

New developments in the GDIS simulation package: integration of VASP and USPEX.

Okadome Valencia Hubert*, Busheng Wang†, Frapper Gilles‡, Rohl Andrew ‡

March 17, 2021

Abstract

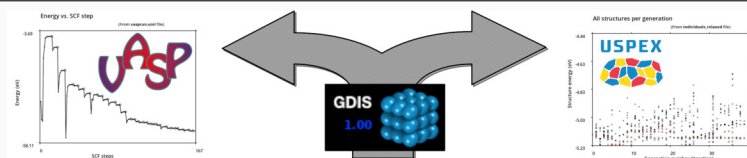
A popular first principles simulation code, the Vienna *Ab initio* Simulation Package (VASP), and a crystal structure prediction package, the Universal Structure Predictor: Evolutionary Xtallography (USPEX) have been integrated into the GDIS visualization software. The aim of this integration is to provide users with a unique and simple interface through which most of the steps of a typical crystal optimization or prediction work. This involved, for the latter, not only setting up a crystal structure prediction (CSP) calculation with complete support for the latest version of USPEX, but also displaying the many structure results by linking each structure geometry and its energy via interactive graphics. For the optimization part, any structure displayed by GDIS can now be the starting point for VASP calculations, with support for its most commonly used parameters. Atomic and electronic structures can be displayed as well as dynamic properties such as total energy, force, volume, and pressure for each ionic step. It is not only possible to start calculations from the GDIS visualization software, using an in-place task manager, but a running calculation can also be followed, allowing a greater control of the simulation process. The GDIS software is available under the GNU public license (GPL) in its second version.

Keywords: Visualization Software, Computational Chemistry, GDIS, VASP, USPEX. ■

*Institute of Materials and Systems for Sustainability, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan. Email: hubert.valencia@imass.nagoya-u.ac.jp

†Institut de Chimie des Milieux et Matériaux de Poitiers, UMR 7285, Université de Poitiers - CNRS 4 rue Michel Brunet TSA 51106, Poitiers 86073 CEDEX 9, France.

‡Curtin Institute for Computation and School of Electrical Engineering, Computing, and Mathematical Science, Kent Street, Bentley, Perth, Western Australia 6102



The GDIS software allows the visualization of chemical structures from the input and output (about 50 different file formats are supported) of 28 computational chemistry programs. Additionally, for 6 of them, GDIS can also be use as a graphical interface to setup calculations. In this work, the integration of a popular first principles simulation software, the Vienna *Ab initio* Simulation Package (VASP) and an equally popular crystal structure prediction package, the Universal Structure Predictor: Evolutionary Xtallography (USPEX) is detailed. The integration allows both the displaying of results and the setting of new calculations, making GDIS an all-in-one interface for 3D periodic system computation and prediction.

INTRODUCTION

The GDIS software¹ was developed to visualize molecular and periodic chemical systems from a wide variety of file formats (See Table 1). It provides a wide range of functionality including the ability to display a trajectory with a movie-like animation of the chemical system, measurements either on the display or in a table, isosurfaces (molecular, Hirshfeld, electron density, or promolecule surfaces), and diffraction patterns (X-ray, neutron or electron diffraction). Dynamical figures (pair count, radial distribution function, or measurements) can also be plotted as the result of, for example, a molecular dynamics (MD) calculation. GDIS also interfaces with the simulation software GULP², GAMESS³, MONTY⁴, and SIESTA⁵. Additionally, powerful tools allow the user to prepare structures to study surface dislocations, docking, equilibrium crystal shape, and to setup solvent / solute MD calculations.

Since the last release of GDIS (version 0.91 in 2015), there has been an increasing need for visualization software that is able to i) quickly display the thousands of structures generated by high-throughput simulation software, and ii) act as a graphical user interface (GUI) to start new calculations from any displayed structure. To this end, GDIS is first interfaced with the 3D-periodic density functional theory (DFT) software VASP^{6,7}. This provides an efficient way of optimizing periodic systems at the *ab initio* level, and VASP is a popular solution among the materials modelling community. We also choose to interface GDIS with USPEX⁸⁻¹⁰, a crystal structure prediction (CSP) package. It relies on an evolutionary algorithm to predict the crystal structures of a given chemical system. The validity of USPEX in rapidly finding new phases has been demonstrated by applications in various systems ranging from elements¹¹ to alloys¹², especially for high-pressure CSP.

The problem that high-throughput software, like USPEX, poses to visualization software is that they produce a considerable number of structural candidates, corresponding to the geometry trials of the CSP algorithm. Such results can be divided into several categories, ranging from the most stable structures, meta-stable ones – which may be stable under specific conditions, and the unreasonable ones, usually disregarded by the user. Unfortunately, there is no *absolute* definition of the barrier between the last two types of results, which is why the possibility to display all structures is necessary.

OVERVIEW

GDIS is written in C language, using the GLib¹³ library for portability, and interfaces with GTK¹⁴ and GTKGLext¹⁵ for the GUI part. Its source code is freely available¹⁶, under a GNU general public license v2.0¹⁷.

Installation is straightforward for UNIX-like systems: once the source package has been downloaded and unpacked, an installation script provides the necessary steps for compiling and installing GDIS. The GDIS software has been successfully compiled with three major compiler brands: GNUs GCC¹⁸, LLVM Developer Groups CLANG¹⁹, and Intels ICC²⁰. Due to the binding of the GDIS regular display operation to the graphical processing unit (GPU), the performance difference between the three compilers is negligible.

GDIS itself consists of 126 source and 79 header files, totaling about 180,000 lines of C code (excluding icon artwork). It can be typically compiled within a minute on desktop computers. Except for the window widget elements, which are rendered using GTK, the main area where the structures and graphs are displayed is rendered with OpenGL²¹ through a GTK widget dedicated to drawing. This means that all structures and graphs elements are rendered using the low-level OpenGL application programming interface (API). Text is rendered as an OpenGL object using the CAIRO²² library.

The user interface of GDIS is shown in Fig. 1-a, with a plain GTK theme and without any window decoration for clarity. The GUI is described below and a schematic of its different parts is available in the Figure S1 of the Supplementary Materials. On the top are the menu and tool bars, whose functions are detailed in Tables S1 and S2 of the Supplementary Materials, respectively. Below the tool bar, the GDIS window is divided vertically in two, with the main render area on the right and a tree view of the currently loaded chemical models on the left. The former is used to render chemical structures and graphics such as plots. The tree view area keeps track of all models currently opened by GDIS, using its name and an icon that represents the corresponding type of chemical system. It also displays a title and icon for the results of several analyses on a given model, as a child document of that model in the tree view area. The available icons and description are listed in the Table S3 of the Supplementary Materials. Below the tree view area is a combo-box that allows switching

between several utilities for displaying and editing the currently selected model. The default is “Model: Content” which displays some information about the chemical structure. Another combo-box immediately below is the select area. It allows the selection mode of the main render area from a list of possible modes to be set. The utilities and selection modes of both combo-box areas are described in the first section of the Supplementary Materials. The bottom of the GDIS user interface is divided into the logo and version part on the left and a status text output on the right. The latter is aimed at general information about the GDIS environment as well as reporting errors and warnings. The basic functionality of GDIS was detailed in a previous paper¹.

In the present work, to avoid confusion in the text, the elements that refer to GDIS GUI will be quoted. Similarly, the parameter word will always be reserved for 3rd party computational software and never be used to describe a GDIS function or GUI element.

RESULTS

VASP

The VASP software is a popular, 3D-periodic, first principles simulation software that began development in the start of the 1990s and quickly became popular initially due to its inclusion of ultrasoft (US) pseudopotentials²³ that resulted in a very efficient way to perform DFT calculation for metals²⁴. By the end of this decade, thanks to a quick 3D Fast Fourier Transform (FFT) implementation and efficient parallelization using the Message Passing Interface (MPI), VASP become to be seen as an essential piece of simulation software. Latter, the addition of the Projector-Augmented Wave (PAW) framework for pseudopotential^{25,26} and continuous integration of up-to-date methodology has kept it at the forefront of the field.

VASP functioning is quite simple to access: it uses several text files as inputs, and produces the corresponding results also as several text files. Each input has a format that is defined in a strict way, and 4 main files are usually necessary for a VASP calculation, namely: i) INCAR, containing a list of tags defining the methodology; ii) POSCAR, containing the

chemical system geometry; iii) KPOINTS, containing the k-points definition - for example using a Monkhorst and Pack (MP) grid^{27,28} - or the points geometry; iv) POTCAR, containing a pseudopotential definition, concatenated for each chemical species. The main output of VASP is contained in the OUTPUT file, together with specialized outputs, for example, CONTCAR, which gives the structure geometry of the chemical system after optimization, in the same format as the POSCAR input; or the CHGCAR and WAVECAR, which contain a grid representation of the charge density and wavefunction, respectively.

Displaying results

Since version 4.5, VASP has implemented a unified output file using the extensible markup language, XML²⁹. This means that each VASP calculation now issues a file, vasprun.xml, that contains a detailed summary of the calculation setting and results. We chose this format for interfacing with GDIS. When opening the VASP XML output, GDIS parses the whole file and sets a record of i) self-consistent field (SCF) energy for each SCF step; ii) geometry, energy, force resultant, and external pressure for each ionic step; iii) band structure and density of states of the final structure, when available; and iv) in case of a vibrational optimization, a table of each calculated frequency.

To be able to display an animated sequence of an ionic optimization, GDIS records the positions in the output file where each step geometry is defined. When rendering of a specific step is required, the output file is read from that point. Such a strategy proved to be especially efficient with dealing with long output files for large chemical structures, since only one step at a time must be loaded into memory. On modern operating systems, the smart caching of output files from a hard drive actually speeds this process up and no noticeable delay is observed. For small output files, the latter caching means that the whole output file ends up in the cache memory, which also speeds up reading significantly.

When opening a VASP output, GDIS displays its reference name, as given in the SYSTEM tag, on the tree view area together with the 3D-periodic system icon (see Table S3 of the Supplementary Materials for a list of all icons). Differently from other output files, a graph is also automatically attached as a child to the previous icon. An example of this graph can be seen in the Figure 1-b. It contains the energy values (in eV) for each SCF step.

To keep a good view of the SCF energy convergence, the initial SCF energy value of a new ionic step is always omitted. Additionally, a diamond symbol represents the last SCF step of a given ionic optimization. That symbol is clickable and will display the energy and ionic step number in the “Model: Content” part of the combo-box area. Switching from the plot child view to the parent 3D display will therefore show the geometry corresponding to that ionic step in the main render area.

To facilitate the interpretation of VASP results, some plotting functions have been added to the GDIS software. They can be found in the “tools/Analysis/Plots” submenu, which opens a small dialog shown in Figure 2. The energy, force, volume, and pressure for each ionic step can be plotted as a function of its index, using the functions from the first tab, “Dynamics”, as shown in Fig. 2-a. The total number of ionic steps is also provided in that tab for reference. The energy plot (see Fig. 3-a) is similar to the SCF one previously described, except that only the energy for the last SCF of each ionic step is displayed. The figure (eV/Å) plotted in the force graph (see Fig. 3-b) is actually calculated as the maximum norm resultant F on an atom of the chemical structure, that is:

$$F = \max_a^{N_{atoms}} \left(\sqrt{(f_a^x)^2 + (f_a^y)^2 + (f_a^z)^2} \right) \quad (1)$$

where max gives the maximum value for all atoms a indexed up to N_{atoms} , which is the total number of atoms. The f_a^x , f_a^y , and f_a^z terms are the force components on atom a in the x, y, and z direction, respectively. The volume (in Å³) of the cell containing the 3D-periodic chemical system can also be plotted for each ionic step (see Fig. 3-c). Finally, an external pressure value (in kB) is calculated from VASP triple vectors of the section called “stress” in the output. Averaging the norm of these 3 vectors leads to a pressure value that can be plotted by GDIS (see Fig. 3-d). For all graphs that can be plotted from this tab, each value indicated by a diamond symbol is clickable and will load the corresponding chemical structure in the parent render screen. While initially developed for the interpretation of VASP results, these new plotting functions also apply to optimization results from any package that can be read by GDIS and includes energy, force, volume, or pressure.

The second tab, “Electronic”, shown in Fig. 2-b deals with the plotting of electronic structures: DOS, band structure, or a combination of both. The numbers of DOS and band

values available in the output file are displayed for reference. The DOS is given as the calculated density (in state/eV) function of the energy (in eV), but the density values are shifted so that the origin of the energy axis (0 eV) corresponds to the calculated Fermi level for the material. An additional vertical axis (in orange, see Fig. 4-a) is provided to mark the origin of the energy at the Fermi Level. The band levels (Fig. 4-b) are given as energy values (in eV) as a function of k-point distance (in \AA^{-1}). Note that, due to the absence of Brillouin zone (BZ) definition in the output file, the k-points labels are not given, and it is up to the user to ensure that the correct BZ description was used and is plotted (in the KPOINTS file). In VASP, the band levels are given for each k-point, relative to the eigenvalues. It means that each n^{th} line that is displayed in the band diagram corresponds to the lowest n^{th} eigenvalues for each k-point. This might not correspond to a “natural” *ie.* symmetry related view of what a band usually refers to in a band diagram. Especially, band crossing is impossible; when two theoretical bands are expected to cross, the representation will display a first line corresponding to the lowest points of each band, and a second corresponding to their maximum. Finally, a combination graph with DOS on the left side, rotated by 90° anticlockwise, and the previous band diagram on the right is produced. The energy axis is then shared between the DOS and the band diagram, as shown in Figure 4-c. For clarity, the latter axis is added between the two parts of the graph in the same color as the Fermi level for DOS and band diagrams, which is also plotted. Note that as of version 1.0 of GDIS, only the sum of high and low spin values is displayed, in the case of a spin-polarized calculation.

The DOS and band structure can also be plotted for output from software other than VASP, with the same limitation for spin. The extension of GDIS electronic structure plotting to other software is an ongoing effort.

The last tab, “Frequency”, is shown in Fig. 2-c and allows for a simplified vibrational frequency graph to be plotted. Similarly to the other tabs, the number of vibrational and Raman modes present in the VASP output is given for reference. In VASP XML output, the frequency of a given vibration mode, calculated either by finite differences or by linear perturbation methods, is given as an array of eigenvalues. Such eigenvalues e_i are given relative to atomic unit for mass, \AA for distance, and eV for energy and have to be converted to s^{-1} or Hz, which is a frequency unit. Then, the value is scaled by $(2\pi c)^{-1}$ to cm^{-1} which

is the classical unit given in experimental infrared spectra. The overall conversion is done using the following formula:

$$\nu = \frac{1}{2\pi c} \sqrt{e_i \frac{e}{lm_A}} \quad (2)$$

with e the elementary charge, m_A the atomic mass, l a length conversion factor (10^{20}), and c the speed of light in vacuum (in cm/s). The Raman functionality is unused for VASP calculation at present, and no frequency intensity is calculated. This is due to the fact that there is no method, whether based on dipole moment or on dielectric calculation, that has been accepted in VASP (as of version 5) for the calculation of intensity. Both functionalities are therefore reserved for other codes and possibly future VASP use. At present, the plotting of vibrational frequencies in GDIS is limited to VASP calculations. Extending this to other computational software is under development.

Owing to the new plotting capabilities of GDIS, an additional “Plot Control” was added to the GDIS toolbar to manage the graph limits and tics, as well as text (title, sub-title, and each axis title). It is also possible to modify a value attribute (symbol and color), or a line type and color using that tool. Note that the text of a plot can be modified with that tool by using the GTK markup language¹⁴. It allows a broader set of text formatting using a HTML-like syntax.

Going beyond the simple displaying of the VASP results, we also add to GDIS the ability to follow a calculation dynamically by adding a “Track Mode” button to the tool bar. When opening an ongoing calculation’s output file, GDIS will produce a warning for “incomplete calculation”. Using the new tracking system, GDIS will start updating the SCF graph periodically (by default every 2 seconds) and register every new ionic step result it encounters. For the user, this is quite useful to determine when a geometry optimization diverges or when some problems were encountered during the electronic minimization. Since the tracking interval is not too short, it also allows a user to track several calculations at the same time. Finally, when calculations are run serially so that the same output file is rewritten, GDIS can detect that and restart the tracking each time. This allows a user to follow a multiple VASP calculations run without having to reopen the output file in GDIS at each step.

Setting up calculations

A GUI to configure and launch VASP calculations was added to GDIS. It consists of an 8 tab dialog that can be launched from the “Tools/Computation/VASP” submenu starting from any chemical system that is already selected in the tree view area of GDIS. This includes 3D-periodic systems, but also 2D-, 1D-, and molecular ones as well. Since VASP requires a strictly 3D-periodic system, for structures with lower periodicity, a vacuum is added within the missing dimensions. Its thickness defaults to 10 Å, but it can be adjusted afterwards. On top of the tabs are two fields for the system name and for using a previous VASP XML output. If the latter is used, the parameters will be taken from the latest set of parameters of that output. All parameters, except for those from a previous output, are given a default value, taken from the VASP guidelines (when available). All fields and elements of the VASP GUI tabs contain a short contextual help (tooltip) that is displayed whenever the mouse cursor hovers on it. The tooltips that relate to an existing VASP keyword display its exact name on its top left, followed by the relevant chapter of the VASP online help³⁰. The default or recommended value is then given after the “DEFAULT” keyword on the top right. The next lines of the tooltip contain a short description of the element. In most cases, setting a default value for a field will have it removed from the GDIS-generated VASP input. Note that when the contextual help does not follow this pattern, it means that the corresponding element does not correspond directly to an existing VASP keyword. In the following, most parameters presented by the GUI are already detailed in the VASP guide³⁰, so we will focus only on parameters that are GDIS additions and the ones that deserve special attention.

The first tab of the VASP GUI, called “PRESET”, is presented in Fig. 5. It contains a rough input generator for the most basic calculations: “Energy” for a single-point energy calculation, “DOS/BANDS” for an electronic structure calculation, “Lattice Dynamics” for a vibrational modes optimization, “Geometry” for a geometric optimization, and “Molecular Dynamics” for a MD calculation. The first two are single-point calculations, while the latter 3 are optimizations. The “ROUGH GEOM.” switch is for starting a calculation from an un-optimized geometry, supposedly far from the groundstate one. The dimension of the system, from 3D- to 0D-periodic (ie. molecular), is given in “DIM”. The electronic properties can

be chosen in “SYSTEM”, characterizing the material as metal, semiconducting, insulator, or unknown. The “POSCAR” field is a read-only indication of the detected species. The “KPT GRID” gives an idea of how dense the k-points grid will be chosen from a available “COARSE”, “MEDIUM”, “FINE”, and “ULTRA-FINE” selection. In the case of a molecular system, that parameter is ignored and only the Γ point is used, treating band dispersion as being unnecessary. For the other dimensions, we use the automatic procedure of VASP to generate a γ -centered MP grid using a length variable l . That l value is determined as $4d(c + 1)$ based on the dimension number d (ie. $d=3$ for 3D), and a density figure c corresponding to the user selection (0 for “COARSE” to 3 for “ULTRAFINE”). This number is then converted to the MP grid dimension $N_1 \times N_2 \times N_3$ by VASP, using the following:

$$N_i = \text{MAX}(1, l \|\vec{b}_i\| + 0.5) \quad (3)$$

where N_i and $\|\vec{b}_i\|$ correspond to a MP grid index number and the norm of the corresponding reciprocal lattice vector, respectively, and MAX is a function that gives the maximum integer between two inputs. The “POTCAR” read-only field indicates the detected species from the pseudopotential (POTCAR) files. The “POTCAR FILE” field and corresponding button allow the user to enter a specific pseudopotential file, which can also be prepared in the “POTCAR” tab of the VASP GUI. The unique file that can be selected in the latter field must be a concatenation of each chemical species pseudopotential in the order that appears in the “POSCAR” field. The latter field must therefore have the same species list in the same order as the “POTCAR” one, once a pseudopotential has been defined. Another limitation is that the pseudopotential methods and relative functional must be identical for all species. The local density approximation (LDA) cannot be mixed with the general gradient approximation (GGA), and pseudopotentials using the Perdew and Wang functional³¹ cannot be mixed with the one using Perdew, Burke, and Ernzerhof one³². Similarly, US²³ and PAW^{25,26} pseudopotential types cannot be mixed. Note that selecting a pseudopotential at that stage is optional. The pseudopotential file preparation can be more conveniently done by using the dedicated “POSCAR” tab of the VASP GUI. Finally, in the “EXEC” line, the NP, NCORE, and KPAR fields correspond to the total number of MPI tasks, the number of MPI tasks that will work on one band, and the number of k-points

that will be worked on by the previous number of MPI tasks simultaneously, respectively. The latter two fields correspond to the NCORE and KPAR keywords of the VASP input (INCAR). The three numbers are of course linked as NP must be lower or ideally equal to the product of NCORE and KPAR. Finally, clicking on the apply button will estimate, taking into account the information from all fields above, the recommended calculation parameters. The exact generated parameter list will be displayed in the “MESSAGE” text field. After that (optional) step, all parameters should be reviewed in the other tabs.

The second tab, “CONVERGENCE”, is presented in Fig. 6. It is dedicated to the parameters that define the algorithm for the convergence of electronic steps, the mixing of electronic density, and the real-space projector. Those are divided into 3 field sections: “General”, “Mixing”, and “Projector”, respectively. In the first field section, “General”, the “USE PREC” switch is a GDIS addition that the user can use to default the values of ENCUT and ENAUG tags in the VASP input (INCAR). When the switch is set, the latter two values are omitted and VASP will automatically retrieve the correct values from the pseudopotentials and the selected precision, given the table that is given for the PREC tag in VASP guide³⁰. The “AUTO” switch has a similar effect: when switched on, the NSIM, NBANDS, TIME, NELEC, and IWAVPR tags are omitted in the VASP input, and the default will be determined from VASP itself. Special attention should be given to “TIME” and “NELEC”, which should be set manually when dealing with damped algorithm (IALGO=5X) and charged systems, respectively. It is always possible to set only one of these values manually and set the other ones to default, in which case the corresponding tags will be removed from the VASP input. Because the default value for the “BAND” field can be difficult to evaluate, we also provide 0 as a value that will cause its removal from the VASP input. Note that the list of available elements in the ALGO field contains a special value added by GDIS which is “USE_IALGO”. Using that value will unlock the IALGO list and the IALGO tag will be selected in place of the ALGO one. Most of the mixing parameters in the next field section can be deactivated using the “AUTO MIXER” switch. The reason for such a switch is that, generally, the electronic density mixer is good enough to achieve the energy convergence for most systems, in most calculations scenarii. One notable exception is some meta-GGA calculations on a large system, for which custom AMIX/BMIX values might be necessary, while

eventually changing to IMIX=1. Special attention must be taken for the NELM, NELMIN, and NELMDL parameters which define the maximum and minimum number of electronic steps, and the number of initial non-SCF steps. When a previous VASP output is loaded, the 3 numbers might be far from ideal because they will correspond to the last optimization steps. In such a case, setting the values to their default: 60, -5, and 2, respectively, is a safe option. For the projector definition, the “LREAL” list must be set carefully. It determines whether the projection operators are evaluated in real or reciprocal space. When restarting from a previous VASP output, the ROPT field can be safely deleted, as it will most likely be reevaluated in the next calculation. The “AUTO GRID” switch allows to manually enter the number of FFT-mesh grid points NGX, NGY, and NGZ along the first, second, and third vector of the reciprocal lattice, respectively, and not in the x, y, or z direction as the names suggest. Similarly, the fine FFT-grid values NGXF, NGYF, and NGZF can be entered.

The “ELECT-I” tab is presented in Fig. 7. It deals with the setup of the electronic methods and general calculation settings. The first field section, “Smearing” deals with the occupancy filling method for each orbital. The “KGAMMA” and “KSPACING” fields of the same section deal with k-point spacing in the reciprocal space, in the absence of a k-point definition. Note that these 2 fields are usually ignored when the corresponding VASP input file (KPOINTS) is provided. The second field section, “Spin”, deals with spin coupling, that is spin-polarized and non-collinear spin calculations. In the event that a spin calculation is started from a previous VASP output, special attention must be given to the MAGMOM and SAXIS values. A safe way is to delete the previous values and let VASP recalculate them, unless some specific calculation settings are used. In the next field section, the exchange-correlation function is selected. Note that the “GGA” field must be consistent with the pseudopotential file provided; this assertion is not checked by GDIS at present. The “LMETAGGA” switch, which corresponds to the undocumented LMETAGGA tag in VASP, will allow the selection of a meta-GGA. The on-site Coulomb interaction L(S)DA+U (or GGA+U) methods^{33,34} are parametrized by the “METAGGA” field within the same section field. Note that for the Dudarev approach³³ of these methods, the “LDAUU” and “LDAUJ” fields values are not independent and only their difference is meaningful. Since for each of the latter tag, a value for each chemical species must be provided, we recommend

not setting the LDAUJ value to avoid confusion. The final section field of the first tab deals with dipole calculation. In the “IDIPOL” field, the u, v, and w axes correspond to the first, second, and third vectors of the structure lattice, respectively. Moreover, note that adding an electric field value by using the “EFIELD” field is only possible when a specific axis has been selected (ie. for IDIPOL values of 1 to 3), which is a limitation of the VASP software.

The “ELECT-II” tab is presented in Fig. 8. It is dedicated to advanced calculation methods and will be the topic of future improvements. For this version, it only contains the “Density of States” and the “Linear Response” fields. Special attention should be paid to the “LORBIT” list when a DOS method is requested, as the 10, 11, and 12 methods are only available when a PAW method and pseudopotential are used. The automatic PAW detection is just an indication and is only relevant when a single pseudopotential file is provided. Note that when a calculation is started from a previous output, the “RWIGS” field contains the Wigner Seitz radius for the sphere decomposition of the projected wavefunction for each species. These values should be carefully checked, to make sure they remain consistent between calculations.

The “IONIC” tab is shown in Fig. 9. It contains parameters for the ionic optimizations and is divided into 4 sections. The “General” section is used to set the main geometry optimization algorithm. Of note, the “ISIF” list field and the 3 other switches on the same line, “atom positions”, “cell shape”, and “cell volume” are interconnected. Selecting the required optimization with the switch will update the “ISIF” field to the correct value automatically. Note that some combinations are not allowed by VASP and therefore cannot be selected. For example, a calculation with relaxed atom positions and a variable cell volume, but a fixed cell shape. When setting a VASP calculation from an existing output file, the “NSW” field, corresponding to the maximum number of ionic optimization steps, must be carefully checked. The “Molecular Dynamics” section contains the parameters necessary for setting up a molecular dynamics (MD) calculation. All fields except “SMASS” are locked when the “IBRION” field is not set to 0. Since “SMASS” can be also used to set the damping factor when “IBRION” is set to 3, it is also unlocked at that time. The “POSCAR & SYMMETRY” section deals with both the chemical structure geometry input (POSCAR), and the parameters relative to the symmetry. The “DYN” field list is a GDIS

parameter that will either lock the position of all atoms, let them all move freely, or allow the user to select which atom will be allowed to move, and along which direction. In the corresponding VASP geometry input file (POSCAR), this will control the 3 T or F tags at the end of each atom coordinate. Note that the “DYN” field is disabled when the “Sel. Dynamics” switch is unset. The latter correspond to the eponymous VASP tag in POSCAR. Similarly, the “Direct Coord.” switch, also linked to the same VASP tag, will control if the coordinate of each atom will be entered in lattice or absolute coordinates, depending on the switch being set or unset, respectively. The “A0” and “U”, “V”, and “W” series of fields correspond to the lattice parameter and the X, Y, and Z components of each lattice vector in Å, respectively. They are written at the beginning of the VASP geometry input file, after the name of the structure, which is given by the “MODEL NAME” field at the top of the dialog. The “ATOMS” list switch is a GDIS field that allows to quickly select and modify the coordinates of an atom in the system selected by the “Direct Coord.” switch. The coordinates can be entered in the X, Y, and Z fields, while the atom index in the structure and its chemical symbol are also presented as an indication. The “Tx”, “Ty”, and “Tz” switches for each atom are only enabled when the “DYN” field list is set to “Manual selection”. They correspond to the T and F tags discussed above when the corresponding switch is set or unset, respectively, for each of the 3 directions given by the lattice vectors. All changes to an atom are only recorded when the “Apply” button is pushed. Using the last entry of the “ATOMS” list, “ADD atom”, it is possible to add a new atom by manually entering its position and, when relevant, Tx, Ty, and Tz tags. The “Delete” button can also be used to remove an atom from the list.

The “KPOINTS” tab is presented in Fig. 10. It is used to define the k-points in the corresponding VASP input file (KPOINTS). The first section of this tab, “from INCAR”, is a repetition of the parameters from the “ELECT-I” tab which have an influence on the k-points definition. For example, the tetrahedron method and its corresponding last section are only available when the “ISMEAR” field is -5. The “General” section defines the method that will be used to generate the k-points. Selecting in that field the “Automatic (M&P)” value will use the method that is described in the “PRESET” tab for the input generator. In that case, only the KX value needs to be entered, and it corresponds to the l value of

the Eq. 3. The “Gamma (M&P)” value corresponds to a MP grid that will be shifted to ensure that it contains the Γ point. It is especially useful when dealing with hexagonal lattices, for which a MP grid might not necessarily contain the Γ point. Conversely, the “Monkhorst-Pack classic” value is a MP grid which will not ensure it contains the Γ point. In both cases, the “KX”, “KY”, and “KZ” fields represent the index of the MP grid, and the “SX”, “SY”, and “SZ” ones an additional shift that can be added to the grid. The “Basis set definition” is a rarely used setting, in which a user will enter the definition of a basis set manually. Whenever this method is used, an equivalent MP could have been given, owing for the seldom use of this setting. Selecting the “Manual Entry” value means that all k-points and their weights will have to be defined individually. The number of k-points is then given by the “NKPTS” field. This mode is often used in combination with the tetrahedron method, to manually enter all tetrahedra. However, it can be used as is, while in that case the interest may be solely academic. Finally, the “Line Mode” method allows the user to easily prepare a band diagram calculation. In this mode, k-points are entered in pairs, and the “NKPTS” field corresponds to the number of equally spaced k-points that will be generated along each segment defined by a pair. Depending on the method, the k-points can be entered in the “Coordinate” section. That section functions similarly to entering the atom coordinates using the “ATOMS” field described above. The X, Y, and Z fields correspond to the lattice or reciprocal lattice coordinates, as chosen by the “Cartesian” switch in the previous field. The W value corresponds to the weight of a given k-point. In the last section of this tab, the tetrahedron can be entered. As indicated above, this section is only enabled if ISMEAR is set to -5 and “GEN” points to “Manual Entry”. The tetrahedra are entered as before, except that 4 k-points indices are needed to define the geometry using “PtA”, “PtB”, “PtC”, and “PtD”. The weight index of the tetrahedron should be given normalized with respect to the volume entered in the “V=” field.

The “POTCAR” tab is shown in Fig. 11. It allows to setup a proper pseudopotential file (POTCAR) for the VASP calculation. The first field section, “General”, only contains a reminder of the chemical species that are present in the structure. A pseudopotential is ready when the same species are detected in the “POTCAR results” section. Two methods are possible for this setup, the file or path method in the “get POTCAR information” section.

With the former method, a pseudopotential file that has been previously prepared must be provided. This has the advantage of being able to simply reuse a previous pseudopotential definition from a previous calculation. However, since this method is not really suited for newer systems or for using a different pseudopotential, the path method can be used instead. In the path method, the user has to provide the path where the POTCAR database from VASP was unpacked. Since a folder is available for each method, the user can easily choose a consistent method for the resulting POTCAR file. For example, in the VASP-provided database of pseudopotentials, the PAW-GGA, PAW-PBE, and US-GGA directories correspond to the pseudopotentials generated from PW91, PBE, and PW91 functional using the PAW, PAW, and US framework within the GGA, respectively. Once a folder has been selected, the last section should display the identical “SPECIES” and “FLAVOR” field as the one from the first section. Using the path method, one can also select a variation of the pseudopotential (called flavor here), which is given in the “FLAVOR” list. In the list, the value after the chemical symbol, when present, will determine the flavor. The possible flavors are `_h` and `_s`, for hard and soft pseudopotentials, respectively, and `_pv`, `_sv`, and `_d` for the ones that will let VASP treat s, p, and d semicore states as valence states, respectively. When only the chemical species appear, it means that the regular pseudopotential is selected, when available. For some pseudopotentials, a terminal suffix `_GW` can be found for use with the VASP implementation of the GW method³⁵. Note that in this version of GDIS, the GW calculation method is not supported and therefore the corresponding POTCAR should not be used.

The last tab, “EXEC”, is presented in Fig. 12. It deals with the execution of the VASP code. The “General” field section shows the detected path for the VASP executable. To be detected by GDIS, the latter must be called “vasp” and be in the system executable path. Next is the folder path in which the calculation will take place. It is important to note that the input files: INCAR, POSCAR, KPOINTS, and POTCAR will all be written in that path and the calculation will start there. In case some other files exist in the folder, for example from a previous calculation, they might or might not be used, depending on the user parameters. The “Parallel Optimization” field section contains the detected mpirun program, when relevant, which allows a parallel VASP calculation. The “NP”, “NCORE”,

and “KPAR” fields are a copy of the ones in the first “PRESET” tab. A remote execution of VASP is planned for a later version of GDIS.

When all settings have been reviewed, the user has the option to save the input files, execute a VASP calculation, or close the dialog losing all the parameters, using the three buttons at the bottom of the VASP calculation dialog. When the “Execute” button is selected, several actions are taken:

- the input files are saved;
- VASP calculation is launched in a GDIS task;
- the VASP calculation dialog is closed;
- a new model is created, connected to the new VASP output;
- the tracking mode is set on that new model.

Immediately after the VASP dialog is closed, the user will see nothing but an empty structure display. This is because the model will be only updated after the first ionic step of the VASP calculation. However, it is possible to use the GDIS integrated task manager, located in the “View/Task manager” menu. An example of the view given by the task manager of a running VASP calculation is given in the Figure 13. In this view, if the calculation is running, the process identification number (PID), name, and status of the calculation are given in the top grid, while an indication of CPU, memory, and time usage follows. The latter three are given only as an estimate and will only correspond to real figures when the UNIX command “ps” is properly configured. It is also possible to abort a calculation using the task manager, by selecting a job and using the “Kill selected” button. The “Kill all” button will abort all calculations present in the task manager, and the “Remove completed” will remove non-running calculations. The “Close” button will only close the task manager dialog, without any consequences on the task list. When the first geometry optimization step (ionic step) is finished, the main render area will be updated with the current geometry, and a graph displaying the SCF energy versus the step number will be displayed and updated for each new SCF value. The main render area will also be updated at the end of each

ionic step. When the calculation is finished, the job status in the task manager will switch to “Completed” and the PID, CPU, memory, and time figures will be removed. At that point, the graph and display will stop being updated, but the current tracking of the output file will only be stopped if the tracking button is pressed or if that file is removed. This allows for subsequent VASP jobs that will rewrite the same output file (vasprun.xml) to be automatically tracked with no need for opening the same file at each time.

USPEX

USPEX is a software for the prediction of crystal structures that uses an evolutionary algorithm designed to tackle the exponential complexity of Crystal Structure Prediction (CSP) based on randomly generated structures. Owing to its success in the field, especially for the discovery of new high-pressure crystal phases, USPEX was later enriched with other methods and algorithms. Initially interfaced only with VASP, it is now, as of version 10.4, interfaced with over a dozen computational software engines, including SIESTA⁵, GULP², QuantumEspresso³⁶, CP2K³⁷, CASTEP³⁸, LAMMPS³⁹, ATK⁴⁰, MOPAC⁴¹, Gaussian⁴², ORCA⁴³, FHI-aims⁴⁴, and ABINIT⁴⁵. The USPEX methods now include, in the approximate order made available to users, a variable-composition CSP method^{46,47}, a corrected particle swarm optimization (PSO) algorithm⁴⁸, the CSP of molecular crystals⁴⁹, an evolutionary metadynamics approach^{50,51}, a variable-cell nudged elastic band (VC-NEB) method^{52,53} for finding crystal transition states, the CSP of 2D crystals^{54,55} and surfaces⁵⁶, the CSP of 1D-polymer crystals⁵⁷, and a transition path sampling (TPS) method⁵⁸⁻⁶⁰. The original USPEX evolutionary algorithm itself has also evolved and notably includes advanced features such as fingerprinting⁶¹, seeding (and the use of antiseeds)¹⁰, and space group determination⁴⁶. USPEX can optimize a number of different objective functions, for example enthalpy and volume minimization, or the maximization of hardness, band-gap, or magnetization values. It is also capable of counter-optimizing all these functions, ie. maximizing the volume instead of the default minimization. Recently, a multi-objective Pareto optimization⁶² was added to deal with several objectives at once, and the optimization of thermoelectric properties was made possible using the BoltzTraP⁶³ code. Additionally, random structures can be obtained

using a new topology-based algorithm⁶⁴.

USPEX can be largely viewed as a collection of MATLAB⁶⁵ scripts – despite a recent move to the Python⁶⁶ language. These scripts are developed to provide USPEX core algorithms, whilst using external calculations of individual structure properties. For example, USPEX main evolutionary algorithm can be fed the results of first-principle calculations performed by VASP to determine the best candidate for a crystal at a given composition.

Displaying results

One difficulty in visualizing the results of an USPEX calculation is the large number of structures that are generated at each iteration (also called a generation) by the evolutionary algorithm. A typical small calculation, using the default USPEX parameters, would allow as many as 20 different structures per generation, for a maximum of 100 generations. This would mean a grand total of 2000 structures should be visualized to represent the full extent of the USPEX calculation. In GDIS, we choose to access each structure through a series of graphs. The first graph, called ALL, is provided regardless of the USPEX calculation type. It plots the energy of each structure as a function of the generation number. This was the visualization strategy used by the STM4 visualization toolkit, a module developed by M. Valle⁶⁷ for AVS/Express⁶⁸. Another strategy would have been a serialized display of structures, similar to the implementation in the VESTA code⁶⁹, in which a structure can be selected and displayed among an ordered listing of all structures. An example of the ALL graph is given in Fig. 19-a for an USPEX calculation of the possible phases of Sn_xN_y from our previous work⁷⁰. In this figure, each square represents the energy of a given structure. The red colored squares indicate the structures marked as best by USPEX for a given generation. Note that in the case of variable composition optimization, there is more than one best structures per generation. Clicking on any square will display the corresponding structure in the main render area. In the “Model: Content” panel, the data is also updated: the total atom number, total number of molecules, total number of electrons (reported as a total charge), ground-state energy, lattice volume, crystal symmetry, chemical formula, and structure index. All information is calculated by GDIS except for the ground-state energy and crystal symmetry which are provided by VASP (through USPEX results) and USPEX,

respectively. While this makes little sense in a regular USPEX calculation, all structures can also be displayed as an animation in the parent display. Each structure then accounts for a frame of the animation, in the same order as they are generated by USPEX, similarly to the way they are displayed using the VESTA code⁶⁹.

For VC-NEB calculations, the animation can actually be seen as the visual representation of the reaction path, starting from a minimum, reaching a transition state, and relaxing to another minimum. In this case the ALL graph is called PATH. However, this trajectory representation is only valid when both extrema have been carefully selected and the VC-NEB calculation has a sufficient number of intermediary points (structures). The new Transition Path Sampling (TPS) method⁷¹ available in newer USPEX versions increases the confidence in such representation. Unfortunately, at present, the TPS method cannot be displayed by GDIS, owing to its new output format. When an USPEX job is restarted, either to seek additional generations or due to an unexpected termination of the program, the structure energies from the previous runs are preserved by USPEX while the corresponding geometries are not. Since the previous calculation might not be available, we have chosen in GDIS to replace the usual square symbols with crosses, meaning that the energy is available but the chemical structure cannot be displayed in the parent view. Another possibility is to concatenate the old and new files containing the structure geometries (files ending with POSCARs) to retrieve a full set of structures available for GDIS to display. The choice is left to the user whether to concatenate the output files or not.

For the USPEX, metadynamics, MINHOP, and PSO methods of USPEX, a second graph called BEST is added to the first one. The graph also represents the energy of structures depending on the generation number but only for a restricted set of candidate structures marked as best by USPEX (from its output files starting with BEST). The corresponding BEST graph from our previous example is shown in Fig. 16-b. This data corresponds to the red-colored symbols on the previous ALL graph, but it is more readable as other structures are not presented. In the case of metadynamics and MINHOP methods, GDIS will display relaxed cell energies whenever available. Usually, a search for a single chemical composition will lead to a simple, one-line BEST graph. With variable composition optimization, there will be many best structures per generation; a line will then join the lowest energy structure

for each generation. In that case, the line is only a guide for the eye and does not bear any physical or chemical meaning.

When a variable composition result is visualized in GDIS, a supplementary graph appears for each atomic element, called COMP_ a , where a is the chemical symbol of the corresponding atom. In this graph, the structure energies are given as a function of the atomic ratio X_a , which is simply:

$$X_a = \frac{N_a}{N_T} \quad (4)$$

with N_a the number of a atoms and N_T the total atom number of the structure's chemical formula. When the variable composition optimization is performed on two elements, the resulting graph is similar to that in Fig. 19-c and 19-d of our previous example. In such binary systems, the lines given by GDIS correspond to the convex hull, as defined in Ref. 8. When the references for X_a extrema are chosen appropriately, the relative stability of a structure can be inferred relative to its position toward the convex hull. That meaning is however lost for ternary systems and systems with a higher number of elements. For those systems, each COMP graph represents only a single atom point of view: for a ternary ABC system, the COMP_A graph gives a lowest energy plot value for a single A composition, whereas several combinations of B and C compositions in ABC can be stable with different energies. Another representation for these systems is currently under development. In GDIS, the calculation of a convex hull relies on the same lowest reference energies for X_a extrema as those provided by the USPEX calculation. The exact algorithm for the calculation is described in the section S2 of the Supplementary Materials. For binary systems, the choice of a reference is important for the relative stability consideration of a given composition. Ideally, the reference for an AB binary system would be the most stable form of pure A and B containing materials, whilst other references might provide insights, for example, in the synthesis of metastable materials. However, such values may be unavailable due to the restriction that pure A and B containing materials must be part of the USPEX calculation to be included in the USPEX convex hull. In the GDIS case, the possibility of using external references was added by including a simple text file (chem.in) in the USPEX directory. The syntax of this file and the use of this functionality are described in the section S2 of the Supplementary Materials.

GDIS is able to visualize the output of USPEX calculations for versions ranging from 9.4.4, which is released under a modified BSD license, to the latest (at the time of writing) version 10.4 which requires registration. Note that across these versions, the format of the files containing geometric structures (POSCAR) has changed from VASP version 4 format to that of version 5. This change is silently dealt with in GDIS.

Setting up calculations

The USPEX GUI in GDIS consists of 5 tabs, each dedicated to a different aspect of the calculation. It can be launched from the “Tools/Computation/USPEX” submenu, and its design is similar to the VASP GUI described in the previous section. To launch the GUI, a model must have been selected in the tree view area. However, in contrast to launching a VASP calculation, the USPEX GUI can be opened from an empty model, created from the “Add Model” tool (see Table S2 of the Supplementary Materials). When the USPEX GUI is launched from an existing model, its chemical formulae will be used to initialize the corresponding atom types and number definitions. On the top of the GUI, similar to the previous VASP GUI, two fields are present. The first one, “MODEL NAME”, allows a name to the USPEX calculation to be provided, which is a GDIS feature. The second field, “CONNECTED OUTPUT”, and its associated button, allows the user to load a previously defined set of parameters to the GUI. Contrary to what its name suggests, the valid files that can be loaded are either the ones created by USPEX that summarize the parameters of the calculation (Parameters.txt) or any input file that contains parameters for the calculation (INPUT.txt). Contextual help similar to that of the VASP GUI is also available for all fields of the GUI. It provides the exact keyword, the relevant chapter in the latest version of the USPEX manual⁷², the default value when relevant, and a short description. The GUI allows the setting of all documented USPEX keywords (plus some undocumented ones), which, to our knowledge, makes it the most complete USPEX interface to date.

The first tab of the GUI, “SYSTEM”, is presented in Fig. 14. It deals with the most general parameters of USPEX calculations. Choosing the calculation method in the “Method” field has an impact on many other fields in all tabs. It is therefore recommended to set it as soon as possible within the use of the GUI. The metadynamics “META”, VC-NEB

“VCNEB”, and PSO “PSO” selectable elements of the “METHOD” combo-box will actually bring up the last tab of the GUI: “SPECIFIC”, which will be described later. The “DIM”, “MAG”, “MOL”, and “VAR” fields will set the dimension, magnetism, molecular, and variable composition part of the unique “Type” field. It corresponds to the calculationType parameter of USPEX. Note that only supported combinations will be allowed, it is not possible to select a molecular calculation together with magnetism, for example, as that type of calculation is not yet available in USPEX (as of version 10.4). The “Ver 10.4” checkbox is a GDIS addition that allows the use of either of the supported USPEX versions: 9.4.4 or 10.4, depending on if the field is unchecked or checked, respectively. The “atomType” field corresponds to several USPEX parameters: the atomType parameter, which gathers the chemical element symbols used in the calculation, the numSpecies parameter, which defines the number of each element, and the valence parameter, which sets the valence of each species. Each parameter can be set by the “@Sym”, “@Num”, and “@Val” fields, respectively. The protected “@Z” field gives the atomic number of each species as an indication. When applying the modification using the neighbour “Apply” button, the fields will be applied to the selected “atomType” species. If “ADD ATOMTYPE” is selected, a new species will be created, unless the species is already in the list, in which case it will become selected. The “atomType” field is initially filled with values coming from the model that was selected in the tree view area. It can also be filled by loading a parameter file in the above “CONNECTED OUTPUT” field. In the case of variable composition optimization, the field “numSpecies” will be unlocked for modification. Several species blocks can be entered or modified to control the stoichiometry of each building block. For example, if the field consists of 0, 1, and 1 in the first line and 2, 1, and 0 in the second, a $A_x B_y C_z$ variable composition search will be done using BC and $A_2 B$ blocks so that $A_x B_y C_z = nBC + mA_2 B$; with m and n two positive or null integers. The “AUTO_BONDS” checkbox is a GDIS addition that allows switching between the implicit values for the goodBonds parameters generated by USPEX, and a manual entry of the matrix. The “NEW” checkbox can only be modified when the “Ver 10.4” switch is unchecked. It allows the optimization parameters to be entered by a single line (the legacy way) or using a block type notation (the new way) when the “NEW” switch is checked. The reason for this switch is that some USPEX version, between 9.4.4

and 10.4, might accept either way. The “AUTO_LAT” switch allows the user to manually define the USPEX Latticevalues parameter. This parameter predefines a cell parameter in which the first structures will be generated. There are three ways to define the Latticevalues block in USPEX. The first one is by entering the volume of a cell – or of several cells when performing a variable composition optimization. Another way is to provide the three vectors (2 for 2D materials) of an ideal lattice. Finally, crystallographic notation ($a, b, c, \alpha, \beta, \gamma$), or reduced 2D notation (a, b , and α) can be used within a single line. All 3 methods of entry are supported in GDIS and can be entered using the dedicated “VALUES” field. The “AUTO_ION” similarly allows for the manual entry of the IonDistances parameter. This parameter contains a matrix of the minimum allowed distances between species in a structure. Each line of the matrix can be modified using the dedicated “DIST” field. The “Mol” field, which defines the MolCenters parameter, is only unlocked when a molecular calculation has been set (in the above “MOL” field). This allows the definition of a minimum distance allowed between molecule centers. It is modified using the “CENTER” field.

The “STRUCTURE” tab is presented in Fig. 15. It gathers information about the population, the variation, and the compositions of generated structures within the USPEX calculation. In the first area, the fields of the second and third lines will only be unlocked if the “MAG” switch was checked in the previous tab. To avoid unnecessary navigation between the tabs, a copy of the “MAG” switch is also present in this tab. Note that, if a molecular system is selected, checking the “MAG” switch will not unlock the fields of the second and third lines, but it will uncheck the “MOL” switch of the first tab. In that case, “MAG” needs to be unset and set again, which will allow the setting of magnetic properties, but with the molecular setting being lost. Among the fields dealing with the structure and variation, the “D_LATMUT” is linked to the now undocumented DisplaceInLatmutation parameter. This parameter was removed from USPEX documentation around version 10.1 and should be carefully used, if possible, only with lower versions. In the fingerprint part of the tab, the second line “Max” and “sigma” fields, corresponding to the antiSeedsMax and antiSeedsSigma parameters, respectively, have different default and recommended values, as mentioned in the short description of the corresponding tooltip. The default values are actually for cases where the anti-seeding technique is not used, while the recommended ones

are for use with antiseeds. Finally, the fields of the variable composition will only be unlocked if the “VAR” switch was checked in the previous tab. In that case, the “min@” and “max@” fields, corresponding to the minAt and maxAt parameters, respectively, should be carefully reviewed. In a variable composition optimization, there is no default for these parameters and both are mandatory.

The “CALCULATION” tab is shown in Fig. 16. It allows the setting of the execution parameters for the external energy calculation software used by USPEX and for its own execution. To give the calculation parameters to external software, USPEX relies on a directory named Specific. Since USPEX generated structures are usually far from the ones optimized using the external software, they can take an unusually long time for the optimization to converge – and it can also fail. To address this issue, USPEX can divide the optimization into several steps of increasing accuracy. This strategy has proved quite effective, especially with *ab initio* software. Consequently, the Specific directory generally contains one input file (with an optional complementary one) for each calculation step. An added underscore and index at the end of each input file identifies which step this file belongs to. Using the GUI, there are 3 ways to create the files of the Specific folder. In the first, simplest version, the user is assumed to have already prepared a Specific folder. The folder can be copied by using the “USE Specific Folder” radio button. In that case, the folder should be provided in the “SPE” field or using the neighbouring “Open” button. In the second version, the “Set Specific Folder” radio button must be selected and the “AUTO STEP” switch must be checked. In that case, for the computation software supported by GDIS, at present VASP and GULP, GDIS will actually generate a ‘good enough’ set of inputs. This set of inputs should lead to a proper USPEX calculation when the “LIB”, “FLAVOR”, and “EXE” fields are properly filled (see below). Using GDIS generation also allows GULP and VASP steps to be mixed, for example, starting with 3 steps of GULP, then 4 steps of VASP. Finally, the last version corresponds to the case where “Set Specific Folder” radio button is selected and “AUTO STEP” is unchecked. In such cases, all input and complementary files must be provided in the “INP” and “OPT” fields, respectively, for each step. Regardless of the chosen solution, the recommended number of steps should be 4 to 6, while higher and lower values are possible. Note that for VC-NEB or TPS calculations, the number of steps should be 1,

since geometry optimization is not needed. The “LIB” field has a special meaning: while it is copied to the Specific folder, it will not be indexed by the step number. This allows the use of a common pseudopotential set for VASP or a forcefield library definition for GULP. Most computational software supported by USPEX also relies on external files for the definition of potentials, which should then be provided in that field. For VASP and GULP, the “LIB” field should correspond to the directory in which the potentials are extracted for VASP (similarly to the VASP GUI) and in which the force field definition files are kept for GULP. The “FLAVOR” field then allows to select a force field definition file for GULP, which can be changed at each step, or a set of pseudopotentials for VASP, which should be constant through all VASP steps. Note that the “EXE” field should contain the command to launch the computational software in the `commandExecutable` parameter format, with the parallel launcher and file redirection where applicable. In the USPEX launch part, the “Ver 10.4” field is a copy of the one in the first tab, provided here for convenience. Note that the “CPU” field is linked to the undocumented `numProcessors` parameter. The “OCTAVE” checkbox is only unlocked when the “Ver 10.4” is unchecked. It allows USPEX to be launched within the Octave or Matlab interpreter. The “Folder” field defines where the USPEX input will be generated and the corresponding optimization run. It should be carefully checked for existing files that will, by default, not be deleted. The “RESTART” field corresponds to the `pickUpYN` parameter. It has been deprecated as of version 10.0 of USPEX. It was used to indicate that a restart of USPEX from a specific generation had been requested. It is now superseded by simply providing the restart generation in the `pickUpGen` parameter, which is available in the neighbour “GEN” field. The “RESTART” field is then provided for compatibility and is silently ignored in the latest USPEX versions. When restarting a USPEX calculation or starting one anew, the “CLEANUP” switch, a GDIS addition, will help to remove the files that would prevent USPEX from starting.

The “ADVANCED” tab is shown in Fig. 17. It is dedicated to parameters that are less commonly used or that refer to the recently introduced thermoelectric (the BoltzTraP interface) and TPS methods. The parameters for developers and seldom used parameters should only be modified by experienced users. Of the latter category, the “DYN_HM” and the “SoftMut” fields, corresponding to the `dynamicalBestHM` and `softMutOnly` parameters,

respectively, have disappeared from the USPEX manual starting from the 10.1 version. Since it is unclear if the latter parameters have been deprecated, GDIS still allows the user to modify the values in the corresponding fields. However, we do not recommend to set these parameters out of their default values. In the BoltzTraP interface, the “cmd” field is intended to let the user externally process the results. Since BoltzTraP inclusion is still experimental in GDIS, this field is silently discarded at present. In the TPS part of the tab, the fields “cmdOP” and “cmdET”, corresponding to the cmdOrderParameter and cmdEnthalpyTemperature parameters, respectively, must be provided by the user. They correspond to the scripts (or small programs) that are used to calculate the order parameter and to extract the enthalpy and temperature from the results of the analysis, respectively. Both have no default value and must be provided. When the fingerprint method is used, which means the “OP_TYPE” switch is checked, only the latter “cmdET” is mandatory. Note that GDIS defaults to using the fingerprint method, while USPEX does not recommend a default method. Even though GDIS does not support displaying the results of a TPS optimization at present, such calculations can still be launched. However, the model will fail to load shortly after the USPEX calculation is launched.

The last tab of the USPEX GUI, “SPECIFIC”, is presented in Fig. 18. It deals with the parameters required for the metadynamics, the PSO, and the VC-NEB methods. Specific parameters for molecular and surface optimization are also provided here. For metadynamics calculations, a good starting point structure is required. It can be either provided by an external file, which should be called POSCAR_1, or by a previously loaded GDIS model. In the latter case, a model can be selected by its name in the “MODEL” combo-box. The latter should have been filled by the available GDIS models when the USPEX GUI was launched. Note that this feature is still experimental: models converted to the POSCAR_1 file will require careful examination before use. For VC-NEB calculation, the vcnebType parameter defining the method will be set by choosing the corresponding “Method”, “Var_Image”, and “Var_Spring” fields. The “Format” field, which corresponds to the FormatType parameter, can be set to output formats other than the default VASP format. It corresponds to the optimised trajectory, and GDIS can read in any of the suggested formats. However, for consistency, the VASP format is recommended. The “Model” field functions similarly to the

“MODEL” format above. One exception is that when a model is chosen, instead of providing the required Images file, such a model should at least contain 2 frames representing the initial and final images of the VC-NEB path optimization. Unfortunately, GDIS at present will not check this prerequisite and therefore selecting a previously loaded model is not recommended. When a molecular crystal simulation is selected, each molecule must be provided as a MOL_i file, where *i* is the index of the molecule. At present, the USPEX GUI supports 3 different methods to provide these molecule fields. The first method is triggered when the user selects the “Already provided” element from the “MODEL” combo-box. In that case, the molecular files are supposed to be already present in the USPEX calculation directory. This is the recommended method. In the second method, when “From MOL_ folder” is selected in the combo-box, a folder must be supplied using the “Open” button and the number of molecules indicated in the “N_MOLS” field will be copied. When a folder has been selected, it will replace the “From MOL_ folder” element in the combo-box. This is an experimental feature, mostly used when similar calculations are performed. The last possibility is to select some already loaded models from GDIS, similarly to the above, which is also experimental. For surface optimization, the “MODEL” field functions similarly to the other “Model” or “MODEL” fields. It is used to provide the substrate model for the calculation. As before, it is recommended to provide a VASP formatted input rather than using the experimental GDIS model selection. The “Stoichio” field represents the undocumented StochiometryStart parameter. This parameter, to our knowledge, defines the formula unit that will be used in a 1x1 dimension cell.

CONCLUSIONS

The GDIS visualization software has now been extended to interface with the VASP software. This means that it is now capable of quickly displaying the results of a VASP calculation. In addition, the dynamic properties, the electronic structure, and vibration frequencies of a VASP calculation can be displayed. While these functions were initially developed for use with VASP, they are still the focus of many improvements. Notably, their extension to the many programs that generate this kind of data. GDIS was also able to monitor VASP calculations. While this is no different than examining the many figures and numbers that are created by VASP during its execution, it is arguably more convenient to be able to *visually* assert that the electronic convergence of a calculation is going wrong. Finally, the ability to launch a VASP calculation from a loaded model, regardless of its origin, was added in GDIS. While the GUI for VASP calculations currently misses some of the most advanced features such as Hartree-Fock and GW calculations, the current interface covers all DFT-related keywords, which is (at the time of writing) sufficient for most routine calculations. Some work is currently undergoing to extend GDIS support to all VASP keywords.

On the other hand, displaying the results of high throughput software such as USPEX is anew to GDIS capability. For this, a strategy of interactive graphs was developed which proved to be an efficient solution. Similarly to VASP, the following of an optimization from the USPEX software is a convenient addition. Admittedly though, it is more difficult to visually determine when such optimizations are failing. The USPEX GUI actually covers *more* than 100% of the documented USPEX keywords, due to the undocumented and GDIS-specific additions. Consequently, most if not all possible USPEX jobs can be launched using GDIS. There is still some room for improvement since the USPEX team is always adding new features.

Combining both additions makes GDIS a unique tool for computational scientists. The ability to simply start a USPEX calculation from a new, empty model can be combined with performing a VASP calculation on any loaded model – including a resulting USPEX structure – to provide a uniform user experience. From this, with the GDIS tools for interpreting the results and its ability to follow calculations, the goal of performing regular computation work

without having to switch back and forth between programs seems within reach.

ACKNOWLEDGMENTS

This work was supported by the MEXT “Program for Research and Development of Next-Generation Semiconductors to Realize an Energy-Saving Society” under the Contract No. JPJ005357.

References

- [1] S. Fleming and A. Rohl, Zeitschrift Fur Kristallographie - Z KRISTALLOGR **220**, 580 (2005).
- [2] J. D. Gale and A. L. Rohl, Molecular Simulation **29**, 291 (2003).
- [3] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, et al., J. Comput. Chem. **14**, 1347 (1993).
- [4] S. X. M. Boerrigter, G. P. H. Josten, J. van de Streek, F. F. A. Hollander, J. Los, H. M. Cuppen, P. Bennema, and H. Meekes, The Journal of Physical Chemistry A **108**, 5894 (2004).
- [5] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, Journal of Physics: Condensed Matter **14**, 2745 (2002).
- [6] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).
- [7] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).
- [8] A. R. Oganov and C. W. Glass, The Journal of Chemical Physics **124**, 244704 (2006).
- [9] A. R. Oganov, A. O. Lyakhov, and M. Valle, Accounts of Chemical Research **44**, 227 (2011).
- [10] A. O. Lyakhov, A. R. Oganov, H. T. Stokes, and Q. Zhu, Computer Physics Communications **184**, 1172 (2013).
- [11] Y. Ma, M. I. Eremets, A. R. Oganov, Y. Xie, I. A. Trojan, S. A. Medvedev, A. O. Lyakhov, M. Valle, and V. B. Prakapenka, Nature **458**, 182 (2009).
- [12] A. Bilic, J. D. Gale, M. Gibson, N. Wilson, and K. McGregor, Scientific reports **5**, 9909 (2015).
- [13] *GLib*, <https://wiki.gnome.org/Projects/GLib>.

- [14] *GTK+*, <https://www.gtk.org>.
- [15] *gtkglext*, <https://projects-old.gnome.org/gtkglext/index.html>.
- [16] *GDIS*, <https://github.com/arohl/gdis>.
- [17] *GNU GPL v. 2.0*, <https://www.gnu.org/licenses/old-licenses/gpl-2.0.txt>.
- [18] *GCC*, The GNU Compiler Collection, a set of compiler which include a compiler for C language; <https://gcc.gnu.org/>.
- [19] *Clang*, Clang is a front-end set of compilers which include a compiler for C language compatible with GCC; <https://clang.llvm.org/>.
- [20] *ICC*, The Intel C++ compiler is a C-compatible compiler profided by Intel; <https://software.intel.com/en-us/intel-compilers>.
- [21] M. Segal and K. Akeley, *The opengl graphics system: A specification (version 4.6 - core profile)* (2019).
- [22] C. D. Worth and K. Packard, *Cairo: Cross-device rendering for vector graphics* (2003).
- [23] D. Vanderbilt, *Phys. Rev. B* **41**, 7892 (1990).
- [24] G. Kresse and J. Furthmller, *Computational Materials Science* **6**, 15 (1996).
- [25] P. E. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
- [26] G. Kresse and D. Joubert, *Phys. Rev. B* **59**, 1758 (1999).
- [27] H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- [28] J. D. Pack and H. J. Monkhorst, *Phys. Rev. B* **16**, 1748 (1977).
- [29] T. Bray, J. Paoli, C. M. Sperberg-McQueen, E. Maler, and F. Yergeau, *Extensible Markup Language (XML) 1.0 5th edn.* (World Wide Web Consortium (W3C) Recommendation, 2008).
- [30] *VASP the guide*, <https://cms.mpi.univie.ac.at/vasp/vasp/vasp.html>.

- [31] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
- [32] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
- [33] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, Phys. Rev. B **57**, 1505 (1998).
- [34] A. I. Liechtenstein, V. I. Anisimov, and J. Zaanen, Phys. Rev. B **52**, R5467 (1995).
- [35] F. Aryasetiawan and O. Gunnarsson, Reports on Progress in Physics **61**, 237 (1998).
- [36] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, et al., Journal of Physics: Condensed Matter **21**, 395502 (2009).
- [37] T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, et al., The Journal of Chemical Physics **152**, 194103 (2020).
- [38] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. Payne, Z. Kristall. **220**, 567 (2005).
- [39] S. Plimpton, Journal of Computational Physics **117**, 1 (1995).
- [40] S. Smidstrup, T. Markussen, P. Vancraeyveld, J. Wellendorff, J. Schneider, T. Gunst, B. Verstichel, D. Stradi, P. A. Khomyakov, U. G. Vej-Hansen, et al., Journal of Physics: Condensed Matter **32**, 015901 (2019).
- [41] *MOPAC*, J.J.P. Stewart, Stewart Computational Chemistry, Colorado Springs, CO, USA; <http://OpenMOPAC.net/>.
- [42] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, et al., *Gaussian 16 Revision B.01* (2016), gaussian Inc. Wallingford CT.
- [43] F. Neese, WIREs Computational Molecular Science **2**, 73 (2012).

- [44] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, *Computer Physics Communications* **180**, 2175 (2009).
- [45] X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, et al., *Computational Materials Science* **25**, 478 (2002).
- [46] A. O. Lyakhov, A. R. Oganov, and M. Valle, *Crystal Structure Prediction Using Evolutionary Approach* (John Wiley & Sons, Ltd, 2010), chap. 7, pp. 147–180, ISBN 9783527632831.
- [47] A. R. Oganov, Y. Ma, A. O. Lyakhov, M. Valle, and C. Gatti, *Reviews in Mineralogy and Geochemistry* **71**, 271 (2010).
- [48] Y. Wang, J. Lv, L. Zhu, and Y. Ma, *Phys. Rev. B* **82**, 094116 (2010).
- [49] Q. Zhu, A. R. Oganov, C. W. Glass, and H. T. Stokes, *Acta Crystallographica Section B* **68**, 215 (2012).
- [50] Q. Zhu, A. R. Oganov, and A. O. Lyakhov, *CrystEngComm* **14**, 3596 (2012).
- [51] Q. Zhu, A. R. Oganov, A. O. Lyakhov, and X. Yu, *Phys. Rev. B* **92**, 024106 (2015).
- [52] G.-R. Qian, X. Dong, X.-F. Zhou, Y. Tian, A. R. Oganov, and H.-T. Wang, *Computer Physics Communications* **184**, 2111 (2013).
- [53] X. Dong, X.-F. Zhou, G.-R. Qian, Z. Zhao, Y. Tian, and H.-T. Wang, *Journal of Physics: Condensed Matter* **25**, 145402 (2013).
- [54] X.-F. Zhou, X. Dong, A. R. Oganov, Q. Zhu, Y. Tian, and H.-T. Wang, *Phys. Rev. Lett.* **112**, 085502 (2014).
- [55] Z. Wang, X.-F. Zhou, X. Zhang, Q. Zhu, H. Dong, M. Zhao, and A. R. Oganov, *Nano Letters* **15**, 6182 (2015).
- [56] Q. Zhu, L. Li, A. R. Oganov, and P. B. Allen, *Phys. Rev. B* **87**, 195317 (2013).

- [57] Q. Zhu, V. Sharma, A. R. Oganov, and R. Ramprasad, *The Journal of Chemical Physics* **141**, 154102 (2014).
- [58] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, *The Journal of Chemical Physics* **108**, 1964 (1998).
- [59] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, *Annual Review of Physical Chemistry* **53**, 291 (2002).
- [60] S. E. Boufelfel, A. R. Oganov, and S. Leoni, *Scientific Reports* **2**, 471 (2012).
- [61] A. R. Oganov and M. Valle, *The Journal of Chemical Physics* **130**, 104504 (2009).
- [62] Z. Allahyari and A. R. Oganov, *Multi-objective Optimization as a Tool for Material Design* (Springer International Publishing, 2020), pp. 2777–2790, ISBN 978-3-319-44680-6.
- [63] G. K. Madsen and D. J. Singh, *Computer Physics Communications* **175**, 67 (2006).
- [64] P. V. Bushlanov, V. A. Blatov, and A. R. Oganov, *Computer Physics Communications* **236**, 1 (2019).
- [65] D. J. Higham and N. J. Higham, *MATLAB guide*, vol. 150 (Siam, 2016).
- [66] M. Lutz and G. V. Rossum, *Programming Python: Object-Oriented Scripting* (O’Reilly & Associates, Inc., USA, 2001), 2nd ed., ISBN 0596000855.
- [67] M. Valle, *Zeitschrift für Kristallographie* **220**, 585 (2005).
- [68] J. Favre and M. Valle, in *The Visualization Handbook*, edited by C. Hansen and C. Johnson (Academic Press, 2004), pp. 655–672.
- [69] K. Momma and F. Izumi, *Journal of Applied Crystallography* **41**, 653 (2008).
- [70] B. Wang, R. Larhlimi, H. Valencia, F. Guégan, and G. Frapper, *The Journal of Physical Chemistry C* **124**, 8080 (2020).
- [71] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, *The Journal of Chemical Physics* **108**, 1964 (1998).

- [72] A. Oganov, C. Glass, A. Lyakhov, Q. Zhu, G.-R. Qian, H. Stokes, P. Bushlanov, Z. Allahyari, P. Graf, V. Stevanovic, et al., *Universal Structure Predictor: Evolutionary Xtallography, MANUAL version 10.4*, November 11, 2020, https://uspex-team.org/static/file/uspex_manual_10.4.pdf.
- [73] B. Wang, R. Larhlimi, H. Valencia, F. Gugan, and G. Frapper, *The Journal of Physical Chemistry C* **124**, 8080 (2020).
- [74] B. Wang, I. Braems, S. Sasaki, F. Gugan, L. Cario, S. Jobic, and G. Frapper, *The Journal of Physical Chemistry Letters* **11**, 8861 (2020).
- [75] *BIOGRAPH*, BGF and XTL formats were originally developed for BIOGRAF software, Molecular Simulations Inc.; circa 1990.
- [76] *BIOVIA*, This file formats are part of the software distributed by BIOVIA (formerly Accelrys); Dassault Systèmes BIOVIA, Material Studio, latest version (5.5), San Diego: Dassault Systèmes, (2017).
- [77] S. R. Hall, F. H. Allen, and I. D. Brown, *Acta Crystallographica Section A* **47**, 655 (1991).
- [78] *CSSR*, Crystal Structure Search Retrieval Database; SERC Chemical Databank System; CSE Applications Group; CCL Daresbury Laboratory, circa 1980.
- [79] M. M. J. Treacy, J. M. Newsam, and M. W. Deem, *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* **433**, 499 (1991).
- [80] B. Delley, *The Journal of Chemical Physics* **92**, 508 (1990).
- [81] W. Smith and T. Forester, *Journal of Molecular Graphics* **14**, 136 (1996).
- [82] M. Phillips, S. Levy, and T. Munzner, *Notices of the American Mathematical Society* pp. 985–988 (1993), computers and Mathematics Column.
- [83] H. Berendsen, D. van der Spoel, and R. van Drunen, *Computer Physics Communications* **91**, 43 (1995).

- [84] P. Csizmadia, *Marvinsketch and marvinview: molecule applets for the world wide web.* (1999), the Third International Electronic Conference on Synthetic Organic Chemistry.
- [85] M. Valiev, E. Bylaska, N. Govind, K. Kowalski, T. Straatsma, H. V. Dam, D. Wang, J. Nieplocha, E. Apra, T. Windus, et al., *Computer Physics Communications* **181**, 1477 (2010).
- [86] PDB, *Nature New Biology* **233**, 223 (1971).
- [87] W. Kraus and G. Nolze, *Journal of Applied Crystallography* **29**, 301 (1996).
- [88] C. J. Howard, B. A. Hunter, and D. A. J. Swinkels, *IUCR Powder Diffr.* **22**, 21 (1997).
- [89] *TRIPOS*, The Tripos mol2 file format specifications are available from archive at http://archive.ambermd.org/200704/att-0298/mol2_2pg_113.pdf.
- [90] *XYZ*, XYZ format is widely used, even though its specifications have never been subject to publication. Its origin is associated with the XMol visualization software, Minnesota Computer Centre, Inc; in 1991.

Figure 1: Graphical user interface of the GDIS software (a) displaying a crystal structure, here the $\text{Sn}(\text{N}_5)_4$ *I*-4 molecular crystal at 50 GPa⁷³ with the Sn site view inside of a polygon, (b) child plot document that is created automatically when opening a VASP output; currently displaying is the SCF steps of the $\text{Sn}(\text{N}_5)_4$ calculation (note the last SCF of each step has a higher energy due to external pressure correction).

Figure 2: Dialog for additional plotting function of GDIS on a loaded VASP output. The Dynamics, Electronic, and Frequency related plots are shown in a), b), and c), respectively.

Figure 3: Examples of GDIS plots produced when a VASP output file is loaded. The energy, force, volume and pressure values are reported for each ionic step in a), b), c), and d), respectively. This example is an optimization of cell parameters, cell volume, and atomic positions of pyrite FeS_2 ⁷⁴.

Figure 4: Example of the plotting of the electronic structure in GDIS. The density of state (DOS), band diagram, and a combination of both graph is presented in a), b), and c), respectively. The example electronic structure is obtain for a zincblende GaAs lattice at a meta-GGA level, using the modified Becke-Johnson (mBJ) functional. The band structure is shown along the Γ -X-K- Γ -L special k-points path.

Figure 5: Preset tab of the VASP GUI in the GDIS software.

Figure 6: Convergence tab of the VASP GUI in the GDIS software.

Figure 7: Elec-I tab of the VASP GUI in the GDIS software.

Figure 8: Elec-II tab of the VASP GUI in the GDIS software.

Figure 9: Ionic tab of the VASP GUI in the GDIS software.

Figure 10: Preset tab of the VASP GUI in the GDIS software.

Figure 11: Potcar tab of the VASP GUI in the GDIS software.

Figure 12: Exec tab of the VASP GUI in the GDIS software.

Figure 13: Example of a running VASP task displayed by the GDIS task manager on a geometry optimization of pyrite FeS_2 ⁷⁴.

Figure 14: SYSTEM tab of the USPEX GUI in the GDIS software.

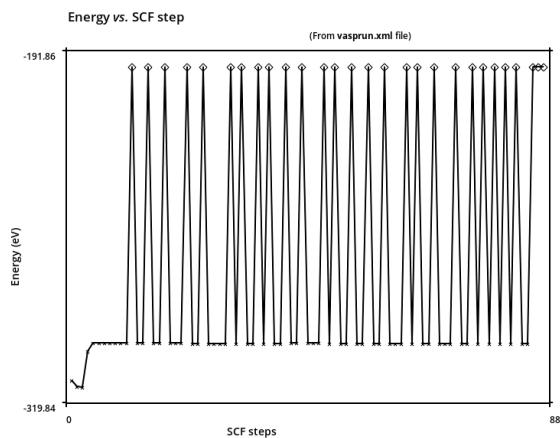
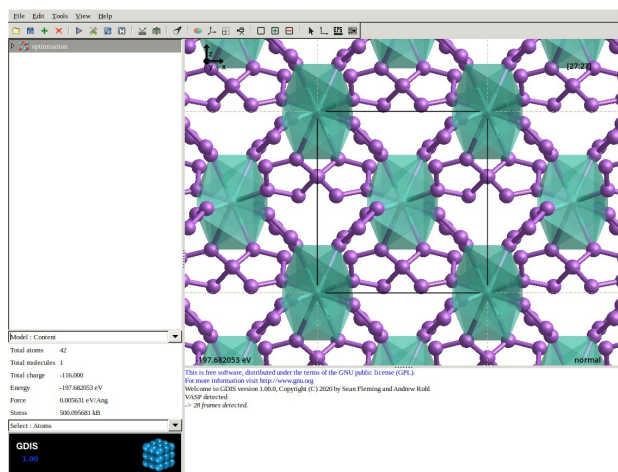
Figure 15: STRUCTURE tab of the USPEX GUI in the GDIS software.

Figure 16: CALCULATION tab of the USPEX GUI in the GDIS software.

Figure 17: ADVANCED tab of the USPEX GUI in the GDIS software.

Figure 18: SPECIFIC tab of the USPEX GUI in the GDIS software.

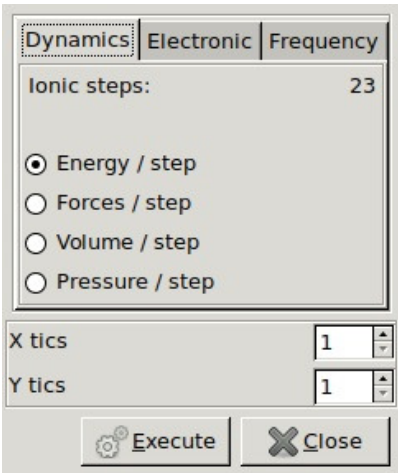
Figure 19: Automatically generated graphs during USPEX visualization. (a) All calculated structures energies as a function of the iteration (generation) of a Sn_xN_y variable composition USPEX search at 200 GPa⁷³. The structures chosen by USPEX as the best one of a generation are represented in red, and a graphic that only accounts for the latter is presented in (b). Energies of crystalline phases versus the nitrogen atomic ratio ($n_{Sn}/(n_{Sn}+n_N)$) and ($n_N/(n_{Sn}+n_N)$), from the same Sn_xN_y optimization, are shown in (c) and (d), respectively. The line that links the lowest stable phases of Sn_xN_y on the latter two graphs is the so-called thermodynamical convex-hull.



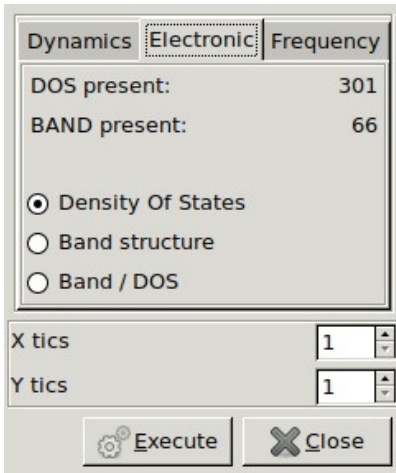
a)

b)

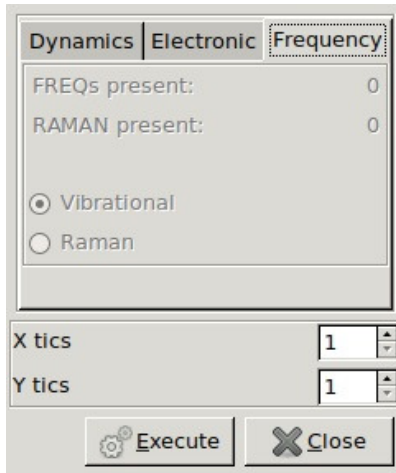
Figure 1
 Okadome Valencia H.,
 B. Wang, Frapper G., Rohl A.
 J. Comput. Chem.



a)

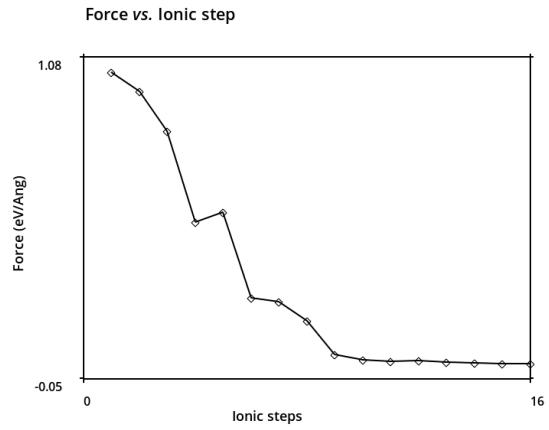
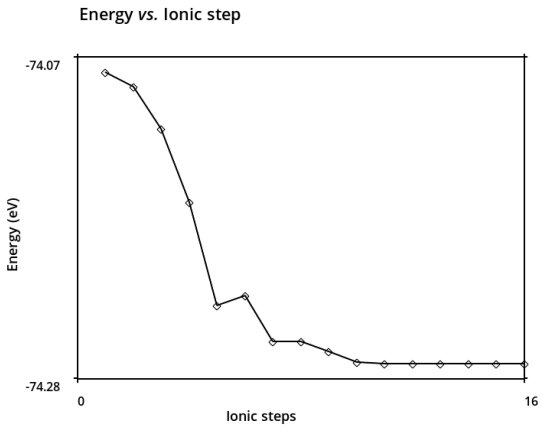


b)



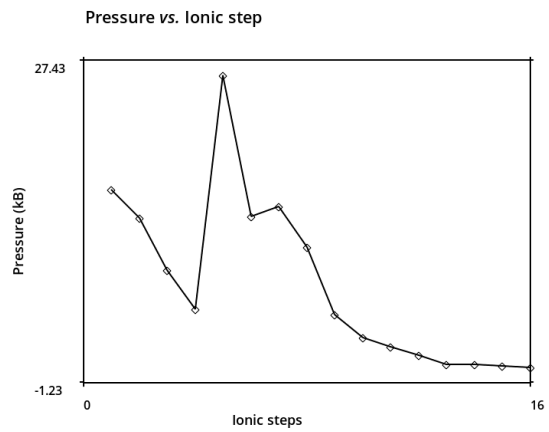
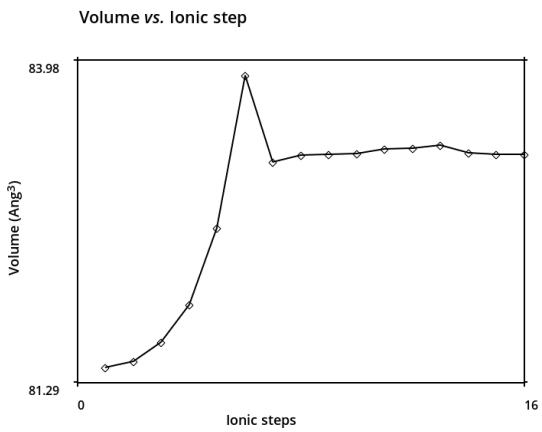
c)

Figure 2
Okadome Valencia H.,
B. Wang, Frapper G., Rohl A.
J. Comput. Chem.



a)

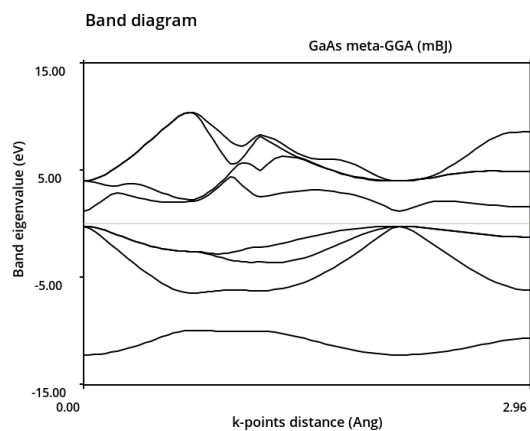
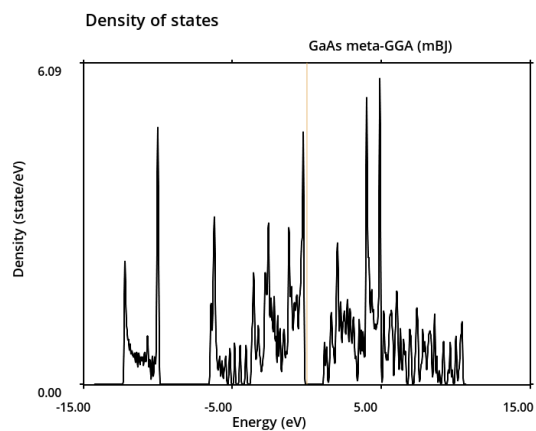
b)



c)

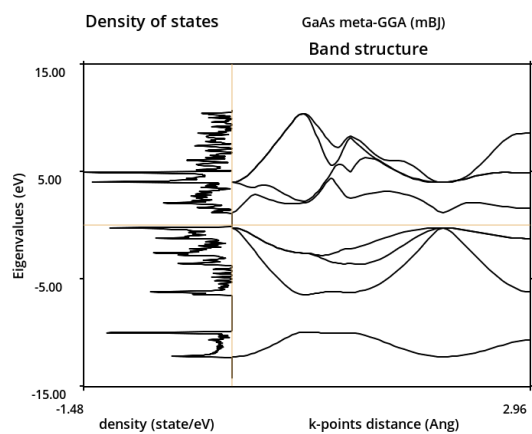
d)

Figure 3
 Okadome Valencia H.,
 B. Wang, Frapper G., Rohl A.
 J. Comput. Chem.



a)

b)



c)

Figure 4
 Okadome Valencia H.,
 B. Wang, Frapper G., Rohl A.
 J. Comput. Chem.

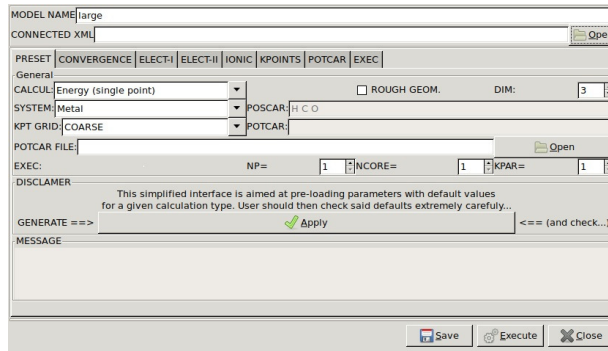


Figure 5
Okadome Valencia H.,
B. Wang, Frapper G., Rohl A.
J. Comput. Chem.

MODEL NAME|large
CONNECTED XML Open

PRESET CONVERGENCE ELECT-I ELECT-II IONIC KPOINTS POTCAR EXEC

General

PREC=Normal USE PREC ENCUT= 0.00 ENAUG= 0.00 EDIFF= 1.00E-04
ALGO=NORMAL LDIAG NSIM= 4 TIME= 0.4000 HWAVPR= 10
IALGO=38.KOSUGI AUTO NBANDS= 0 NELECT= 0.0000
 INIWAV ISTART= -1 ICHARG= -1

Mixing

IMIX=4-BROYDEN AUTO MIXER NELM= 60 NELMDL= -12 NELMIN= 2
MIXPRE=1.F=20 INVERSE KERKE AMIX= 0.4000 BMIX= 1.0000 AMIN= 0.1000
INIMIX=1.KERKER MAXMIX=-45 AMIX_MAG= 1.6000 BMIX_MAG= 1.0000 WC= 1000.0

Projector

LREAL=FALSE ROPT=
 ADDGRID LMAXMIX= 2 LMAXPAW= -100
 AUTO GRID NGX= 0 NGY= 0 NGZ= 0
NGXF= 0 NGYF= 0 NGZF= 0

Save Execute Close

Figure 6
Okadome Valencia H.,
B. Wang, Frapper G., Rohl A.
J. Comput. Chem.

MODEL NAME	large		
CONNECTED XML			
Open			
PRESET	CONVERGENCE	ELECT-I	ELECT-II
	IONIC	XPOINTS	POTCAR
	EXEC		
Smearing			
ISMEAR=	1	SIGMA=	0.2000
		<input checked="" type="checkbox"/> KGAMMA	KSPACING=
			0.5000
FERMWE=	FERMDO=		
Spin			
<input type="checkbox"/> SPIN POLARIZED	<input type="checkbox"/> NON COLLINEAR	<input type="checkbox"/> LSORBIT	<input checked="" type="checkbox"/> GGA_COMPAT
			NUPDOWN=
			-1.0
MAGMOM=	SAXIS=		
Exchange Correlation			
GGA=	PE:PBE	<input type="checkbox"/> VOSKOWN	
METAGGA=	MBJ	<input type="checkbox"/> LMETAGGA	<input type="checkbox"/> LASPH
			<input type="checkbox"/> LMIXTAU
			LMAXTAU=
			0
CMBJ=		CMBJA=	-0.0120
			CMBJB=
			1.0230
LDAUTYPE=	2:LSDA+U Dudarev	<input type="checkbox"/> LDAU	LDAUL=
LDAUPRINT=	0:Silent	<input type="checkbox"/> LDAUU=	LDAUj=
Dipole			
IDIPOL=	0:no calcul	<input type="checkbox"/> LDIPOL	<input type="checkbox"/> LMONO
		DIPOL=	
		EPSILON=	1.0000
		EFIELD=	0.0000
Save Execute Close			

Figure 7
 Okadome Valencia H.,
 B. Wang, Frapper G., Rohl A.
 J. Comput. Chem.

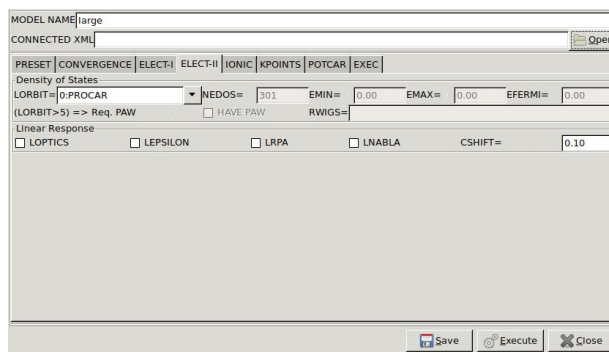


Figure 8
Okadome Valencia H.,
B. Wang, Frapper G., Rohl A.
J. Comput. Chem.

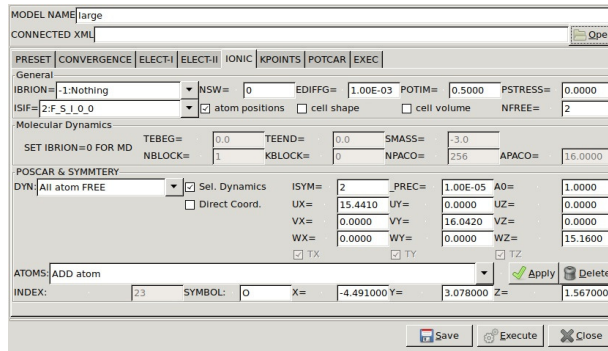


Figure 9
Okadome Valencia H.,
B. Wang, Frapper G., Rohl A.
J. Comput. Chem.

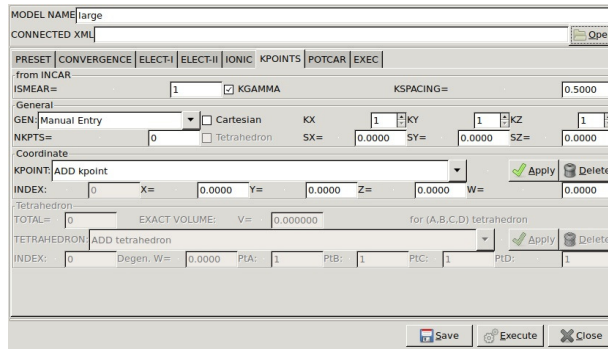


Figure 10
 Okadome Valencia H.,
 B. Wang, Frapper G., Rohl A.
 J. Comput. Chem.

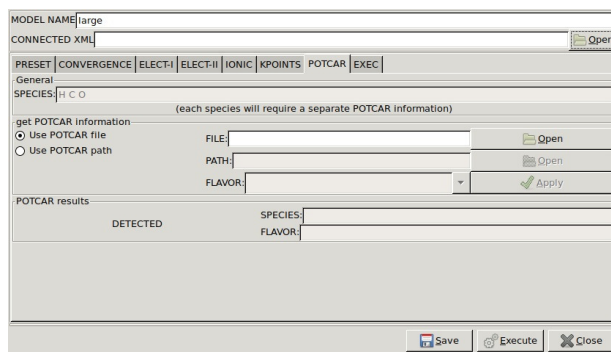


Figure 11
Okadome Valencia H.,
B. Wang, Frapper G., Rohl A.
J. Comput. Chem.

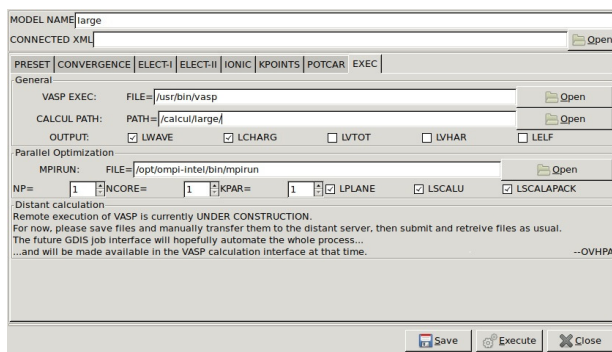


Figure 12
Okadome Valencia H.,
B. Wang, Frapper G., Rohl A.
J. Comput. Chem.

PID	Job	Status	% CPU	% Mem	Time
1149078	VASP	Running	100.0	9.3	00:00:52


```

POSCAR found type information on POSCAR C H
POSCAR found : 2 types and 8 ions
scaLAPACK will be used
LDA part: xc-table for Pade appr. of Perdew
POSCAR, INCAR and KPOINTS ok, starting setup
FFT planning ...
WAVECAR not read
entering main loop
  N   E           dE      d eps   ncg   rms   rms(c)
DAV: 1  0.128215231326E+02  0.12822E+02 -0.32616E+03  40  0.468E+02
DAV: 2 -0.400871058721E+02 -0.52909E+02 -0.52755E+02  28  0.131E+02
DAV: 3 -0.457940067327E+02 -0.57069E+01 -0.56987E+01  32  0.480E+01
DAV: 4 -0.460040394269E+02 -0.21003E+00 -0.21001E+00  36  0.101E+01
DAV: 5 -0.460048143434E+02 -0.77492E-03 -0.77491E-03  36  0.567E-01  0.110E+01
RMM: 6 -0.415897338060E+02  0.44151E+01 -0.83670E+00  24  0.151E+01  0.572E+00

```

Figure 13
 Okadome Valencia H.,
 B. Wang, Frapper G., Rohl A.
 J. Comput. Chem.

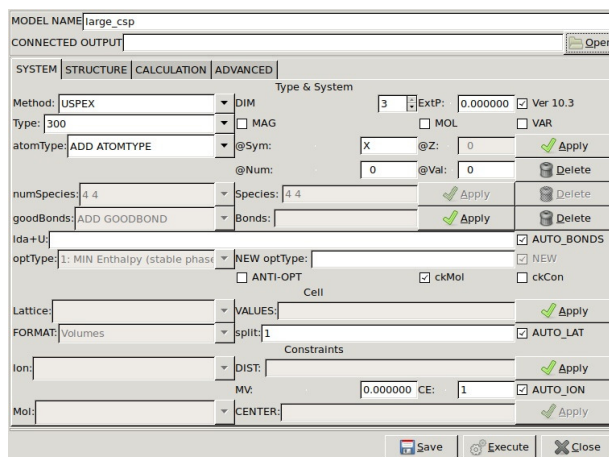


Figure 14
 Okadome Valencia H.,
 B. Wang, Frapper G., Rohl A.
 J. Comput. Chem.

MODEL NAME: large_csp

CONNECTED OUTPUT Open

SYSTEM STRUCTURE CALCULATION ADVANCED

Population & selection

SIZE:	60	INIT:	60	NGEN:	100	STOP:	8
N.M.:	0.1000	FM-LS:	0.2250	AFM-L:	0.2250	FM-LH:	0.0000
<input type="checkbox"/> MAG		FM-HS:	0.2250	AFM-H:	0.2250	AF-LH:	0.0000
Best:	0.700000	BestHM:	9	<input type="checkbox"/> reopt		fitLimit:	0.000000

Structure & Variation

symmetries: 2-230

Heredity:	0.500000	Random:	0.200000	TOPRand:	0.200000	Permutation:	0.100000
AtmMut:	0.100000	RotMut:	0.000000	LatMut:	0.100000	SpinMut:	0.100000
NSwaps:	2	Swaps:					
mutationDegree:	0.000000	mutationRate:	0.500000	D_LATMUT:	1.000000	<input type="checkbox"/> AutoFrac	

Fingerprint, antiseed, & spacegroup

Fingerprint	RMax:	10.00000	delta:	0.080000	sigma:	0.030000
Antiseed	Activation:	5000	Max:	0.000000	sigma:	0.001000
Space group		<input checked="" type="checkbox"/> Active			TOL:	0.100000

Variable-composition

1st_Gen:	11	min@:	0	max@:	0
rTrans:	0.100000	rTrans:	0.200000	Trans:	

Save Execute Close

Figure 15
 Okadome Valencia H.,
 B. Wang, Frapper G., Rohl A.
 J. Comput. Chem.

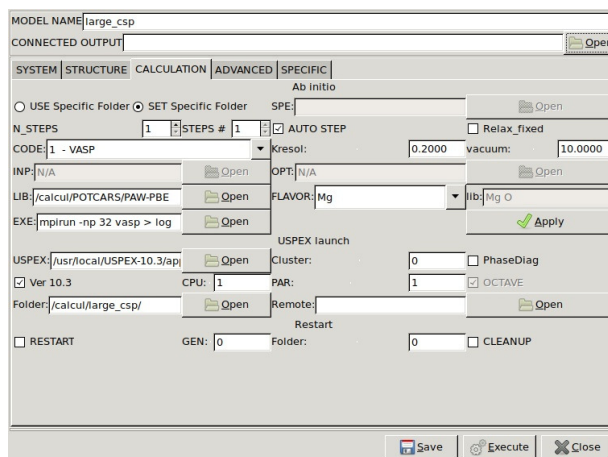


Figure 16
Okadome Valencia H.,
B. Wang, Frapper G., Rohl A.
J. Comput. Chem.

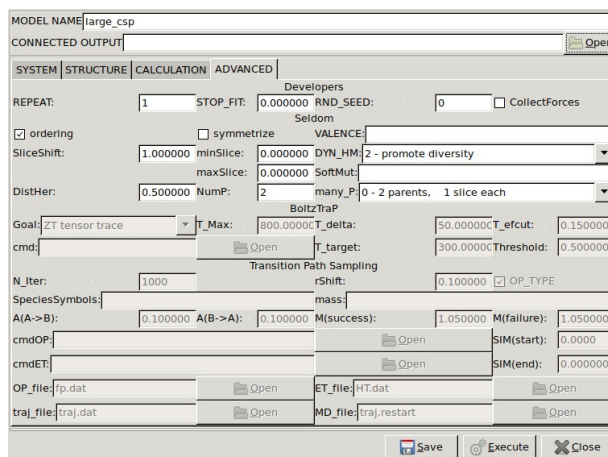


Figure 17
 Okadome Valencia H.,
 B. Wang, Frapper G., Rohl A.
 J. Comput. Chem.

MODEL NAME: large_csp

CONNECTED OUTPUT Open

SYSTEM | STRUCTURE | CALCULATION | ADVANCED | SPECIFIC

Metadynamics

Relax: 2 - Relax all different struct MaxV: 0.000000 GaussW: 0.000000 GaussH: 0.000000

MODEL: From VASP5 POSCAR file Open

Particles Swarm optimization

SoftMut: 1.000000 BestStruct: 1.000000 BestEver: 1.000000

Variable-cell nudged elastic band

Method: 1 - VC-NEB method VC-NEB: 110 Var_Image Var_Spring

Img: 2 - Initial and final + Intermed N_Img: 9 N_Step: 200 Freeze_Img

Opt: 1 - Steepest Descent dt: 0.050000 Conv: 0.003000 PathLength: 0.000000

Relax: 3 - full, cell and positions re Kmin: 5.000000 Kmax: 5.000000 Kcte: 5.000000

CJ/DI: 0 - No CJ/DI method will be u Start CJ/DI: 100 Pickup: PrintStep: 1

Format: 2 - VASP v5 (POSCAR) form MODEL: From VASP5 POSCAR file Open

Molecules

MODEL: Already provided N_MOLS: 1 Open

GDIS_MOL: 0: model_0 MOL_#: 1 GULP_FORM Apply

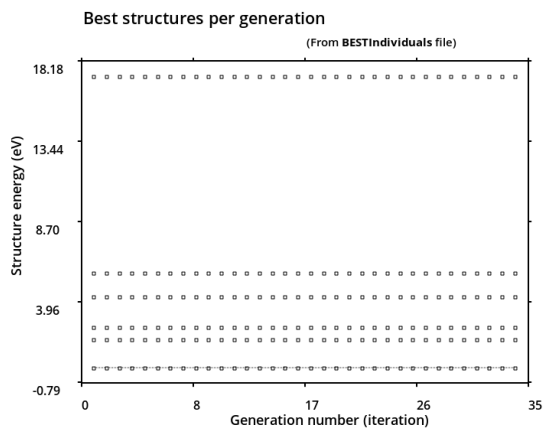
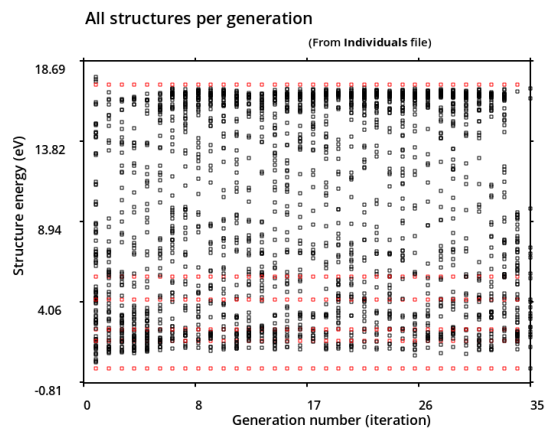
Surfaces

MODEL: From VASP5 POSCAR file Open S_thick: 2.000000

Stoichio: N_Surf: 1 B_thick: 3.000000

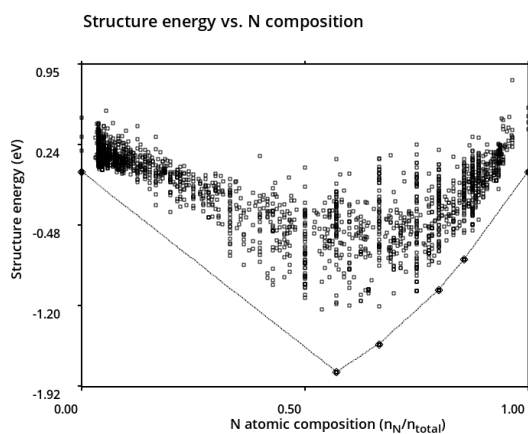
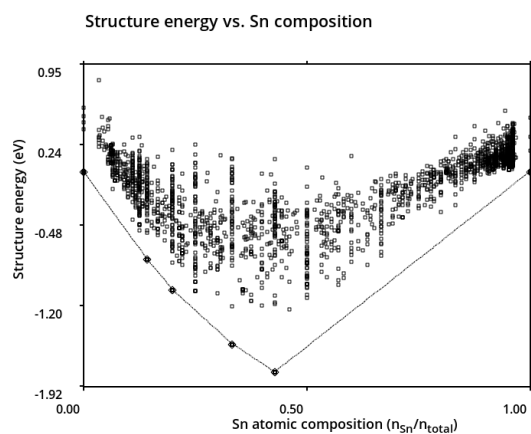
Save Execute Close

Figure 18
 Okadome Valencia H.,
 B. Wang, Frapper G., Rohl A.
 J. Comput. Chem.



a)

b)



c)

d)

Figure 19
Okadome Valencia H.,
B. Wang, Frapper G., Rohl A.
J. Comput. Chem.

Software	Ref.	Read file format	Write file format
ABINIT	45	.about, .abot	N/A
BIOGRAPH	75	.bgf, .xtl	.bgf, .xtl
BIOVIA	76	.arc, .car, .cor	.car
CASTEP	38	.castep	N/A
CIF	77	.cif	.cif
CSSR	78	.cssr	.cssr
DIFFaX	79	.dfx	.dfx
DMOL	80	.dmol	.dmol
DLPOLY	81	.dlpoly	.dlpoly
GAMESS	3	.inp, .gmot, .gmout	.inp
GAUSSIAN	42	.log	.com
GDIS	1	.gmf, .pcf, .xml	.gmf, .pcf, .xml
GEOMVIEW	82	.off	N/A
GROMACS	83	.gro	.gro
GULP	2	.gin, .res, .got, .gout	.gin, .res
MARVIN	84	.mar, .mvn(-r), .mot, .mv(n)out	.mar, .mvn(-r)
MONTY	4	.cgf	.cgf
MOPAC	41	.out	N/A
NWCHEM	85	.nwin, .nw	.nwout, .nwot, .nwo
PDB	86	.pdb	.pdb
PowderCell	87	.cel	N/A
QuantumEspresso	36	.qein, .qeout	.qein
Rietica	88	.inp	-
SIESTA	5	.fdf, .out	.fdf
TRIPOS	89	.mol2	N/A
XYZ	90	.xyz, .ani	.xyz

Table 1: Table of the currently supported formats for reading or writing in GDIS. This list excludes the newly introduced VASP and USPEX formats, which are detailed in the text.