) MEP	[31]	**MEPmin**
$\bar{V}_+$	Mean of the positive MEP values	[31]	**meanMEP+**
$\bar{V}_-$	Mean of the negative MEP values	[31]	**meanMEP-**
$\bar{V}$	Mean of all MEP values	[31]	**meanMEP**
ΔV	MEP-range	[31]	**MEP-range**
σ_+^2	Total variance in the positive MEP values	[31]	**MEPvar+**
σ_-^2	Total variance in the negative MEP values	[31]	**MEPvar-**
σ_{tot}^2	Total variance in the MEP	[31]	**MEPvartot**
ν	MEP balance parameter	[31]	**MEPbalance**
$\sigma_{tot}^2 \nu$	Product of the total variance in the MEP and the balance parameter	[31]	**var*balance**
γ_1^V	Skewness of the MEP-distribution	$\gamma_1^{\alpha_L} = \frac{\sum_{i=1}^{N}\left(\alpha_L^i - \bar{\alpha}_L\right)^3}{(N-1)\sigma^3}$	**MEPskew**

Descriptor	Description	Formula/ Reference	Symbol in CSV file
γ_2^V	Kurtosis of the MEP-distribution	$\gamma_2^V = \frac{\sum_{i=1}^{N}\left(V_i - \bar{V}\right)^4}{(N-1)\sigma^4} - 3$	**MEPkurt**
$\int_V$	Integrated MEP over the surface	$\int_V = \sum_{i=1}^{N} V_i a_i$	**MEPint**
$IE_L^{\max}$	Maximum value of the local ionization energy		**IELmax**
$IE_L^{\min}$	Minimum value of the local ionization energy		**IELmin**
$\overline{IE_L}$	Mean value of the local ionization energy	$\overline{IE_L} = \frac{1}{N}\sum_{i=1}^{N} IE_L^i$	**IELbar**
ΔIE_L	Range of the local ionization energy	$\Delta IE_L = IE_L^{\max} - IE_L^{\min}$	**IELrange**
σ_{IE}^2	Variance in the local ionization energy	$\sigma_{IE}^2 = \frac{1}{N}\sum_{i=1}^{N}\left[IE_L^i - \overline{IE_L}\right]^2$	**IELvar**
$\gamma_1^{IE_L}$	Skewness of the local ionization energy distribution	$\gamma_1^{IE_L} = \frac{\sum_{i=1}^{N}\left(IE_L^i - \overline{IE_L}\right)^3}{(N-1)\sigma^3}$	**IELskew**
$\gamma_2^{IE_L}$	Kurtosis of the local ionization energy distribution	$\gamma_2^{IE_L} = \frac{\sum_{i=1}^{N}\left(IE_L^i - \overline{IE_L}\right)^4}{(N-1)\sigma^4} - 3$	**IELkurt**
$\int_{IE_L}$	Integrated local ionization energy over the surface	$\int_{IE_L} = \sum_{i=1}^{N} IE_L^i a_i$	**IELint**
$EA_L^{\max}$	Maximum of the local electron affinity		**EALmax**
$EA_L^{\min}$	Minimum of the local electron affinity		**EALmin**
$\overline{EA_{L+}}$	Mean of the positive values of the local electron affinity	$\overline{EA_{L+}} = \frac{1}{N^+}\sum_{i=1}^{N^+} EA_{L+}^i$	**EALbar+**
$\overline{EA_{L-}}$	Mean of the negative values of the local electron affinity	$\overline{EA_{L-}} = \frac{1}{N^-}\sum_{i=1}^{N^-} EA_{L-}^i$	**EALbar-**
$\overline{EA_L}$	Mean value of the local electron affinity	$\overline{EA_L} = \frac{1}{N}\sum_{i=1}^{N} EA_L^i$	**EALbar**
ΔEA_L	Range of the local electron affinity	$\Delta EA_L = EA_L^{\max} - EA_L^{\min}$	**EALrange**

Descriptor	Description	Formula/ Reference	Symbol in CSV file
σ^2_{EA+}	Variance in the local electron affinity for all positive values	$\sigma^2_{EA+} = \frac{1}{m}\sum_{i=1}^{m}\left[EA_i^+ - \overline{EA^+} \right]^2$	**EALvar+**
σ^2_{EA-}	Variance in the local electron affinity for all negative values	$\sigma^2_{EA-} = \frac{1}{n}\sum_{i=1}^{n}\left[EA_i^- - \overline{EA^-} \right]^2$	**EALvar-**
σ^2_{EAtot}	Sum of the positive and negative variances in the local electron affinity	$\sigma^2_{EAtot} = \sigma^2_{EA+} + \sigma^2_{EA-}$	**EALvartot**
ν_{EA}	Local electron affinity balance parameter	$\nu_{EA} = \frac{\sigma^2_{EA+} \cdot \sigma^2_{EA-}}{\left[\sigma^2_{EA}\right]^2}$	**EALbalance**
δA^+_{EA}	Fraction of the surface area with positive local electron affinity	$\delta A^+_{EA} = \frac{A^+_{EA}}{A}$, A = total surface area	**EALfraction+**
A^+_{EA}	Surface area with positive local electron affinity		**EALarea+**
$\gamma_1^{EA_L}$	Skewness of the local electron affinity distribution	$\gamma_1^{EA_L} = \frac{\sum_{i=1}^{N}\left(EA_L^i - \overline{EA}_L \right)^3}{(N-1)\sigma^3}$	**EALskew**
$\gamma_2^{EA_L}$	Kurtosis of the local electron affinity distribution	$\gamma_2^{EA_L} = \frac{\sum_{i=1}^{N}\left(EA_L^i - \overline{EA}_L \right)^4}{(N-1)\sigma^4} - 3$	**EALkurt**
$\int_{EA_L}$	Integrated local electron affinity over the surface	$\int_{IE_L} = \sum_{i=1}^{N} EA_L^i a_i$	**EALint**
$\alpha_L^{\max}$	Maximum value of the local polarisability		**POLmax**
$\alpha_L^{\min}$	Minimum value of the local polarisability		**POLmin**
$\overline{\alpha_L}$	Mean value of the local polarisability	$\overline{\alpha_L} = \frac{1}{N}\sum_{i=1}^{N}\alpha_L^i$	**POLbar**
$\Delta\alpha_L$	Range of the local polarisability	$\Delta\alpha_L = \alpha_L^{\max} - \alpha_L^{\min}$	**POLrange**
σ^2_α	Variance in the local polarisability	$\sigma^2_\alpha = \frac{1}{N}\sum_{i=1}^{N}\left[\alpha_L^i - \overline{\alpha_L} \right]^2$	**POLvar**
$\gamma_1^{\alpha_L}$	Skewness of the local polarisability distribution	$\gamma_1^{\alpha_L} = \frac{\sum_{i=1}^{N}\left(\alpha_L^i - \bar{\alpha}_L \right)^3}{(N-1)\sigma^3}$	**POLskew**

Descriptor	Description	Formula/ Reference	Symbol in CSV file
$\gamma_2^{\alpha_L}$	Kurtosis of the local polarisability distribution	$\gamma_2^{\alpha_L} = \frac{\sum_{i=1}^{N}\left(\alpha_L^i - \bar{\alpha}_L\right)^4}{(N-1)\sigma^4} - 3$	**POLkurt**
$\int_{\alpha_L}$	Integrated local polarisability over the surface	$\int_{\alpha_L} = \sum_{i=1}^{N} \alpha_L^i a_i$	**POLint**
$\chi_L^{\max}$	Maximum value of the local electronegativity		**ENEGmax**
$\chi_L^{\min}$	Minimum value of the local electronegativity		**ENEGmin**
$\overline{\chi_L}$	Mean value of the local electronegativity	$\overline{\chi_L} = \frac{1}{N}\sum_{i=1}^{N}\chi_L^i$	**ENEGbar**
$\Delta\chi_L$	Range of the local electron electronegativity	$\Delta\chi_L = \chi_L^{\max} - \chi_L^{\min}$	**ENEGrange**
σ_χ^2	Variance in the local electronegativity	$\sigma_\chi^2 = \frac{1}{N}\sum_{i=1}^{N}\left[\chi_L^i - \overline{\chi_L}\right]^2$	**ENEGvar**
$\gamma_1^{\chi_L}$	Skewness of the local electronegativity distribution	$\gamma_1^{\chi_L} = \frac{\sum_{i=1}^{N}\left(\chi_L^i - \bar{\chi}_L\right)^3}{(N-1)\sigma^3}$	**ENEGskew**
$\gamma_2^{\chi_L}$	Kurtosis of the local electronegativity distribution	$\gamma_2^{\chi_L} = \frac{\sum_{i=1}^{N}\left(\chi_L^i - \bar{\chi}_L\right)^4}{(N-1)\sigma^4} - 3$	**ENEGkurt**
$\int_{\chi_L}$	Integrated local electronegativity over the surface	$\int_{\chi_L} = \sum_{i=1}^{N}\chi_L^i a_i$	**ENEGint**
$\eta_L^{\max}$	Maximum value of the local hardness		**HARDmax**
$\eta_L^{\min}$	Minimum value of the local hardness		**HARDmin**
$\overline{\eta_L}$	Mean value of the local hardness	$\overline{\eta_L} = \frac{1}{N}\sum_{i=1}^{N}\eta_L^i$	**HARDbar**
$\Delta\eta_L$	Range of the local electron hardness	$\Delta\eta_L = \eta_L^{\max} - \eta_L^{\min}$	**HARDrange**
σ_η^2	Variance in the local hardness	$\sigma_\eta^2 = \frac{1}{N}\sum_{i=1}^{N}\left[\eta_L^i - \overline{\eta_L}\right]^2$	**HARDvar**
$\gamma_1^{\eta_L}$	Skewness of the local hardness distribution	$\gamma_1^{\eta_L} = \frac{\sum_{i=1}^{N}\left(\eta_L^i - \bar{\eta}_L\right)^3}{(N-1)\sigma^3}$	**HARDskew**

Descriptor	Description	Formula/ Reference	Symbol in CSV file
$\gamma_2^{\eta_L}$	Kurtosis of the local hardness distribution	$\gamma_2^{\eta_L} = \frac{\sum_{i=1}^{N}\left(\eta_L^i - \bar{\eta}_L\right)^4}{(N-1)\sigma^4} - 3$	**HARDkurt**
$\int_{\eta_L}$	Integrated local hardness over the surface	$\int_{\eta_L} = \sum_{i=1}^{N} \eta_L^i a_i$	**HARDint**
$F_N^{\max}$	Maximum value of the electrostatic field normal to the surface		**FNmax**
$F_N^{\min}$	Minimum value of the field normal to the surface		**FNmin**
$\overline{F_N}$	Mean value of the field normal to the surface	$\overline{F_N} = \frac{1}{N}\sum_{i=1}^{N} \chi_L^i$	**FNmean**
σ_F^2	Variance in field normal to the surface	$\sigma_F^2 = \frac{1}{N}\sum_{i=1}^{N}\left[F_N^i - \overline{F_N}\right]^2$	**FNvartot**
σ_{F+}^2	Variance in the field normal to the surface for all positive values	$\sigma_{F+}^2 = \frac{1}{m}\sum_{i=1}^{m}\left[F_N^{i+} - \overline{F_N^+}\right]^2$	**FNvar+**
σ_{F-}^2	Variance in the field normal to the surface for all negative values	$\sigma_{F-}^2 = \frac{1}{n}\sum_{i=1}^{n}\left[F_N^{i-} - \overline{F_N^-}\right]^2$	**FNvar-**
ν_F	Normal field balance parameter	$\nu_F = \frac{\sigma_{F+}^2 \cdot \sigma_{F-}^2}{\left[\sigma_F^2\right]^2}$	**FNbal**
$\gamma_1^{F_N}$	Skewness of the field normal to the surface	$\gamma_1^{F_N} = \frac{\sum_{i=1}^{N}\left(F_N^i - \bar{F}_N\right)^3}{(N-1)\sigma^3}$	**FNskew**
$\gamma_2^{F_N}$	Kurtosis of the field normal to the surface	$\gamma_2^{F_N} = \frac{\sum_{i=1}^{N}\left(F_N^i - \bar{F}_N\right)^4}{(N-1)\sigma^4} - 3$	**FNkurt**
$\int_{F_N}$	Integrated field normal to the surface over the surface	$\int_{F_N} = \sum_{i=1}^{N} F_N^i a_i$	**FNint**
$\int_{F_N}^{+}$	Integrated field normal to the surface over the surface for all positive values	$\int_{F_N}^{+} = \sum_{i=1}^{N} F_N^i a_i \text{ if } F_N^i \geq 0$	**FN+**
$\int_{F_N}^{-}$	Integrated field normal to the surface over the surface for all negative values	$\int_{F_N}^{-} = \sum_{i=1}^{N} F_N^i a_i \text{ if } F_N^i < 0$	**FN-**

Descriptor	Description	Formula/ Reference	Symbol in CSV file
$\int_{\lvert F_N \rvert}$	Integrated absolute field normal to the surface over the surface	$\int_{F_N} = \sum_{i=1}^{N} \lvert F_N^i \rvert a_i$	**FNabs**
Additionally if the Shannon Entropy is calculated			
$H_{in}^{\max}$	Maximum value of the internal Shannon Entropy		**SHANImax**
$H_{in}^{\min}$	Minimum value of the internal Shannon Entropy		**SHANImin**
$\overline{H_{in}}$	Mean value of the internal Shannon Entropy	$\overline{H_{in}} = \frac{1}{N} \sum_{i=1}^{N} H_{in}^i$	**SHANIbar**
$\sigma_{H_{in}}^2$	Variance in the internal Shannon Entropy	$\sigma_{H_{in}}^2 = \frac{1}{N} \sum_{i=1}^{N} \left[H_{in}^i - \overline{H_{in}} \right]$	**SHANIvar**
$\int_{H_{in}}$	Integrated internal Shannon Entropy over the surface	$\int_{H_{in}} = \sum_{i=1}^{N} H_{in}^i a_i$	**SHANItot**
And if the external Shannon Entropy is available			
$H_{ex}^{\max}$	Maximum value of the external Shannon Entropy		**SHANEmax**
$H_{ex}^{\min}$	Minimum value of the external Shannon Entropy		**SHANEmin**
$\overline{H_{ex}}$	Mean value of the external Shannon Entropy	$\overline{H_{ex}} = \frac{1}{N} \sum_{i=1}^{N} H_{ex}^i$	**SHANEbar**
$\sigma_{H_{ex}}^2$	Variance in the external Shannon Entropy	$\sigma_{H_{ex}}^2 = \frac{1}{N} \sum_{i=1}^{N} \left[H_{ex}^i - \overline{H_{ex}} \right]$	**SHANEvar**
$\int_{H_{ex}}$	Integrated internal Shannon Entropy over the surface	$\int_{H_{ex}} = \sum_{i=1}^{N} H_{ex}^i a_i$	**SHANEtot**

1.10 Surface-integral models (polynomial version)*

1.11 Binned surface-integral models*

* Only available in the full version.

1.12 Spherical harmonic “hybrids”

Once the molecular shape or a local property have been fitted to a spherical-harmonic expansion, [13] the shape or property can be described succinctly as a series of spherical-harmonic “hybridization” coefficients analogous to the concept of hybrid atomic orbitals. Thus, for each value of l in **Eq. 1** the “hybridization” coefficient H_l is given by:

$$H_l = \sum_{i=-m}^{m} \left(c_l^m \right)^2 \tag{11}$$

The hybridization coefficients H_l can be used as additional descriptors for fast QSPR screening.

1.13 Descriptors and moments based on polynomial surface-integral models*

1.14 Shannon entropy*

1.15 Surface autocorrelations*

1.16 Standard Rotationally Invariant Fingerprints (RIFs)

Mavridis et al. [32] introduced standard rotationally invariant fingerprints (RIFs) based on the spherical-harmonic hybridization coefficients defined above. These fingerprints provide a detailed description of the molecular shape, electrostatics, donor/acceptor properties and polarisability as a standard series of 54 floating point numbers.

1.17 Maxima and Minima of the Local Properties*

1.18 Atom-centred descriptors*

1.19 Fragment analysis*

* Only available in the full version.

2 PROGRAM OPTIONS

2.1 Command-line options

ParaSurf™ program options are given as command-line arguments. Arguments are separated by blanks, so that no single argument may contain a blank character. Arguments may be written in any combination of upper and lower case. The options are:

Table 2: ParaSurf™ command-line options

<name>		Base name for the input file (must be the first argument).**<name>** is not required if the first argument is **–version** (see below) Using this option, the input file is assumed to be **<name>_v.sdf** if a file with this name exists. Otherwise the file **<name>.sdf** will be used as input. If neither of these files are found, the program will use an .sdf file written by the Cepos version of Mopac 6. These files are called **<name>_m.sdf** The output files are **<name>_p.out** **<name>_p.sdf** **<name>.psf** (optional)* **<name>.asd** (optional)* **<name>_p.vmp** (optional)*
surf=	**wrap**	Shrink-wrap surface (default)
	cube	Marching-cube surface
contour=	**isoden**	The surface is defined by the electron density
	solvex	A solvent-excluded surface is used.
fit=	**sphh**	Spherical-harmonic fitting (default for **surf=wrap**)
	isod	Smooth to preset isodensity value (default for **surf=cube**)
	none	No fitting
iso=	***n.nn***	Isodensity value set to **n.nn** $e^{-}Å^{-3}$ (default for shrink-wrap surface = **0.00002**; default for marching-cube surface = **0.0003**; minimum possible value = **0.00001**)

* Only available in the full version.

rsol=	***n.nn***	A solvent-probe radius of **n.nn** Å is used for calculating the solvent-excluded or solvent-accessible surface (default=**1.0**, allowed range is from **0.0** to **2.0** Å)
mesh=	***n.nn***	The mesh size used to triangulate the surface is set to **n.nn** Å (default value = **0.2** Å, allowed range is from **0.1** to **1.0** Å)
grid=	**<filename>** **auto**	Read the Cartesian coordinates at which to calculate a grid of the four properties (MEP, IE_L, EA_L, α_L). See **3.8.1** ParaSurf™ calculates an automatic grid (see **3.8.2**)
lattice=	***n.nn***	Sets the lattice spacing for the **grid=auto** option (see **3.8.2**)
-version		Must be the first argument. Requests that ParaSurf™ prints the version number to the standard output channel and then stops without performing a calculation.
eal09		Do not use the selection procedure for virtual orbitals [**11**] when calculating the local electron affinity. This option provides continuity with earlier versions of ParaSurf™

Examples:

```
parasurf test surf=wrap fit=sphh iso=0.03 estat=naopc
```

Use the input file **test_v.sdf**, **test.sdf** or **test_m.sdf** to calculate a shrink-wrap surface with an isodensity value of 0.03 $e^{-}Å^{-3}$, perform a spherical-harmonic fit, use NAO-PC electrostatics and write the spherical-harmonic coefficients to **test_P.sdf**.

```
parasurf test surf=cube fit=none
```

Use the file **test_v.sdf**, **test.sdf** or **test_m.sdf** as input to perform a marching-cube surface determination without fitting and to calculate the descriptor set.

2.2 Options defined in the input SDF-file

2.2.1 Defining the centre for spherical-harmonic fits

The automatic determination of the molecular centre for spherical-harmonic fitting can be overridden by adding a field to the Input (usually VAMP) SDF-file with the tag:

<SPHH_CENTER>

The centre can be defined using Cartesian coordinates using an input line (immediately after the **SPHH_CENTER** tag) of the format:

Cartesian x.xx y.yy z.zz

where ***x.xx***, ***y.yy*** and ***z.zz*** are the *x*, *y*, and *z*-coordinates, respectively. The capitalization of "**Cartesian**" is required.

Alternatively, a list of atoms can be given using the format

Atoms n1 n2 n3 n4 n5 n6 ….

where ***n1*** etc. are the numbers of the atoms to be used to calculate the centre of gravity. The capitalization of "**Atoms**" is required and the list of atoms is limited to one line.

2.2.2 Defining fragments[*]

[*] Only available in the full version.

3 INPUT AND OUTPUT FILES

ParaSurf™ uses the following files for input and output:

Table 3: ParaSurf™ input and output files

File	Name	Description
Input	**<filename>_v.sdf** or **<filename>.sdf** (if available) or **<filename>_m.sdf**	VAMP .sdf file output. VAMP must be run with the **ALLVECT** option to be able to calculate all the properties. The VAMP version used must be able to calculate AO-polarisabilities. If no VAMP .sdf file is found, ParaSurf™ defaults to a Cepos Mopac 6 .sdf file. It is strongly recommended to use the **EF** option for geometry optimizations in Mopac.
Hamiltonian	**Vhamil.par**	The VAMP parameters file (also found in the VAMP executable directory). This file must be copied to the ParaSurf™ executable directory.
Output	**<filename>_p.out**	Always written.
SD-file	**<filename>_p.sdf**	Always written.

3.1 The VAMP .sdf file as input

VAMP .sdf files, an extension of the MDL .sdf file format, [33] are the primary communication channel between VAMP and ParaSurf™. The atomic coordinates and bond definitions are given in the MDL format as shown in **Figure 8**.The remaining fields are indicated by tags with the form:

<FIELD_NAME>
FIELD_NAME is a predefined text tag used to locate the relevant data within the .sdf file.

Only the important fields for a ParaSurf™ calculation will be described here:

<HAMILTONIAN>
The Hamiltonian field defines the semiempirical Hamiltonian (model and parameters) used for the calculation. The Hamiltonian must be defined for ParaSurf™ to be able to calculate the electrostatics and the local polarisabilities. NAO-PC electrostatics and the local polarisability are not available for all methods. Quite generally, the multipole electrostatics model is to be preferred over the NAO-PC model, which can only be used if the VAMP .sdf file contains a block with the tag:

<NAO-PC>
NAO-PCs cannot be calculated for methods with *d*-orbitals. The local polarisability calculation has not yet been extended to these methods, but will be in a future release.

```
1-Bromo-3,5-difluorobenzene
OMVAMP81A04250313563D 1   0.00000     0.00000     0

 12 12  0  0  0  0              1 V2000
   -2.6274    0.2410    0.0003 F
   -1.2738    0.2410    0.0003 C
   -0.5810    1.4623    0.0003 C
    0.8231    1.4389    0.0003 C
    1.5096    2.6055    0.0004 F
    1.5266    0.2198    0.0001 C
    0.8142   -0.9793    0.0001 C
    1.7431   -2.6055   -0.0004 Br
   -0.5805   -0.9840    0.0002 C
   -1.1264    2.4167   -0.0003 H
    2.6274    0.2339    0.0003 H
   -1.1515   -1.9253    0.0001 H
  1  2  1
  2  3  4
  3  4  4
  4  5  1
  4  6  4
  6  7  4
  7  8  1
  2  9  4
  7  9  4
  3 10  1
  6 11  1
  9 12  1
M  END
```

Figure 8 The headers and titles, atomic coordinates and bond definitions from a VAMP .sdf file. The format follows the MDL definition [25]

The following table gives an overview of the methods and their limitations:

Table 4: Hamiltonians and the available electrostatic and polarisability models.

Hamiltonian	Reference	Electrostatics		Local Polarisability
		NAO-PC	Multipole	
MNDO	[18]	YES	YES	YES
AM1	[19]	YES	YES	YES
PM3	[20]	YES	YES	YES
MNDO/c	[34]	YES	YES	NO
MNDO/d	[21]	NO	YES	NO
AM1*	[22]	NO	YES	NO

<VAMPBASICS>

The VAMPBASICS block contains the following quantities (`FORTRAN format 6f13.6`):

Heat of Formation	kcal mol^{-1}
HOMO energy	eV
LUMO energy	eV
Dipole moment	

x-component	Debye
y-component	Debye
z-component	Debye

<TOTAL COULSON CHARGE>
The total charge of the molecule.

<DENSITY MATRIX ELEMENTS>
The DENSITY MATRIX ELEMENTS block contains the one-atom blocks of the density matrix for the non-hydrogen atoms. For an *sp*-atom, there are ten elements, for an *spd*-atom 45. The squares of the diagonal elements for hydrogen atoms are included in the **<CHARGE ON HYDROGENS>** block that follows the density matrix. The density-matrix elements are used in ParaSurf™ to calculate the local properties and are essential.

<ORBITAL VECTORS>
The ORBITAL VECTORS block contains the MO-eigenvectors and related information and is essential for calculating the local properties. VAMP must be run with the keyword **ALLVECT** in order to write all the MO vectors to the SDF file.

The entire SDF input file is echoed to the **<filename>_p.sdf** output file and the properties calculated by ParaSurf™ are added in additional blocks at the end.

*3.1.1 Multi-structure SD-files**

3.2 The Cepos MOPAC 6.sdf file as input

Cepos Mopac 6 writes an .sdf file containing the above blocks with the exception that the MOPACBASICS block replaces VAMPBASICS. No additional keywords are required to request the correct .sdf output for ParaSurf™.

3.3 The Vhamil.par file

The file Vhamil.par is used by VAMP to define the available Hamiltonians and elements and supply the parameters. This file is also used by ParaSurf™ for the same purpose. A Vhamil.par file for standard Hamiltonians and elements is supplied with the ParaSurf™ program. In order to be sure that all Hamiltonians and elements available to VAMP can also be handled by ParaSurf™, however, the

* Only available in the full version.

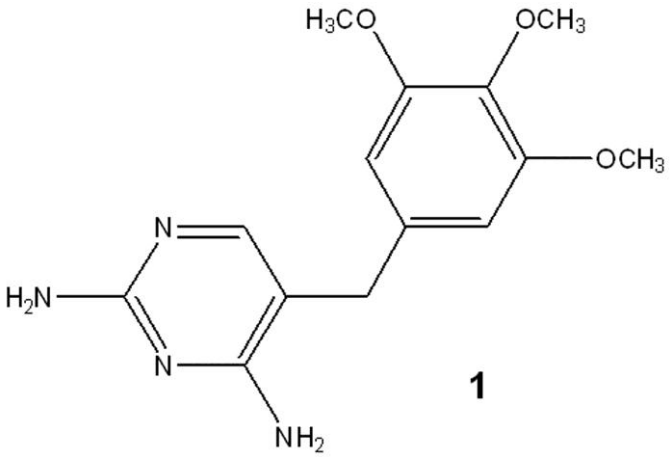
Vhamil.par file from the VAMP executable directory should be copied into the ParaSurf™ executable directory.

3.4 The ParaSurf™ output file

The ParaSurf™ output file provides the user with information about the calculation and the results. It is, however, not intended as the primary means of communication between ParaSurf™ and other programs. Thus, the essential information contained in the output file is also available from the ParaSurf™ output .sdf file.

3.4.1 For a spherical-harmonic surface

Figure 9 shows the output for a calculation using the options **`surf=wrap fit=sphh translate`** for trimethoprim, **1**.




```
<> ParaSurf'10 Academic Version
 <> Copyright (c) 2006,2007,2008,2009,2010 Friedrich-Alexander-Universitaet
                  Erlangen-Nuernberg and Cepos InSilico Ltd.
                  All rights reserved.

 <> Input = trimethoprim.sdf

 <> Program options :

    Using shrink-wrap isocontour surface
    Fitting surface to spherical harmonics
    Using an isodensity surface contour
    Isodensity value = 0.2000E-04 electrons/Angstrom**3
    Triangulation mesh =    0.20 Angstrom
    Using multipole electrostatics
```

Figure 9 ParaSurf™ output for trimethoprim, **1**, using a spherical-harmonic surface

Figure 9 continued

```
<> AM1   calculation for Trimethoprim
<> Fitting surface to spherical harmonics
<> Order(l)    RMSD

      0       1.92523465
      1       1.96112137
      2       1.55522693
      3       1.10610407
      4       0.93107725
      5       0.70604021
      6       0.63660643
      7       0.57075868
      8       0.52397917
      9       0.50047987
     10       0.47258810
     11       0.44413396
     12       0.41917572
     13       0.40317390
     14       0.39305683
     15       0.38338670

<> Spherical harmonic fit for MEP:
<> Order(l)    RMSD

      0      11.06679887
      1      11.02896096
      2       8.63328473
      3       6.86212353
      4       5.49789123
      5       4.58515348
      6       4.17133704
      7       3.45021190
      8       3.12551976
      9       2.77785314
     10       2.36026963
     11       2.07229809
     12       1.90436905
     13       1.72364959
     14       1.64559721
     15       1.46827862
     16       1.27844169
     17       1.07448912
     18       0.93472662
     19       0.88271611
     20       0.82766903

<> Spherical harmonic fit for IE(l):
<> Order(l)    RMSD

      0      56.96267883
      1      50.08995849
      2      45.41506823
      3      43.50398342
      4      40.07275301
      5      35.50489297
      6      32.77872159
      7      26.87765240
      8      23.10395662
      9      19.60420788
     10      17.97895259
     11      16.20677848
     12      15.12291903
     13      14.61861765
     14      13.86580806
     15      13.43494797
     16      13.22461242
     17      12.62315996
     18      12.26642783
     19      12.26642783
     20      12.26642783
```

Figure 9 continued

```
<> Spherical harmonic fit for EA(l):
<> Order(l)    RMSD

       0       12.18546245
       1       11.86493015
       2       11.74652208
       3        9.50375558
       4        8.72724156
       5        7.28928968
       6        7.13942468
       7        6.79112856
       8        6.47917898
       9        6.02700721
      10        5.73179346
      11        5.46853066
      12        5.18673696
      13        4.51767269
      14        4.11406977
      15        3.92082708
      16        3.68234845
      17        3.60371852
      18        3.40232841
      19        3.23619071
      20        3.10610529

<> Spherical harmonic fit for Field(N):
<> Order(l)    RMSD

       0        7.95868149
       1        7.94096053
       2        7.17401521
       3        6.65225208
       4        5.99511525
       5        5.55916448
       6        5.41073978
       7        5.06133734
       8        4.86238016
       9        4.56150784
      10        4.11032910
      11        3.75302455
      12        3.52058064
      13        3.43143087
      14        3.35170488
      15        3.08389445
      16        2.71006812
      17        2.33242842
      18        2.10812470
      19        2.03521010
      20        1.94087735

<> Spherical harmonic fit for Alpha(l):
<> Order(l)    RMSD

       0        0.02367600
       1        0.01660891
       2        0.01369187
       3        0.01110787
       4        0.00911432
       5        0.00818939
       6        0.00769335
       7        0.00722758
       8        0.00695207
       9        0.00643405
      10        0.00588639
```

Figure 9 continued

```
      11       0.00574100
      12       0.00532006
      13       0.00531482
      14       0.00521057
      15       0.00514790
      16       0.00514790
      17       0.00514790
      18       0.00514790
      19       0.00514790
      20       0.00514790

<> Property ranges:
   Density   :   0.3567E-05 to   0.9969E-04
   IE(l)     :       391.03 to       671.23
   EA(l)     :      -108.54 to       -38.27
   MEP       :       -48.51 to        16.80
   Alpha(l)  :       0.2369 to       0.3374
   Field(N)  :       -41.48 to        16.61

<> Descriptors :

   Dipole moment        :       1.2486 Debye
   Dipolar density      :   0.1936E-02 Debye.Angstrom**-3
   Molecular pol.       :      31.2348 Angstrom**3
   Molecular weight     :       290.32
   Globularity          :       0.7689
   Total surface area   :       469.52 Angstrom**2
   Molecular volume     :       644.95 Angstrom**3

   Most positive MEP    :        16.80 kcal/mol
   Most negative MEP    :       -48.51 kcal/mol
   Mean +ve MEP         :         5.59 kcal/mol
   Mean -ve MEP         :       -10.79 kcal/mol
   Mean MEP             :        -3.13 kcal/mol
   MEP range            :        65.31 kcal/mol
   MEP +ve Variance     :        10.80 (kcal/mol)**2
   MEP -ve Variance     :        94.42 (kcal/mol)**2
   MEP total variance   :       105.22 (kcal/mol)**2
   MEP balance parameter:       0.0921
   MEP balance*variance :       9.6913 kcal/mol
   MEP skewness         :      -1.1819
   MEP kurtosis         :       1.3877
   Integral MEP         : -1166.29     kcal.Angstrom**2/mol

   Maximum IE(l)        :       671.23 kcal/mol
   Minimum IE(l)        :       391.03 kcal/mol
   Mean IE(l)           :       475.67 kcal/mol
   IE(l) range          :       280.20 kcal/mol
   IE(l) variance       :      3233.34 (kcal/mol)**2
   IE(l) skewness       :       0.6768
   IE(l) kurtosis       :      -0.2286
   Integral IE(l)       :  9649.93     eV.Angstrom**2

   Maximum EA(l)        :       -38.27 kcal/mol
   Minimum EA(l)        :      -108.54 kcal/mol
   Mean +ve EA(l)       :         0.00 kcal/mol
   Mean -ve EA(l)       :       -93.87 kcal/mol
   Mean EA(l)           :       -93.87 kcal/mol
   EA(l) range          :        70.28 kcal/mol
   EA(l) +ve variance   :         0.00 (kcal/mol)**2
   EA(l) -ve variance   :       142.43 (kcal/mol)**2
   EA(l) total variance :       142.43 (kcal/mol)**2
   EA(l) skewness       :       1.7819
   EA(l) kurtosis       :       4.1713
   Integral EA(l)       : -1913.50     eV.Angstrom**2
   EA(l) balance param. :       0.0000
   Fraction pos. EA(l)  :       1.0000 ( =  469.52 Angstrom**2)
```

Figure 9 continued

```
    Max. local Eneg.      :       299.61 kcal/mol
    Min. local Eneg.      :       143.17 kcal/mol
    Mean local Eneg.      :       190.90 kcal/mol
    Local Eneg. range     :       156.44 kcal/mol
    Local Eneg. variance  :       958.85 (kcal/mol)**2
    Local Eneg. skewness  :         0.82
    Local Eneg. kurtosis  :         0.01
    Integral local Eneg.  :  3868.21     eV.Angstrom**2

    Max. local hardness   :       371.63 kcal/mol
    Min. local hardness   :       247.42 kcal/mol
    Mean local hardness   :       284.77 kcal/mol
    Local hard. range     :       124.21 kcal/mol
    Local hard. variance  :       729.04 (kcal/mol)**2
    Local hard. skewness  :         0.58
    Local hard. kurtosis  :        -0.48
    Integral local Hard.  :  5781.71     eV.Angstrom**2

    Maximum alpha(l)      :   0.3374     Angstrom**3
    Minimum alpha(l)      :   0.2369     Angstrom**3
    Mean alpha(l)         :   0.2821     Angstrom**3
    Alpha(l) range        :   0.1005     Angstrom**3
    Variance in alpha(l)  :   0.5455E-03 Angstrom**6
    Alpha(l) skewness     :     -0.7808
    Alpha(l) kurtosis     :     -0.6145
    Integral Alpha(l)     :  132.761     Angstrom**5

    Maximum field normal  :    16.61 kcal/mol.Angstrom
    Minimum field normal  :   -41.48 kcal/mol.Angstrom
    Mean field            :    -0.45 kcal/mol.Angstrom
    Field range           :    58.09 kcal/mol.Angstrom
    Total field variance  :    63.25 (kcal/mol.Angstrom)**2
    +ve field variance    :     6.65 (kcal/mol.Angstrom)**2
    -ve field variance    :    66.54 (kcal/mol.Angstrom)**2
    Field balance param.  :     0.08
    Field skew            :     2.34
    Field kurtosis        :   4.278
    Integral F(N)         :   51.51     kcal.Angstrom/mol
    Integral F(N +ve)     :   1332.     kcal.Angstrom/mol
    Integral F(N -ve)     :  -1281.     kcal.Angstrom/mol
    Integral |F(N)|       :   2613.     kcal.Angstrom/mol

<> Spherical-Harmonic Hybridization:

    Shape hybrids         :
  17.576151    1.110861    3.450758    2.848775    1.410632    1.602015
   0.752068    0.688202    0.462321    0.389758    0.387721    0.318766
   0.254783    0.212811    0.209028    0.200477

    MEP hybrids           :
  13.217889    4.729529   26.188350   18.489609   13.992226   10.724730
   7.220981    8.124247    4.622006    4.919896    4.395234    3.580830
   2.541562    2.738907    1.629285    2.033261    1.920217    1.871438
   1.474428    1.035562    1.017963

    IE(l) hybrids         :
  1698.1619     90.2420     81.3581     62.4739     59.3195     54.6003
    53.6317     54.9183     40.0437     44.6963     25.5245     24.8963
    19.8380     15.9993     18.4093     15.7391     14.4282     16.4144
    14.7771      0.0000      0.0000

    EA(l) hybrids         :
    324.0302      6.1868     12.8590     25.9665     13.9365     16.5139
      8.5527      9.0223      8.0923      7.5131      6.2313      5.3970
      5.1376      7.5750      5.6376      4.4820      4.8386      3.7336
      3.7363      3.6150      3.4503

    Alpha(l) hybrids      :
 1.01371172  0.05270984  0.03338628  0.02985029  0.02186603  0.01485299
 0.01197223  0.00890841  0.00764054  0.00889521  0.00767265  0.00638097
 0.00682204  0.00625396  0.00851990  0.00656432  0.00000000  0.00000000
 0.00000000  0.00000000  0.00000000
```

Figure 9 continued

```
   Field(N) hybrids     :
     2.6423      1.8631     13.5263     10.4406     10.4941      8.0651
     6.0754      7.7827      5.4503      5.8168      5.6348      5.4245
     4.7299      4.2978      2.4961      3.4376      3.8431      4.0455
     3.3698      2.1922      1.9336

<> Standard rotationally invariant fingerprint:(L. Mavridis, B. D. Hudson
   and D. W. Ritchie, J. Chem. Inf. Model., 2007, 47, 1787-1796.)

     4.19239     1.05397     1.85762     1.68783     1.18770
     1.26571    0.867219     3.63564     2.17475     5.11746
     4.29995     3.74062     3.27486     2.68719     2.85031
     2.14989     2.21808     2.09648     1.89231     1.59423
     41.2088     9.49958     9.01987     7.90404     7.70191
     7.38920     7.32337     18.0008     2.48732     3.58595
     5.09574     3.73317     4.06373     2.92450     1.00683
    0.229586    0.182719    0.172772    0.147872    0.121873
    0.109418     1.62552     1.36495     3.67782     3.23120
     3.23946     2.83992     2.46484     2.78974     2.33458
     2.41180     2.37377     2.32906     2.17484

<> Atomic surface properties:


  Atom    Area         MEP            IE(l)           EA(l)          mean      Field(N)
                     max    min     max     min     max     min     pol.    max     min
C     1   0.000
O     2   0.073  -38.63  -41.63  535.57  533.00  -77.38  -79.30   0.260  -21.49  -27.99
C     3   3.380   -6.37  -46.93  578.39  474.06  -41.34  -89.55   0.307   -2.87  -30.99
C     4   1.259   -5.18  -15.61  573.38  498.15  -67.78  -90.31   0.322   -4.24   -8.83
C     5   0.699   -8.34  -13.91  573.56  531.88  -83.99  -92.50   0.320   -1.60   -7.33
C     6   0.000
C     7   0.803   -9.81  -15.04  559.06  516.48  -71.63  -91.20   0.319   -2.88   -9.76
C     8   4.155   -1.37  -21.94  585.08  484.62  -48.40  -95.76   0.295    2.87  -18.92
N     9   4.183  -14.72  -32.41  535.62  452.64  -79.87 -104.94   0.279  -13.08  -41.48
C    10  10.329   -1.83  -27.19  633.65  532.84  -38.83  -86.77   0.284    3.35  -26.41
N    11   0.000
N    12   1.441  -14.68  -33.29  537.97  472.79  -74.67  -98.55   0.268  -16.13  -37.17
C    13   6.263   -8.51  -26.50  636.94  512.98  -38.27  -82.53   0.286   -0.32  -26.69
N    14   0.000
C    15   2.104   -7.01  -15.40  589.73  496.26  -64.25  -92.16   0.316   -3.39   -9.83
C    16   3.889  -11.80  -43.40  569.47  479.95  -40.62  -85.85   0.310   -4.76  -35.29
O    17   0.000
C    18   0.000
C    19   6.191  -17.34  -48.51  583.26  465.76  -46.51  -85.87   0.315   -6.05  -34.36
O    20   2.039  -31.11  -44.39  532.30  445.12  -75.33  -93.15   0.247  -13.38  -34.51
C    21   0.000
H    22  32.635   13.06  -39.51  560.91  405.93  -82.22  -99.65   0.297    8.87  -25.66
H    23  21.584   14.07  -20.11  561.95  408.24  -83.73  -96.09   0.294    9.65  -28.49
H    24  24.164   14.04  -30.22  567.47  406.95  -68.87  -95.30   0.293    9.71  -17.83
H    25   6.870   11.16   -5.33  527.22  425.98  -83.09  -95.66   0.288    7.63   -5.92
H    26  18.506    7.90   -7.44  535.28  400.17  -87.74 -100.67   0.300    5.64   -5.92
H    27  17.885    7.35  -20.61  543.51  399.68  -68.57 -100.82   0.303    4.52  -21.74
H    28  26.817    8.86  -27.87  536.72  413.57  -60.29 -103.43   0.284    5.88  -32.28
H    29  33.031   16.80  -28.67  671.23  474.66  -72.02 -107.86   0.248   16.61  -32.31
H    30  32.854   16.44  -28.19  669.05  476.80  -72.56 -107.92   0.241   16.30  -32.17
H    31  33.925   11.44  -28.94  657.13  471.46  -67.51 -108.53   0.244   12.43  -35.12
H    32   9.936   11.30  -27.10  641.75  472.59  -73.11  -99.95   0.258   12.66  -34.70
H    33   5.297    9.88  -13.59  504.99  434.61  -73.78  -96.58   0.295    7.73   -9.33
H    34  29.288   11.33  -40.38  564.99  406.94  -80.27  -99.69   0.291    7.86  -33.46
H    35  23.557   11.45  -33.33  567.57  407.56  -72.90  -95.67   0.293    8.06  -23.22
H    36  23.013   11.40  -13.11  563.83  407.55  -81.57  -99.02   0.295    8.02  -12.33
H    37  33.967    4.59  -38.60  554.53  394.86  -87.16 -108.54   0.295    4.15  -30.30
H    38  18.879    4.19  -40.99  558.76  396.53  -84.56 -108.22   0.298    3.76  -25.13
H    39  26.447    3.97  -34.03  550.40  391.03  -86.19 -107.60   0.296    3.63  -26.93

Total   465.463

<> ParaSurf used         6.86 seconds CPU time
```

ParaSurf'10™ first fits the calculated shrink-wrap surface at full resolution for each of the local properties. It lists the root-mean-square deviations (RMSDs) for the surface points as a function of the order of the spherical-harmonic expansion, first for the geometry of the surface and then for each of the five local properties. The RMSD values give an idea of how well each order of the spherical-harmonic expansion fits the calculated shrink-wrap surface or the relevant property. The highest order used by ParaSurf™ is 15 for the surface itself and 20 for each property.

The descriptor table is then printed. For molecules with no surface areas with positive EA_L, $\sigma^2_{EA_{L+}}$ is set to zero. The descriptors are those described in Table 1.

The spherical-harmonic hybridization coefficients are then listed for the shape and the five local properties. The coefficients are listed by increasing l starting from zero. The standard rotationally invariant fingerprint (RIF) [**32**] is printed. Note that the individual RIF-values correspond to the square roots of the hybridization coefficients from the tables above and that the RIF definition has been expanded to include hybridization coefficients of the field normal to the surface (the last 13 elements).

The table of atomic surface properties is derived by first finding the atom that contributes most (according to a Coulson analysis) to the electron density for each surface point. The point is then assigned to this atom and the maxima and minima in the MEP, IE_L, EA_L and F_N as well as the mean local polarisability for the points assigned to each atom are calculated. Note that, because of the fitting procedure, the values reported in this table may contain spurious ones if the fitted surface comes particularly close to an atom (or does not approach it). This situation is generally recognisable from the RMSD values printed for the fit. The surface used to calculate the descriptors and atomic-surface properties is the fitted spherical-harmonic surface of order 15.

3.4.2 For a marching-cube surface

Figure 10 shows the output for a calculation using the options **surf=cube** for trimethoprim.

```
<> ParaSurf'10 Academic Version
 <> Copyright (c) 2006,2007,2008,2009,2010 Friedrich-Alexander-Universitaet
                  Erlangen-Nuernberg and Cepos InSilico Ltd.
                  All rights reserved.

 <> Input = trimethoprim.sdf

 <> Program options :

    Using marching-cube isodensity surface
    Surface fitting turned off
    Using an isodensity surface contour
    Isodensity value = 0.3000E-03 electrons/Angstrom**3
    Triangulation mesh =    0.20 Angstrom
    Using multipole electrostatics


 <> AM1   calculation for Trimethoprim
 <> Number of triangles =  15024
 <> Number of unique points :   7517

 <> Property ranges:
    Density   :   0.2879E-03 to   0.3111E-03
    IE(l)     :       392.33 to       654.93
    EA(l)     :      -109.82 to       -28.98
    MEP       :       -69.91 to        24.82
    Alpha(l)  :       0.2288 to       0.3302
    Field(N)  :      -106.67 to        72.35

 <> Descriptors :

    Dipole moment        :       1.2486 Debye
    Dipolar density      :   0.3160E-02 Debye.Angstrom**-3
    Molecular pol.       :      31.2348 Angstrom**3
    Molecular weight     :       290.32
    Globularity          :       0.7042
    Total surface area   :       369.81 Angstrom**2
    Molecular volume     :       395.14 Angstrom**3

    Most positive MEP    :        24.82 kcal/mol
    Most negative MEP    :       -69.91 kcal/mol
    Mean +ve MEP         :         9.04 kcal/mol
    Mean -ve MEP         :       -18.73 kcal/mol
    Mean MEP             :        -4.94 kcal/mol
    MEP range            :        94.73 kcal/mol
    MEP +ve Variance     :        31.62 (kcal/mol)**2
    MEP -ve Variance     :       240.01 (kcal/mol)**2
    MEP total variance   :       271.63 (kcal/mol)**2
    MEP balance parameter:       0.1028
    MEP balance*variance :      27.9357 kcal/mol
    MEP skewness         :      -1.0235
    MEP kurtosis         :       0.6089
    Integral MEP         : -1674.03     kcal.Angstrom**2/mol

    Maximum IE(l)        :       654.93 kcal/mol
    Minimum IE(l)        :       392.33 kcal/mol
    Mean IE(l)           :       486.25 kcal/mol
    IE(l) range          :       262.60 kcal/mol
    IE(l) variance       :      3586.57 (kcal/mol)**2
    IE(l) skewness       :       0.4205
    IE(l) kurtosis       :      -0.7628
    Integral IE(l)       :  7764.79     eV.Angstrom**2
```

Figure 10 ParaSurf™ output for trimethoprim using a marching-cube surface

Figure 10 continued

```
Maximum EA(l)         :       -28.98 kcal/mol
Minimum EA(l)         :      -109.82 kcal/mol
Mean +ve EA(l)        :         0.00 kcal/mol
Mean -ve EA(l)        :       -89.08 kcal/mol
Mean EA(l)            :       -89.08 kcal/mol
EA(l) range           :        80.83 kcal/mol
EA(l) +ve variance    :         0.00 (kcal/mol)**2
EA(l) -ve variance    :       276.48 (kcal/mol)**2
EA(l) total variance  :       276.48 (kcal/mol)**2
EA(l) skewness        :       1.4612
EA(l) kurtosis        :       1.5705
Integral EA(l)        : -1438.96     eV.Angstrom**2
EA(l) balance param.  :       0.0000
Fraction pos. EA(l)   :       1.0000 ( =  369.81 Angstrom**2)

Max. local Eneg.      :       290.20 kcal/mol
Min. local Eneg.      :       143.77 kcal/mol
Mean local Eneg.      :       198.59 kcal/mol
Local Eneg. range     :       146.43 kcal/mol
Local Eneg. variance  :      1206.49 (kcal/mol)**2
Local Eneg. skewness  :         0.52
Local Eneg. kurtosis  :        -0.78
Integral local Eneg.  :  3162.92     eV.Angstrom**2

Max. local hardness   :       371.14 kcal/mol
Min. local hardness   :       247.89 kcal/mol
Mean local hardness   :       287.67 kcal/mol
Local hard. range     :       123.25 kcal/mol
Local hard. variance  :       725.03 (kcal/mol)**2
Local hard. skewness  :         0.45
Local hard. kurtosis  :        -0.66
Integral local Hard.  :  4601.87     eV.Angstrom**2

Maximum alpha(l)      :   0.3302     Angstrom**3
Minimum alpha(l)      :   0.2288     Angstrom**3
Mean alpha(l)         :   0.2830     Angstrom**3
Alpha(l) range        :   0.1014     Angstrom**3
Variance in alpha(l)  :   0.4904E-03 Angstrom**6
Alpha(l) skewness     :     -0.8023
Alpha(l) kurtosis     :     -0.3815
Integral Alpha(l)     :  104.504     Angstrom**5

Maximum field normal  :    72.35 kcal/mol.Angstrom
Minimum field normal  :  -106.67 kcal/mol.Angstrom
Mean field            :    -0.35 kcal/mol.Angstrom
Field range           :   179.02 kcal/mol.Angstrom
Total field variance  :   398.24 (kcal/mol.Angstrom)**2
+ve field variance    :    35.02 (kcal/mol.Angstrom)**2
-ve field variance    :   585.33 (kcal/mol.Angstrom)**2
Field balance param.  :     0.05
Field skew            :     2.74
Field kurtosis        :   6.578
Integral F(N)         :   69.53     kcal.Angstrom/mol
Integral F(N +ve)     :   2364.     kcal.Angstrom/mol
Integral F(N -ve)     :  -2295.     kcal.Angstrom/mol
Integral |F(N)|       :   4659.     kcal.Angstrom/mol
```

Figure 10 continued

```
<> Atomic surface properties:

  Atom    Area         MEP            IE(l)           EA(l)        mean      Field(N)
                     max    min     max     min     max     min     pol.    max     min
C     1   0.255   -24.20  -46.87  569.21  546.16  -81.57  -92.67   0.268   -7.45  -44.94
O     2   3.663   -15.51  -69.27  594.82  456.70  -63.60  -81.53   0.269   23.64 -106.67
C     3   6.489    -7.34  -64.83  642.96  499.79  -30.30  -99.31   0.304    1.34  -84.49
C     4   2.172    -3.32  -19.78  631.80  493.54  -40.13 -100.66   0.316   -2.82  -20.02
C     5   1.621    -2.14  -18.29  633.57  546.99  -53.30 -100.26   0.313    3.29  -11.55
C     6   0.000
C     7   2.050    -4.01  -22.32  605.59  512.50  -50.08  -91.07   0.317    0.14  -26.55
C     8   5.598     4.10  -28.25  638.27  488.28  -35.86  -88.17   0.288    9.86  -29.36
N     9   6.696   -19.85  -58.79  571.47  417.77  -54.64 -103.25   0.260  -14.68  -93.22
C    10   9.420    -0.80  -46.25  654.93  543.27  -41.16  -81.01   0.280   13.28  -61.33
N    11   0.536   -46.54  -53.02  615.74  593.92  -60.29  -79.02   0.276  -60.55  -81.55
N    12   6.115   -16.63  -55.65  571.35  417.33  -51.08  -98.91   0.247  -15.21  -91.12
C    13   7.551   -10.11  -44.06  644.97  527.96  -37.44  -82.79   0.284   10.94  -58.77
N    14   0.713   -41.89  -57.14  618.69  590.74  -64.83  -81.20   0.287  -40.63  -95.16
C    15   4.193    -8.08  -22.32  640.33  494.10  -30.85 -100.79   0.314    0.50  -22.28
C    16   5.831   -15.48  -61.39  641.08  507.48  -28.98  -94.37   0.307    0.56  -75.79
O    17   1.251   -18.69  -69.91  566.92  464.96  -64.24  -87.63   0.252  -11.13 -104.49
C    18   0.290   -16.74  -56.74  573.31  531.64  -74.36  -94.92   0.267   -7.21  -55.54
C    19   5.583   -15.50  -60.30  617.67  492.45  -39.17  -96.41   0.314   -0.58  -50.34
O    20   3.968   -31.04  -63.58  581.05  438.59  -68.75  -94.46   0.265   -3.30  -76.06
C    21   0.545   -26.28  -54.85  563.65  530.77  -90.80 -106.12   0.269   -0.65  -62.76
H    22  20.849    22.24  -44.01  561.09  407.48  -83.43  -97.55   0.297   19.34  -39.51
H    23  16.018    22.20  -47.62  566.04  408.23  -70.44  -97.22   0.293   19.83  -89.24
H    24  16.231    22.21  -45.82  567.70  407.67  -66.85  -96.73   0.291   19.67  -51.22
H    25   7.150    16.35   -8.07  537.75  430.17  -70.34  -97.50   0.287   14.85  -11.83
H    26  13.485    13.09   -5.84  570.04  401.19  -85.82 -100.62   0.299    9.80  -10.05
H    27  13.124    11.65  -38.34  610.80  400.49  -74.20 -100.86   0.301   13.51  -51.85
H    28  17.543    13.34  -29.83  533.67  415.37  -53.69 -100.29   0.282   18.51  -31.14
H    29  20.142    24.45  -44.08  639.93  488.10  -72.39 -107.70   0.247   33.70  -70.29
H    30  20.327    24.82  -48.23  644.32  488.09  -70.46 -107.77   0.241   29.18  -71.86
H    31  20.033    22.74  -51.84  642.48  483.81  -67.40 -108.33   0.245   32.52  -81.11
H    32  10.797    22.33  -49.42  644.51  478.85  -79.21 -102.14   0.258   26.11  -87.40
H    33   7.985    15.01  -21.60  523.96  429.37  -65.63  -98.60   0.294   10.93  -18.28
H    34  20.354    17.51  -31.88  560.04  408.30  -87.92  -99.03   0.290   18.83  -16.57
H    35  16.207    17.81  -47.27  565.70  408.17  -66.88  -96.52   0.291   15.15  -40.08
H    36  16.208    17.77  -37.50  557.34  408.12  -69.21  -96.79   0.294   15.10  -55.02
H    37  20.706     8.06  -41.68  545.25  396.03  -95.88 -109.82   0.294   10.46  -27.25
H    38  16.212     8.06  -54.60  594.17  394.19  -82.67 -109.53   0.296   72.35  -73.80
H    39  18.662     7.90  -40.12  586.35  392.33  -74.88 -109.41   0.295   26.55  -32.47

Total   366.571

<> ParaSurf used         5.75 seconds CPU time
```

The table of RMSD values is no longer printed and the range of the electron-density values for the surface points (a test for the quality of the surface) is closer to the target isodensity value (in this case 0.0003 $e^{-}Å^{-3}$) than for the fitted surface. The internal precision used by the program is ± 2% of the target isodensity value. The values of the descriptors and the atomic-surface properties are more consistent using the marching-cube surface and are recommended for QSPR and surface-integral applications.

*3.4.3 For a job with Shannon entropy**
*3.4.4 For a job with autocorrelation similarity**

* Only available in the full version.

3.5 ParaSurf™ SDF-output

The SDF output file (a fixed-format file) contains additional blocks with the information generated by ParaSurf™. These are:

<ParaSurf OPTIONS>

The ParaSurf™ OPTIONS block consists of one line giving the options used in the ParaSurf™ calculation. These are:

```
<surface> <fit> <electrostatic model> <isodensity level>   (a4,2x,a4,2x,a5,2x,f8.3)
```

Where the individual variables can be:

<surface>	**WRAP**	Shrink-wrap surface
	CUBE	Marching-cube surface
<fit>	**NONE**	No fitting, unsmoothed marching-cube surface
	ISO	Marching-cube surface corrected to ± 2% of the preset isodensity value
	SPHH	Spherical-harmonic surface fit
<electrostatic model>	**NAOPC**	NAO-PC electrostatics
	MULTI	Multipole electrostatics
<isodensity level>	***n.nn***	The target isodensity value in $e^{-}Å^{-3}$
<solvent probe radius>		The radius of the solvent probe used to calculate the SES or SAS
<triangulation mesh>		The mesh size used to triangulate the Surface

<MOLECULAR_CENTERS>

The molecular centres block appears only for calculations that use spherical harmonic fits. It includes two lines of the form:

```
"Spherical harmonic center = ", 3f12.6
"Center of gravity         = ", 3f12.6
```

These blocks give the *x*, *y* and *z* coordinates of the centre of the molecule used for the spherical-harmonic fit and the centre of gravity, respectively. These two centres are usually identical, but may be different if the centre of gravity lies outside the molecule (e.g. for U-shaped molecules).

<SPHERICAL_HARMONIC_.....>

The spherical harmonic fits are described in **<SPHERICAL_HARMONIC_…..>** blocks. These blocks all have the same format and vary only in the property described. Each block has the form:

The spherical harmonic fits are described in <SPHERICAL_HARMONIC_…..> blocks. These blocks all have the same format and vary only in the property described. Each block has the form:

Order = nn	("Order = ",i4)
l(c_l^m)m = -l to l	(I5, 10f8.4/5x,10f8.4/5x,10f8.4/5x,10f8.4) (One set of coefficients each for l = 1 to 15)
RMSDs: **l, RMSD[1], RMSD[2]**	("RMSDs:") (i8, 2f12.8) (One line for each l for l = 1 to 15, where RMSD[1] is the area-weighted RMSD and RMSD[2] the simple RMSD)

There are six such blocks, indicated by the tags:

<SPHERICAL_HARMONIC_SURFACE> The fitted molecular surface (radial distances) in Ångstrom

<SPHERICAL_HARMONIC_MEP> The MEP values at the spherical-harmonic surface (*l* = 20) in kcal mol^{-1}

<SPHERICAL_HARMONIC_IE(l)> The IE_L values at the spherical-harmonic surface (*l* = 20) in kcal mol^{-1}

<SPHERICAL_HARMONIC_EA(l)> The EA_L values at the spherical-harmonic surface (*l* = 20) in kcal mol^{-1}

<SPHERICAL_HARMONIC_ALPHA(l)> The α_L values at the spherical-harmonic surface (*l* = 20) in kcal mol^{-1}

<SPHERICAL_HARMONIC_FIELD(N)> The FN values at the spherical-harmonic surface (l = 20) in kcal mol-1 Å-1

<ParaSurf Descriptors>

The ParaSurf™ descriptors block lists the calculated descriptors in the following groups:

Molecular:	μ, μ_D, α, MW, G, A, VOL ("Molecular ",5f10.4,2f10.2)
MEP:	V_{max}, V_{min}, $\bar{V}_+$, $\bar{V}_-$, $\bar{V}$, ΔV, σ_+^2, σ_-^2, σ_{Tot}^2, ν, $\sigma_{tot}^2\nu$, γ_1^V, γ_2^V, $\int_V$ ("MEP ",7f10.2/10x, f10.2,5f10.4,2x,g12.6)
IE(l):	IE_L^{max}, IE_L^{min}, $\overline{IE_L}$, ΔIE_L, σ_{IE}^2, γ_1^{IE}, γ_2^{IE}, $\int_{IE}$ ("IE(l) ",5f10.2,2f10.4/12x,g12.6)
EA(l):	EA_L^{max}, EA_L^{min}, $\overline{EA_{L+}}$, $\overline{EA_{L-}}$, $\overline{EA_L}$, ΔEA_L, σ_{EA+}^2, σ_{EA-}^2, σ_{EA}^2, ν_{EA}, δA_{EA}^+, A_{EA}^+, γ_1^{EA}, γ_2^{EA}, $\int_{EA}$ ("EA(l) ",7f10.2/2f10.2,2f10.4,f10.2,2f10.4/12x,g12.6)
Eneg(l):	χ_L^{max}, χ_L^{min}, $\overline{\chi_L}$, $\Delta\chi_L$, σ_χ^2, γ_1^χ, γ_2^χ, $\int_\chi$ ("Eneg(l) ",5f10.2,2f10.4/12x,g12.6)
Hard(l):	η_L^{max}, η_L^{min}, $\overline{\eta_L}$, $\Delta\eta_L$, σ_η^2, γ_1^η, γ_2^η, $\int_\eta$ ("Hard(l) ",5f10.2,2f10.4/12x,g12.6)
Alpha(l):	α_L^{max}, α_L^{min}, $\overline{\alpha_L}$, $\Delta\alpha_L$, σ_α^2, γ_1^α, γ_2^α, $\int_\alpha$ ("Alpha(l) ",5f10.2,2f10.4/12x,g12.6)
F_N	F_N^{max}, F_N^{min}, $\overline{F_N}$, σ_F^2, σ_{F+}^2, σ_{F-}^2, ν_F, $\gamma_1^{F_N}$, $\gamma_2^{F_N}$, $\int_{F_N}$, $\int_{F_N}^+$, $\int_{F_N}^-$, $\int_{F_N}$ ("Field desc",7f10.4/" ",6f10.4)

For calculations using a spherical-harmonic fit, the hybridization coefficients are printed to the .sdf file as follows (tag line followed by as many lines with the coefficients as necessary):

<SHAPE HYBRIDS>	(15 coefficients, 6f12.6)
<MEP HYBRIDS>	(20 coefficients, 6f12.6
<IE(L) HYBRIDS>	(20 coefficients, 6f12.2)

<EA(L) HYBRIDS>	(20 coefficients, 6f12.2)
<ALPHA(L) HYBRIDS>	(20 coefficients, 6f12.8)
<FIELD(N) HYBRIDS>	(20 coefficients, 6f12.4)

The hybridization coefficients are listed in order of increasing *l* from zero, exactly as in the output file.

The atomic surface properties are listed in the atomic order according to the following headings (tag line followed by as many lines with the surface properties as necessary):

<ATOMIC SURFACE AREAS>	**Areas** (10f8.4)
<ATOMIC SURFACE MEP MAXIMA>	**MEP maxima** (10f8.2)
<ATOMIC SURFACE MEP MINIMA>	**MEP minima** (10f8.2)
<ATOMIC SURFACE IE(L) MAXIMA>	**IE(l) maxima** (10f8.2)
<ATOMIC SURFACE IE(L) MINIMA>	**IE(l) minima** (10f8.2)
<ATOMIC SURFACE EA(L) MAXIMA>	**EA(l) maxima** (10f8.2)
<ATOMIC SURFACE EA(L) MINIMA>	**EA(l) minima** (10f8.2)
<ATOMIC SURFACE MEAN POL>	**Mean pol.** (10f8.4)
<ATOMIC SURFACE FIELD(N) MAXIMA>	**FN maxima** (10f8.2)
<ATOMIC SURFACE FIELD(N) MINIMA>	**FN minima** (10f8.2)

<STANDARD RIF>

The rotationally invariant fingerprint [32] is printed as a list of 54 floating point numbers (5g12.6). The first 41 are those defined in reference [32] and the last 13 are the square roots of the hybridization coefficients for the normal field from *l*=0-12.

Hard(l)	$\eta_L^{\max}$, $\eta_L^{\min}$, $\overline{\eta_L}$, $\Delta\eta_L$, σ_η^2, γ_1^η, γ_2^η, $\int_\eta$ ("Hard(l) ",5f10.2,2f10.4/12x,g12.6)
Alpha(l)	$\alpha_L^{\max}$, $\alpha_L^{\min}$, $\overline{\alpha_L}$, $\Delta\alpha_L$, σ_α^2, γ_1^α, γ_2^α, $\int_\alpha$ ("Alpha(l) ",5f10.2,2f10.4/12x,g12.6)
F_N	$F_N^{\max}$, $F_N^{\min}$, ΔF_N, $\overline{F_N}$, σ_F^2, σ_{F+}^2, σ_{F-}^2, ν_F, $\gamma_1^{F_N}$, $\gamma_2^{F_N}$, $\int_{F_N}$, $\int_{F_N}^+$, $\int_{F_N}^-$ ("Field desc",7f10.4/" ",6f10.4)

*3.5.1 Optional blocks in the SDF-output file**

3.6 The surface (.psf) file*

3.7 Anonymous SD (.asd) files*

*3.7.1 Optional blocks**

* Only available in the full version.

3.8 Grid calculations with ParaSurf™

3.8.1 User-specified Grid

The command

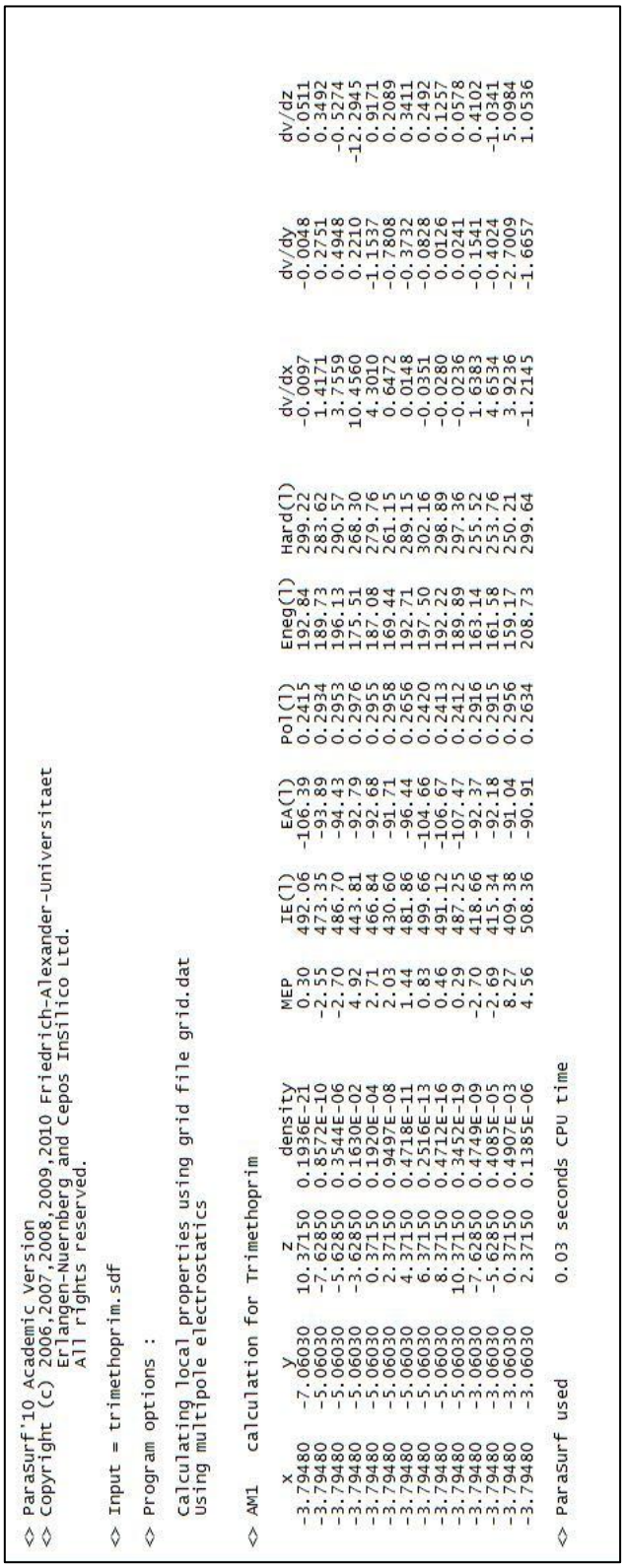
```
parasurf <filename> grid=grid.dat
```

instructs ParaSurf™ to read a set of Cartesian coordinates from the file grid.dat and to calculate the four local properties (MEP, IEL, EAL, αL). The format of the file grid.dat (which must be in the same directory as the input) is one line per atom containing the x, y and z coordinates in free format, comma-separated, maximum line length 80. For instance, the following grid file:

```
-3.79480 , -7.06030 , 10.37150
-3.79480 , -5.06030 , -7.62850
-3.79480 , -5.06030 , -5.62850
-3.79480 , -5.06030 , -3.62850
-3.79480 , -5.06030 ,  0.37150
-3.79480 , -5.06030 ,  2.37150
-3.79480 , -5.06030 ,  4.37150
-3.79480 , -5.06030 ,  6.37150
-3.79480 , -5.06030 ,  8.37150
-3.79480 , -5.06030 , 10.37150
-3.79480 , -3.06030 , -7.62850
-3.79480 , -3.06030 , -5.62850
-3.79480 , -3.06030 ,  0.37150
-3.79480 , -3.06030 ,  2.37150
```

Figure 11 Sample grid file

gives the output shown in Figure 12.



```
<> ParaSurf'10 Academic Version
<> Copyright (c) 2006,2007,2008,2009,2010 Friedrich-Alexander-Universitaet
                  Erlangen-Nuernberg and Cepos InSilico Ltd.
                        All rights reserved.

<> Input = trimethoprim.sdf

<> Program options :

   Calculating local properties using grid file grid.dat
   using multipole electrostatics

<> AM1  calculation for Trimethoprim

      x          y          z          density      MEP      IE(1)     EA(1)    Pol(1)    Eneg(1)   Hard(1)    dv/dx     dv/dy     dv/dz
   -3.79480   -7.06030   10.37150   0.1936E-21     0.30    492.06   -106.39    0.2415    192.84    299.22   -0.0097   -0.0048    0.0511
   -3.79480   -5.06030   -7.62850   0.8572E-10    -2.55    473.35    -93.89    0.2934    189.73    283.62    1.4171    0.2751    0.3492
   -3.79480   -5.06030   -5.62850   0.3544E-06    -2.70    486.70    -94.43    0.2953    196.13    290.57    3.7559    0.4948   -0.5274
   -3.79480   -5.06030   -3.62850   0.1630E-02     4.92    443.81    -92.79    0.2976    175.51    268.30   10.4560    0.2210  -12.2945
   -3.79480   -5.06030    0.37150   0.1920E-04     2.71    466.84    -92.68    0.2955    187.08    279.76    4.3010   -1.1537    0.9171
   -3.79480   -5.06030    2.37150   0.9497E-08     2.03    430.60    -91.71    0.2958    169.44    261.15    0.6472   -0.7808    0.2089
   -3.79480   -5.06030    4.37150   0.4718E-11     1.44    481.86    -96.44    0.2656    192.71    289.15    0.0148   -0.3732    0.3411
   -3.79480   -5.06030    6.37150   0.2516E-13     0.83    499.66   -104.66    0.2420    197.50    302.16   -0.0351   -0.0828    0.2492
   -3.79480   -5.06030    8.37150   0.4712E-16     0.46    491.12   -106.67    0.2413    192.22    298.89   -0.0280    0.0126    0.1257
   -3.79480   -5.06030   10.37150   0.3452E-19     0.29    487.25   -107.47    0.2412    189.89    297.36   -0.0236    0.0241    0.0578
   -3.79480   -3.06030   -7.62850   0.4749E-09    -2.70    418.66    -92.37    0.2916    163.14    255.52    1.6383   -0.1541    0.4102
   -3.79480   -3.06030   -5.62850   0.4085E-05    -2.69    415.34    -92.18    0.2915    161.58    253.76    4.6534   -0.4024   -1.0341
   -3.79480   -3.06030    0.37150   0.4907E-03     8.27    409.38    -91.04    0.2956    159.17    250.21    3.9236   -2.7009    5.0984
   -3.79480   -3.06030    2.37150   0.1385E-06     4.56    508.36    -90.91    0.2634    208.73    299.64   -1.2145   -1.6657    1.0536

<> ParaSurf used       0.03 seconds CPU time
```

Figure 12 Sample grid output file

The name and the extension (if any) of the grid file are free. Only the output file is written. The units of the local properties are those used in the normal output (i.e. V, IEL, and EAL in kcal mol^{-1}, αL in Ångstrom^3.

3.8.2 Automatic grids

ParaSurf™ can generate grids automatically for lead compounds in ComFA®-like procedures. The **grid=auto** option generates a grid around the molecule (with a 4 Å margin around the positions of the atoms in each direction) and includes all points for which the electron density is lower than 10^{-2} (i.e. for points outside the molecule). The spacing of the grid is set to a default value of 1.0 Å, but can be set to any value up to a maximum of 2.0 Å by the command-line argument **lattice=*n.n***, which sets the lattice spacing to *n.n* Å. The grid thus generated is output (with the values of the local properties analogously to a calculation that uses an predefined grid and can be used for other molecules that have been aligned with the lead.

3.9 The SIM file format*
3.10 Output tables*
3.11 Binned SIM descriptor tables*
3.12 Autocorrelation similarity tables*

3.13 Shared files

The Vhamil.par and SIM files are accessed in shared, read-only mode so that multiple ParaSurf™ jobs can access the same files.

* Only available in the full version.

4 TIPS FOR USING PARASURF´10™

4.1 Choice of surface

ParaSurf™ was originally written to use isodensity surfaces. However, calculations that use a solvent-excluded surface are very much faster than their equivalents with isodensity surfaces and will usually give comparable results. Surface-integral models may benefit from using a solvent-excluded surface with a solvent radius of 0.5-1.0 Å as this appears to be the most relevant surface for many physical properties. Surfaces fitted to spherical-harmonic expansions require more CPU-time than marching-cube surfaces but are essential for fast numerical applications such as ParaFit™. Again, solvent-excluded shrink-wrap surfaces are faster to calculate than their isodensity equivalents.

4.2 ParaSurf™ and ParaFit™

ParaFit™ is Cepos InSilico's very fast shape-matching program that is based on spherical-harmonic expansions generated by ParaSurf™. ParaFit™ can be used to overlay molecules with a common scaffold by defining the centre to be used for generating the spherical-harmonic fit in ParaSurf™ in the input SDF-file (see 2.2)

4.3 QSAR using grids

As outlined in 3.8.2, ParaSurf™ can generate a grid for the lead molecule automatically that can then be used for a set of aligned (e.g. with ParaFit™) molecules for grid-based QSAR. This procedure has proven to be especially effective for test datasets, especially if the molecules are aligned to a common scaffold, as outlined in 4.2.

5 SUPPORT

5.1 Error reporting

Some of the routines in ParaSurf™ may detect error conditions that have not yet been encountered in our tests. In this case, an error message will be printed requesting that the input and output files be sent to the programming team at the above e-mail address. We realize that this will not always be possible for confidentiality reasons, but if the details can be sent, we will be able to treat the exception and improve the program.

5.2 CEPOS InSilico GmbH

Computer-Chemie-Centrum (CCC)
Nägelsbachstr. 25
91052 Erlangen
Germany

6 LIST OF TABLES

7 LIST OF FIGURES

8 REFERENCES

1 T. Clark, A. Alex, B. Beck, F. Burkhardt, J. Chandrasekhar, P. Gedeck, A. H. C. Horn, M. Hutter, B. Martin, G. Rauhut, W. Sauer, T. Schindler, and T.Steinke, VAMP 8.2, Erlangen 2002; available from Accelrys Inc., San Diego, USA (**http://www.accelrys.com/products/datasheets/vamp.pdf**).

2 J. J. P. Stewart, MOPAC2000, 1999, Fujitsu, Ltd, Tokyo, Japan. MOPAC 6.0 was once available as: J. J. P. Stewart, QCPE # 455, Quantum Chemistry Program Exchange, Bloomsville, Indiana, 1990.

3 J. H. Van Drie, *"Shrink-wrap" surfaces: A new method for incorporating shape into pharmacophoric 3D database searching*, *J. Chem. Inf. and Comp. Sci.*, **1997**, *37*, 38-41; J. H. Van Drie und R. A. Nugent, *Addressing the challenges of combinatorial chemistry: 3D databases, pharmacophore recognition and beyond*, *SAR and QSAR in Env. Res.*, **1998**, *9*, 1-21; J. Erickson, D. J. Neidhart, J. Van Drie, D. J. Kempf, X. C. Wang, D. W. Norbeck, J. J. Plattner, J. W. Rittenhouse, M. Turon, N. Wideburg, et al., *Design, activity, and 2.8 A crystal structure of a C2 symmetric inhibitor complexed to HIV-1 protease*, *Science*, **1990**, *249*, 527-533.

4 W. Heiden, T. Goetze, and J. Brickmann, *Fast generation of molecular surfaces from 3D data fields with an enhanced "marching cube" algorithm. J. Comput. Chem.* **1993**, *14*, 246-50.

5 D. W. Ritchie und G. J. L. Kemp, *Fast computation, rotation, and comparison of low resolution spherical harmonic molecular surfaces, J. Comput. Chem.*, **1999**, *20*, 383.

6 *Chemical Applications of Atomic and Molecular Electrostatic Potentials. Reactivity, Structure, Scattering, and Energetics of Organic, Inorganic, and Biological Systems*, Politzer P; Truhlar DG; (Eds), Plenum Press, New York, NY, **1981**.

7 P. Sjoberg, J. S. Murray, T. Brinck and P. A. Politzer , *Average local ionisation energies on the molecular surfaces of aromatic systems as guides to chemical reactivity, Can. J. Chem.* **1990**, *68*, 1440-3.

8 B. Ehresmann, B. Martin, A. H. C. Horn and T. Clark, *Local molecular properties and their use in predicting reactivity*, *J. Mol. Model.,* **2003**, *9*, 342-347.

9 B. Ehresmann, M. J. de Groot, A. Alex and T. Clark, *New Molecular Descriptors Based on Local Properties at the Molecular Surface and a Boiling-Point Model Derived from Them*., *J. Chem. Inf. Comp. Sci.*, **2004**, *43*, 658-668.

10 B. Ehresmann, M. J. de Groot and T. Clark, *A Surface-Integral Solvation Energy Model : The Local Solvation Energy, J. Chem. Inf. Comp. Sci.,* **2005**, *45,* 1053-1060.

11 T. Clark, *The Local Electron Affinity for Non-Minimal Basis Sets*, *J. Mol. Model.*, **2010**, *16*, 1231-1238.

12 J. J. P. Stewart, *J. Mol. Model.*, **2007**, *13*, 1173-1213.

13 L. M. Loew and W. R. MacArthur, *A molecular orbital study of monomeric metaphosphate. Density surfaces of frontier orbitals as a tool in assessing reactivity, J. Am. Chem. Soc.,***1977**, *99*,1019-25.

14 B. S. Duncan and A. J. Olson, *Approximation and Characterization of Molecular Surfaces*; Scripps Insitute, San Diego, California, **1995**.

15 J.-H. Lin and T. Clark, *An analytical, variable resolution, complete description of static molecules and their intermolecular binding properties*, *J. Chem. Inf. Model.*, **2005**, *45,* 1010-1016.

16 G. Rauhut and T. Clark, *Multicenter Point Charge Model for High Quality Molecular Electrostatic Potentials from AM1 Calculations, J. Comput. Chem*., **1993**, *14*, 503 – 509.

17 B. Beck, G. Rauhut and T. Clark, *The Natural Atomic Orbital Point Charge Model for PM3: Multipole Moments and Molecular Electrostatic Potentials, J. Comput. Chem.*, **1994**, *15*, 1064 – 1073.

18 M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc*., **1977**, *99*, 4899 - 4907; 4907- 4917; *MNDO*, W. Thiel, *Encyclopedia of Computational Chemistry*, P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer,III and P. R. Schreiner (Eds), Wiley, Chichester, **1998**, *3*, 1599.

19 M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model, J. Am. Chem. Soc*. **1985**, *107*, 3902-3909;
A. J. Holder, AM1, *Encyclopedia of Computational Chemistry*, Schleyer, P. v. R.; Allinger, N. L.; Clark, T.; Gasteiger, J.; Kollman, P. A.; Schaefer, H. F., III; Schreiner, P. R. (Eds), Wiley, Chichester, **1998**, *1*, 8-11.

20 J. J. P. Stewart, *J. Comput. Chem.*, **1989**, *10*, 209 - 220; 221 - 264; PM3, J. J. P. Stewart, *Encyclopedia of Computational Chemistry*, P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer,III and P. R. Schreiner (Eds), Wiley, Chichester, **1998**, *3*, 2080.

21 W. Thiel and A. A. Voityuk, *Extension of the MNDO formalism to d orbitals: integral approximations and preliminary numerical results*, *Theoret. Chim. Acta*, **1992**, *81*, 391 - 404;
W. Thiel and A. A. Voityuk, *Extension of MNDO to d orbitals: parameters and results for the halogens*, **1996**, *93*, 315 - 315;
W. Thiel and A.A. Voityuk, *Extension of MNDO to d orbitals: parameters and results for silicon***,** *J. Mol. Struct*., **1994**, *313*, 141 - 154;
W. Thiel and A. A. Voityuk, *Extension of MNDO to d Orbitals: Parameters and Results for the Second-Row Elements and for the Zinc Group*, *J. Phys. Chem*., **1996**, *100*, 616 - 626;
MNDO/d, W. Thiel, *Encyclopedia of Computational Chemistry*, P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer,III and P. R. Schreiner (Eds), Wiley, Chichester, **1998**, *3*, 1604.

22 P. Winget, C. Selçuki, A. H. C. Horn, B. Martin and T. Clark, *AM1* Parameters for Phosphorous, Sulfur and Chlorine, J. Mol. Model*. **2003**, *9*, 408-414.

23 A. H. C. Horn, J.-H. Lin and T. Clark, *A Multipole Electrostatic Model for NDDO-based Semiempirical Molecular Orbital Methods, Theor. Chem. Accts,* **2005**, *113*, 159-168.

24 G. Schürer, P. Gedeck, M. Gottschalk and T. Clark, *Accurate Parametrized Variational Calculations of the Molecular Electronic Polarizability by NDDO-Based Methods, Int. J. Quant. Chem*., **1999**, *75*, 17.

25 D. Rinaldi and J.-L. Rivail, Calculation of molecular electronic polarizabilities. Comparison of different methods, *Theor. Chim. Acta* 1974, *32*, 243-251; J.-L. Rivail and D. Rinaldi, *Variational calculation of multipole electric polarizabilities*, Comptes Rendus, Serie B: Sciences Physiques **1976**, *283*, D. Rinaldi and J.-L. Rivail, *Molecular polarizabilities and dielectric effect of the medium in the liquid state. Theoretical study of the water molecule and its dimers*, *Theor. Chim. Acta* **1973**, *32*, 57.

26 B. Martin, P. Gedeck and T. Clark, *An Additive NDDO-Based Atomic Polarizability Model. Int. J. Quant. Chem*., **2000**, *77*, 473-497.

27 R. J. Abraham, B. D. Hudson, M. W. Kermode and J. R. Mines, *J. Chem. Soc., Faraday Trans. I*, **1988**, *84*, 1911-1917.

28 Tomasi, B. Mennucci and R. Cammi, *Quantum Mechanical Continuum Solvation Models, Chem. Rev.,* **2005,** 105, 2999.

29 C. J. Cramer, G. R. Famini and A. H. Lowrey, *Acc. Chem. Res.*, **1993**, *26*, 599 – 605.

30 A. Y. Meyer, *The Size of Molecules*, *Chem. Soc. Rev.,* **1986**, *15*, 449-475.

31 J. S. Murray and P. Politzer. *Statistical analysis of the molecular surface electrostatic potential: an approach to describing noncovalent interactions in condensed phases, J. Mol. Struct. (Theochem),* **1998**, *425*, 107-114;
J. S. Murray, S. Ranganathan and P. Politzer, *Correlations between the solvent hydrogen bond acceptor parameter β and the calculated molecular electrostatic potential, J. Org. Chem*., **1991**, *56*, 3734-3747;

P. Politzer, P. Lane, J. S. Murray and T. Brinck, *Investigation of relationships between solute molecule surface electrostatic potentials and solubilities in supercritical fluids, J. Phys. Chem*., **1992**, *96*, 7938-7643;
J. S. Murray, P. Lane, T. Brinck, K. Paulsen, M. E. Grice and P. Politzer, *Relationships of critical constants and boiling points to computed molecular surface properties, J. Phys. Chem*., **1993**, *97*, 9369-9373.;
T. Brinck, J. S. Murray and P. Politzer, *Quantitative determination of the total local polarity (charge separation) in molecules, Mol. Phys.,* **1992**, *76*, 609.

32 L. Mavridis, B. D. Hudson and D. W. Ritchie, *Toward high throughput virtual screening using spherical harmonic surface representations*, *J. Chem. Inf. Model.,* **2007**, *47*, 1787-1796.

33 **http://www.mdl.com/solutions/white_papers/ctfile_formats.jsp**

34 W. Thiel, *The MNDOC method, a correlated version of the MNDO model, J. Am. Chem. Soc.,* **1981**, 103, 1413-1420.