Pocket descriptors

extrait de l'aide du logiciel Fpocket (https://fpocket.sourceforge.net/manual_fpocket2.pdf)

In order to discriminate an interesting pocket from a lot of uninteresting ones, fpocket usesdescriptors for each pocket. A scoring function, using these descriptors, was trained to well identifywhat we generally call "binding site". Here are set together all descriptors implemented in fpocket.The ones that are currently used for scoring are marked with a *, and the one having the tag. Advanced features normalized associated with have a normalized (ie. scaled to a [0, 1] range, the highest (resp thelowest) value of a given descriptor being set to 1 (resp 0)) equivalent descriptor.

Number of alpha spheres (normalized)

As the title says, this is surely the most simple descriptor. The number of alpha spheres reflectsgenerally more or less proportionally the size of the cavity.

Density of the cavity (normalized)

This descriptor tends to measure the density and "buriedness" of a pocket. It is nothing else thanthe mean value of all alpha sphere pair to pair distances in the binding pocket. Thus, a small valueindicates a rather big compactness of the binding pocket and thus a rather burried pocket. Largervalues give indication about more extended and exposed cavities.

Polarity Score (normalized)

In the contrary to hydrophobicity this descriptor tries to measure the hydrophilicity character ofa binding pocket. To each residue of the binding pocket a polarity score is assigned (as published onhttp://www.info.univangers.fr/~gh/ldas/proprietes.htm). The final polarity score is the mean of allpolarity scores of all residues in the binding pocket. This is extremely approximative, so should notbe overestimated. Each residue is evaluated only once.

Mean local hydrophobic density (normalized)

This descriptor tries to identify if the binding pocket contains local parts that are ratherhydrophobic. For each apolar alpha sphere the number of apolar alpha sphere neighbors is detectedby seeking for overlapping apolar alpha spheres. The sum of all apolar alpha sphere neighbors isdivided by the total number of apolar alpha spheres in the pocket. Last this score is normalizedcompared to other binding pockets.

Proportion of apolar alpha spheres (normalized)

This descriptor, returned as percentage, reflects the proportion of apolar alpha spheres among allalpha spheres of one pocket identified by fpocket. This can reflect somehow the hydrophobic/philiccharacter of a binding pocket.

Druggability Score

The druggability score is a numerical value between 0 and 1 associated to each pocket using alogistic function. This scores intends to assess the likeliness of the pocket to bind a small drug likemolecule. A low score indicates that drug like molecules are likely to not bind to this pocket. A druggability score at 0.5 (the threshold) indicates that binding of prodrugs or druglike moleculescan be possible. 1 indicates that binding of druglike molecules is very likely. The theoretical basis of the score is currently in the lengthy process of scientific publication.

Maximum distance between two alpha sphere (normalized)

This descriptor store the maximum distance found between two alpha sphere in a given pocket.

Hydrophobicity Score

This descriptor is based on a residue based hydrophobicity scale published by Monera & al. inthe Journal of Protein Science 1, 319329 (1995). For all residues implicated in the binding site themean hydrophobicity score is calculated and is used as descriptor for the whole pocket. Eachresidue is evaluated only once.

Charge Score

According to (http://www.info.univangers.fr/~gh/ldas/proprietes.htm) the charge of each aminoacid in the binding site is tracked. The mean charge for all amino acids in contact with at least onealpha sphere of the pocket is calculated to form this charge score. Each residue is evaluated onlyonce.

Volume Score

Similarly to other descriptors, this one is based on data published on (http://www.info.univangers.fr/~gh/ldas/proprietes.htm). This data resumes relative volume of different amino acids. Inorder to calculate this descriptor the mean volume score of all amino acids in contact with at leastone alpha sphere of the pocket is calculated. Each residue is evaluated only once.

Composition of amino acids

As the name indicates, fpocket tracks the composition in amino acids of binding pockets. If atleast one atom of a residue is in contact with at least one alpha sphere of a binding pocket it isaccounted to be part of the binding site. This descriptor is returned as cumulative list, for instanceyou can find 2 valines, 3 glutamates etc... in the binding site. Occurences of amino acids in different descriptor outputs are given in the following order : Ala,Cys, Asp, Glu, Phe, Gly, His, Ile, Lys, Leu, Met, Asn, Pro, Gln, Arg, Ser, Thr, Val, Trp, Tyr.

Pocket volume

As indicated by the name, this descriptor tries to evaluate the volume of a binding pocket using a Monte-Carlo algorithm that calculates full volume occupied by all alpha sphere in a given pocket.The number of iteration of this algorithm can be controlled using fpocket input parameters.

Polar Surface Area

This descriptor provides an estimation of the polar surface area of the pocket based oninformation of the receptor atoms. The method used to calculate the area only provides anapproximation, but should be good enough to get some rather relevant estimates.

Apolar Surface Area

See polar surface area in the previous point, only for apolar atoms.

Total Surface Area

The sum of the polar and apolar surface area of the pocket, that is to say the receptor side surfacearea of the pocket.

B-factor score (normalized)

Please handle with a lot of care this score with native crystal structures. This score is based onthe mean B-factor of all atoms of the binding pocket (atoms that are contacted by at least one alphasphere). As the B-factor does not necessarily reflect flexibility in crystal structures, this score issomehow abusive. However, one could imagine performing molecular dynamics or other in order todetermine relative flexibility of atoms and store this information in the B-factor column of the PDBfile format. This descriptor is normalized with other pockets of the same protein.