

F²Dock Parameters

Rezaul Alam Chowdhury

September 26, 2010

In this documentation the following symbols are used to denote data type:

C: single character/letter
S: string of characters/letters
I: integer
F: floating point
B: Boolean (true/false)

1 Input/Output

For each file name parameter below the name of the file must be specified with full path if it is not in the current folder.

- ***staticMolecule* (S)**: Name of the F2D file for the molecule that is to be held stationary during docking. Typically the static molecule is the larger of the two molecules to be docked.

- ***staticMoleculePQR* (S)**: Name of the PQR [4] file for the static molecule. Current version of F²Dock only works on PQR files without chain information (fix is easy, and will be done soon). However, we also plan to remove this parameter from future versions of F²Dock because the corresponding F2D file already contains all information F²Dock extracts from the PQR file except that the residues in the F2D file are not listed in increasing order of residue numbers (some F²Dock computations depend on this ordering). F²Dock will be updated to internally reorder the residues read from the F2D files whenever necessary.

- ***staticMoleculeQUAD* (S)**: Name of the quadrature points [1, 3] file for the static molecule. Each line of the quadrature points file contains the description of an integration point on the surface of the molecule. Each such line consists of 7 floating point numbers: the first three are the coordinates of the point followed by three floats giving the x , y and z components, respectively, of the unit outward surface normal at that point, while the last one is the weight given to that point.

- ***movingMolecule* (S)**: Name of the F2D file for the molecule that is to be moved around during docking. Typically the smaller of the two molecules to be docked is treated as the moving molecule for performance reasons.

- ***movingMoleculePQR* (S)**: Name of the PQR file for the moving molecule.

- ***movingMoleculeQUAD* (S)**: Name of the quadrature points file for the moving molecule.

- ***rmsdAtoms* (S)**: Name of a text file containing the coordinates of the atoms of the moving molecule that must be included in the RMSD calculation. If this parameter is not specified RMSD

Parameter	complexType		
	<i>A</i>	<i>E</i>	<i>G</i>
<i>peaksPerRotation</i>	3	2	4

Table 1: Default values for *peaksPerRotation* based on complexType.

values will not be calculated. The first line of the file contains the number of atoms in the moving molecule, and the 2nd line contains the number of atoms (say, n) to be included in the RMSD computation. Each of the next n lines contains four numbers. The first number is the atom number/id (an integer) from the F2D file of the moving molecule, and the next three numbers are floating point values giving the co-ordinates of that atom. Observe that the co-ordinate values are redundant as they can be obtained from the F2D file once we know the atom number, and so should be removed from the future versions of F²Dock. The first line of the file (i.e., the number of atoms in the moving molecule) can also be removed.

- **outFile (S):** Name of the file to which F²Dock will output the potential docking solutions (poses).
- **numSolutions (I):** A positive integer giving the upper bound on the number of top docking poses F²Dock should output. Default value is 20,000.

2 Rotations

For each file name parameter below the name of the file must be specified with full path if it is not in the current folder.

- **rotFile (S):** Name of a text file containing the rotations to be applied to the moving molecule during docking. Each line of the file contains one rotation specified by three Euler angles.
- **numRot (I):** The top *numRot* rotations from *rotFile* are used. If left unspecified all rotations will be used.
- **rotateVolume (B):** If set to *true*, the original orientation of the moving molecule is used for gridding to FFT grid, and then the nonzero grid points are rotated and the rotated grid is computed through interpolation. If set to *false* the atoms are gridded from scratch for each rotation. Default value is *true*.
- **randomRotate (B):** If set to *true* a random rotation matrix is applied on the moving molecule initially. Default value is *false*.
- **peaksPerRotation (I):** Only the top *peaksPerRotation* peaks will be retained for each rotation. Default values are given in Table ??.

3 FFT Grid

- **effGridFile (S):** Name of a text file containing the grid sizes for efficient FFT computation. Each line of the file contains an integer (positive even) n indicating a 3D $n \times n \times n$ grid for FFT computation. The list of integers is machine-specific, and F²Dock uses this list as follows for

improved performance. For any given grid size n (implicitly or explicitly) requested by the user F²Dock scans the list of integers in this file, and finds the first integer m that is not smaller than n . The grid sizes in the file are ordered so that FFT on an $m \times m \times m$ grid runs at least as fast as that on an $n \times n \times n$ grid on the machine for which the listing is generated. F²Dock can generate this file automatically if run with the *-effGridFile* option preceding the name of the input parameters file.

- ***gridSpacing* (F)**: A positive floating point value giving an upper bound on the spacing (in Å) between adjacent grid points in the 3D spatial FFT grid. F²Dock consults the *effGridFile* file in order to compute a grid size so that the grid spacing does not exceed the upper bound provided by the user, and at the same time FFT can be computed as efficiently as possible. Thus *numFreq* parameter depends on this approximated grid spacing. The default upper bound is 1.2 Å.
- ***enforceExactGridSpacing* (B)**: If set to *true* the the FFT grid is scaled up to match the user-specified grid-spacing upper bound (i.e., the grid cells are expanded without changing the number of cells in the grid). The default value is *false*.
- ***numFreq* (I)**: A positive integer n specifying the number of frequencies n^3 to be used during FFT computations. If this parameter is not specified it is computed from the *gridSpacing* parameter.
- ***sparseFFT* (B)**: If set to *true* the sparsity of the input and the output grids will be exploited for faster computation. Default value is *true*.
- ***narrowBand* (B)**: If set to *true* only the positions of the moving molecule that lie within a narrow band around the static molecule will be considered for finding potential solutions, otherwise the entire 3D grid is searched. Default value is *true*.
- ***minEffGridSize* (I)**: A positive integer specifying the minimum grid size to consider when creating *effGridFile* by running F²Dock with the *-effGridFile* option.
- ***maxEffGridSize* (I)**: A positive integer specifying the maximum grid size to consider when creating *effGridFile* by running F²Dock with the *-effGridFile* option.

4 Complex Type

- ***complexType* (C)**: Type of the docked complex that has four possible values: *A* (antibody-antigen), *E* (enzyme-inhibitor or enzyme-substrate), *G* (neither A nor E), and *U* (unknown). The default value is *U* in which case F²Dock tries to identify the complex type itself (very successfully for antibody-antigen, and with moderate success for enzyme-inhibitors and enzyme-substrates).

5 Shape Complementarity

- ***blobbiness* (F)**: Specifies the value controlling the blobbiness of the Gaussians used for representing the atoms. Default value is -2.3.
- ***pseudoAtomRadius* (F)**: Radius of each pseudo (skin) atom of the static molecule. Default value is 1.1 Å.

Parameter	complexType		
	<i>A</i>	<i>E</i>	<i>G</i>
<i>skinSkinWeight</i>	0.73	0.78	0.57
<i>skinCoreWeight</i>	-0.31	-0.08	-0.23
<i>coreCoreWeight</i>	31.00	5.00	5.00
<i>curvatureWeightingRadius</i>	4.5	6.0	4.5

Table 2: Default values for various shape complementarity related parameters based on complexType.

- ***singleLayerLigandSkin* (B):** If set to *false* the core atoms adjacent to the skin atoms of the moving molecule are also considered part of the skin. Default value is *false*.
- ***curvatureWeightedStaticMol* (B):** If set to *true* the skin atoms of the static molecule are weighted based on curvature. Default value is *true*.
- ***curvatureWeightedMovingMol* (B):** If set to *true* the skin atoms of the moving molecule are weighted based on curvature. Default value is *false*.
- ***curvatureWeightingRadius* (F):** The radius of the spherical region around each atom center used for curvature computation. Default value is given in Table 2.
- ***bandwidth* (F):** The width of the onion shell type bands of core atoms constructed for computing depth-dependant core weights. Default value is 2 Å.
- ***gradFactor* (F):** The weight of the core atoms in any given shell is a factor of *gradFactor* more than the weights of the core atoms in the shell just outside of it. Default value is 1.1.
- ***skinSkinWeight* (F):** A positive floating point value specifying the reward given to unit skin-skin overlap during shape-complementarity scoring. See Table 2 for default value.
- ***skinCoreWeight* (F):** Specifies the wight (reward or penalty) given to skin-core overlaps during shape-complementarity scoring. Default value is given in Table 2.
- ***coreCoreWeight* (F):** A positive floating point value specifying the penalty given to unit core-core overlaps during shape-complementarity scoring. See Table 2 for default value.

6 Electrostatics

- ***elecWeight* (F):** The weight given to the electrostatics score computed by F²Dock. See Table 3 for default value.
- ***elecKernelVoidRad* (F):** Specifies the d_0 distance in computing the distance-dependant dielectric constant $E(\mathbf{x})$ using Equation 1 given below (a generalization of the Gabb et al. expression [5]). Default value is given in Table 3.

$$E(\mathbf{x}) = \begin{cases} 0 & \text{if } ||\mathbf{x}|| \leq d_0, \\ v_l & \text{if } d_0 < ||\mathbf{x}|| \leq d_l, \\ c_1||\mathbf{x}|| + c_2 & \text{if } d_l < ||\mathbf{x}|| \leq d_h, \\ v_h & \text{if } d_h < ||\mathbf{x}||, \end{cases} \quad (1)$$

Parameter	complexType		
	<i>A</i>	<i>E</i>	<i>G</i>
<i>elecWeight</i>	0.72	0.15	0.72
<i>elecKernelVoidRad</i>	0.0	3.0	0.0
<i>elecKernelValLow</i>	4.0	1.0	4.0

Table 3: Default values for various electrostatics related parameters based on complexType.

where, $c_1 = \frac{v_h - v_l}{d_h - d_l}$, and $c_2 = v_l - d_l c_1$.

- ***elecKernelDistLow* (F)**: Specifies the d_l distance in Equation 1. Default value is 6 Å.
- ***elecKernelDistHigh* (F)**: Specifies the d_h distance in Equation 1. Default value is 8 Å.
- ***elecKernelValLow* (F)**: Specifies the v_l value in Equation 1. See Table 3 for default value.
- ***elecKernelValHigh* (F)**: Specifies the v_h value in Equation 1. Default value is 80.
- ***elecRadiusInGrids* (F)**: A positive floating point value specifying the constant radius of each atom within which its charge is diffused using a Gaussian during electrostatics scoring. Default value is 2.9.

7 Hydrogen Bonding

- ***hbondWeight* (F)**: The weight given to the hydrogen bonding score computed by F²Dock. Default value is 0.0.
- ***hbondDistanceCutoff* (F)**: A positive floating point value specifying the surface to surface distance upper bound between any pair of donor and acceptor atoms. The default value is 2 Å.

8 Hydrophobicity Scoring

- ***hydrophobicityWeight* (F)**: Weight given to the ratio $r_h = \frac{\text{interface hydrophobic score}}{\text{interface hydrophilic score}}$ computed by F²Dock. Default value is given in Table 4.
- ***hydrophobicityProductWeight* (F)**: Weight given to the product $r_h \times \text{interface hydrophobic score}$ computed by F²Dock. Default value is 0.001.
- ***hydroRatioTolerance* (F)**: A positive floating point value giving the upper bound on r_h . If r_h exceeds this value r_h is set to *hydroPenalty* (*hydroPenalty* = -10 in the current version). See Table 4 for default value.
- ***hydroMinRatio* (F)**: A nonnegative floating point value giving the lower bound on r_h . If r_h is below this value r_h is set to *hydroPenalty*. Table 4 gives the default value.
- ***hydroRatioNumeratorLow* (F)**: Lower bound on *interface hydrophobic score*. If this value is below the lower bound r_h is set to *hydroPenalty*. See Table 4 for default value.

Parameter	complexType		
	<i>A</i>	<i>E</i>	<i>G</i>
<i>hydrophobicityWeight</i>	8.5	9.0	8.5
<i>hydroRatioTolerance</i>	8.0	8.0	10.0
<i>hydroMinRatio</i>	1.5	1.22	0.5
<i>hydroRatioNumeratorLow</i>	1.25	2.0	2.0
<i>hydroRatioDenominatorLow</i>	0.45	1.0	0.2
<i>hydroRatioDenominatorHigh</i>	2.5	7.0	6.0

Table 4: Default values for various hydrophobicity related parameters based on complexType.

- ***hydroRatioNumeratorHigh* (F)**: Upper bound on *interface hydrophobic score*. If this value is above the upper bound r_h is set to *hydroPenalty*. Default value is 100.0.
- ***hydroRatioDenominatorLow* (F)**: Lower bound on *interface hydrophilic score*. If this value is below the lower bound r_h is set to *hydroPenalty*. See Table 4 for default value.
- ***hydroRatioDenominatorHigh* (F)**: Upper bound on *interface hydrophilic score*. If this value is above the upper bound r_h is set to *hydroPenalty*. Table 4 gives the default value.
- ***twoWayHydrophobicity* (B)**: If set to *false* only the hydrophobicity of the atoms of the static molecule is considered, otherwise both molecules are considered. Default value is *true*.
- ***useInterfacePropensity* (B)**: If set to *true* interface propensity values from [7] are used, otherwise hydrophobicity values from [2] are used. Default value is *true*.
- ***perResidueHydrophobicity* (B)**: If set to *true* per residue hydrophobicity values from [2] are used, per atom hydrophobicity values [8] are used. Default value is *true*. If *useInterfacePropensity* is set to *true* this parameter is ignored (i.e., always assumed to be *true*).
- ***staticMolHydroDistCutoff* (F)**: A non-negative floating point value giving the maximum distance from the surface of a core atom of the static molecule to the center of any skin atom in order to consider that core atom for hydrophobicity calculation. Default value is 4.0 Å.

9 Simple Charge Complementarity

- ***simpleChargeWeight* (F)**: Weight given to the simple charge complementarity score computed by F²Dock. Default value is given in Table 5.
- ***simpleRadExt* (F)**: The radius of each atom is extended by this value for simplified charge complementarity computation. Default value is 1.5 Å.

10 Clash Filter

- ***applyClashFilter* (B)**: If set to *true* clash filter is applied. Default value is *true*.

Parameter	complexType		
	<i>A</i>	<i>E</i>	<i>G</i>
<i>simpleChargeWeight</i>	0.1	5.5	2.0

Table 5: Default values for various simple charge complementarity related parameters based on complexType.

Parameter	complexType		
	<i>A</i>	<i>E</i>	<i>G</i>
<i>clashTolerance</i>	2	9	10
<i>clashWeight</i>	-30	-0.5	-0.5

Table 6: Default values for various clash filter related parameters based on complexType.

- ***eqmDistFrac* (F)**: Two atoms are considered to be in a clash if the distance between the atom centers is less than *eqmDistFrac* fraction smaller than the sum of their radii. Default value is 0.5.
- ***clashTolerance* (I)**: Upper bound on the number of atomic clashes permitted. Default value is given in Table 6.
- ***clashWeight* (F)**: Weight given to each clash when added to the total score. See Table 6 for default value.

11 Lennard-Jones Filter

- ***applyVdWFilter* (B)**: If set to *true* Lennard-Jones filter is applied. Default value is *true*.
- ***vdWCutoffLow* (F)**: If $\#clashes < clashTolerance / 2$, poses with vdW potential $\geq vdWCutoffLow$ are penalized. Default value is 0.0.
- ***vdWCutoffHigh* (F)**: If $\#clashes \geq clashTolerance / 2$, poses with vdW potential $\geq vdWCutoffHigh$ are penalized. Default value is given in Table 7.
- ***vdWEqmRadScale* (F)**: All r_{eqm} values are multiplied by this factor. Default value is 0.3.
- ***epsilonLJ* (F)**: Error control parameter ($\in (0, 1]$) for Lennard-Jones potential. Default value is 0.5.
- ***useSSE* (B)**: Is set to *true* SSE (Streaming SIMD Instructions) will be used for faster execution. Default is *false*.

Parameter	complexType		
	<i>A</i>	<i>E</i>	<i>G</i>
<i>vdWCutoffHigh</i>	0.0	20.0	5.0

Table 7: Default values for various Lennard-Jones filter related parameters based on complexType.

12 Hydrophobicity Filter

- ***applyPseudoGsolFilter* (B)**: If set to *true* hydrophobicity filter will be applied. Default value is *true*.
- ***pseudoGsolWeight* (F)**: Weight given to the value $\frac{(\text{interface hydrophobic score})^2}{\text{interface hydrophilic score}}$ computed by the hydrophobicity filter which is then added to the total score. Default value is 0.0. This parameter is now redundant and will be removed.
- ***useInterfacePropensity* (B)**: Same as in Section 8.
- ***perResidueHydrophobicity* (B)**: Same as in Section 8.

13 Dispersion Filter

- ***applyDispersionFilter* (B)**: If set to *true* dispersion filter will be applied. Default value is *false*.
- ***dispersionEnergyLimit* (F)**: If the computed dispersion energy has value larger than *dispersionEnergyLimit* it is set to *dispersionEnergyLimit*, and if it has value smaller than *-dispersionEnergyLimit* it is set to *-dispersionEnergyLimit*. Default value is 1000.0.
- ***dispersionMinAtomRadius* (F)**: Smallest permitted atom radius during dispersion energy calculation. Default value is 0.1 Å.
- ***dispersionWeight* (F)**: Weight given to the dispersion energy value computed which is then added to the total score. Default value is 0.0.
- ***epsilonBR* (F)**: Error control parameter ($\in (0, 1]$) for dispersion energy approximation. Default value is 0.3.

14 Reranking

- ***rerank* (B)**: If set to *true* the output of F²Dock will be reranked based on various filters. Default value is *false*.
- ***numRerank* (I)**: Number of top solutions to rerank. Default value is 2000.
- ***applyAntibodyFilter* (B)**: If set to *true* antibody filter will be applied which filters based on the active sites on the antibody. Default value is *true*.
- ***applyEnzymeFilter* (B)**: If set to *true* enzyme filter will be applied which filters based on the abundance of Glycine residues. Default value is *true*.
- ***applyResidueContactFilter* (B)**: If set to *true* filtering will be done based on residue contact preferences (from Table III, page 94 of [6]). Default value is *true*.
- ***rerankerF2DockScoreWeight* (F)**: Weight given to the original F²Dock score. Default value is 100.0.
- ***rerankerPseudoGsolWeight* (F)**: Weight given to the value $\frac{(\text{interface hydrophobic score})^2}{\text{interface hydrophilic score}}$ computed by the hydrophobicity filter. Default value is 1.0.

15 Clustering

- ***clusterTransRad* (F):** The radius of the level $i \in [1, 3]$ cluster is $i \cdot \text{clusterTransRad}$. Default value is 1.2 Å.
- ***clusterTransSize* (F):** Maximum number of docking poses (i.e., geometric center of the moving molecule) in cluster $i \in [1, 3]$ is $i \cdot \text{clusterTransSize} + c_i$, where $c_1 = 0$, $c_2 = 1$ and $c_3 = 3$. Default value is 1.

16 Parallelization

- ***numThreads* (I):** A positive integer specifying the number of concurrent threads to use. Default value is 4.

17 Misc

- ***scoreScaleUpFactor* (F):** F²Dock scores are scaled up by this factor. Default value is 10,000.

References

- [1] BAJAJ, C., AND ZHAO, W. Fast molecular solvation energetics and forces computation. *SIAM Journal on Scientific Computing* 31, 6 (2010), 4524–4552.
- [2] BLACK, S., AND MOULD, D. Development of hydrophobicity parameters to analyze proteins which bear post- or cotranslational modifications. *Anal. Biochem.* 193 (1991), 72–82.
- [3] CHOWDHURY, R. A., AND BAJAJ, C. Algorithms for faster molecular energetics, forces and interfaces. ICES report 10-32, Institute for Computational Engineering & Science, The University of Texas at Austin, Austin, TX, USA 78712., August 2010.
- [4] DOLINSKY, T., NIELSEN, J., MCCAMMON, J., AND BAKER, N. Pdb2pqr: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Research* 32 (2004), 665–667.
- [5] GABB, H. A., JACKSON, R. M., AND STERNBERG, M. J. E. Modelling protein docking using shape complementarity, electrostatics and biochemical information. *Journal of Molecular Biology* 272, 1 (1997), 106–120.
- [6] GLASER, F., STEINBERG, D. M., VAKSER, I. A., AND BEN-TAL, N. Residue frequencies and pairing preferences at proteinprotein interfaces. *PROTEINS: Structure, Function, and Genetics* 43 (2001), 89–102.
- [7] JONES, S., AND THORNTON, J. M. Analysis of protein-protein interaction sites using surface patches. *Journal of Molecular Biology* 272, 1 (1997), 121–132.
- [8] KAPCHA, L., AND ROSSKY, P. Personal Communication, University of Texas at Austin.