User’s Manual

WRAP 'N' SHAKE

Systematic exploration of multiple drug binding sites

Version 1.0

www.wnsdock.xyz

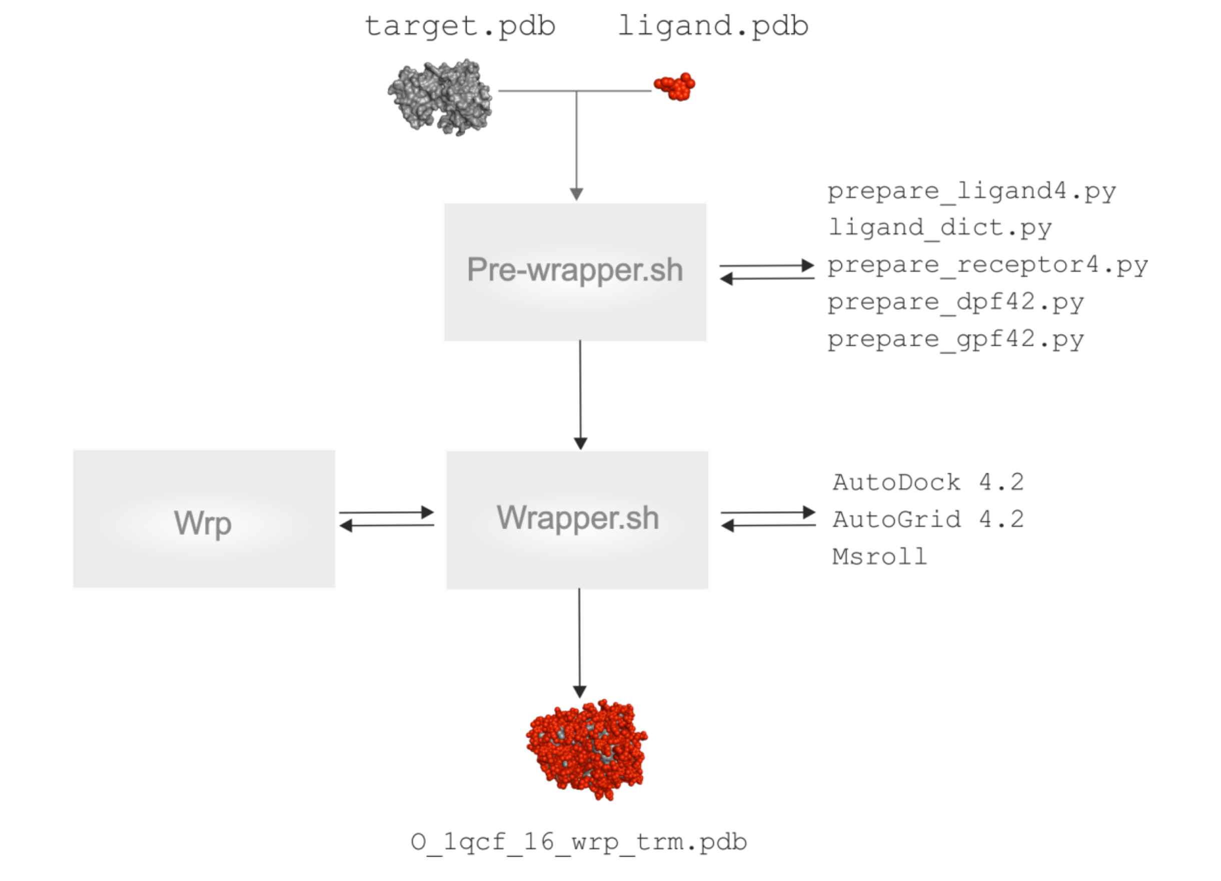
2017

Table of Contents

1 Introduction

The Wrap 'n' Shake (WnS) method was designed for systematic exploration of multiple drug binding sites. WnS starts with Wrapper (**Fig. 1**), which systematically covers the entire surface of the target with a monolayer of ligand copies. Wrapper performs several fast blind docking (BD) cycles using AutoDock 4.2.3. In the monolayer, ligand-ligand interactions are minimized, as the ligand copies mostly interact with the target. Wrapper includes two shell scripts (Pre-wrapper.sh and Wrapper.sh) and a C program (Wrp).

Fig. 1



Components of Wrapper (grey boxes) and their connection with external shell scripts and programs

2 Pre-wrapper.sh

Overview

Pre-wrapper.sh is a bash shell script for preparation of AutoDock 4.2 and AutoGrid 4.2(Morris, et al., 2009) inputs such as pdbqt files for the ligand and target molecules, docking and grid parameter files.

Inputs

Energy-minimized structures of the target and ligand molecules are the main inputs. The commonly used Protein Databank (PDB) file format is accepted. Hydrogen atoms are not added to the input PDB files. It is assumed that input structures are complete to have an integer total charge on both molecules (target and ligand).

Algorithm

The script generates *.pdbqt files and the corresponding parameter files (*.dpf and *.gpf) for each target and ligand pair. External python scripts used by Pre-wrapper.sh are listed in Table 1. The scripts are distributed as part of the open source software AutoDockTools and downloadable at

```
http://autodock.scripps.edu
```

Pre-wrapper.sh adds the actual path to the modified AD4_parameters.dat (**Fig. 2**) file as an additional first line to *.dpf and *gpf. This tells to the programs of AutoDock 4.2.3 to use the modified AD4_parameters.dat file instead of the default one normally stored in source code folder

```
src/autodock
```

Pre-wrapper.sh also adds new entries of excluded atom types LL and YY (commonly marked as X in our original publication(Mónika Bálint, 2017)) to the *.dpf and *.gpf files. This step is performed only once, as the same parameter files can be used in all wrapping cycles later. This step is necessary for generation of the corresponding new map files (*LL.map and *YY.map).

The use of Pre-wrapper.sh is not mandatory for Wrapper. Generation of the input *.pdbqt and parameter files can be performed with the AutoDockTools and the above modifications can be inserted manually. However, to avoid human mistakes, the use of Pre-wrapper.sh is recommended especially if multiple target files, or a library of ligand structures are handled.

Usage

Pre-wrapper.sh can be launched by the following command, where target.pdb, and ligand.pdb files and the path of external scripts are obligatory inputs.

```
$pre-wrapper.sh -t target.pdb -l ligand.pdb -p /home/user/bin
```

Table 1 Pre-wrapper.sh

Input files and arguments (mandatory)		
`-t`	`target.pdb`	Target structure minimized and prepared for docking
`-l`	`ligand.pdb`	Ligand structure minimized and prepared for docking
`-p`	`/home/user/bin`	Path for external scripts (below), and for the modified AD4_parameters.dat file
Output files		
`1qcf_target.pdbqt`		Target file in pdbqt format used as docking input
`1qcf_ligand.pdbqt`		Ligand file in pdbqt format used as docking input
`1qcf_target.dpf`		Docking parameter file
`1qcf_target.gpf`		Grid parameter file
External scripts		
`prepare_ligand4.py`		External Python scripts are available as part of AutoDockTools.
`ligand_dict.py`		
`prepare_receptor4.py`		
`prepare_dpf42.py`		
`prepare_gpf4.py`		

Fig. 2

```
atom_par H       2.00  0.020   0.0000   0.00051  0.0  0.0  0  -1  -1  3       #Non H-bonding Hydrogen
atom_par HD      2.00  0.020   0.0000   0.00051  0.0  0.0  2  -1  -1  3       #Donor 1 H-bond Hydrogen
atom_par HS      2.00  0.020   0.0000   0.00051  0.0  0.0  1  -1  -1  3       #Donor S Spherical Hydrogen
atom_par C       4.00  0.150  33.5103  -0.00143  0.0  0.0  0  -1  -1  0       #NonH-bonding Aliphatic Carbon
atom_par A       4.00  0.150  33.5103  -0.00052  0.0  0.0  0  -1  -1  0       #Non H-bonding Aromatic Carbon
atom_par N       3.50  0.160  22.4493  -0.00162  0.0  0.0  0  -1  -1  1       #Non H-bonding Nitrogen
atom_par NA      3.50  0.160  22.4493  -0.00162  1.9  5.0  4  -1  -1  1       #Acceptor 1 H-bond Nitrogen
atom_par NS      3.50  0.160  22.4493  -0.00162  1.9  5.0  3  -1  -1  1       #Acceptor S Spherical Nitrogen
atom_par OA      3.20  0.200  17.1573  -0.00251  1.9  5.0  5  -1  -1  2       #Acceptor 2 H-bonds Oxygen
atom_par OS      3.20  0.200  17.1573  -0.00251  1.9  5.0  3  -1  -1  2       #Acceptor S Spherical Oxygen
atom_par F       3.09  0.080  15.4480  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Fluorine
atom_par Mg      1.30  0.875   1.5600  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Magnesium
atom_par MG      1.30  0.875   1.5600  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Magnesium
atom_par P       4.20  0.200  38.7924  -0.00110  0.0  0.0  0  -1  -1  5       #Non H-bonding Phosphorus
atom_par SA      4.00  0.200  33.5103  -0.00214  2.5  1.0  5  -1  -1  6       #Acceptor 2 H-bonds Sulphur
atom_par S       4.00  0.200  33.5103  -0.00214  0.0  0.0  0  -1  -1  6       #Non H-bonding Sulphur
atom_par Cl      4.09  0.276  35.8235  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Chlorine
atom_par CL      4.09  0.276  35.8235  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Chlorine
atom_par Ca      1.98  0.550   2.7700  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Calcium
atom_par CA      1.98  0.550   2.7700  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Calcium
atom_par Mn      1.30  0.875   2.1400  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Manganese
atom_par MN      1.30  0.875   2.1400  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Manganese
atom_par Fe      1.30  0.010   1.8400  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Iron
atom_par FE      1.30  0.010   1.8400  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Iron
atom_par Zn      1.48  0.550   1.7000  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Zinc
atom_par ZN      1.48  0.550   1.7000  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Zinc
atom_par Br      4.33  0.389  42.5661  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Bromine
atom_par BR      4.33  0.389  42.5661  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Bromine
atom_par I       4.72  0.550  55.0585  -0.00110  0.0  0.0  0  -1  -1  4       #Non H-bonding Iodine
atom_par YY      3.60  1E-04  00.0000   0.00000  0.0  0.0  0   0   0  0       #Excluded target atom
atom_par LL      3.60  1E-04  00.0000   0.00000  0.0  0.0  0   0   0  0       #Excluded ligand atom
```

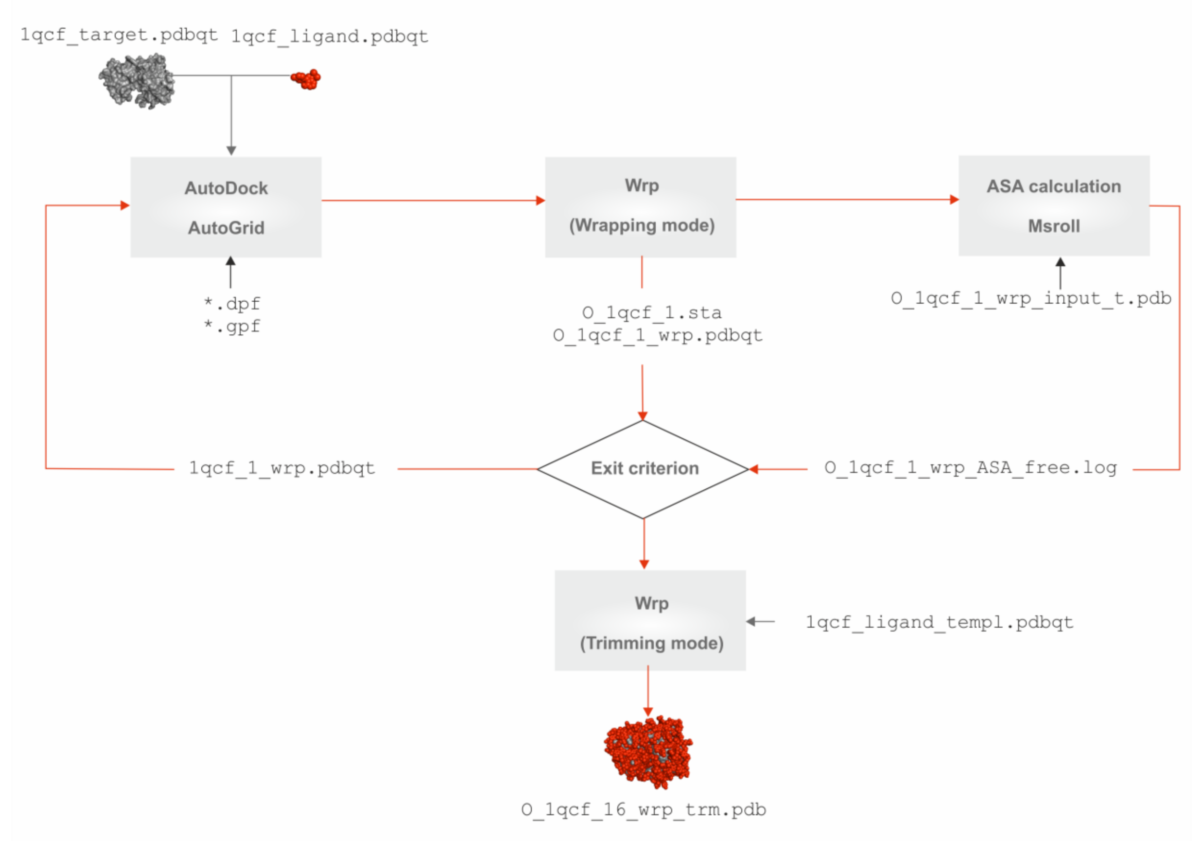
End part of the modified AD4_parameters.dat. The modifications included in the file are highlighted in yellow. Note that atom types LL and YY have the same parameters and mentioned as X in the original publication of WnS(Mónika Bálint, 2017).

3 Wrapper.sh

Overview

Wrapper.sh (**Fig. 3**) is bash shell script, the central engine of Wrapper. It performs multiple cycles of blind docking, and automatically creates a monolayer of ligand copies systematically covering the entire surface of the target molecule.

Fig. 3



Wrapper.sh Input and output files names correspond to the example files, presented on web site of WnS (www.wnsdock.xyz). In the presented files, 1qcf stands for the System name, and 16 for the total number of wrapper cycles performed

Inputs

Wrapper.sh requires the files generated with Pre-wrapper.sh. An additional input, the ligand template file (*ligand_templ.pdbqt) is also necessary in the final stage of Wrapping to re-name all atoms according to their names in the molecular dynamics topology. The user also has to provide the maximum number of docking cycles.

Algorithm

Wrapper.sh calls external programs AutoDock 4.2, AutoGrid 4.2(Morris, et al., 2009), Msroll (Connolly, 1993), and Wrp, an open source C program developed specially for WnS (see Section 4 for a detailed description of Wrp). The main steps of the algorithm are described here.

1) A function named cycle is created, in order to perform the following steps in multiple docking cycles. A docking cycle starts with creating the map files for each atom type specified in the grid parameter file (*.gpf), using AutoGrid 4.2. In the next step, docking is performed by AutoDock 4.2, using the docking parameters specified in the docking parameter file (*.dpf). Besides the docking and grid parameter files, reading the AD4_parameters.dat file form the *.dpf and *.gpf is also necessary in this step as the parameters of the new atom types LL and YY are defined in AD4_atom parameters.dat as described in Section 2. Function cycle has three arguments: the number of the current cycle, the name of the current target, and the number of the next cycle.

2) Calling program wrp. The docked conformations from the *.dlg file produced by AutoDock 4.2 are clustered and ranked, based on their interaction energy (E_{inter}) values with the target. E_{inter} is the AutoDock 4.2 free energy of binding, without the intramolecular energy terms. The *wrp.pdbqt file produced by Wrp is renamed from O_1qcf_1_wrp.pdbqt into 1QCF_1_wrp.pdbqt, and this will be used as input target file for the next wrapper cycle, if the two exit criterions are not met. Wrp also provides a O_1qcf_1_wrp_input_t.pdb file for the next step.

3) Program Msroll is called for calculation of accessible surface area (ASA) using the O_1qcf_1_wrp_input_t.pdb file. The output of the Msroll is shortened to a log file (O_1qcf_1_wrp_ASA_free.log), containing the free target ASA not covered by ligands.

4) The output file resulted from wrp (O_1qcf_1_wrp.pdbqt) contains the new atom types assigned to the ligand (LL) and target (YY). This file is used as input for the next docking cycle if the two exit criterions are not met. The exit criterions (ASA ≤ 1 % or $E_{inter} \geq 0$ kcal/mol) are checked in each wrapper cycle. The latter is checked in O_1qcf_1.sta file, where the E_{inter} of each cluster representative is recorded. The ASA is read from the O_1qcf_1_wrp_ASA_free.log file. If one of the exit criterion is met, then function cycle function will terminate, and a trimming step is started. In the Trimming mode of wrp, ligands that have the distance from the target greater than 3.5 Å, are eliminated.

5) In the folder where the last cycle was performed, and one of the exit criterion was met, track keeping file *surface_exit.log or *energy_exit.log file is generated.

Outcomes

The target structure wrapped in a monolayer of ligand copies is final outcome of Wrapper.sh. The structure is stored in a *.pdbqt file as generated by program Wrp (see Chapter 5) and after the trimming step it can be readily used in the Shaker method.

Usage

Wrapper.sh can be launched by the following command, where each argument described in Table 2 is an obligatory input.

```
$wrapper.sh -n Cycle Number -t Target Name(pdbqt) -l Ligand Name(pdbqt) -b
Binary Path -p Parameter Path -r *ligand templ.pdbqt
```

Table 2 Wrapper.sh		
Input files and arguments		
`-n`	`Maximal count of cycles`	This number is provided by the user. A choice of n=30 will suffice for most of the targets.
`-t`	`1qcf_target.pdbqt`	Target pdbqt file
`-l`	`1qcf_ligand.pdbqt`	Ligand pdbqt file
`-p`	`home/user/bin`	Path for binary files
`-b`	`home/user/bin`	Path for parameter files
`-r`	`*ligand_templ.pdbqt`	Ligand template file, for re-numbering and re-naming atoms and residues after docking. The atoms follow the order of the original ligand.pdbqt file.
Output files ()		
`1qcf_1.dlg`		Docking log file
`O_1qcf_1_ASA_free.log`		Shortened log file, containing the free accessible protein surface area
`*surface_exit.log /*energy_exit.log`		In the exiting cycles directory, one of the two (*surface_exit.log /*energy_exit.log) files will be generated by wrapped.sh The exiting cycle number can be found in the standard output, written out by wrapper.sh.
External programs		
`AuroGrid 4.2`		Calculation of grid maps. (Maximum number of grid points should be set large enough for blind docking on the entire surface of the target. In many cases, 300 grid points are enough for blind docking on common targets.)
`AutoDock 4.2`		The docking engine
`Wrp`		See Section 4
`Msroll`		The surface area of the input target is calculated for each cycle, using the method of Connolly et al 1993.

4 Wrp

Overview

Wrp is an open source C program performing clustering, ranking, and assignation of excluded atoms in Wrapper. Wrp also performs a trimming step. Repeated use of Wrp in Wrapper provides the target structure systematically wrapped in a monolayer of ligand copies.

Inputs

Wrp is called by Wrapper.sh. Depending on the running mode, different input files are required. In wrapping mode, the docking log (*.dlg) is the main input as resulted from AutoDock 4.2 in a wrapping cycle. In the trimming mode, the wrapped target (*wrp.pdbqt), serves as an input.

Algorithm

Wrapping mode. Wrp has various functions in wrapping mode as listed here.

1) Wrp performs clustering and ranking of the 100 docked ligand conformations listed in the *.dlg file by AutoDock 4.2. The clustering step is integrated in the wrapping mode of wrp program. After clustering, the non-overlapping cluster representatives are retained for the next blind docking round. The cluster representative is selected as the conformation with lowest binding energy from the cluster. Cluster representatives are separated from each other using the ranking tolerance –drnk.

2) Wrp is also used to assign atom type LL for all atoms in the generated cluster representatives, and YY for the protein atoms situated at maximum distance from the cluster representative as specified by assignation tolerance –dsgn.

3) Wrp calculates the target–residue heavy atom interface that is situated at distance tolerance set with –difc (Interface tolerance) switch. In all cycles, atom type LL is assigned to all atoms of the docked ligand poses and YY to target atoms situated within a 3.5 Å distance from the docked ligand atoms.

Trimming mode. Ligands with a maximal atomic distance from the target larger than the trimming tolerance –dmax are eliminated.

Outcomes

In wrapping mode the outcome is the target structure (partially) covered by ligand copies in a *wrp.pdbqt file. Additionally to the *wrp.pdbqt file, a statistics file (*.sta) is also generated, that contains the E_{inter} of the cluster representatives. Two types of outputs are generated after trimming. The *trm.pdb file contains the target structure wrapped in a monolayer of N ligand copies and can be readily used in Shaker method.

Usage

By default, Wrp is used with silent verbosity in the wrapping mode by wrapper.sh. An example command line is presented below.

```
$ wrp -f 1qcf_1_wrp.dlg -p *ligand_templ.pdbqt -t 1qcf_1_wrp.pdbqt -c cycle
nr. -difc 3.500 -drnk 2.000 -dsgn 3.500 -m wrapping -v silent
```

Trimming mode of wrp is used only once after the wrapping process, and an example command line is presented below.

```
$ wrp -f 1qcf_1_wrp.trm.pdbqt -p 1qcf_templ.pdbqt -dmax 3.500 -m trimming -
v silent
```

Table 3 Wrp		
Input files (silent or diagnostic)		
`-f`	`1qcf_1_wrp.dlg`	Docking log file, resulted from AutoDock 4.2, containing 100 docked ligand copies
`-f`	`1qcf_1_wrp.pdbqt`	Target with ligand copies in trimming mode
`-p`	`1qcf_ligand_templ.pdbqt`	Template ligand, see Table 2
`-r`	`*ref.pdb (optional)`	Reference pdb file, used for RMSD calculation (not used in the default wrapper.sh)
`-t`	`1qcf_1_wrp.pdbqt`	Target (with ligand copies) in wrapping mode
Parameters and defaults		
`-c`	`1`	Serial number of wrapping cycle
`-difc`	`3.500`	Interface tolerance (Å)
`-dmax`	`3.500`	Trimming tolerance (Å)
`-drnk`	`2.000`	Ranking tolerance (Å)
`-dsgn`	`3.500`	Assignation tolerance (Å)
`-m`	`Wrapping`	Program mode, wrapping/trimming
`-v`	`silent`	Verbosity, silent/diagnostic
Output files		
Silent		
`O_1qcf_1.log`		Log file, stating the used parameters
`O_1qcf_1.sta`		Statistics file
`O_1qcf_1_wrp.pdbqt`		Target-ligand complex file, that will be used in the next wrapper cycle
`O_1qcf_1_input_t.pdb`		Target or target – ligand complex file, that was used as input for the given wrapper cycle
Diagnostic		
`O_1qcf_1_wrp.pdb`		Target-ligand complex file, that will be used in the next wrapper cycle, in pdb format

`O_1qcf_1_rank_1.pdb`
…
`O_1qcf_1_rank_CC.pdb`
		Cluster representative, complexed with target residues that were calculated within the Interface tolerance. (CC stands for cluster count).

5 References

Connolly, M.L. (1993) The molecular surface package, *Journal of molecular graphics*, **11**, 139-141.

Bálint M, *et al.* (2017) Systematic exploration of multiple drug binding sites, *Submitted for publication*.

Morris, G.M.*, et al.* (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, *Journal of computational chemistry*, **30**, 2785-2791.

6 Disclaimer

WRAP 'N' SHAKE (REFERRED TO AS WNS IN THE FORTHCOMING TEXT) IS A FREE SOFTWARE: YOU CAN REDISTRIBUTE IT AND/OR MODIFY IT UNDER THE TERMS OF THE GNU GENERAL PUBLIC LICENSE AS PUBLISHED BY THE FREE SOFTWARE FOUNDATION, EITHER VERSION 3 OF THE LICENSE, OR (AT YOUR OPTION) ANY LATER VERSION. WNS IS DISTRIBUTED IN THE HOPE THAT IT WILL BE USEFUL, BUT WITHOUT ANY WARRANTY; WITHOUT EVEN THE IMPLIED WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. SEE THE GNU GENERAL PUBLIC LICENSE FOR MORE DETAILS. YOU SHOULD HAVE RECEIVED A COPY OF THE GNU GENERAL PUBLIC LICENSE ALONG WITH WNS. IF NOT, SEE AT HTTP://WWW.GNU.ORG/LICENSES THERE IS NO WARRANTY FOR WNS, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE WNS “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION. LIMITATION OF LIABILITY. IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MODIFIES AND/OR CONVEYS THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE WNS (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF WNS TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. DISCLAIMER ON THE DOCUMENTATION AND WEB SITE HTTP://WWW:WNSDOCK.XYZ OF WNS. IN NO WAY CAN ANY RIGHTS BE DERIVED FROM, OR CLAIMS MADE, WITH REGARD TO THE CONTENT OF THIS WEBSITE AND THE DOCUMENTATION THEREIN. ALTHOUGH THE GREATEST POSSIBLE CARE HAS BEEN TAKEN WITH THE COMPILATION OF THE CONTENT OF THIS WEBSITE AND DOCUMENTATION, IT IS POSSIBLE THAT CERTAIN INFORMATION MAY (AFTER A WHILE) BE OUT-OF-DATE OR (NO LONGER) BE CORRECT. WE ARE NOT RESPONSIBLE FOR EVENTUAL DAMAGES ARISING FROM THE USE OF INFORMATION FROM THIS SITE. WE HEREBY REJECT ALL RESPONSIBILITY FOR DAMAGES AS A RESULT OF THE USE OF THIS INFORMATION OR INFORMATION TO WHICH LINKS REFER ON THIS SITE (THESE SITES). THE INFORMATION ON THIS SITE MAY BE CHANGED WITHOUT PRIOR WARNING. WE DO NOT GIVE GUARANTEES WITH REGARD TO THE NATURE AND THE CONTENT OF THIS SITE. ALL RESPONSIBILITY FOR POSSIBLE DAMAGES DUE TO ACCESS TO AND USE OF THE SITE IS EXPLICITLY REJECTED BY US. WE DO NOT ACCEPT ANY RESPONSIBILITY WITH REGARD TO THE CONTENT, ADVERTISEMENTS, PRODUCTS, OR OTHER ISSUES ON SUCH SITES OR SOURCES OR AVAILABILITY. WE ARE NOT RESPONSIBLE FOR ANY KIND OF DAMAGE OR LOSS CAUSED BY OR IN CONNECTION WITH THE USE OF, OR BY RELYING ON THE CONTENT, PRODUCTS OR SERVICES OFFERED ON SUCH SITES OR SOURCES. COPYRIGHT (C) MONIKA BALINT AND CSABA HETENYI, 2017.