Using Cresset's Spark to grow and link distant fragment hits with sensible chemistry

Martin Slater and Tim Cheeseright

Cresset, New Cambridge House, Bassingbourn Road, Litlington, Cambridgeshire, SG8 0SS, UK

martin@cresset-group.com   www.cresset-group.com

Abstract

The main advantage of fragment-based drug discovery (FBDD) is that the starting point, despite having weak potency, often has high efficiency of interaction with biological target. Optimization of the fragment by adding atoms ideally preserves this initial efficiency, to yield a tractable lead with superior properties. One potential pitfall is that modern hit optimization strategies, particularly those involving molecular modeling, can be deficient in the way they handle these quite small molecular fragments.

In an extension to our existing scaffold hopping methodology we show how a highly valuable FBDD task - fragment growing or linking - can be conducted computationally with high probability of success by using Spark, Cresset's scaffold hopping and bioisostere replacement software. The use of databases of structures derived from real compounds gives a high degree of diversity combined with reasonable probability of chemical synthesis. Creation of final 'product' molecules, which are fully energy minimized before computational scoring, ensures that only linking fragments that can truly work are progressed into the final results.

Background

Uracil DNA glycosylase (UDG) is a potentially interesting target for both cancer and anti-viral therapies. In the body it is important as a DNA repair protein. During DNA replication uracil can be mis-incorporated into DNA or formed enzymatically through the action of cytosine deaminase. Uracil removal is an ongoing process necessary for maintaining DNA integrity. UDG actively targets uracil containing DNA (Figure 1), senses uracil and clips the N-glycosidic linkage (Figure 2). Inhibition of UDG has potentially useful clinical consequences through degrading replication efficiency in cells loaded with viral or cancer genes. A recent effort to produce synthetic inhibitors of this protein[1] relied on an active fragment tethering approach, which yielded some interesting albeit weakly potent bis-oxime linked active ligands. We describe an alternative to this, using a molecular modeling technique, which demonstrates the utility of Spark for the efficient elaboration of tractable fragment growing or linking chemistry.

Fragment growing and linking methodology

The methodology is ligand centric, but can be effective using either RNA or protein in addition to a small molecule ligand as a reference template. The protein DNA mimetic for example (Figure 3) was successfully used as a template to score potential fragments. The UDG protein is also employed as an excluded volume.
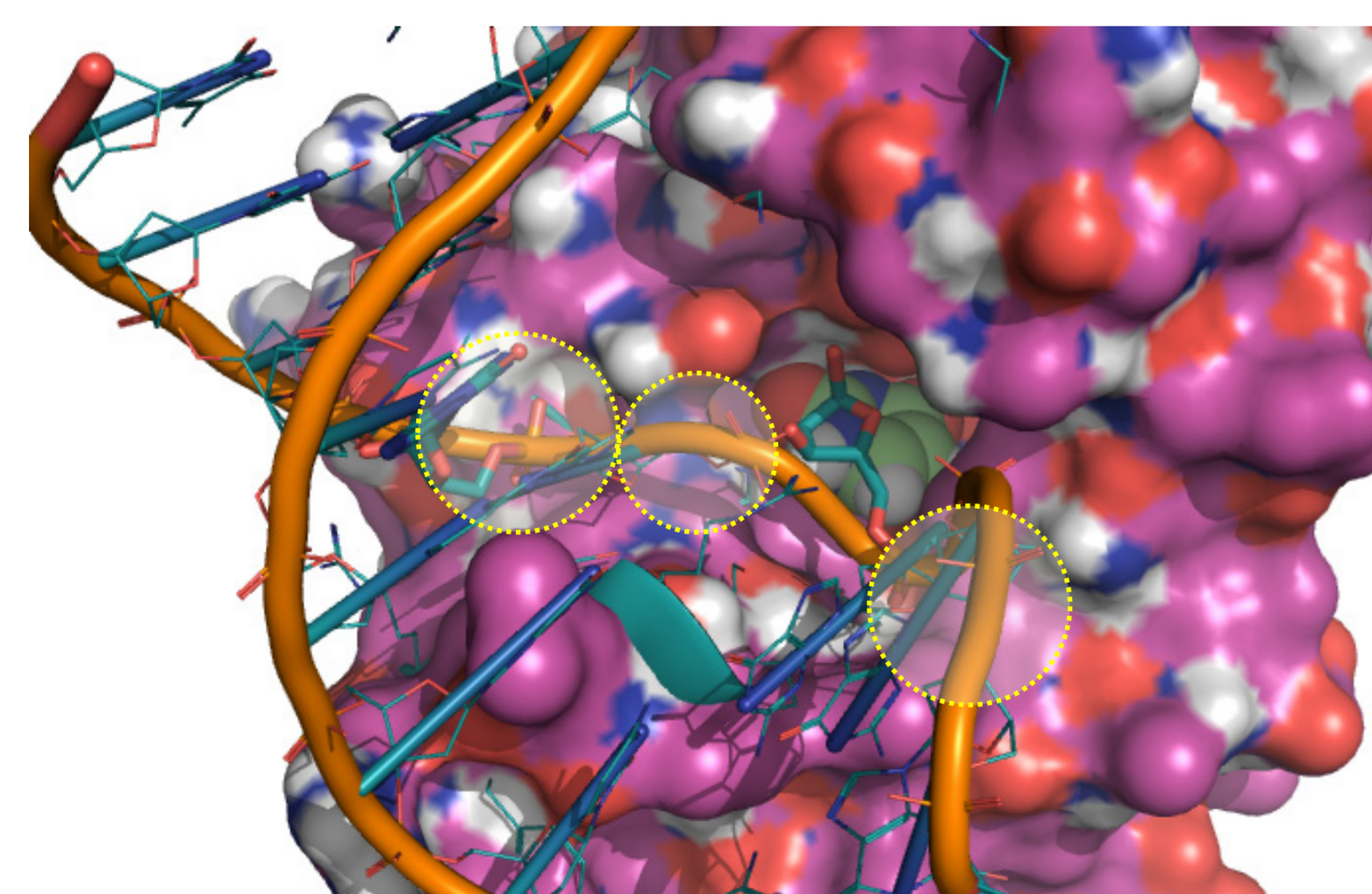

Figure 1. PDB: 4SKN DNA bound UDG.

There are three important phosphate ester recognition sites in UDG (shown as yellow/green circles in Figures 1-5). These correspond to two of the protein's helix dipoles (which stabilize the negative potential) and an important Asn residue (Figure 2). Both a DNA mimicking protein (Figure 3) and subsequent active fragments appear to recognize and bind preferentially to these protein hot spots so are potentially useful sites to target.
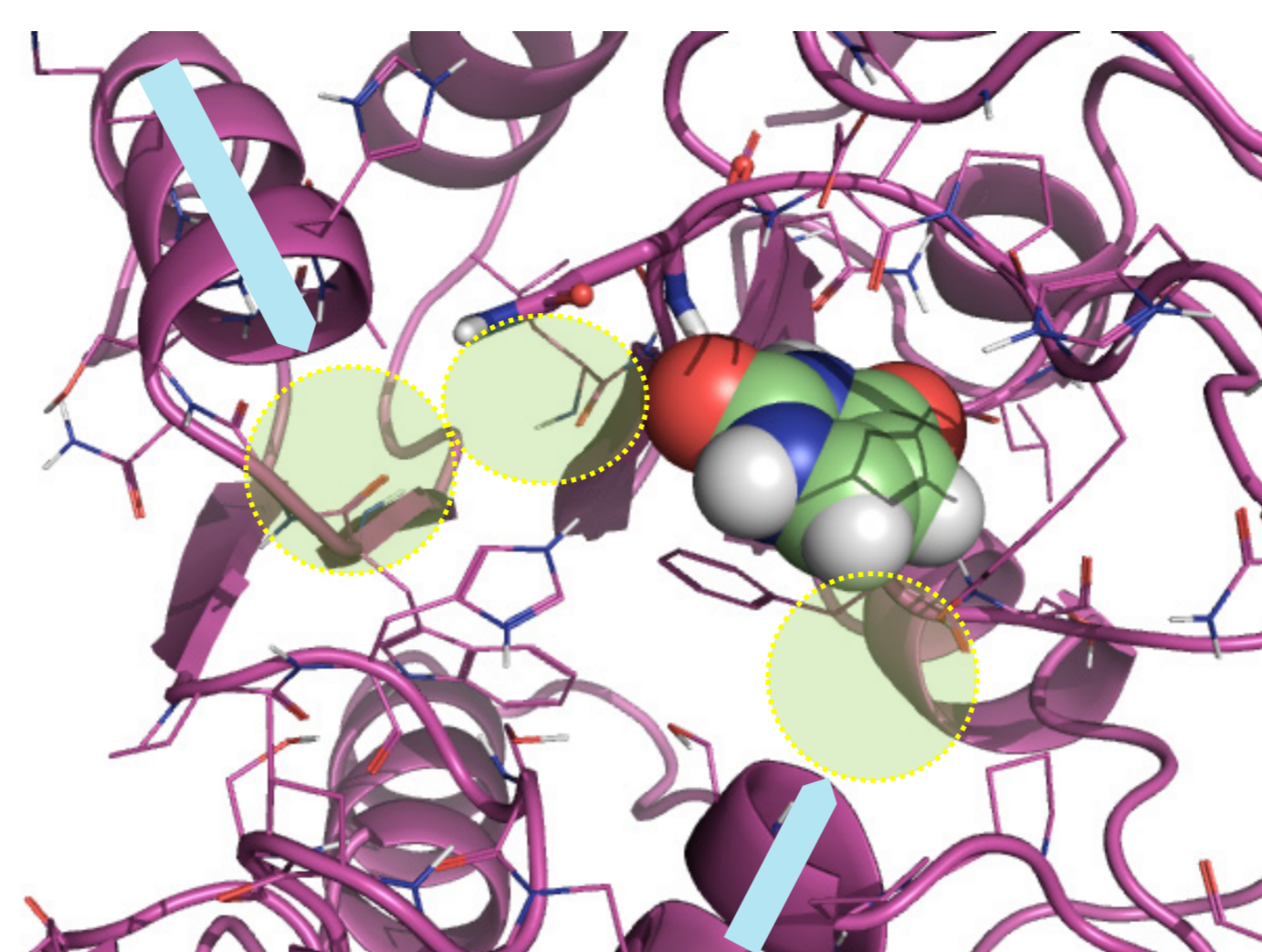

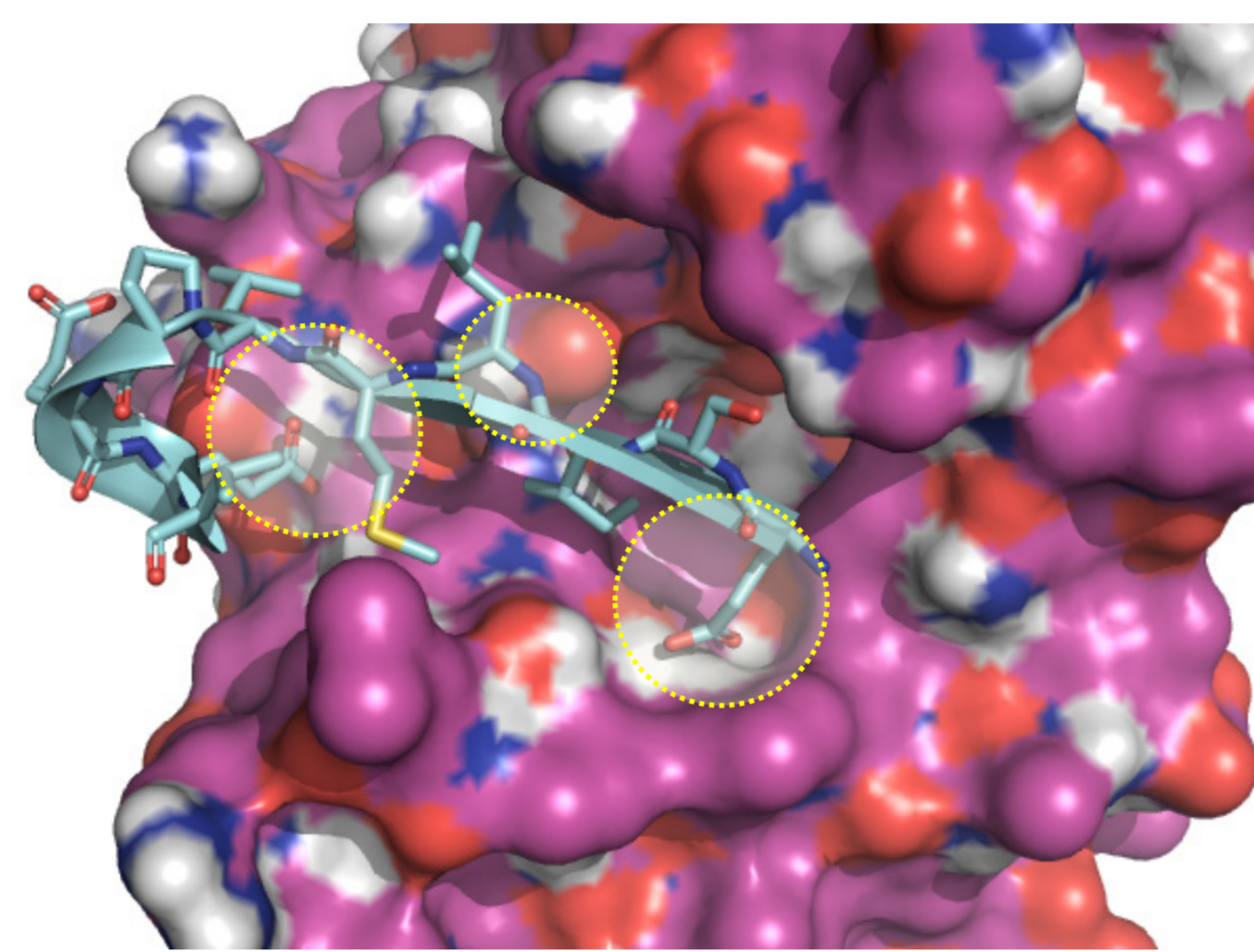Figure 2. PDB: 1UUG, with Uracil (green) dipoles as arrows.

Figure 3. PDB: 1UDI protein DNA mimic.

To illustrate the utility of Spark for the reconnection of distant fragments, we excised the bis-oxime linker from a 1.5µM active (PDB: 3FCI, Figure 4).
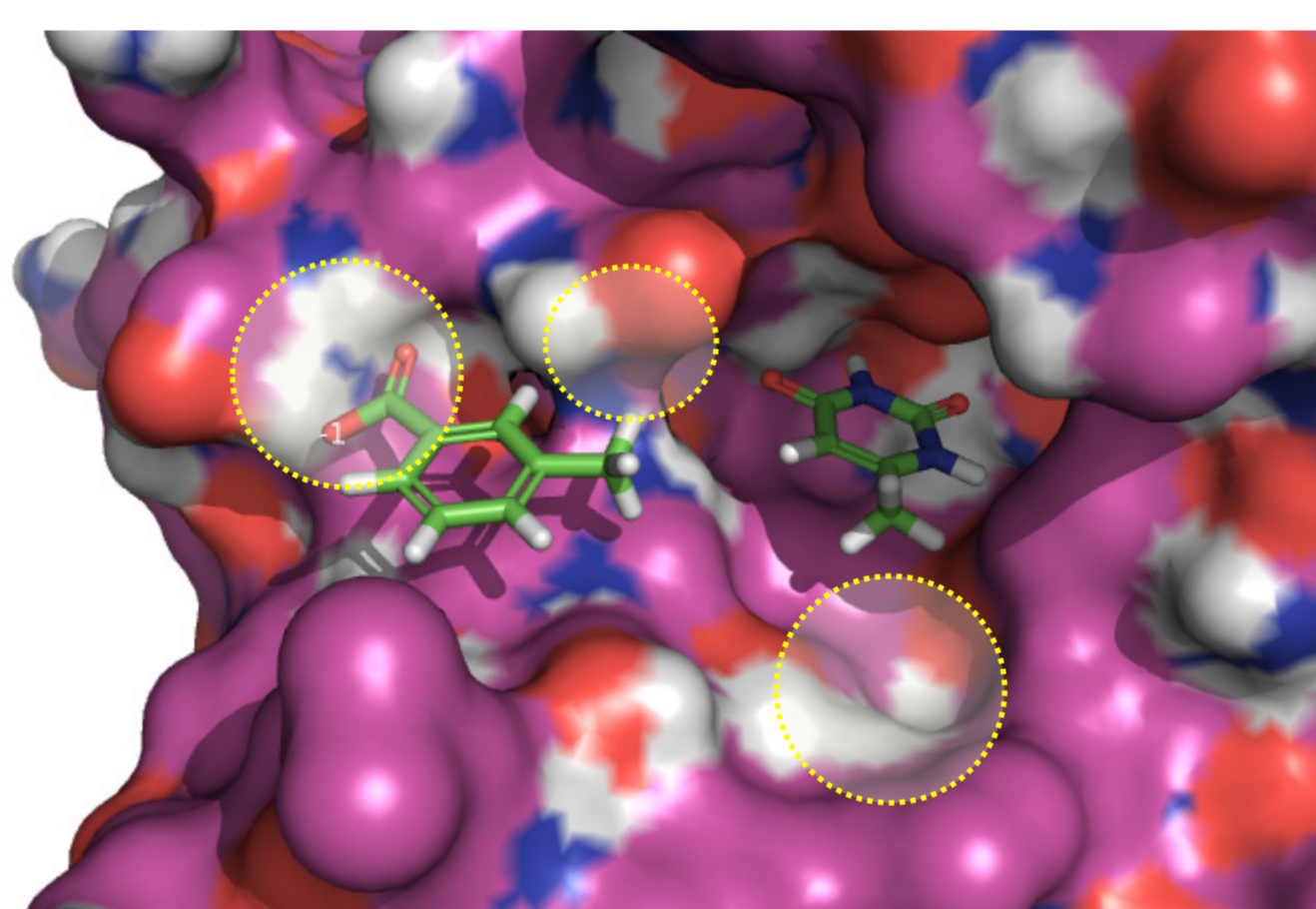

Figure 4. Active fragments originally tethered by a bis-oxime linker (disconnected for this experiment).

In Spark the two fragment atoms to be joined were selected (the methyl groups of m-benzoic acid and uracil) and suitable fragments which were capable of re-joining them were automatically inserted. Fragments were sourced from Spark's large internal databases (derived from Zinc / ChEMBL). The resultant molecules were automatically constructed in-situ, minimized and, in the simplest case, scored against the parent using Cresset's electrostatic- and shape-based 3D similarity metric. Further control was applied here by using other additional molecules (e.g. the protein from Figure 3) to score the output and by applying 'field point constraints' such that important interactions have 'up-weighted' ranking in the output. Additional experiments were also conducted in which one fragment (uracil) was chosen as the starter and grown out towards the benzoic acid fragment. The limit here is the size of fragments in the database and so this works best as a sequential or stepwise workflow.

Results

Example results (Table 1) show the diverse range of output suggestions provided for new linking chemistry for the initial experiments. More importantly, each of the new fragments not only satisfies the geometry and length required for reconnection of the fragments, but also possess features which are consistent with important interactions within the protein.

Additional rounds, involving sequential elaboration of the output, can provide structures which are optimized for both 'protein interactions' and 'physiochemical properties'. Given that computation time is in the order of minutes (not hours) many experiments can be tried e.g. different connection atom substitution (para rather than meta or uracil 5-position).

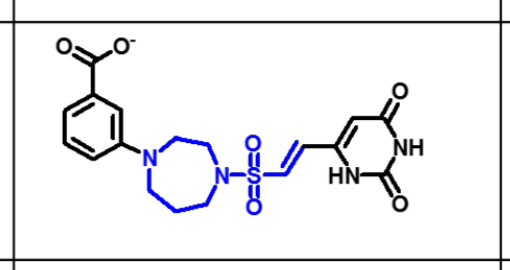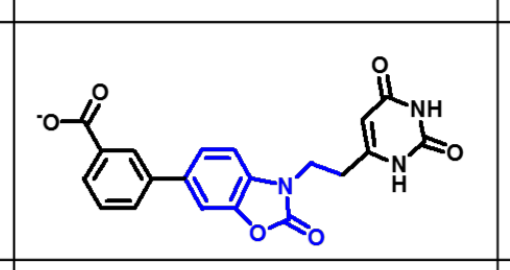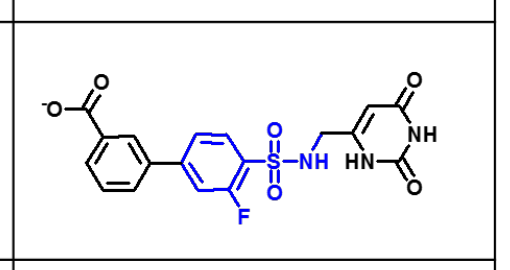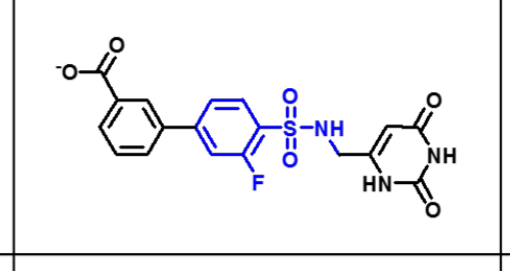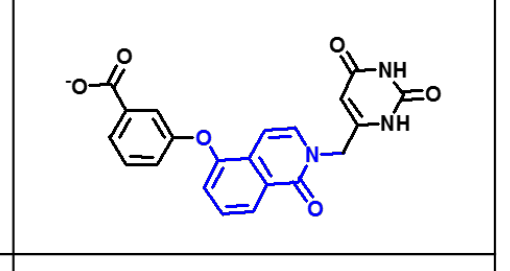Table 1 Example 2D fragment linking structures output by Spark.
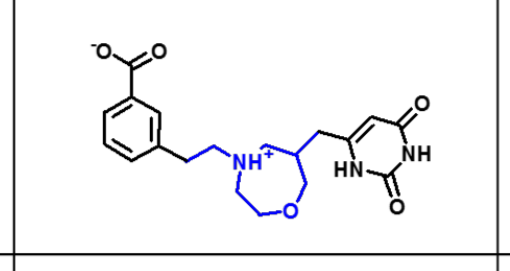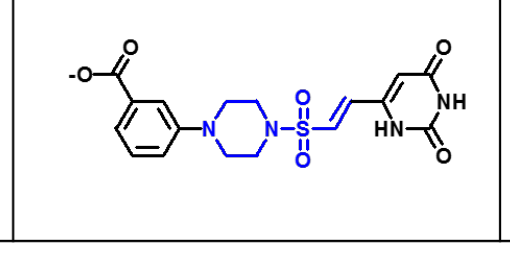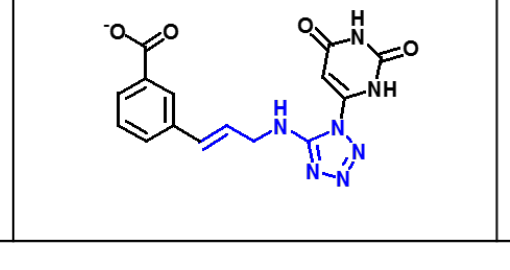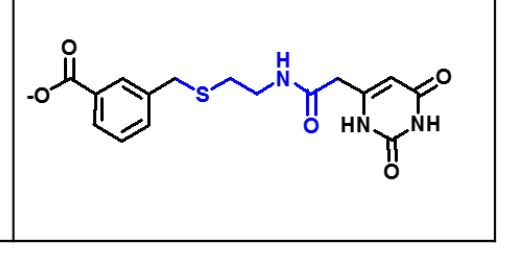
Rank (BIF%)	Structure	Rank (BIF%)	Structure	Rank (BIF%)	Structure
3 (38)		16 (37)		32 (34)	
60 (34)		76 (33)		174 (29)	
196 (28)		203 (28)		860 (22)	

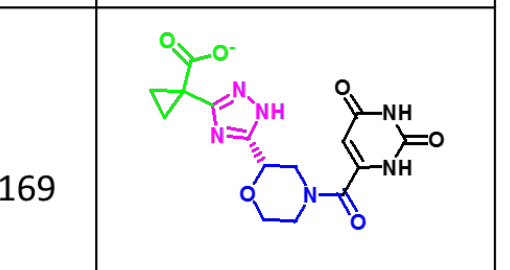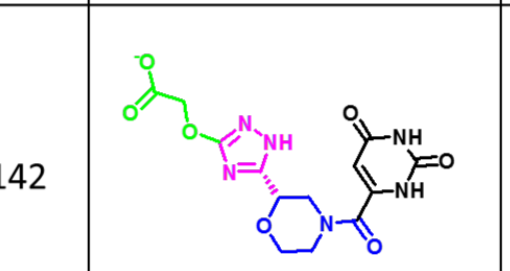Table 2 Example 2D fragment growing structures output by Spark.
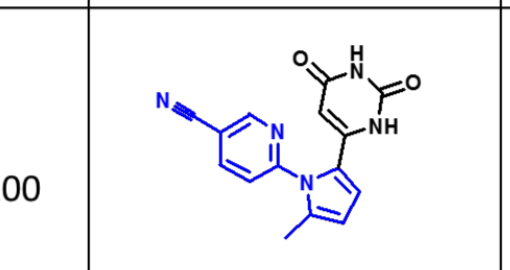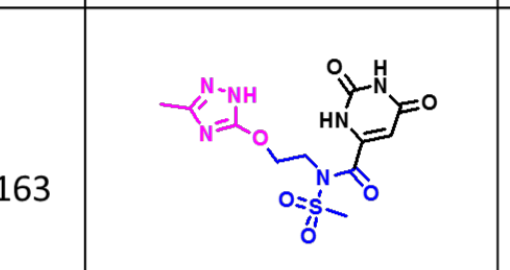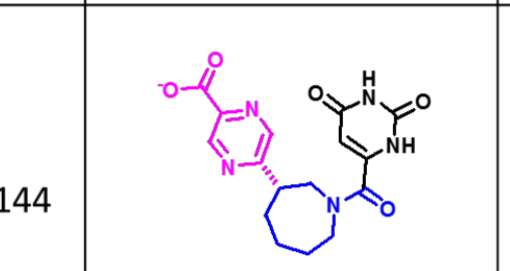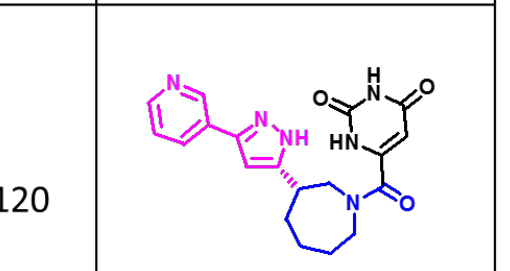
Example round logP / TPSA	Structure	Example round logP / TPSA	Structure	Example round logP / TPSA	Structure
1 R1 1.2 / 100		2 R2 -2.1 / 163		3 R3 -1.5 / 169	
4 R3 -0.8 / 142		5 R4 -0.1 / 144		6 R4 1.2 / 120	

Table 1 (entry 2 'rank 16'), and Table 2 (example 5, R4), are shown below in Figures 5 and 6 in the context of the protein. The results confirm the snug fit of the ligands to the protein and also show how the electrostatic fit of the examples coincide with the positions of the important binding hot spots.
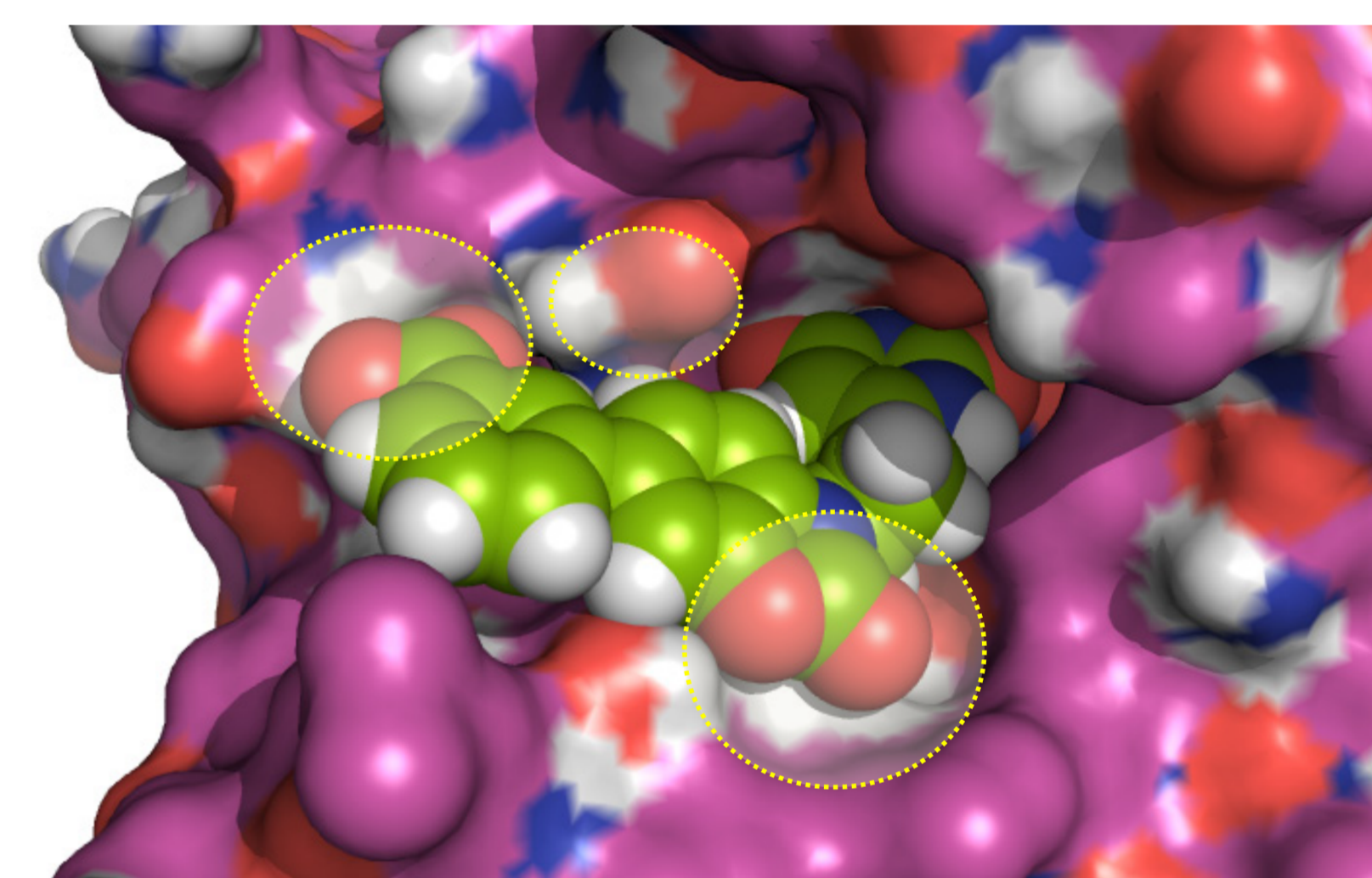

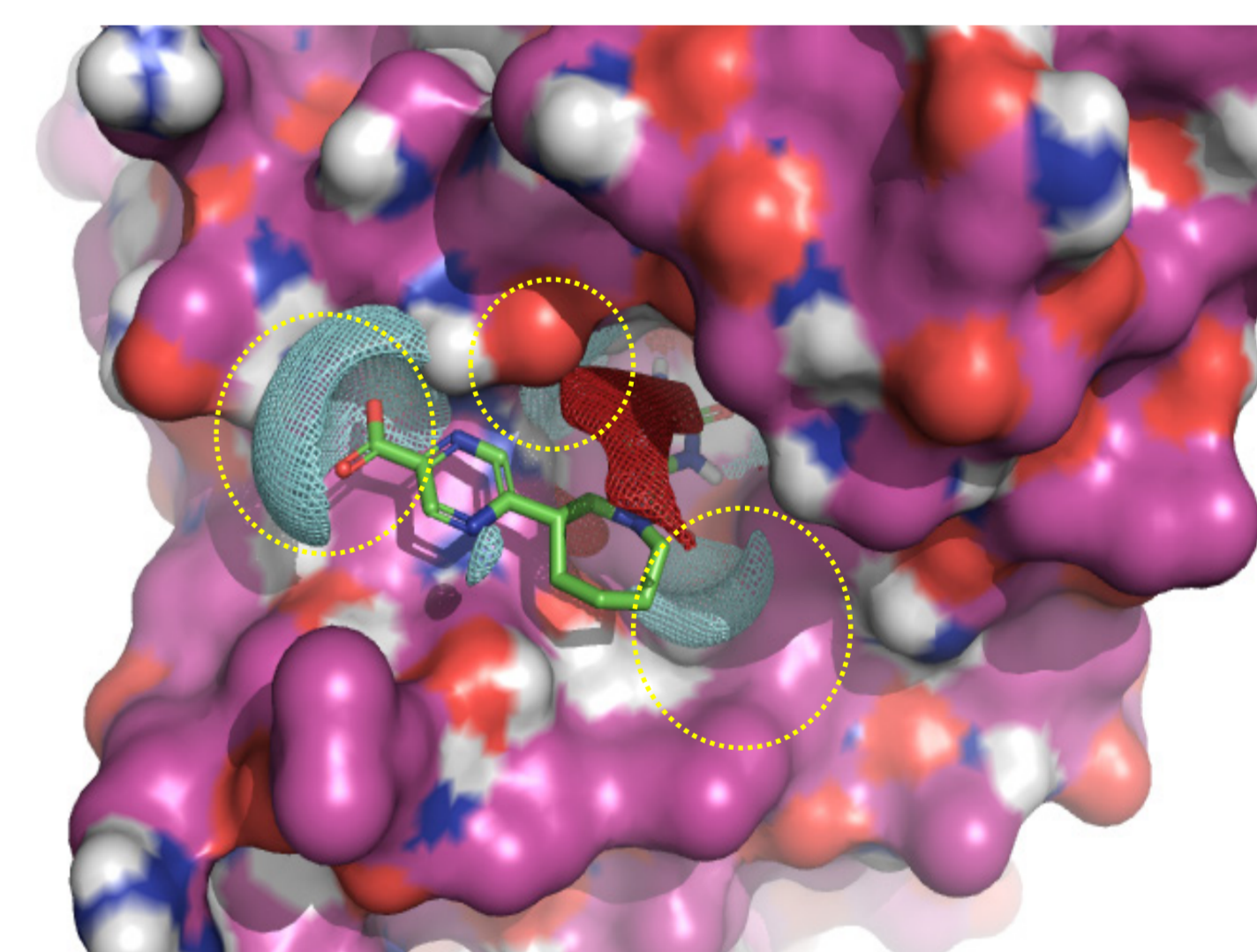Figure 5. Spark output entry 2 'rank 16' with Van der Waals surface.

Figure 6. Spark output example 5, R4 with Cresset's XED force field-derived negative (cyan mesh) and positive (red mesh) electrostatic isopotential surface.

Summary

Spark is a powerful molecular modelling tool for the rapid virtual elaboration of fragment ideas; either in fragment growing, merging or linking experiments.

Applying Spark to fragments bound to the active site of UDG provided a range of interesting and synthetically-feasible suggestions that not only preserve the fragment geometry but also included important protein interactions in the linker.

References

1. Krosky, Nucleic Acids Research, 2006, Vol. 34, No. 20 5872-9, 2006.

Protein pictures were rendered using open source Pymol from Delano Scientific.