Electronic Supplementary Material (ESI) for RSC Medicinal Chemistry.

Drugging the undruggable: A computational chemist's view of KRAS^{G12C}

Michael S. Bodnarchuk*, Doyle J. Cassar, Jason G. Kettle, Graeme Robb, and Richard A. Ward

SUPPORTING INFORMATION

Protein preparation

All structures were prepared using the 'Protein Preparation Wizard' tool in Maestro, retaining all crystallographic waters. Protein and ligand protonation states assigned using PROPKA and Epik respectively at pH 7.0 ± 1. The hydrogen-bonding network was assigned for each complex, followed by a restrained minimisation using a customised version of the OPLS3e forcefield to a RMSD of 0.3 Å.

Torsion scans

Torsion scans were performed in Jaguar using a relaxed coordinate torsion protocol. Scans were run at the B3LYP level of theory using a 631G^{**+} basis set. Scans were performed for 48 iterations in the gas-phase using 12 equally spaced windows between 0^{o} and 360^{o}.

Free Energy Perturbation

All free-energy perturbation (FEP) simulations were run using the crystal structure 6T5U and the FEP^{+} package in Maestro. A customised force-field, run using the Force-Field builder in Maestro, was used to describe the torsions of the molecules, with simulations run for 25 ns per lambda window in both the solvent and complexed legs within the NPT ensemble. Replica exchanges between neighbouring λ windows were attempted every 1.2 ps. 16 λ windows were used for core-hopping, whilst 24 λ windows were used for the charge hopping transformations. Perturbation maps were automatically generated by Maestro, with additional legs incorporated to ensure each compound was connected to at least 3 other compounds in the map. A run size of 25 ns was found to be sufficient to achieve convergence in the free energy estimate, with shorter run times leading to poorer convergence and occasionally hysteresis. Prior to the simulations, validation was performed on a set of 20 compounds to ensure that the predicted binding free energies correlated with experimental pIC_{50} values (demonstrating an R^2 value of 0.75).

Grand Canonical Monte Carlo simulations

GCMC simulations were performed using the ProtoMS software package (v3.4). GCMC was performed for 45 M MC moves around the region of interest (using a 3x 3x 5 Å^3 box) using

20 equally spaced B values between 0 and -20. An initial GCMC equilibration of 5 M MC moves was performed, using a 1:1:1 ratio of insertion, deletion, and GC water sampling moves. 40 M production MC moves across the entire system were then performed. Bulk solvent is prohibited from entering the GCMC region whilst ligand and protein atoms can fully sample the region.