Accessing 3D molecules from fragments *via* non-planar exit vectors – PBFev

Joshua Meyers*[1], Nicholas C. Firth[2], Yi Mok[1], Nathan Brown[1] and Julian Blagg[1]

1a. 3D fragments?

- 3D molecules may have characteristics that are desirable in drug discovery – solubility, cell permeability, increased probability of clinical success and potentially the ability to bind to interesting target classes [1]
- 3D fragments libraries are often claimed to have the potential to generate 3D molecules from fragment hits
- **3D fragments, however, are necessarily more "complex", and therefore the probability of finding leads for drug discovery may be decreased [2]**
- 3D substructures are not necessarily a requisite for 3D molecules – molecules can be 3D despite being comprised of 2D substructures [1]

3D fragments may not generate 3D molecules

1b. Exit vectors

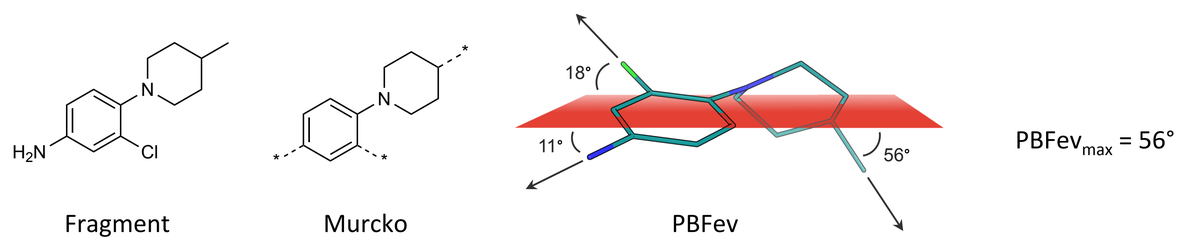
- We sought a method that would quantify the **potential of a fragment to grow into a 3D molecule**
- We hypothesised that the potential for evolution from a fragment to a 3D molecule could be modeled by considering the **angle between a central molecular scaffold and its associated "exit vectors"**
- Current methodologies for analysing "exit vectors" are either **scaffold specific** (*e.g.* biaryls [3]) or **challenging to interpret** [4]

Hypothesis:

The potential for fragments to grow towards 3D molecular space can be predicted by considering scaffold exit vectors

2. PBFev – method

1. Exit vectors are defined as known attachment points to the **Bemis-Murcko framework** [5]
2. The **Plane of Best Fit (PBF)** [6] of the Murcko scaffold atoms is set as a reference plane



3. The **PBFev$_{max}$ is set as the maximum angle** between the Murcko exit vectors and the PBF plane of the Murcko scaffold atoms – pendant linkers do not affect the plane

$$angle = \sin^{-1}\frac{|\vec{n}\cdot\vec{v}|}{\sqrt{|\vec{n}\cdot\vec{n}|}\cdot\sqrt{|\vec{v}\cdot\vec{v}|}}$$

Angle of intersection between a vector (v) defined by two points and a plane represented by its normal (n)

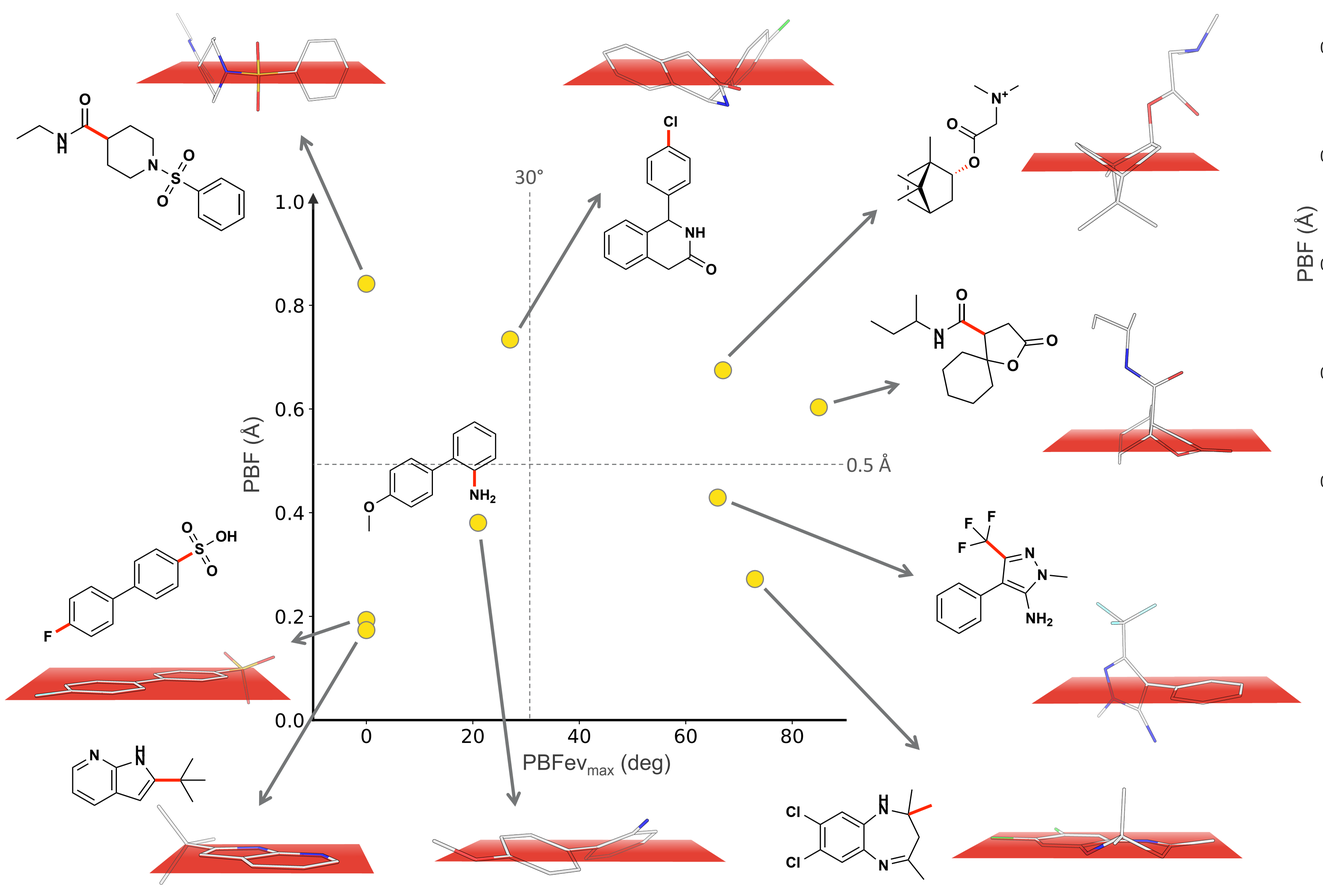
- Consideration of the maximum value of PBFev assesses the potential of a fragment to access 3D molecules
- Possible values for PBFev$_{max}$ range from 0 – 90°
- It would also be possible to consider PBFev$_{min}$ or the entire PBFev array
- The PBFev descriptor is interpretable and generalisable across multiple molecular scaffolds

```python
def PBFev(mol):
    '''returns an array of exit vectors for this mol'''
    # Get murcko SMILES
    murcko = MurckoScaffold.GetScaffoldForMol(mol)

    # Get PBF plane for murcko scaffold only
    confId = -1
    conf = murcko.GetConformer(confId)
    if not conf.Is3D():
        print('This mol is not 3D - all PBFev angles will be 0 degrees')
        return [0]
    pts = np.array([list(conf.GetAtomPosition(i))  # Get atom coordinates
                    for i in xrange(murcko.GetNumAtoms())])
    # GetBestFitPlane is in the RDKit Contrib directory as part of PBF
    # Plane is xyz vector with a c intercept adjustment
    plane = GetBestFitPlane(pts)

    # Map onto parent structure coords (this func adds exit vectors [*])
    murckoEv = Chem.ReplaceSidechains(mol, murcko)

    confId = -1  # embed 3D conf object with EVs (atom indices do not change)
    conf = murckoEv.GetConformer(confId)

    # Where [#0] matches exit vector SMILES [*]
    patt = Chem.MolFromSmarts('[#0]-[*]')
    matches = murckoEv.GetSubstructMatches(patt)
    if len(matches) == 0:
        return None

    # Calculate angles between exit vectors and the murcko plane of best fit
    exitVectors = np.zeros(len(matches))
    denom = np.dot(plane[:3], plane[:3])
    denom = denom**0.5
    for n, match in enumerate(matches):
        evCoords = conf.GetAtomPosition(match[0])
        anchorCoords = conf.GetAtomPosition(match[1])
        v = np.array(((evCoords[0]-anchorCoords[0]),
                      (evCoords[1]-anchorCoords[1]),
                      (evCoords[2]-anchorCoords[2])))
        angle = np.arcsin((np.dot(v, plane[:3])) /
                          ((denom)*((np.dot(v, v))**0.5)))
        angle = np.abs(np.degrees(angle))
        exitVectors[n] = angle

    return exitVectors
```

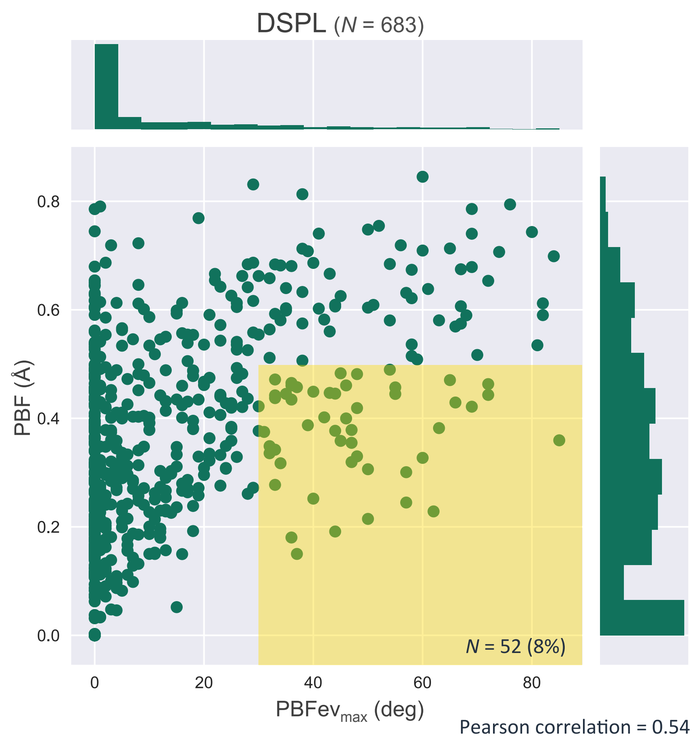
*GitHub: https://github.com/JoshuaMeyers/Publications/blob/master/2017-PBFev/pbfev.py

3. Comparison of fragment PBFev$_{max}$ with fragment three-dimensionality

- Comparison of the **PBFev$_{max}$ of the Murcko scaffold** with the **PBF exhibited by a fragment conformation** allows assessment of both 3D complexity and the potential to grow to 3D molecules
- The Murcko scaffold conformation is maintained from that of the full molecule generated using the ETKDG method in the RDKit [7]
- The four examples with PBFev$_{max}$ > 30° demonstrate the differences between complex 3D scaffolds (PBF > 0.5 Å) and those apparently "simpler", and their similar potential to access 3D molecules according to PBFev$_{max}$



- The **Diamond-SGC Poised Library (DSPL)** comprises of 776 fragments that contain a synthetically accessible scaffold associated with a robust functionalisation reaction [8]. The library contains 423 unique Murcko frameworks.



- Fragments in the DSPL library tend towards a low PBF score, and planar exit vectors
- 93 molecules (12%) were omitted because the Murcko framework is the same as the fragment and so no exit vectors were defined
- **52 molecules (8%) were identified with PBFev$_{max}$ > 30° and PBF < 0.5 Å (yellow area). These molecules are of lower 3D complexity but offer the potential to access 3D molecules by appropriate functionalisation**

4. Summary

- **Fragments with low PBF have the potential to generate 3D molecules by functionalisation along PBFev$_{max}$ exit vectors**
- **PBFev$_{max}$ is a useful descriptor of the potential for generating 3D molecules from synthetically accessible fragment-like scaffolds**

Contact

*joshua.meyers@icr.ac.uk
[1]Cancer Research UK Cancer Therapeutics Unit, Division of Cancer Therapeutics The Institute of Cancer Research, London, SM2 5NG, United Kingdom.

[2]Centre for Medical Image Computing, Department of Computer Science, University College London, London, United Kingdom

Acknowledgments

J.M. is supported by Wellcome Trust Grant 102361/Z/13/Z; N.F. is supported by EPSRC Grant EP/M006093/1); and Y.M., N.B. and J.B. are supported by Cancer Research U.K. Grant C309/A11566.

References

[1] J. Meyers, M. Carter, Y. Mok and N. Brown, *Future Med. Chem.*, (2016) **8**(14), 1753–1767.
[2] M. Hann, A. Leach and G. Harper, *J. Chem. Inf. Comput. Sci.*, (2001) **41**, 856–864.
[3] D. Brown and J. Boström, *J. Med. Chem.*, (2016), **59**(10), 4443–4458.
[4] O. Grygorenko, P. Babenko, D. Volochnyuk, O. Raievskyi, and I. Komarov, *RSC Adv.*, (2016), **6**(21), 17595-17605.
[5] G. Bemis and A. Murcko, *J. Med. Chem.*, (1996), **39**(15), 2887–2893.
[6] N. Firth, N. Brown and J. Blagg, *J. Chem. Inf. Model*, (2012) **52**(10), 2516–2525.
[7] S. Riniker and G. Landrum, *J. Chem. Inf. Model.*, (2015), **55**(12), 2562–2574.
[8] O. Cox, T. Krojer, P. Collins *et al.*, *Chem. Sci.*, (2016) **7**(3), 2322–2330.