



MPhil to PhD Upgrade Report

Drug Discovery: Identification of inhibitors against IL-1-mediated inflammation

Nebangwa Derrick, Neba
Student number: K19072207

Supervisors:
1st: Prof Franca Fraternali
2nd: Dr Gian Felice De Nicola

Randall Division for Cell & Molecular Biophysics; Faculty of Life Science & Medicine; King's College London; New Hunt's House; Guy's Campus; London; SE1 1UL; UK

July 2022

Table of Contents

<i>Abstract</i>	iii
1. <i>Introduction</i>	1
1.1. Interleukin-1 (IL-1) and the inflammatory response in health and disease.....	1
1.2. Regulation and therapeutic implications of IL-1 signaling	3
1.2.1. Therapeutic strategies for targeting IL-1 signaling	4
1.3. Problem statement and preliminary data.....	6
2. <i>Aims</i>	9
2.1. Main objective	9
2.2. Specific Objectives	9
3. <i>Materials and Methods</i>	11
3.1. Study design.....	11
3.2. Acquisition of structures and compound libraries.....	11
3.3. Database and structure preparation, and suitability for docking	11
3.4. Analyses of Protein-Ligand Interaction Fingerprint (PLIF) for generation of pharmacophoric binding restraints.....	12
3.5. Molecular Docking	12
4. <i>Results</i>	14
4.1. Determination of Protein-Ligand Interaction Fingerprints of bound fragments	14
4.2. Screening the FDA approved drugs library identifies potential IL-1 inhibitors	15
4.3. Screening the AI predicted Cyclica Inc. compounds for potential IL-1 inhibitors.....	22
5. <i>Discussion</i>	26
6. <i>Outline of future plans</i>	28
7. <i>Project timetable</i>	31
<i>Acknowledgements</i>	31
<i>Appendix A: Python Script for Pairwise ligand RMSD calculation</i>	35
<i>Supplementary Table 1</i>	36
<i>Supplementary Table 2</i>	38

Abstract

Inflammation is an important mediator of both the innate and adaptive immune responses. However, it can be a “double-edged sword” as dysregulation of IL-1-mediated inflammation, for example, is known to cause several autoimmune and autoinflammatory diseases. A large body of pre-clinical and clinical evidence correlates IL-1 inhibition to decreased disease severity for several tested conditions – highlighting the therapeutic importance of the IL-1 pathway (Dinarello et al., 2012). Methotrexate, Ibuprofen, high-dose aspirin, and prednisone exists as first line anti-inflammatory drugs, however, they are strongly associated with severe adverse effects including bleeding, hypertension, and cancers, amongst others. Protein-based biologics IL-1 blockades are also used to overcome the limitations of first line drugs, but they are also expensive, unstable, and dependent on refrigeration for storage. Small molecules on the other hand, although difficult to design for the large and flat protein-protein interaction surfaces like that of IL1 β – IL1RI complex, do have more benefits over biologics as they are easier to synthesize, cheaper to produce, relatively more stable, and are less dependent on cold chain storage. Unfortunately, to the best of our knowledge, no small molecule anti-IL1 agent currently exist. In this proposal, we aim to identify small anti-IL-1 blockers via virtual screening and biophysical techniques. The Fraternali’s group in collaboration with the De Nicola’s group, solved the crystal structures of a set of fragments with strong affinity for the IL-1 interface. We use the structures as starting point for screening the US FDA approved drugs and an AI (Cyclica Inc) pre-screened datasets against pocket-3 of IL-1. Consequently, 13 FDA drugs and 7 Cyclica Inc. compounds make the hit list as potential IL-1 inhibitors. We now aim to perform NMR spectroscopy and ITC assays to assess binding affinity between the hits and IL-1 target.

1. Introduction

Inflammation is an important mediator of both the innate and adaptive immune responses used by organisms to protect themselves from injury. Inflammation is manifested in tissue and organs through the Cornelius Celsus' four cardinal signs of "redness", "swelling", "heat", and "pain" ("On Medicine, Volume III — Celsus," 1938). Non-steroidal anti-inflammatory drugs (NSIADs) like Methotrexate, Ibuprofen, and high-dose Aspirin are the first line of drugs to treat inflammatory conditions. But they are associated with severe gastrointestinal disturbances and bleeding, especially when administered over long time periods in chronic conditions. Glucocorticoids like prednisone and immunosuppressive therapies are also commonly prescribed for treatments, but they are equally associated with increased risk of opportunistic infections, hypertension, loss of bone and skin integrity, metabolic disturbances and cancers (Dinarello et al., 2012). In inflammatory conditions known to be associated with Interleukin-1 (IL-1) signaling, specifically targeting the signaling pathway can be very effective without causing adverse side-effects (Dinarello et al., 2012; Kaneko et al., 2019). Although several biologics exist as anti-IL-1 blockers, including Anakinra and Canakinumab amongst others, they are quite expensive to produce and are generally. A small molecule drug targeting this signaling pathway could have many medical applications at a reduced cost. As such, my research aims at identifying small molecule inhibitors of the IL-1.

1.1. Interleukin-1 (IL-1) and the inflammatory response in health and disease

Much has been done at the molecular scale to understand the signaling events mediating inflammation, and interleukin-1(IL-1) is implicated as the hallmark regulator of the phenomenon, due to the high homology between the cytoplasmic domains of the IL-1 receptor type I and all toll-like receptors. (Dinarello, 2009; Dinarello et al., 2012; Gery et al., 1972; Kaneko et al., 2019). For IL-1 to be produced, two distinct signaling pathways – the Toll-like receptor (TLR)/IL-1-receptor (IL-RI) and inflammasome pathways – are required. While the TLR/IL-RI signaling pathways induce the production of IL-1 in its inactive pro-form, the inflammasome pathway serves to modulate the activation of caspases which in turn cleave pro-IL-1 (31KDa) to mature secreted IL-1 (17Kda). TLRs expressed on resident monocytes and macrophages are activated to produce IL-1 when infectious molecular patterns (PAMPs) on pathogens or endogenous damage signals (DAMPs) from compromised self-cells are detected (Gabay et al., 2010). Moreso, self-induction when IL-1 binds its receptor (IL-RI) stimulates the production of more IL-1 in more sustained amounts (Dinarello, 2018). Indeed, the TLR and IL-1RI pathways converge in the cytoplasm to form a common downstream signaling cascade which induces the nuclear factor kB (NF- κ B) protein to initiate the transcription and translation of inactive pro-IL-1. (Dinarello, 2018, 1994). However, for IL-1 to exert any effects, it must be matured to its bioactive forms by the second signaling pathway – the inflammasome.

The ‘inflammasome’ is a multimeric protein complex induced by lysosomal stress during phagocytosis (Dinarello, 2018, 2009; Kaneko et al., 2019) and is well known to be the “sensor of injury” – due to infection or damage. However, its main role is the processing of pro-inflammatory caspases which activate pro-form cytokines to their bioactive mature forms (Evavold and Kagan, 2018). When formed the inflammasome cleaves pro-caspase-1 to mature caspase-1 which in turn activates resident cytosolic pre-mature IL-1 to the bioactive forms. The downstream effects of IL-1 are extremely potent and ultimately leads to inflammation. Binding of IL-1 to its targeted cells mainly results to the recruitment of pro-inflammatory immune cells (leukocytes like neutrophils, macrophages, dendritic cells, and T/B cells, etc) to the injured site (Dinarello, 2018, 1992). This is achieved directly and indirectly by inducing effected cells to express more IL-1 and other pro-inflammatory cytokines like IL-6 (Libermann and Baltimore, 1990), chemokines such as CXCL, adhesion molecules like ICAM1, metalloproteinase enzymes and the release of other inflammatory molecules like prostaglandin E2 and NO (Gabay et al., 2010). Indeed, through these networks the inflammatory response is established and tightly regulated, with IL-signalling playing a central role.

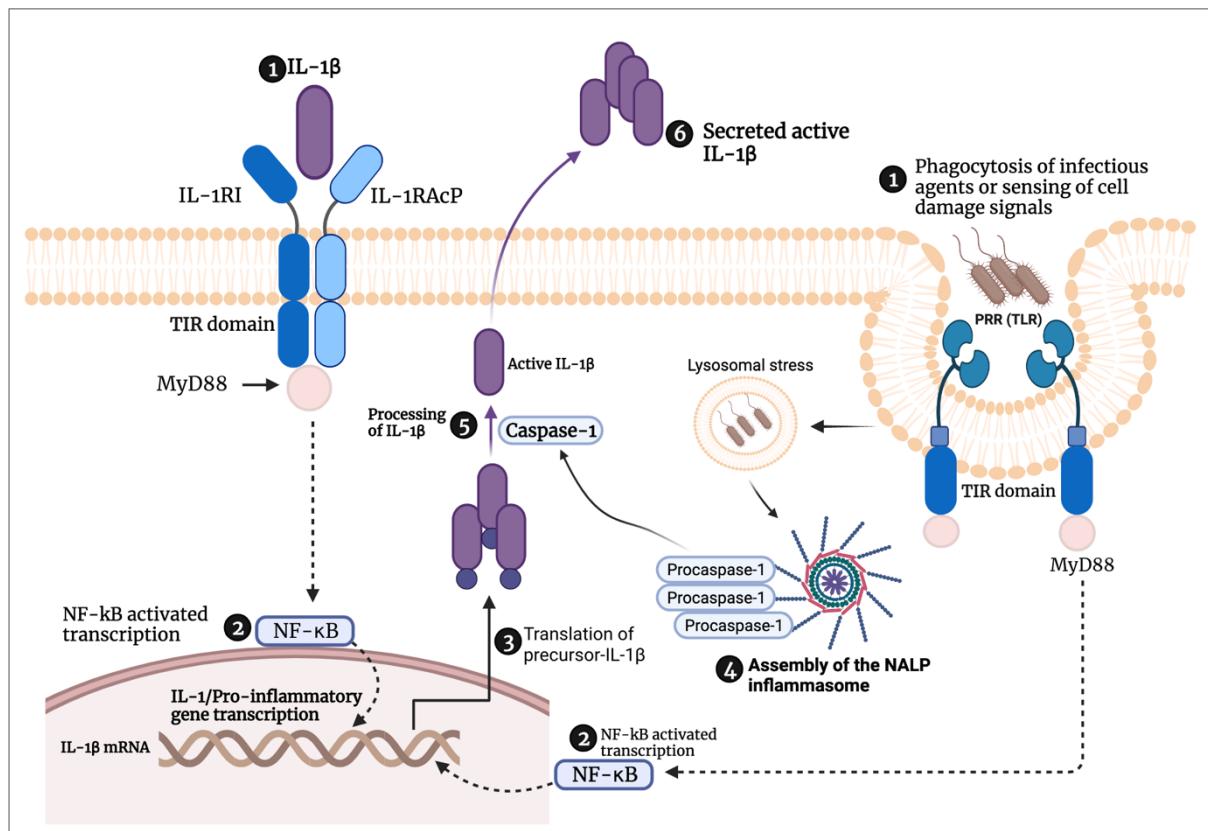


Fig 1: Overview of the steps involved in IL-1 signalling and the role of the inflammasome

Activation by IL-1 β (step 1). Formation of the IL-1 receptor ternary complex (IL1 β -ILRI-ILRAcP) results in activation of the (TIR) domains (dark blue) and recruitment of MyD88 which leads to NF- κ B mediated transcription (step 2), this signalling can occur either via the IL-1RI or the inflammasome pathways indicated with dotted line arrows. Step 3 is the translation of IL-1 β pre-mature protein. Step 4 involves the activation by lysosomal stress due to cell damage or infection and assembly of the inflammasome pathway to produce mature caspase-1. Step 5 illustrates the activation of precursor-IL-1 β by caspase-1 cleavage and subsequent secretion (Step 6) of cleaved and active

IL-1 β into circulation where it induces itself by re-binding its receptor or induces immunity and inflammation by action on diverse cell populations. However, for caspase-1 to be produced and matured, it requires the activation and assembly of the inflammasome (Step 4) via the inflammasome pathway which begins with sensing of PAMPs from infectious agents or DAMPs released by self-cells upon damage. The assembled inflammasome initiate the processing of pro-caspase-1, resulting in the formation of the active caspase-1.

However, inflammation can be a “double-edged sword” – able to rapidly offer protection but sometimes, this comes with extensive collateral damage to healthy cells. A large body of evidence implicates dysregulation of IL-1 signalling to autoimmune, autoinflammatory and chronic diseases such as joint and muscle diseases, neurodegenerative diseases, stroke, diabetes type 2, multiple sclerosis, cancers, and epilepsy, amongst an extensive list of clinical conditions to which inflammation plays a causative role (Dinarello et al., 2012; Mendiola and Cardona, 2018; Shaftel et al., 2008). The mechanisms by which IL-1-mediated-inflammation induces the different diseases are quite distinct and not well understood for all the disease cases. However animal, and clinical studies of anti-inflammatory agents targeting IL-1 signalling provides a wealth of evidence associating IL-1 to these conditions, where inhibiting IL-1 correlates with decreased disease severity for tested diseases (Dinarello et al., 2012). This highlights the importance of IL-1 signalling pathway for therapeutic applications.

1.2. Regulation and therapeutic implications of IL-1 signaling

Since its mouse (Lomedico et al., 1984) and human (Auron et al., 1984) DNA sequence were derived, IL-1 research has rapidly continued to expand. Based on shared sequence and structural features, 11 members of the IL-1 superfamily were classified and include IL-1 α , IL-1 β , IL-1Ra, IL-18, IL-33, IL-36 α , IL-36 β , IL-36 γ , IL-36Ra IL-37, and IL-38 (Rivers-Auty et al., 2018). However, only IL-1 α , IL-1 β , and IL-1 receptor antagonist (IL-1Ra) can engage the major IL-1 type I receptor (IL-1RI) signaling pathway (Dinarello, 2018). While IL-1 α and IL-1 β induce this receptor, binding of IL-1Ra does not transmits downstream signals. Unlike other cytokine families, the biology of IL-1 is quite unique in many aspects. At the receptor level, signaling starts when either IL-1 α or IL-1 β binds the IL-1 type I receptor. The affinity of IL-1 for this receptor is known to be amongst the most potent (picomolar scale) in the cytokine families, and this is thought to be a reason why the IL-1 signaling is the only cytokine pathway for which there exist a naturally occurring inhibitor, IL-1Ra antagonist. The affinity of IL-1Ra is even stronger and able to competitively displace both IL-1 α and IL-1 β from the receptor. A second IL-1 decoy receptor (IL-1R2) which does not signal when IL-1 binds due to lack of a cytoplasmic domain, also exists as an evolutionary “sink” for regulating IL-1 levels (Kaneko et al., 2019). When either, IL-1 α or IL-1 β binds the IL-1RI receptor recruits an accessory protein termed the IL-1 receptor accessory protein (IL-1RAcP or IL-1R3) to form a ternary complex which signals downstream cascade. However, this complex is not formed upon bidding the natural antagonist IL-1Ra. As with PRR (toll-like receptor) signaling, IL-1 receptor signals via the cytoplasmic Toll/IL-1 receptor (TIR) which recruits the myeloid differentiation primary

response protein 88 (MYD88) in the cytoplasm. MYD88 in turn triggers the TIR downstream cascade which ends in the nucleus with the NF- κ B-mediated transcription of pro-inflammatory cytokines including IL-1.

Tissue expression of IL-1 α and IL-1 β is roughly similar and has been detected in lymphoid organs like the bone marrow, thymus, lymph nodes, and spleen; and non-lymphoid organs including the liver, gut, and lungs, to name a few (Kaneko et al., 2019) while cellular expression mostly involves monocytes, macrophages, and dendritic cells. The pathological expression patterns largely reflect those in cell and tissue described above, and so supports IL-1 involvement in several diseases and provides a strong premise for its therapeutic applications as a drug target. While IL-1 α is rarely detected in peripheral blood and thought to be mostly localised in the cytosol when produced (Dinarello et al., 2012; Gabay et al., 2010), IL-1 β on the other hand is secreted in its mature form into circulation where it acts on effector cells both locally and systemically. As such, IL-1 β has pulled significant attention as an anti-inflammatory target of choice over IL-1 α . A decade of animal breeding research involving IL-1 β -deficient mice showed no signs of disease, although when challenged with oil these mice could not manifest acute phase inflammation within 24 hours, and had no circulating IL-6 (Dinarello, 2018) compared to the wildtype. However, IL-1 β -deficient mice had elevated responses to LPS, IL-1 β or IL-1 α compared to the wildtype (Alheim et al., 1997). Another study of IL-1 receptor 1 knockout mice reported that IL-1 signalling induced many processes of significant benefits to the development of collagen synthesis during atherosclerosis (Alexander et al., 2012). Taken together, these data from IL-1-knockout animals not only validates the target, but also confirms the feasibility and safety of inhibition for therapeutic benefits. However, given its important role in inflammation and innate immunity, therapeutic strategies must find balance between inhibiting the detrimental effects of IL-1 signalling while also retaining and promote its useful ones.

1.2.1. Therapeutic strategies for targeting IL-1 signaling

The biological and pharmacological importance of IL-1 signalling cannot be overemphasised. Currently, the IL-1 family has two agonists (IL-1 α or IL-1 β) and one antagonist (IL-1Ra) which all bind the same surfaces on the receptor (IL-1RI) (Fields et al., 2019). The binding of IL-1 α or IL-1 β is the first step in IL-1 signal transduction. Therefore, blocking this interaction is a crucial target for development of new anti-inflammatory drugs (Kaneko et al., 2019; Dinarello, 2018). Several therapeutic inhibitors targeting this signaling are currently in clinical use (Fig 2) and are mainly biologics (Dinarello et al., 2012; Kaneko et al., 2019). These recombinant proteins can broadly be classified as either IL-1 receptor blockers, IL-1-antibody neutralisers, and other strategies including therapeutic vaccines, amongst others. (Kaneko et al., 2019; Dinarello et al., 2012)(Fig. 2). One successful example of the receptor blockers is Anakinra. Anakinra was the first IL-1 receptor blocker to be approved for clinical use by the US Food and Drug Administration (FDA) to treat

Rheumatoid arthritis (RA). As a recombinant form of the naturally occurring IL-1 receptor antagonist, anakinra can competitively inhibit binding of either IL-1 α or IL-1 β . Since its approval in 2001, the drug has been used to treat a broad spectrum of diseases associated with inflammation, and so has played a central role to prove the link between IL-1-mediated inflammation and these diseases (Dinarello et al., 2012). In five different efficacy and safety trials against arthritis in humans, significant improvements with anakinra (50–150 mg) dosing versus the placebo was achieved. In addition, a remarkable safety profile was established as the treatments were performed daily with high doses for up to six months with no significantly serious adverse effects recorded. (Mertens and Singh, 2009). A long history of safety with anakinra has been observed even at blood levels from intravenous administration 100-folds higher than when cutaneous dosing was performed, (Dinarello et al., 2012). These data validate the rational that therapeutically inhibiting IL-signaling can be achieved without necessarily sacrificing the benefits of this pathway in the innate and adaptive immunity.

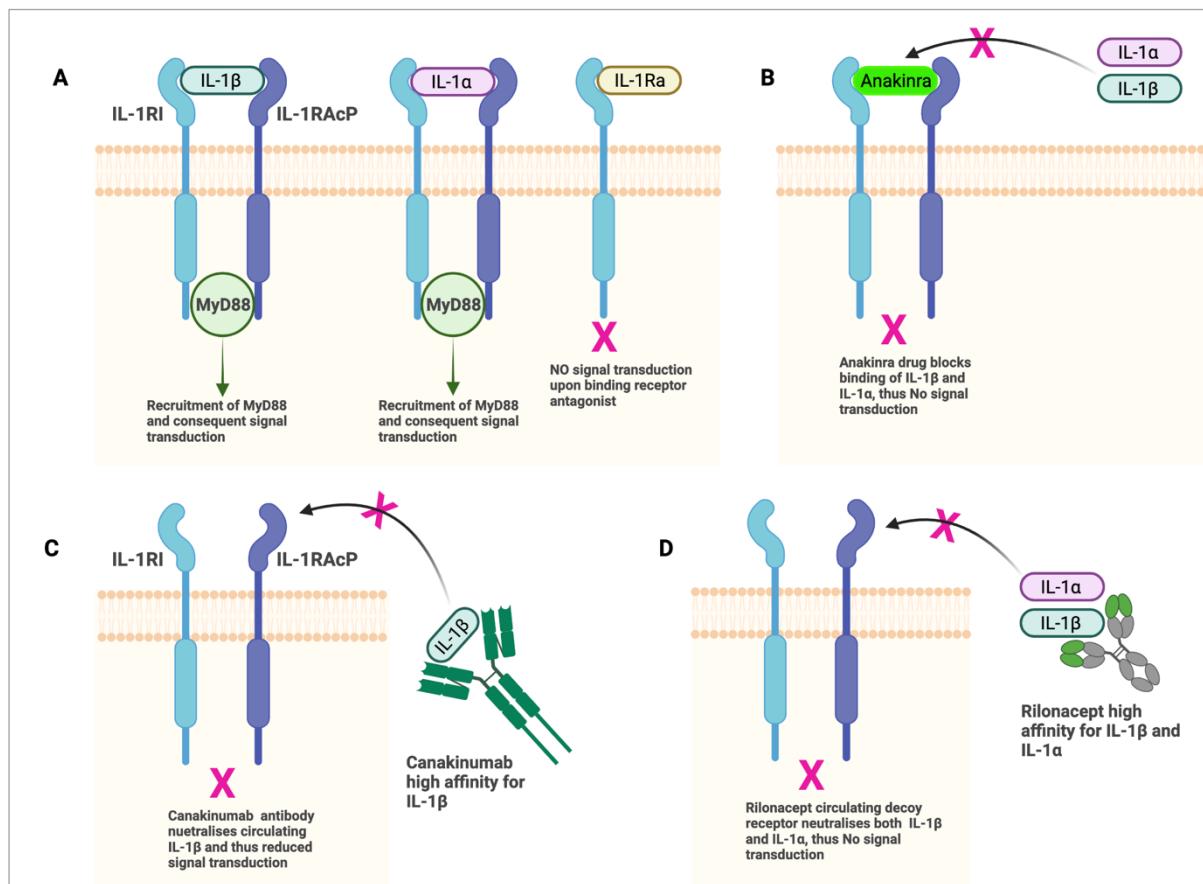


Fig. 2: Strategies by which current clinically approved drugs target IL-1 signaling. A) IL-1 type I receptor (IL-1RI) interacts with both IL-1 α and IL-1 β and together with the co-receptor accessory protein IL-1R3(IL-RAcP). IL-1Ra the naturally occurring receptor antagonist can bind to and inhibit IL-1R1 signaling but not IL-1R3. IL-1R2, the decoy receptor. B) Anakinra is the recombinant form of the human receptor antagonist (IL-1Ra) that can bind IL-1R1 and inhibit both IL-1 α and IL-1 β . C) Canakinumab is a monoclonal antibody able to neutralise

IL-1 β with high affinity and specificity. D) Rilonacept is a recombinant protein of human IL-1R1 and IL-1R3 (IL-1RAcP) combined with the Fc portion of IgG1, which neutralizes both IL-1 α or IL-1.

One disadvantage of receptor blockers like anakinra is that very frequent (daily) high amounts of dosing is required for extended periods (months) to achieve clinical benefits. As such, monoclonal antibodies able to specifically neutralized IL-1 α or IL-1 β has gained more research traction over the years. This is due to the fact that very low concentrations (in pg per ml range) of IL-1 are produced and circulate in blood during diseases states – only five folds higher than levels circulating in healthy individuals (Dinarello et al., 2012) Therefore, only minute doses of antibodies would be required to titrate against the low blood levels of IL-1, compared to anakinra which requires high doses to titrate against the large concentrations of IL-1 receptor expressed on almost every cell surface. Research efforts to develop anti-IL-1 neutralisers have focused more on anti-IL-1 β compared to anti-IL-1 α . One major reason for this is observed in studies in humans in which blood monocytes released more IL-1 β from patients with various autoinflammatory diseases than healthy individuals, when compared to IL-1 α in the same samples (Goldbach-Mansky et al., 2006; Pascual et al., 2005; Dinarello et al., 2012). Owing to this and other reasons, Canakinumab, an anti-IL-1 β monoclonal antibody was developed by Novartis and approved in 2009 by the FDA for treating arthritis. Since then, the drug has been in several human trials for the treatment of a broad spectrum of other diseases. The CANTOS trial, the largest trial of any anti-cytokine drug, examined 17,200 individuals and established that neutralising IL-1 β with canakinumab reduced cardiovascular events in patients with atherosclerotic disease (Ridker et al., 2017). Moreover, interesting secondary outcome observations in the same study showed significant reduction in lung cancer mortality, providing the first support for the hypotheses that IL-1 β inhibition has a beneficial effect on cancer incidence and mortality. Currently, at least 60 distinct diseases (Dinarello et al., 2012) are responsive to anakinra and canakinumab, with many more undergoing clinical trials. Several other anti-IL-1 β inhibitors exist including Rilonacept which is a recombinant protein of human IL-1R1 and IL-1R3 (IL-1RAcP) combined with the Fc portion of IgG1, that neutralises both IL-1 α or IL-1.

1.3. Problem statement and preliminary data

Despite remarkable advances in developing blockades of IL-1 signaling, successes so far have been on the use of biologics. While these therapeutic agents are quite specific, they do face several challenges which include, but are not limited to, high costs of production, complete dependence on cold chain refrigeration for storage, and instability of the molecular make-up. Small molecule ligands on the other hand, although difficult to design for large PPI surfaces, do have more benefits over biologics because they are easier to synthesize, cheaper to produce, relatively more stable, and are less dependent on cold chain for storage. Nonetheless, and to the best of our knowledge, no small molecule agent is available for the specific inhibition of IL-1R signaling. In this work, we use both computational and biophysical methods to identify small molecule inhibitors of IL-1

signaling. In this research, we focus on IL-1 β inhibitors due to several advantages over IL-1 α mentioned in section 1.2.1 above. The IL1 β – IL1RI complex is a protein-protein interaction (PPI) complex with larger and shallower pockets and grooves than the deep and narrow binding pockets of classical enzyme-type drug targets. It is hence, challenging to design inhibitors against the former. However, recent advances imply that with the right tools, certain classes of PPI can yield to the efforts of medicinal chemists to develop small molecule inhibitors (Scott et al., 2016). In fact, the first PPI inhibitors have reached clinical development.

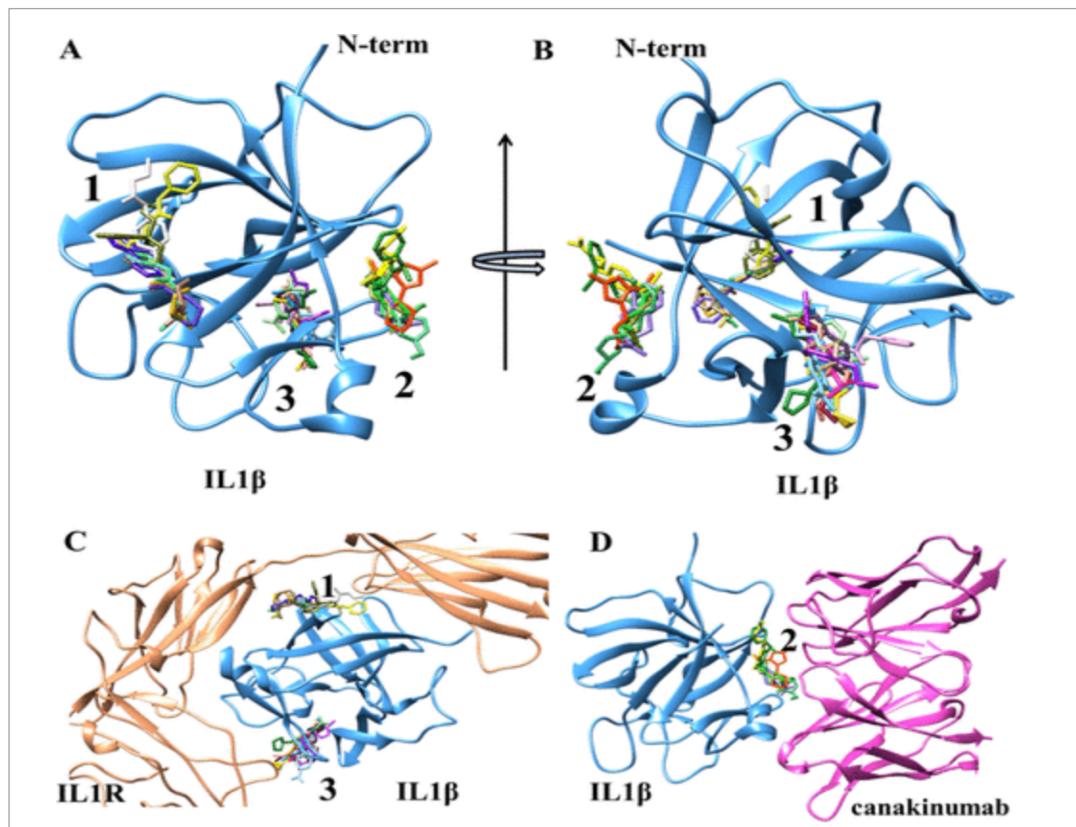


Figure 3: Fragment screening reveals 3 binding hotspots on IL1 β (from Nichols et al., 2020)

A) Shows an overlay of the fragments in stick representation bound to IL1 β in blue ribbon representation, and B) rotated by 90°. The fragments cluster in three sites, labeled 1, 2, and 3. C) Illustration of superimposed fragments in stick representation bound to IL1 β superimposed to the IL1 β -receptor in bronze ribbon representation D) and the IL1- β -canakinumab antibody complex in magenta ribbon representation. E) A zoom-in view of binding pocket-3 showing superimposed fragments.

The Fraternali's group in collaboration with Dr. De Nicola's group used bioinformatics analyses and X-ray-based fragment screening (FBDD) to identify small molecule fragments which bind interfaces localized at the interface between IL-1 β and its receptor (IL-1RI) (Nichols et al., 2020). This approach proves effective and highly sensitive at identifying even weak binders and its main strength is probing key binding hotspots on protein surfaces with large flat pockets. The IL-1 receptor has three immunoglobulin-like (Ig-like) ectodomains (D1, D2 and D3) at 2.5A resolution that form two binding sites D1/D2 and D3, which together drive interaction with IL-1 cytokine

(Vigers et al., 1997). These three ectodomains adopt a 3D-architecture that resembles a “grasping hand” or “question mark” sign which wraps around IL-1 β , upon binding. IL-1 β on the other hand, has two binding regions (sites A and B) as identified by site directed mutagenesis (Labriola-Tompkins et al., 1991; Evans et al., 1995) which respectively bind the two pockets of the receptor, as in receptorD1/D2 – siteA-IL1 β and receptorD3 – siteB-IL1 β . The outcome of the Fraternali and De Nicola’s study provided solved X-ray crystallographic structures at ~ 1.46 Å resolution, of chemical scaffolds clustered on three surface hotspots on IL-1 β (Nichols et al., 2020). Interestingly, two of these hotspots turn out to overlap with the receptor–IL1 β binding domains (receptorD1/D2–siteA-IL1 β and receptorD3–siteB-IL1 β) with the third hotspot overlapping with the pocket of the clinical antibody, canakinumab (Fig 3).

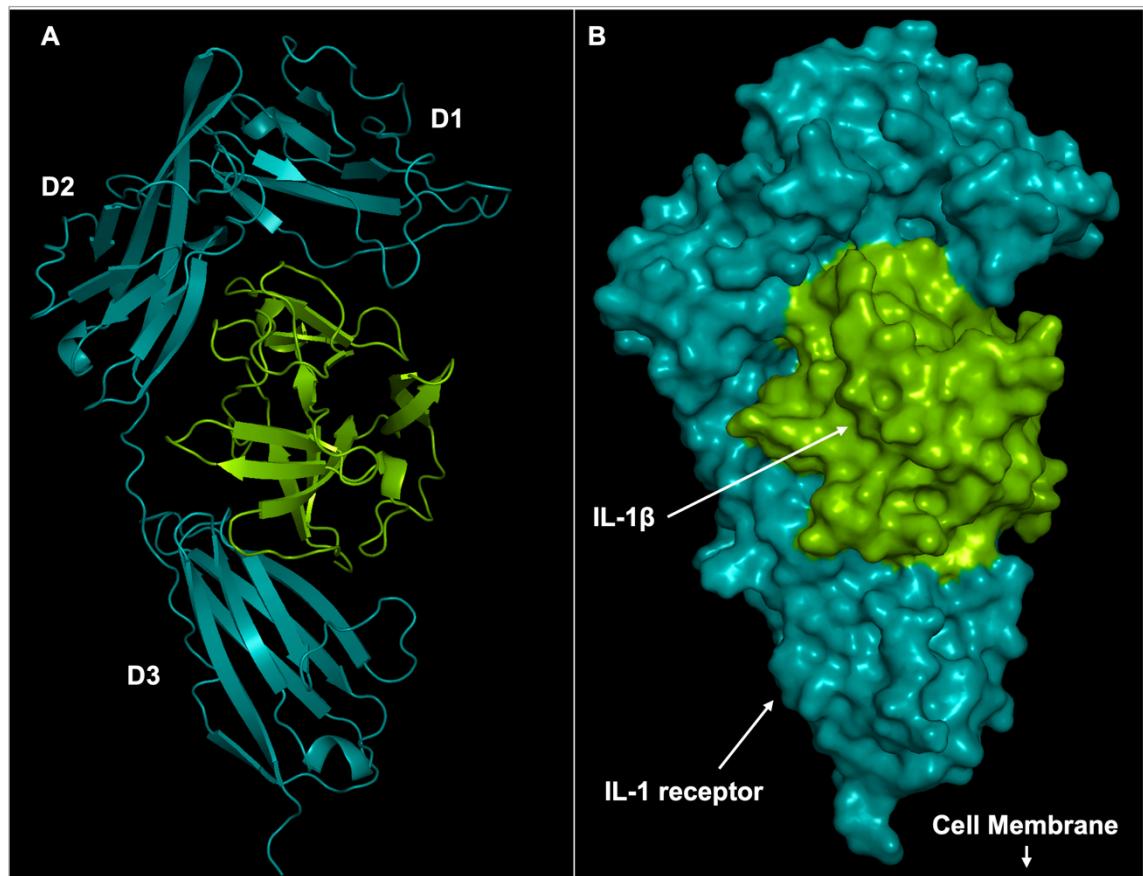


Fig 4: Crystal structure of IL1 β complexed to the receptor IL-1RI (PDB 1ITB). A) Ribbon representation. B) Surface representation. In both panels, IL1 β is colored green while the IL-1RI is colored cyan. The three immunoglobulin-like (Ig-like) ectodomains (D1, D2 and D3) which together drive interaction with ligands adopt a 3D-architecture that resembles a “grasping hand” or “question mark” sign which wraps around IL-1 β , upon binding.

The analysis of interactions made by the fragments bound to receptorD3–siteA-IL1 β (pocket I in De Nicola et al.,) shows that a hydrogen bond acceptor positioned by most of the ligands make interaction with the amino group of Met 148 of the D3-receptor pocket, and a hydrogen bond donor

to interacting with the carboxyl group of Met 148. Also, the hydrophobic grove defined by Leu110, Phe146, and Met 148 of the receptorD3–siteA-IL1 β pocket shows some flexibility in accommodating a five- or six-member ring. Further analyses of the scaffolds clustered on the receptorD1/D2–siteA-IL1 β pocket (pocket III in De Nicola et al.,) shows a hydrophobic patch defined by residues Tyr24, Leu26,80,82, Val132, and Phe133. This site can accommodate a trifluoro-methyl group bound to an aromatic or a saturated six membered ring. Hydrogen bond acceptors on most of the ligands clustered here make contacts with acceptors of the amino groups of Leu 26, and hydrogen bond donors of the ligands engage the carboxyl group of Leu 26, and Tyr 24. Upon binding, some of the ligands cause significant conformational changes to domains on IL-1 β which could result in disrupting the interaction with the receptor. My research, therefore, exploits in fine detail the mechanistic features encoded in these data for pharmacophore-guided searching of available compound libraries, in the design of small molecule inhibitors of signaling between IL-1 β , and its receptor, IL-1RI. We particularly screen the Food and Drug Administration's (FDA) database of clinically approved drugs (1615 drugs) to identify potential IL-1 β binders. In addition, Artificial Intelligence (AI) was used by a data-driven drug discovery company, Cyclica Inc., to search their larger database (~2 million compounds) for binders of IL-1 β and 223 shortlisted hits were sent to us for further screening, and we applied our pharmacophore restraints to identify a few of the compounds for experimental testing.

2. Aims

2.1. Main objective

To the best of our knowledge, no small molecule drug exists as inhibitor of IL1R receptor signaling. We therefore aim herein, to identify and validate a small molecule ligand inhibitor which binds with high affinity at the interface used by IL1 β to signal its receptor, IL1R. It is our goal that identified inhibitors would demonstrate the potential to distort the downstream signaling cascade that results in transcription, translation, and secretion of pro-inflammatory molecules initiated when IL-1 β binds this receptor. To achieve this, we used *in silico* screening to predict potential binders to be validated in a second phase of the project by wetlab-based biophysical techniques including Isothermal titration calorimetry (ITC) and Nuclear Magnetic Resonance Spectroscopy (NMR). In this report, I summarise the first part of the proposal which involves computational screening of virtual libraries against our protein target. We use molecular docking to predict potential binders.

2.2. Specific Objectives

The specific aims of this proposal are:

1. To search the FDA datasets for clinically approved drugs which bind IL-1 β with high affinity.
2. To prioritise potential inhibitors from a set of 223 compounds pre-selected as IL-1 β binders using Artificial Intelligence by a data-driven drug discovery company, Cyclica Inc.
3. To biophysically characterize lead compounds via Nuclear Magnetic Resonance (NMR) Spectroscopy and Isothermal Titration Calorimetry (ITC).

3. Materials and Methods

3.1. Study design

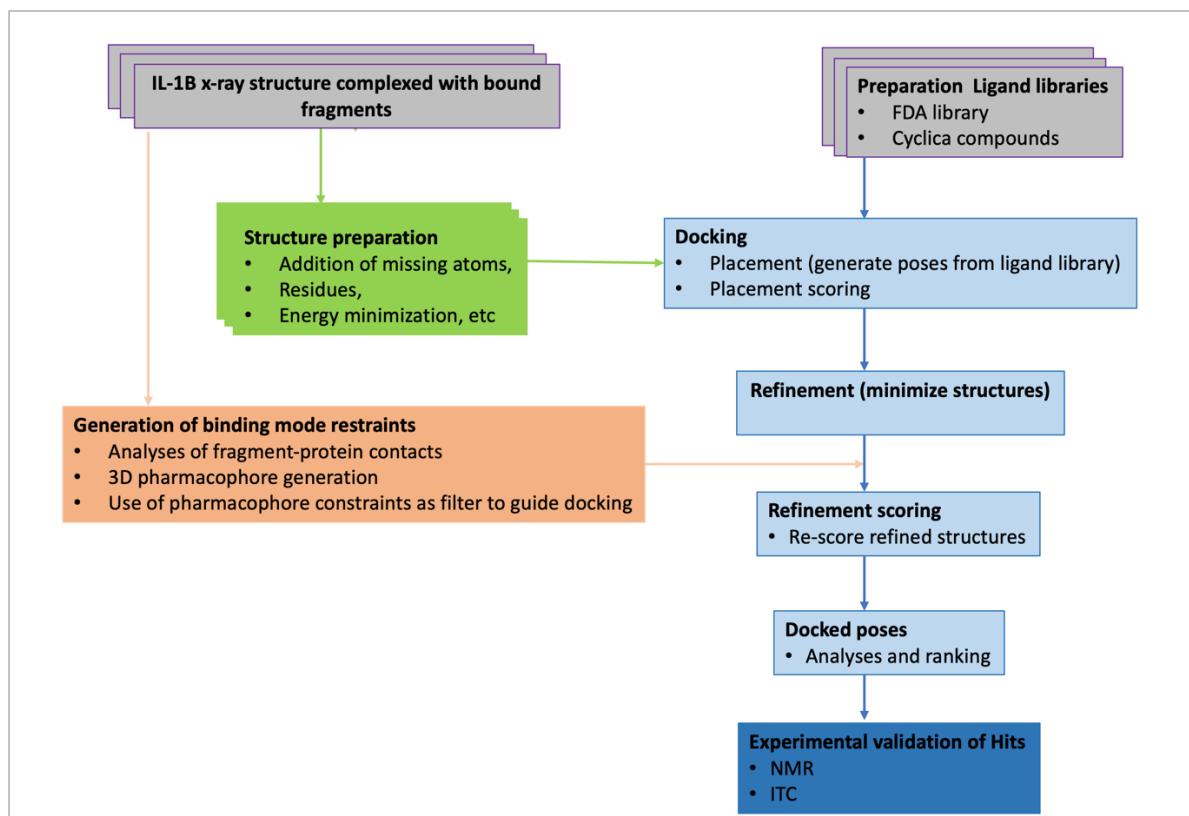


Fig 4: General workflow outlining the steps involved in this proposal. Coloured in grey and green are boxes explaining protein and ligand database's structure preparations and input files; and in light blue are the steps used by the docking algorithm to score affinities between each small molecule compound and the binding site of the target; while the dark blue box highlights the level at which experimental binding assays will be conducted.

3.2. Acquisition of structures and compound libraries

The X-ray crystallographic structures of IL-1 β with bound co-crystallised fragments were retrieved from the [RCSB PDB database](#) (Berman et al., 2000) with an average resolution of ~1.47 Å. The clinically approved drugs were retrieved on the 30/11/2021 from the ZINC database (<https://zinc.docking.org/substances/home/>). While the AI screened compounds were provided by the data-driven drug discovery company Cyclica Inc.

3.3. Database and structure preparation, and suitability for docking

Studies to identify and prioritise approved drugs and compounds as inhibitors of IL-1 β was done using molecular docking. Every docking campaign begins with a suitable high-resolution target with well identified and labelled binding site and a clean compound database containing, in this case, small molecules. In this study, I use the X-ray crystal structure of IL-1 β (resolution 1.4 Å) and two compound datasets including 1615 FDA drugs and 223 Cyclica compounds. It is common

for PDB structures, even when solved at high resolutions, to still contain some structural flaws. To address these, I implemented the “structure preparation” application of the MOE software (*Molecular Operating Environment (MOE)*, 2021) assesses and corrects any errors on crystal structures including missing hydrogens or residues, missing side chains and loops, to name a few. In addition, the “wash” application of the MOE suit was also used to correct physico-chemically related errors of small molecule drugs and compound databases, by re-arranging properties like protonation states. The resulting clean target and compound libraries were converted to multiple different file formats including .sdf and. smi for use as inputs.

3.4. Analyses of Protein-Ligand Interaction Fingerprint (PLIF) for generation of pharmacophoric binding restraints

Although protein structures can be modelled and used for virtual screening, the most suitable starting point, however, is a high-resolution experimentally solved ligand-bound structure (Bender et al., 2021). Structures with bound ligands usually outperform ligand-free structures as the 3-D conformations of binding pockets are better defined in the bound state structures, than in the unbound ones (Rueda et al., 2010). I started by analysing 12 X-ray structures previously solved in our labs(Nichols et al., 2020). From these, I extracted the protein-ligand interaction fingerprints of the 12 fragments whose binding modes to pocket-3 of the IL-1 target are captured and well defined at resolution ~1.4 Å. I then generate a binding constraint file which encodes the contact types and angles made by each fragment with the receptor’s binding pocket-3. This information was applied as a filter to screen placement poses during the docking process.

3.5. Molecular Docking

Docking calculations were run using protein and ligand structures that were already prepared as inputs and was guided by the 3D pharmacophore restraints previously generated as described in section 2.4 above. To begin with however, we assessed and validated the accuracy of our scoring function known as the *generalized-born volume integral/weighted surface area (GBVI/WSA) scoring function* (Labute, 2008). This was accomplished with one of the structures (PDB 5R8D) selected for their relative stability compared with other fragment-bound structures. Before the simulations, we first re-docked the co-crystallised fragment towards pocket-3 of IL-1. The scoring function was said to be correct if the deviation between the original and re-docked poses of co-crystallised fragment was less than 1 Å. Indeed, all the docking workflows used in this study produced RMSD values of less than 0.5 Å, at the top-scoring energy positions. In addition, after manual inspection, we observed that the best re-docked fragment poses were well fitted to the original co-crystallised binding modes, indicating that our scoring function and docking parameters, to a large extend can discriminate correct from wrong binding orientations, as well as strong from weak binders of pocket-3 of IL-1. The *(GBVI/WSA) scoring function* estimates the enthalpic contribution to the free energy of binding

of a given ligand pose using the AMBER99 forcefield (Cornell et al., 1995), trained on 99 different experimentally solved protein-ligand complexes. Structural analyses and visualisation were performed with the PyMOL Molecular Graphics System (*PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC, n.d.*) and statistical analyses was done with Microsoft ™ Excel (version 16.56) and Python.

4. Results

4.1. Determination of Protein-Ligand Interaction Fingerprints of bound fragments

Analyses of the interaction signatures of 12 fragments independently bound to pocket-3 of the X-ray structures of IL-1 reveal important interactions made between pocket residues and fragment scaffolds. The key residues which contact the different fragments include Tyr 24, Glu 25, Leu 26, Lys 74, Leu 80, Pro 131, Val 132, and surrounding water molecules (fig 5A). Analyses of interacting functional groups of the bound fragments show that the amide, amine, and ketone functional groups account for more than 60% of all interactions made with pocket residues (fig 5A). To address which pocket residues are important for drug design and screening campaigns, I constructed a contact distribution plot (fig 5B) and noted that while Tyr 24, Glu 25, and Leu 26 make more than 80% of interactions with the different functional groups, Leu 26 alone reacts with 8 out of the 12 experimental fragments. Details of interaction fingerprints between the fragments and pocket residues are elaborated in fig 5C and table 1 below, and indicate a range of interaction types including polar, ionic, and arene attractions. These data are used later for virtual screening and are encoded in a 3D-pharmacophore (Wermuth et al., 1998) file (PH4 file) applied as a docking filter by which compounds undergoing screening are scored at the binding site of the target, and those meeting conditions of the pharmacophore restraints are relatively highly scored.

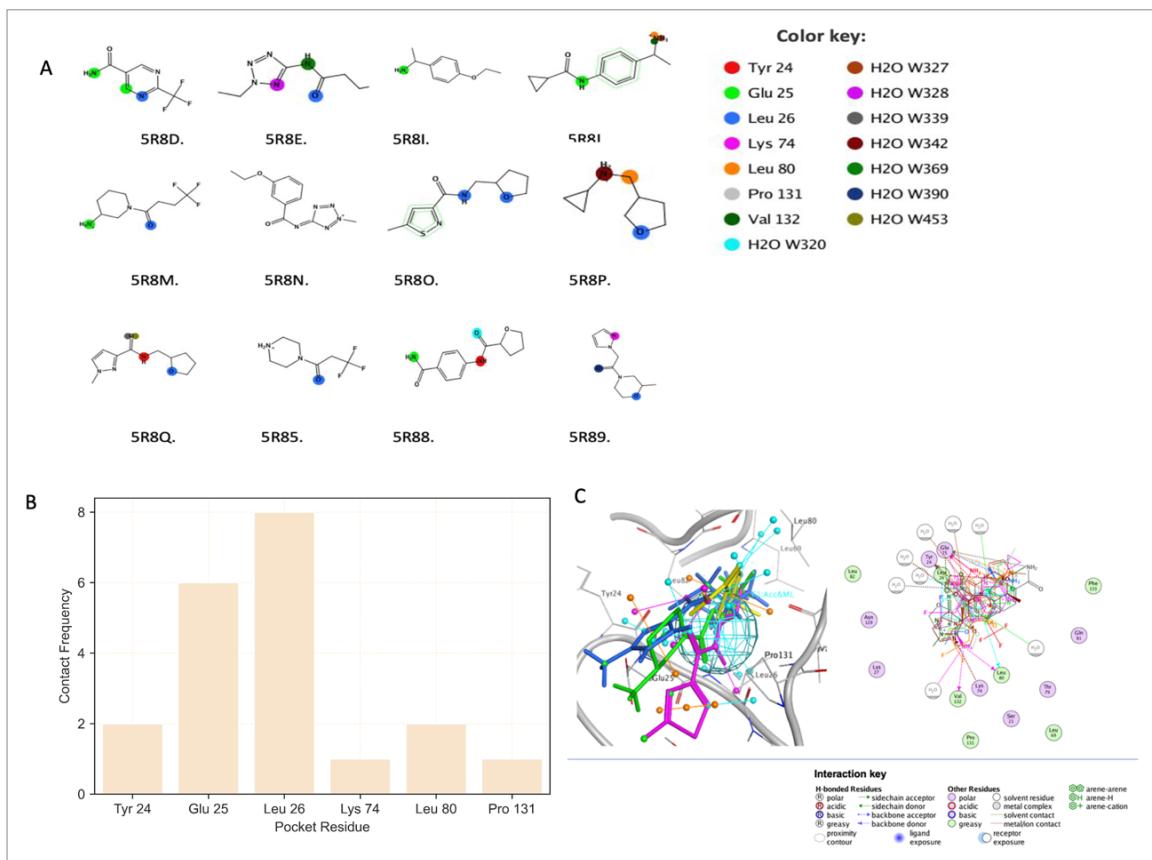


Fig 5: Contact frequency between fragments and pocket residues. A) Structure of fragments showing scaffolds which interact with pocket residues and B) Distribution illustrating number of fragments out of 12 making

contact with key pocket residues. Residues occupy the x-axis of the distribution, and the y-axis represents the frequency of contacts. Color codes maps interactions made between fragment scaffolds and pocket residues. C) 3D-Pharmacophoric illustration of fragments bound to pocket-3 of IL-1, highlighting average binding mode of fragments and key pocket residues making contacts.

Table 1: Protein-Ligand Interaction Fingerprint of fragment-IL1B complexes

Molecule	TYR24	GLU25	LEU26	LYS74	LEU80	PRO131	VAL132
5R8D	--	DD	aa	--	--	--	--
5R8E	--	--	aa	--	--	--	--
5R8I	--	DDII	--	--	--	--	--
5R8L	--	DDRR	--	--	dd	--	dd
5R8M	--	DDIIC	aa	--	--	C	--
5R8N	--	C	--	--	--	--	--
5R8O	--	CRR	ddaa	--	--	--	--
5R8P	--	--	aa	--	d	--	--
5R8Q	d	CC	aa	--	--	--	--
5R85	--	--	aa	--	--	--	--
5R88	dd	DDCC	--	--	--	--	--
5R89	--	--	aa	AA	C	--	--

D=sidechain hydrogen bond donor; d=backbone hydrogen bond donor; A=sidechain hydrogen bond acceptor; a=backbone hydrogen bond acceptor; I=ionic attraction; R=arene attraction; C=total surface contact

4.2. Screening the FDA approved drugs library identifies potential IL-1 inhibitors

To investigate if the FDA clinically approved drugs dataset contains any potential IL-1 inhibitors, we used molecular docking to predict strong binders. The pharmacophore restraints which encode the binding modes of X-ray solved structure of fragments bound to pocket-3 of the protein was used to guide the virtual screening. Fig 6 below outlines how the docking was run and analysed to identify 13 FDA drug Hits now needing experimental validation.

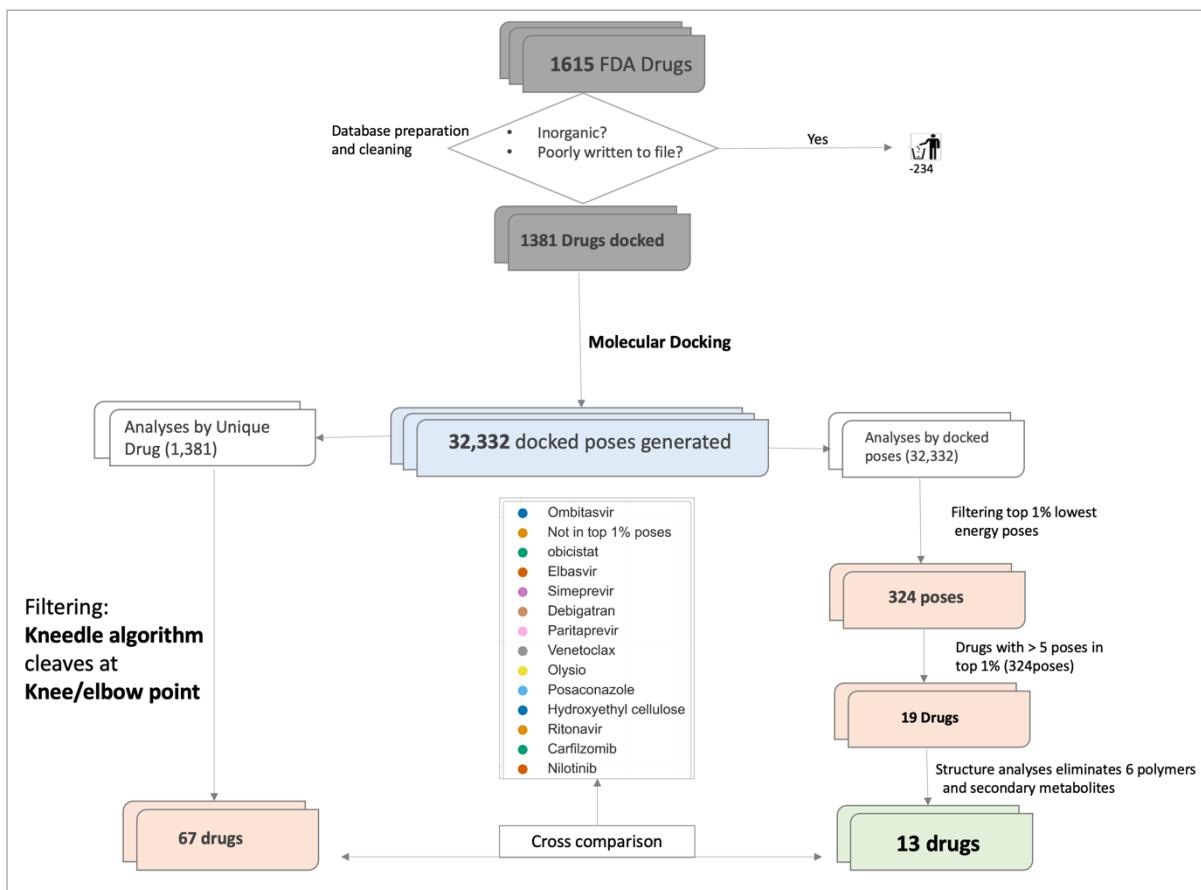


Fig 6: Outline of FDA screening pipeline. A database of 1615 FDA drugs was prepared and docked to generate 32,332 docked poses (retaining 25 best scoring energy poses for each drug) and associated energy values. An integrated approach was used to identify the best scoring hits by analyses of unique drug IDs Vs. docked pose ranks. Thirteen (13) high-scoring hits are finally shortlisted for experimental validation.

After the cleaned dataset of 1381 drugs were docked towards the binding site guided by X-ray-based structural restraints, 32,332 poses were generated and analysed. To identify drugs with significantly meaningful energy scores from the starting database, we employed the *kneedle* algorithm (Satopaa et al., 2011). This algorithm implements the line-of-best-fit for a set of continuous data, in this case the predicted drug energy scores, to calculate the maximum curvature point (“Knee/elbow” point). This point corresponds to the peak point which results when the knee/elbow-shaped part of the curve is rotated by 90° clockwise (Fig 7A). In principle, this point divides the curve in two parts – the useful Vs. less useful energy bits, at the red dashed-line. Ranked from lowest (best) to highest energy, the knee/elbow point occurred at the 67th position of 1381 drug ranks, and this translates to 67 valid energy scores corresponding to 67 FDA drugs (Fig 7C).

Next, we filtered the top 1% (324/32,332 poses) of docked energy poses and extracted from these 19 drugs (Supplementary table 1) which each has more than five (5) different energy

poses in the top 1% of docked poses (Fig 6). Of these, six drugs were eliminated because they are thought to be promiscuous with high false positive rates against several different targets during virtual screening projects (Dantas et al., 2019; Bender et al., 2021). Drug molecules were therefore, eliminated either because they are long flexible polymers with too many atoms and rotatable bonds providing several degrees of conformational freedom which allows fitness to diverse array of pocket geometries or are secondary metabolites (from plants) which are usually difficult to synthesize (Bruns and Watson, 2012). The remaining thirteen (13) drugs also map within the acceptable energy threshold (left-hand side) of the knee-distribution (Fig 7C, Table 2), and so provides a second level of support for their prediction as top scoring resulting Hits of this screening campaign exercise.

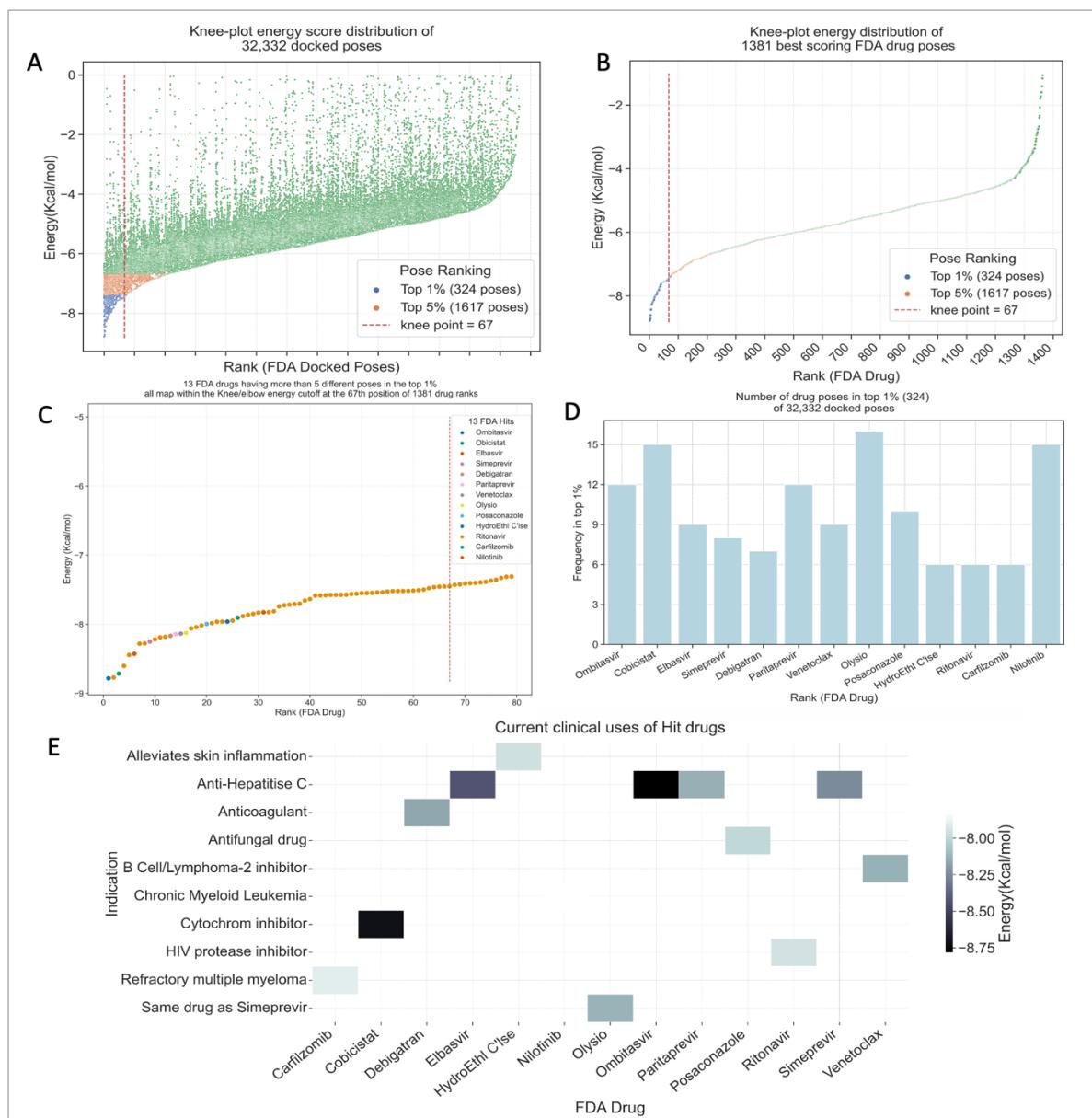


Fig 7: Filtering FDA Drug Hits. A) Knee-plot illustrating the energy distribution of all 32,332 docked poses generated from 1381 drugs. The docking application's scoring function retained ~25 best energy poses for each drug. Color key describes the relative positions of top scoring poses, and the knee/elbow cutoff point is also indicated as red-dashed line. The Kneedle algorithm was used to calculate the knee/elbow point to occur at the 67th position of the 1381 drug ranks. Only the datapoints on the left-hand side of the knee/elbow line are predicted to be valid docked poses. B) Knee-plot distribution of the best energy poses of 1381 drugs. This is a distribution of the best energy scores of each drug. C) Energy scores of 13 FDA Hit drugs which each have greater than five different energy poses in top 1% of the 32,332 poses. All 13 drug energies are within the Knee/elbow energy cutoff at 67th position of 1381 drug ranks providing a two-layer support for their predictions. HydroEthl C'lse = Hydroxylethyl Cellulose. D) Number of different energy-poses of each of the 13 Hit drugs in top 1% of docked poses. E) Current clinical indications of 13 Hit drugs and their relative energy scores

Table 2: Biological, Clinical and Chemical Properties of 13 FDA Hits

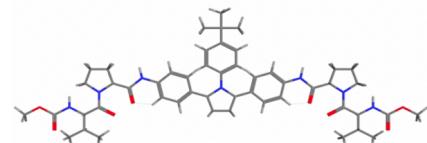
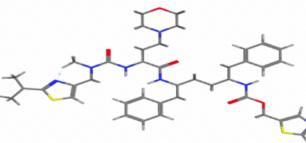
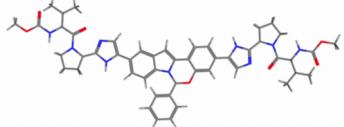
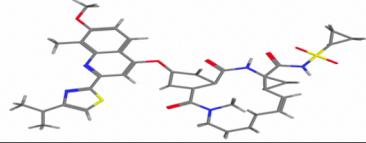
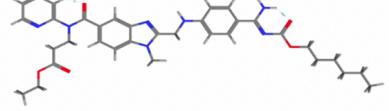
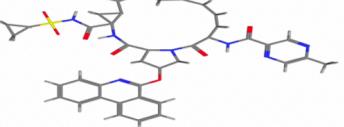
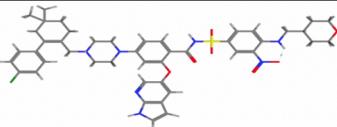
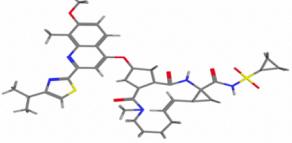
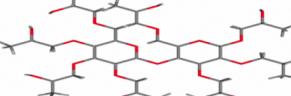
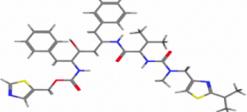
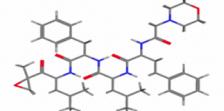
FDA Drug	Target	Indication	Chemical structure	Frequency in top 1% poses	Pose rank /32,332	Drug rank /1381	Energy (Kcal/mol)	% Score
Ombitasvir	Non-structural protein 5A	Anti-Hepatitis C		12	1	1	-8.78	Top 1%
Cobicistat	CYP3A inhibitor	HIV, indirect treatment		15	3	3	-8.71	Top 1%
Elbasvir	Non-structural protein 5A	Anti-Hepatitis C		9	9	6	-8.42	Top 1%
Simeprevir	Non-structural protein 5A	Anti-Hepatitis C		8	15	9	-8.25	Top 1%
Debigatran	Thrombin inhibitor	Anticoagulant		7	23	13	-8.16	Top 1%
Paritaprevir	viral protease NS3/4A	Anti-hepatitis C		12	24	14	-8.14	Top 1%

Table 2: continued.

Venetoclax	Apoptosis regulator Bcl-2	B Cell /Lymphoma-2 inhibitor		9	25	15	-8.13	Top 1%
Olysio	Same drug as Simeprevir	Same drug as Simeprevir		16	27	16	-8.12	Top 1%
Posaconazole	Cytochrome P450 inhibitor	Antifungal drug		10	47	20	-7.99	Top 1%
Hydroxylethyl Cellulose	Topical application	Alleviates skin inflammation		6	56	24	-7.96	Top 1%
Ritonavir	HIV protease	HIV protease inhibitor		6	61	25	-7.94	Top 1%
Carfilzomib	Proteasome inhibitor	Refractory multiple myeloma		6	70	26	-7.90	Top 1%
Nilotinib	Kinase inhibitor	Chronic Myeloid Leukemia		15	93	31	-7.82	Top 1%

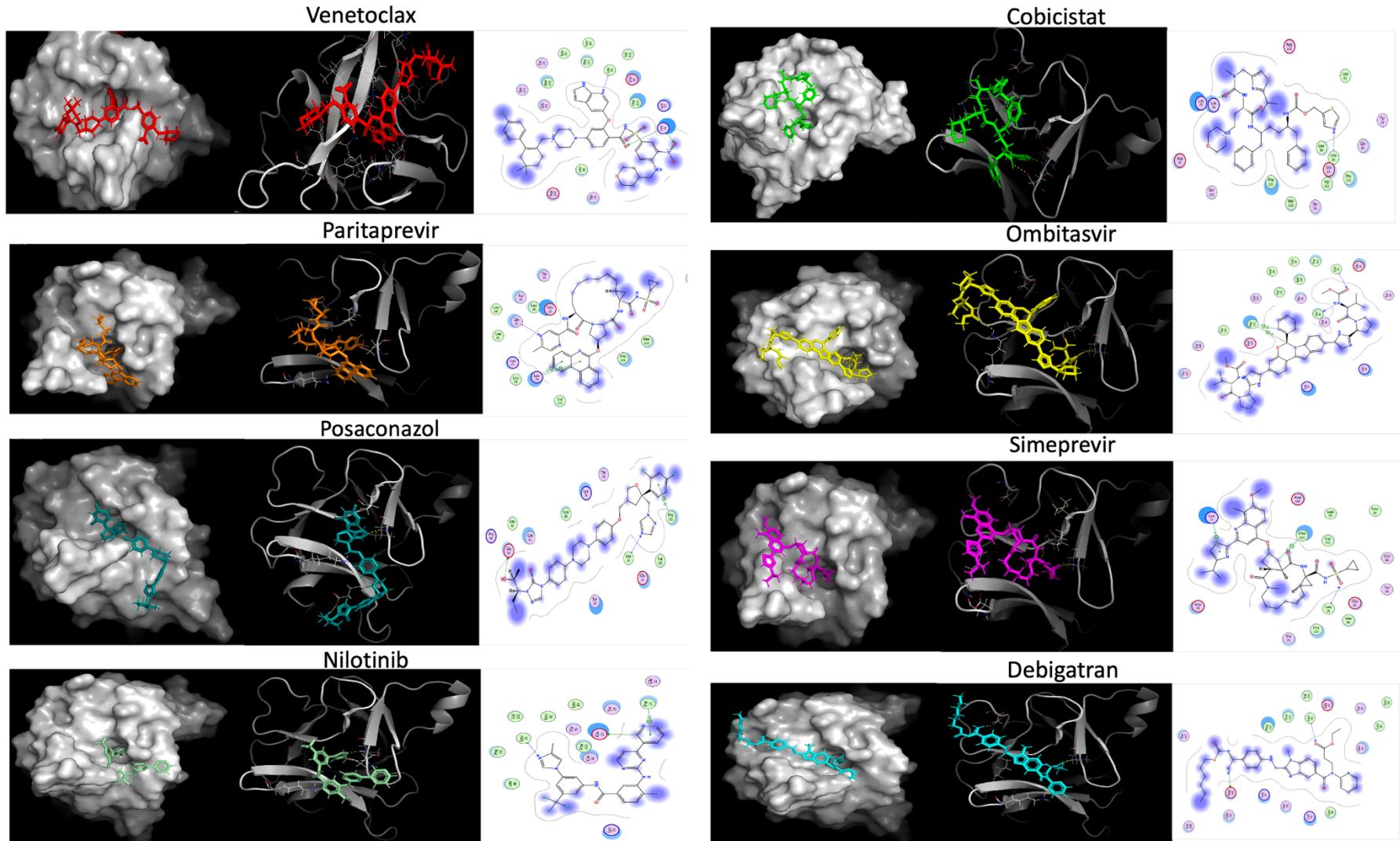


Fig 8: Binding modes of selected FDA Hits. In all panels, we have from left to right a (left) surface representation of IL-1 β and the drug in sticks; (middle) IL-1 β represented in cartoon with bound drugs in sticks; (right) and the 2-D interaction map showing polar contacts made by each molecule with key pocket residues. IL-1 β is colored grey in all panels where the protein is shown. Polar contacts are colored in yellow, for the middle panels with cartoon represented target.

4.3. Screening the AI predicted Cyclica Inc. compounds for potential IL-1 inhibitors

Like the FDA screening campaign, we used molecular docking to predict the best binders from a dataset of 223 compounds which were previously predicted by artificial intelligence as good binders to pocket-3 of IL-1. This was done by a data-driven drug discovery company called Cyclica Inc. Most of the methods used to screen the FDA applies to this screening which I henceforth call “Cyclica screening” or “Cyclica compounds/Hits” and Fig 9 below outlines the analyses pipeline.

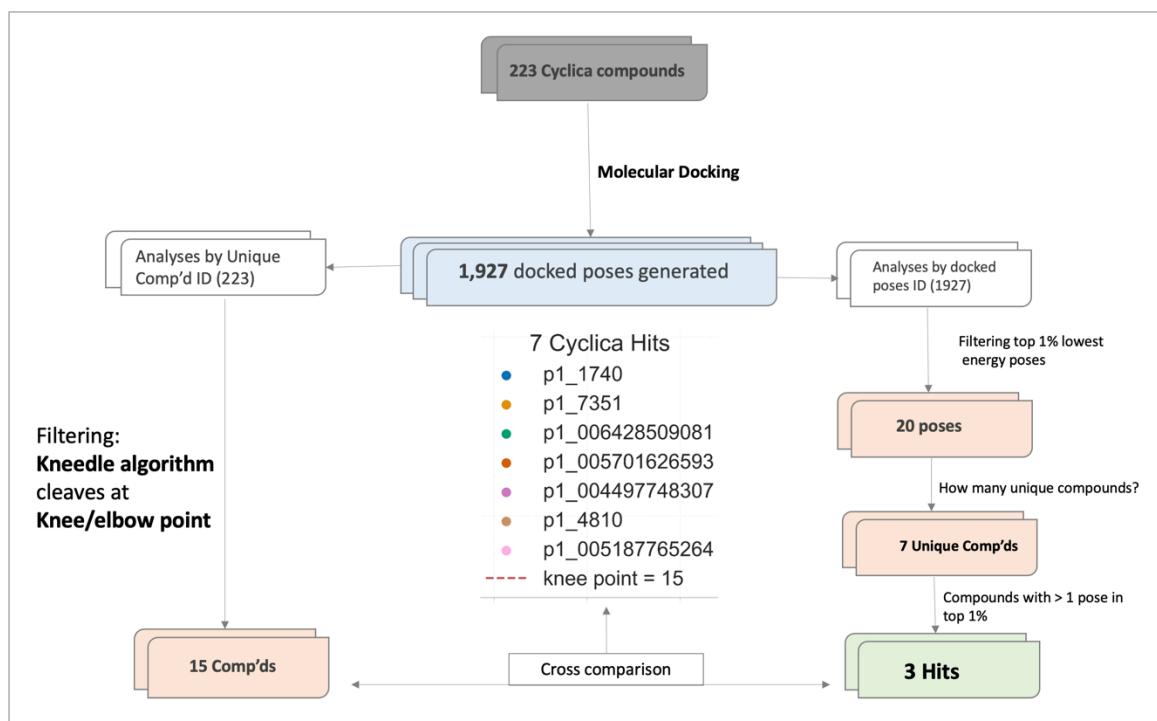


Fig 9: Outline of Cyclica screening workflow. A database of 223 compounds was docked to generate 1,927 docked poses and associated energy values. An integrated approach was used to identify the best scoring hits by analyses of unique Cyclica compound IDs Vs. docked pose ranks. Three (3) high-scoring hits are finally shortlisted for experimental validation.

The Cyclica screening campaign results to Seven (7) Hit compounds needing biophysical characterisation (Fig 10, Table 3). These were first filtered from 1927 docked poses by using the *kneedle* algorithm which isolated 15 compounds (Supplementary Table 2) meeting the knee/elbow point threshold cut-off. Next, I identified the top 1% (20/223 poses) of total docked energy poses and extracted the unique molecules making up the poses. This led to the selection of the seven (7) hits which were found to fall amongst the 15 *kneedle*-shortlisted compounds (Fig 10C, Table 3). Three (3) of these each has more than one docked pose in the top 1% of poses (Fig 10D, 11, and Table 3). Fig 11 below compares the binding modes of our docked hits to the original poses obtained from the Cyclica company, and the results show

that the Cyclica original poses make more polar contacts to the pocket-3 residues of IL-1 than our docked poses.

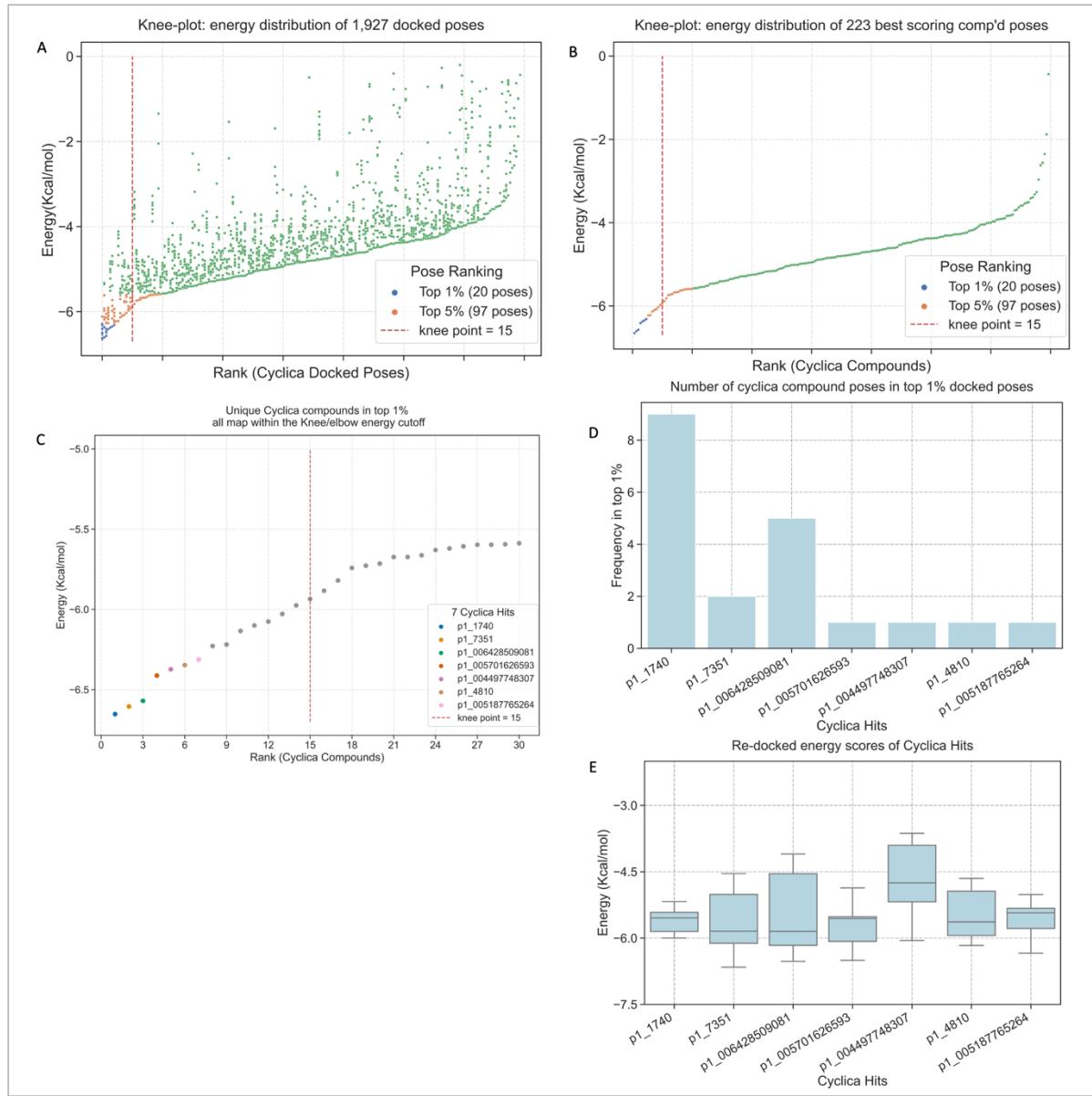


Fig 10: Filtering Cyclica Hit compounds. A) Knee-plot illustrating the energy distribution of all 1,927 docked poses generated from 223 Cyclica compounds. The scoring function retained ~10 best energy poses for each compound. Color key describes the relative positions of top scoring poses, and the knee/elbow cutoff point is also indicated as red-dashed line. The Kneedle algorithm was used to calculate the knee/elbow point to occur at the 15th position of the 223 compound ranks. Only the datapoints on the left-hand side of the knee/elbow line are predicted to be valid docked poses. B) Knee-plot distribution of the best energy poses of 223 compounds. This is a distribution of the best energy scores for each compound. C) Energy scores of 7 Cyclica Hits map within the Knee/elbow energy cutoff at 15th position of 223 compound ranks. This provides a two-layer support for compounds predictions. D) Number of energy-poses of each of the 7 Hits in top 1% of docked poses. E) Re-docked energy distribution of the 7 Hit compounds.

Table 3: Properties of Seven Cyclica Hits

Cyclica Comp'd	Chemical structure	Frequency in top 1% of docked poses	Pose rank/1927	Comp'd rank/223	Energy (Kcal/mol)	% Score
p1_1740		9	1	1	-6.65	Top 1%
p1_7351		2	3	2	-6.60	Top 1%
p1_006428509081		5	4	3	-6.57	Top 1%
p1_005701626593		1	10	4	-6.41	Top 1%
p1_004497748307		1	14	5	-6.37	Top 1%
p1_4810		1	16	6	-6.34	Top 1%
p1_005187765264		1	18	7	-6.31	Top 1%

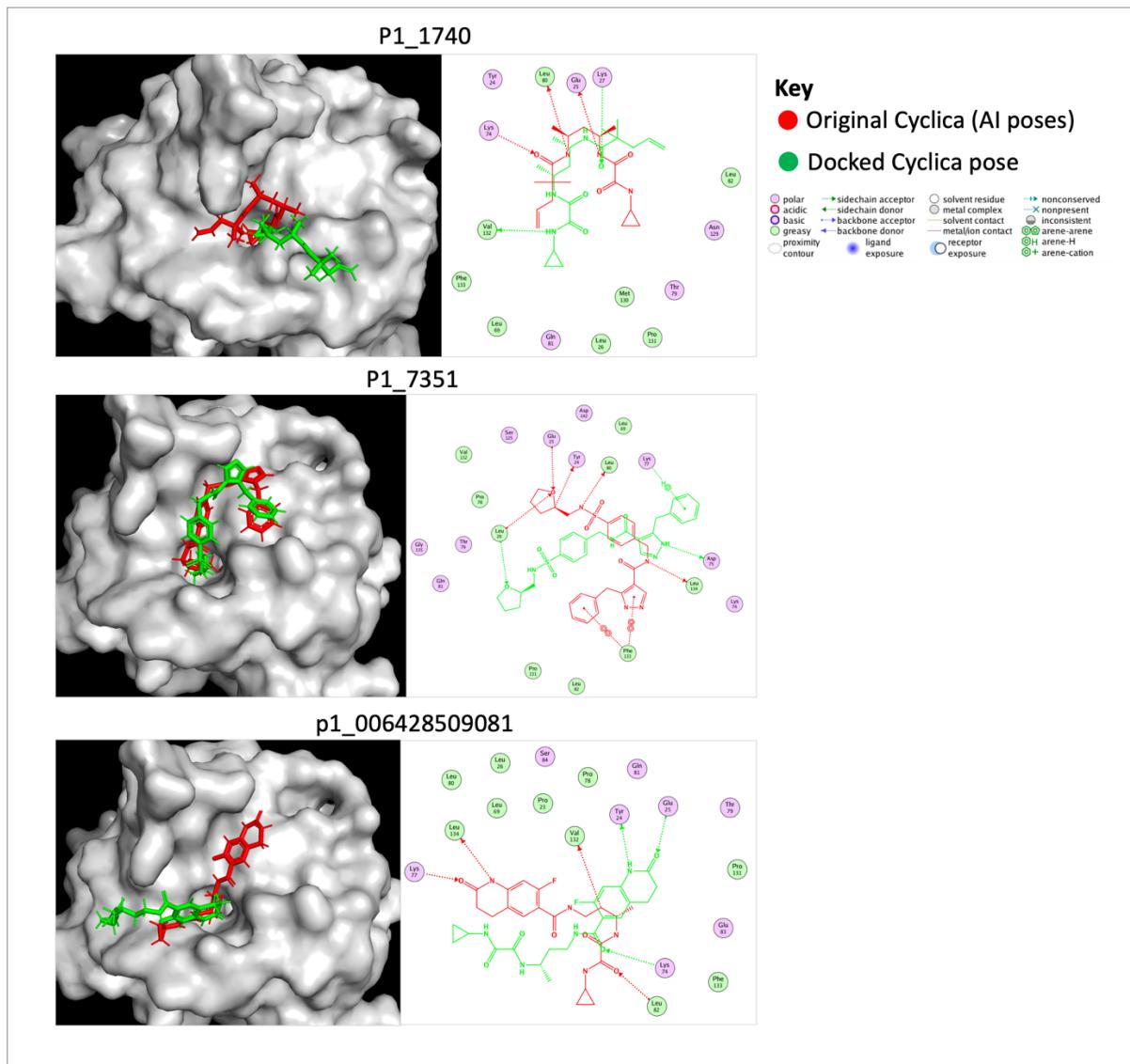


Fig 11: Binding modes of 3 Cyclica Hits (coloured in green) relative to the original AI poses (coloured in red) as obtained from the Cyclica Inc. In all panels, we have from left to right a (left) surface representation of IL-1 β and the bound compounds in sticks; (right) and the 2-D interaction map showing polar contacts made by each hit molecule (green) with key pocket residues, in comparison to the original AI pose (colored in red) as obtained from the Cyclica Inc..

5. Discussion

Here, we show that the FDA clinically approved drug and Cyclica datasets hold a few potential IL-1 inhibitors, with relatively good predicted binding energies, when compared to the rest of the compounds in the respective datasets. Currently, no known small molecule binders to IL-1 β exists (Nichols et al., 2020), against which to compare binding strengths. Using a data-driven approach, we were able to identify 13 FDA drugs and 7 Cyclica Inc. compounds as potential IL-1 inhibitors. We used 3-D pharmacophore restraints to guide all docking calculations. In principle, a pharmacophore abstractly describes the molecular features which are required for molecular recognition to occur between a ligand and macromolecular receptor (Wermuth et al., 1998). First, we extracted such restrain information from the experimentally determined binding mechanisms of 12 compounds (small fragments) co-crystallised to the PDB X-ray structure of the target. Briefly, the pharmacophore restraints incorporates greater than 80% of the important polar and ionic interactions made between key residues (Tyr 24, Glu 25, and Leu 26) and functional groups (amide, amine and ketone) of the 12 co-crystallised fragments described in Table 1 and Figure 5B above (Nichols et al., 2020). Then, we encoded and used this information as a pharmacophore filter by which the binding modes and chemical properties of compounds to be screened are scored.

When all docking calculations were run, we next asked if the 13 FDA and 7 Cyclica inc. top-scoring hits derived from the screening campaign retained binding information found in the starting X-ray crystal structures. Further structural analyses of the quality of interactions reveal a conservation of binding mechanisms between all the 13 FDA drugs and 7 Cyclica hits, and the starting bound fragments of the crystallographic structures. All the hits made contacts with at least one key residue in the predetermined pharmacophore restraints, including but not limited to Tyr 24, Glu 25, Leu 26, Lys 74, Leu 80, Pro 131, and Val 132. Key functional groups of hit compounds, amongst others, frequently include five and six-membered rings containing oxygen and nitrogen atoms making polar contacts; ketone, and amine groups also occur frequently. This interaction signature quite matches that of pharmacophore retrains applied during the screening process.

Screening the FDA approved drugs dataset provides several benefits to our goal of identifying IL-1 inhibitors, as we would only be discovering new modes of actions for existing drugs – with multiple layers of good safety and efficacy records (Oprea et al., 2011) . The most important advantages are the reduced cost and time needed to transfer resulting hits to bedside for treatment of disease conditions. Herein, we identify 13 FDA approved drugs as potential IL-1 binders and note that they are currently indicated for a diverse range of diseases including

anti-cancers (Venetoclax, Carfilzomib, Nilotinib), anti-hepatitis C (Ombistavir, Elbasvir, Simeprevir, Paritraprevir, Olysio), anti-HIV/AIDS (Cobicistat, Ritonavir), anti-fungal (Posaconazole), and anti-coagulants (Debugatran). One interesting observation is that about half (53%) of the FDA hits classifies within the anti-viral drug class, with up to five of these indicated as treatment of Hepatitis C infections. Fig 12A shows how the energy scores for all poses of each FDA hit are relatively distributed. Nonetheless, If one must select fewer compounds, say six (6), in order to minimise bench cost and time during experimental characterisation, this data becomes quite informative. For example, priority may be given to drugs with datapoints skewed towards the better energy (left) side of the figure, including *Ombistavir*, *Cobicistat*, *Paritraprevir*, *Olysio*, *Venetoclax*, and *Nilotinib*. Fig 12B portrays the relative binding modes of each drug at the binding site.

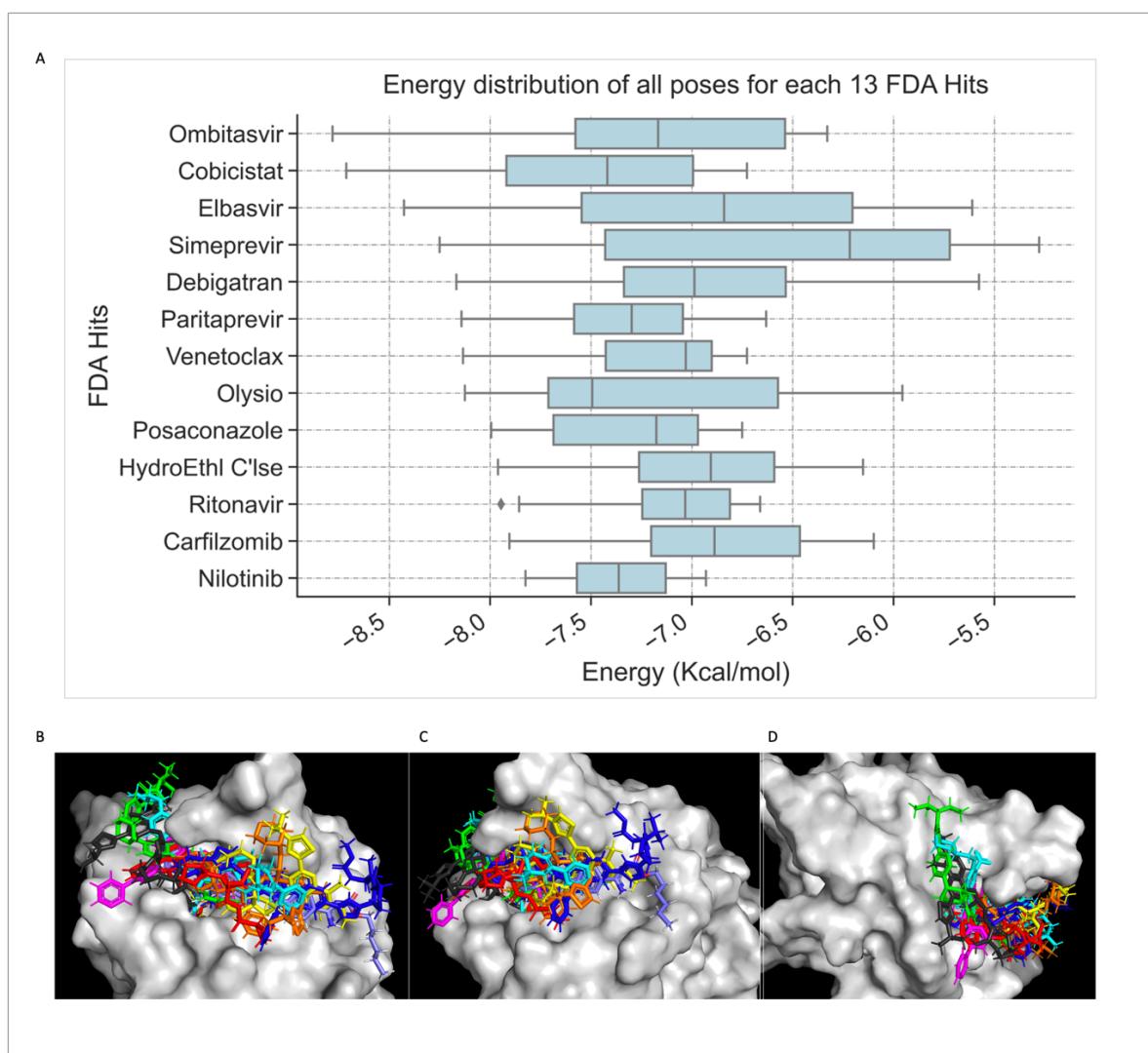


Fig 12: Comparison of 13 FDA Hits. A) Energy distribution of energy scores per FDA hit. Energy scores are represented on x-axis in Kcal/mol, while drug IDs are on the y-axis. **B)** Eight (8) FDA Hits superposed at pocket-3 of IL-1beta (It gets over-crowded when all 13 hits are superposed). Panels **C** is a 45° anti-clockwise rotation of the binding pocket in panel B. **D** is a 45° clockwise rotation of the binding pocket in panel B. **Ombistavir** is coloured

deep blue; Cobicastat coloured Orange; Paritraprevir coloured red; Olysiot coloured black; Venetoclax coloured Cyan; Nilotinib coloured Magenta; Simeprevir coloured Yellow, Posaconazole coloured Green; and Debigatran coloured Purple.

Like the FDA hits, seven (7) Cyclica Inc. compounds are identified in this study as potential binders of IL-1. Three, of these (figure 11, Table 3) not only have more than one energy pose in the top 1% of screens, but also poses interaction signatures of the pharmacophore restraints applied as a filter during the docking calculations. Therefore, they have conserved interaction types as starting X-ray fragments. To understand how the hits obtained from our screening efforts structurally compares to the original structures delivered by the Artificial Intelligence Cyclica company, I superposed the structures of our docked hits to those of the original Cyclica poses onto the binding pocket-3 of the target as illustrated in Fig 11. Interestingly, the AI determined poses show better quality of polar interactions, as more divers contacts are made with the protein. This is not surprising, as AI predictions are modelled based on large amounts of information from experimental data. Although, the docked and original AI poses differ for all 3 top hits, it does not affect the validity of shortlisted hits as the docking process on the one hand is a semi-quantitative tool that only informs which compounds are relatively strong binders compared to weak ones in a database. In conclusion, this study essentially provides a few small molecules which we aim to characterise using biophysical methods to validate the claim that they have some interaction affinity for binding pocket-3 of IL-1.

6. Outline of future plans

1. I will optimise the virtual screening pipeline we use for computational screening to incorporate new selective functionalities, scoring functions, and more robust quality control measures as described in (Bender et al., 2021). More advanced techniques to clean and prepare compound libraries will be incorporated into the new screening workflow.
2. Once optimisation is established, I will apply the improved protocols to continue to computationally screen the complete set of clinically approved drugs database. So far, we have searched the FDA dataset of approved drugs. However, this set accounts for only about ¼ of all globally approved drugs. I will, therefore, continue to explore in fine details, the physicochemical properties of these small molecules with respect to affinity to our target, so that we can be able to answer the question whether this dataset contain any IL-1 inhibitors or not. One key aspect of this would be the use of cheminformatics tools including but not limited to scaffold 2-D and 3-D clustering of compound sets, to identify similarities between clinically untested compounds Vs.

approved drugs. I have in the past written a python script (see appendix A) which calls up and uses the application programming interface (API) of ‘PYMOL align’ to calculate in a high-throughput fashion, the pairwise RMSDs of small ligands. Such data is useful for 3-dimensional clustering to assess compound similarities. Table 4 below exemplifies sample data generated from running my code. The code has the capacity to automatically generate pairwise RMSDs for thousands of compounds, in a single run. We are now looking at improving and applying this to future screening projects, as another level of filtering.

Table 4: Sample Data Generated by RMSD Python Script for Pairwise Compound Clustering

Reference Ligand	Aligned Ligand	RMSD Before Refinement (Å)	RMSD After refinement (Å)	Number of Atoms Aligned	Alignment score
control_fda1	fda1	0	0	8	5
control_cyc1	cyclica1	0	0	38	5
control_cyc2	cyclica2	0	0	57	5
control_cyc1	control_fda1	2.26	2.26	5	0
control_cyc1	fda1	2.26	2.26	5	0
cyclica1	fda1	2.26	2.26	5	0
control_cyc1	control_cyc2	3.32	3.32	30	5

- As it stands, we have identified 13 FDA drugs and 7 Cyclica compounds as potential binders. The priority now, is to conduct *in vitro* binding assays to provide experimental support for the identified hits from our predictions. I will conduct nuclear magnetic resonance (NMR) spectroscopy and Isothermal calorimetric (ITC) assays to assess binding affinities between the hits and the target. In the past, I have learned how to express and purify the protein and was able to severally produce enough quantities of high purity (see Fig 13A).

With the produced protein, we were briefly, able to test if a few in-house designed ligands could bind the target via ITC and NMR (Fig 13B) and attempted to crystallise the protein, and so produced good quality crystals shown in Fig 13C. However, the ITC and NMR screening efforts did not yield any good binders. In this proposal, we may crystallise interesting hits bound to the protein, if the screening campaigns amount to promising findings. The aim would be to identify the binding modes of the ligands to the target. This will serve as a platform or biological tool for lead optimization of promising molecules for treatment of inflammation via inhibiting IL-1 β . ITC is the only thermodynamic method which provides multiple measures of several thermodynamic

properties like entropy, enthalpy, and stoichiometry of ligand-protein interaction in a single experimental run. The 1D, w-logsy, T2, and 2D HSQC NMR assays will be performed to assess in-solution binding affinities on multiple levels.

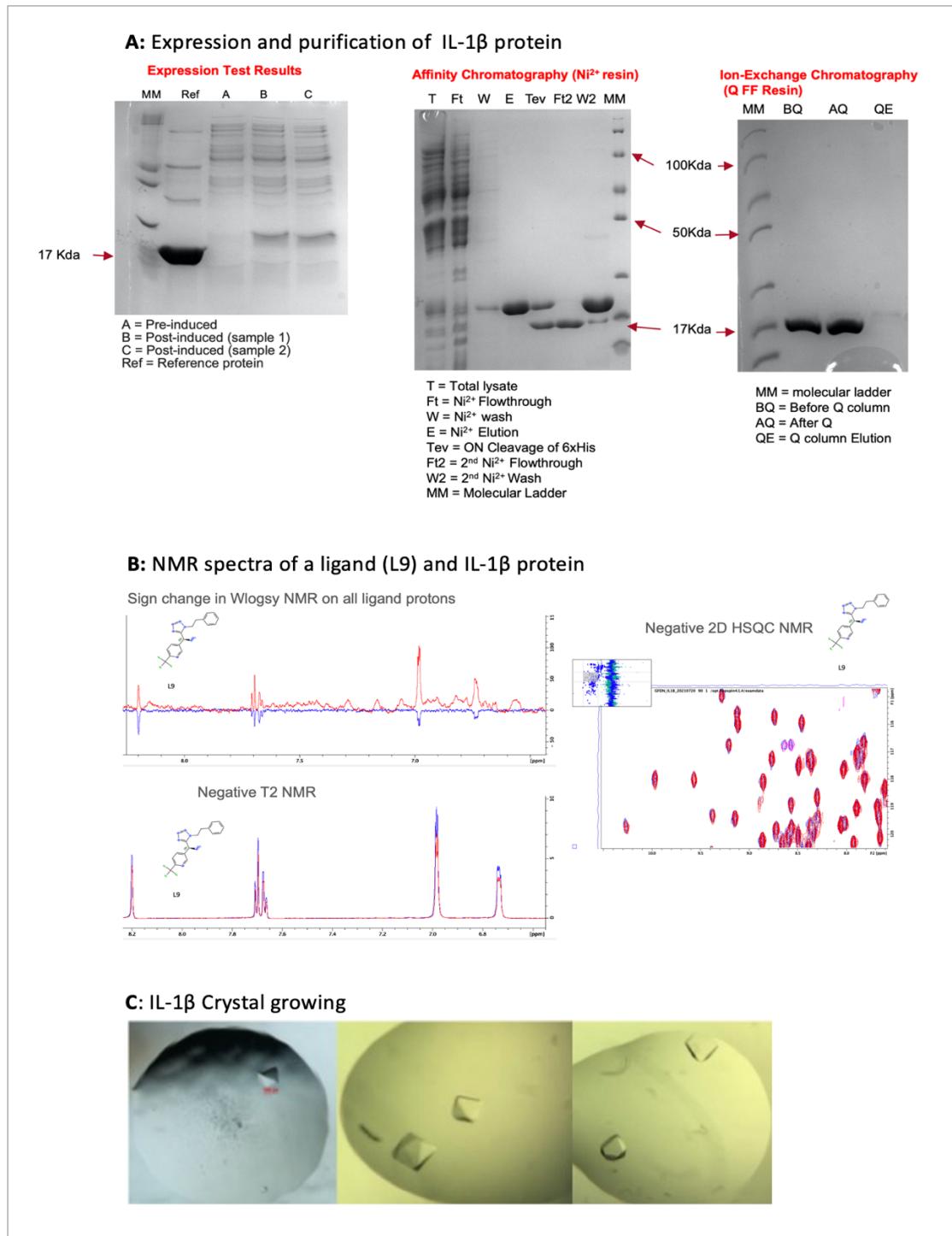


Fig 13: Experimental attempts with NMR and X-ray crystallography **A)** Protein expression and purification of IL-1 β (17.4KDa). *Expression test* (left panel) confirming induction of the protein for production during bacteria growth. *Purification* (middle panel) of the His-tagged protein using affinity chromatography (Ni^{2+} -resin), and *Ion-exchange chromatography* (right panel) (Q FF resin) to eliminate non-specific binders of the Ni^{2+} -resin. **B)** 1D (t_2 , W-logsy) and 2D HSQC NMR assessing the strength of interaction between an L9 (in-house designed ligand) and the protein. **C)** IL-1 β Crystal growing. It takes ~2 to 3 weeks to grow good quality crystals of the protein.

All the panels in the figure indicate ready-made protein crystals poised for reactions (fragment or ligand soaking) with ligands for X-ray diffraction data collection.

7. Project timetable

Objectives	Year 1			Year 2			Year 3		
	1	2	3	1	2	3	1	2	3
Aims 1 and 2: Identification of high affinity ligands via virtual screening									
Optimisation of virtual screening workflow									
Virtual screening (Docking)									
Molecular dynamic simulations									
Cheminformatics analyses of database compounds									
Aim 3: Determination of the binding affinities of predicted hits via <i>in vitro</i> binding assays									
Protein expression and purifications									
Determination of binding affinity by ITC									
Binding affinity determination by NMR spectroscopy									
Possible X-ray Crystallography									
Writing									
Publication of articles									
Writing-up of thesis									

Acknowledgements

I would like to thank my supervisors, Prof. Franca Fraternali and Dr Gian Felice De Nicola for their advice and guidance in shaping this project. I would also like to thank all the members of my Thesis Committee for the active role they have played so far to guide and improve the quality of my work. I also want to especially appreciate the Darwin trust of Edinburgh for funding this research, as well as King's College London for providing me with the appropriate training.

References

- Alexander, M.R., Moehle, C.W., Johnson, J.L., Yang, Z., Lee, J.K., Jackson, C.L., Owens, G.K., 2012. Genetic inactivation of IL-1 signaling enhances atherosclerotic plaque instability and reduces outward vessel remodeling in advanced atherosclerosis in mice. *J. Clin. Invest.* 122, 70–79. <https://doi.org/10.1172/JCI43713>
- Alheim, K., Chai, Z., Fantuzzi, G., Hasanvan, H., Malinowsky, D., Di Santo, E., Ghezzi, P., Dinarello, C.A., Bartfai, T., 1997. Hyperresponsive febrile reactions to interleukin (IL) 1alpha and IL-1beta, and altered brain cytokine mRNA and serum cytokine levels, in IL-1beta-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* 94, 2681–2686. <https://doi.org/10.1073/pnas.94.6.2681>
- Auron, P.E., Webb, A.C., Rosenwasser, L.J., Mucci, S.F., Rich, A., Wolff, S.M., Dinarello, C.A., 1984. Nucleotide sequence of human monocyte interleukin 1 precursor cDNA. *Proc. Natl. Acad. Sci. U. S. A.* 81, 7907–7911. <https://doi.org/10.1073/pnas.81.24.7907>
- Bender, B.J., Gahbauer, S., Luttens, A., Lyu, J., Webb, C.M., Stein, R.M., Fink, E.A., Balius, T.E., Carlsson, J., Irwin, J.J., Shoichet, B.K., 2021. A practical guide to large-scale docking. *Nat. Protoc.* 16, 4799–4832. <https://doi.org/10.1038/s41596-021-00597-z>
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242. <https://doi.org/10.1093/nar/28.1.235>
- Bruns, R.F., Watson, I.A., 2012. Rules for Identifying Potentially Reactive or Promiscuous Compounds. *J. Med. Chem.* 55, 9763–9772. <https://doi.org/10.1021/jm301008n>
- Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., Kollman, P.A., 1995. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* 117, 5179–5197. <https://doi.org/10.1021/ja00124a002>
- Dantas, R.F., Evangelista, T.C.S., Neves, B.J., Senger, M.R., Andrade, C.H., Ferreira, S.B., Silva-Junior, F.P., 2019. Dealing with frequent hitters in drug discovery: a multidisciplinary view on the issue of filtering compounds on biological screenings. *Expert Opin. Drug Discov.* 14, 1269–1282. <https://doi.org/10.1080/17460441.2019.1654453>
- Dinarello, C.A., 2018. Overview of the IL-1 family in innate inflammation and acquired immunity. *Immunol. Rev.* 281, 8–27. <https://doi.org/10.1111/imr.12621>
- Dinarello, C.A., 2009. Immunological and Inflammatory Functions of the Interleukin-1 Family. *Annu. Rev. Immunol.* 27, 519–550. <https://doi.org/10.1146/annurev.immunol.021908.132612>
- Dinarello, C.A., 1994. The interleukin-1 family: 10 years of discovery. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 8, 1314–1325.
- Dinarello, C.A., 1992. The role of interleukin-1 in host responses to infectious diseases. *Infect. Agents Dis.* 1, 227–236.
- Dinarello, C.A., Simon, A., van der Meer, J.W.M., 2012. Treating inflammation by blocking interleukin-1 in a broad spectrum of diseases. *Nat. Rev. Drug Discov.* 11, 633–652. <https://doi.org/10.1038/nrd3800>
- Evans, R.J., Bray, J., Childs, J.D., Vigers, G.P., Brandhuber, B.J., Skalicky, J.J., Thompson, R.C., Eisenberg, S.P., 1995. Mapping receptor binding sites in interleukin (IL)-1 receptor antagonist and IL-1 beta by site-directed mutagenesis. Identification of a single site in IL-1ra and two sites in IL-1 beta. *J. Biol. Chem.* 270, 11477–11483. <https://doi.org/10.1074/jbc.270.19.11477>

- Evavold, C., Kagan, J.C., 2018. How inflammasomes inform adaptive immunity. *J. Mol. Biol.* 430, 217–237. <https://doi.org/10.1016/j.jmb.2017.09.019>
- Fields, J.K., Günther, S., Sundberg, E.J., 2019. Structural Basis of IL-1 Family Cytokine Signaling. *Front. Immunol.* 10.
- Gabay, C., Lamacchia, C., Palmer, G., 2010. IL-1 pathways in inflammation and human diseases. *Nat. Rev. Rheumatol.* 6, 232–241. <https://doi.org/10.1038/nrrheum.2010.4>
- Gery, I., Gershon, R.K., Waksman, B.H., 1972. POTENTIATION OF THE T-LYMPHOCYTE RESPONSE TO MITOGENS : I. THE RESPONDING CELL. *J. Exp. Med.* 136, 128–142. <https://doi.org/10.1084/jem.136.1.128>
- Kaneko, N., Kurata, M., Yamamoto, T., Morikawa, S., Masumoto, J., 2019. The role of interleukin-1 in general pathology. *Inflamm. Regen.* 39, 12. <https://doi.org/10.1186/s41232-019-0101-5>
- Labriola-Tompkins, E., Chandran, C., Kaffka, K.L., Biondi, D., Graves, B.J., Hatada, M., Madison, V.S., Karas, J., Kilian, P.L., Ju, G., 1991. Identification of the discontinuous binding site in human interleukin 1 beta for the type I interleukin 1 receptor. *Proc. Natl. Acad. Sci. U. S. A.* 88, 11182–11186.
- Labute, P., 2008. The generalized Born/volume integral implicit solvent model: Estimation of the free energy of hydration using London dispersion instead of atomic surface area. *J. Comput. Chem.* 29, 1693–1698. <https://doi.org/10.1002/jcc.20933>
- Liebermann, T.A., Baltimore, D., 1990. Activation of interleukin-6 gene expression through the NF-kappa B transcription factor. *Mol. Cell. Biol.* 10, 2327–2334. <https://doi.org/10.1128/mcb.10.5.2327-2334.1990>
- Lomedico, P.T., Gubler, U., Hellmann, C.P., Dukovich, M., Giri, J.G., Pan, Y.-C.E., Collier, K., Semionow, R., Chua, A.O., Mizel, S.B., 1984. Cloning and expression of murine interleukin-1 cDNA in Escherichia coli. *Nature* 312, 458–462. <https://doi.org/10.1038/312458a0>
- Mendiola, A.S., Cardona, A.E., 2018. The IL-1 β phenomena in neuroinflammatory diseases. *J. Neural Transm.* 125, 781–795. <https://doi.org/10.1007/s00702-017-1732-9>
- Mertens, M., Singh, J.A., 2009. Anakinra for Rheumatoid Arthritis: A Systematic Review. *J. Rheumatol.* 36, 1118–1125. <https://doi.org/10.3899/jrheum.090074>
- Molecular Operating Environment (MOE), 2021. . Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7.
- Nichols, C., Ng, J., Keshu, A., Kelly, G., Conte, M.R., Marber, M.S., Fraternali, F., De Nicola, G.F., 2020. Mining the PDB for Tractable Cases Where X-ray Crystallography Combined with Fragment Screens Can Be Used to Systematically Design Protein–Protein Inhibitors: Two Test Cases Illustrated by IL1 β -IL1R and p38 α -TAB1 Complexes. *J. Med. Chem.* 63, 7559–7568. <https://doi.org/10.1021/acs.jmedchem.0c00403>
- On Medicine, Volume III — Celsus [WWW Document], 1938. URL <https://www.hup.harvard.edu/catalog.php?isbn=9780674993709> (accessed 4.7.22).
- Oprea, T.I., Bauman, J.E., Bologa, C.G., Buranda, T., Chigaev, A., Edwards, B.S., Jarvik, J.W., Gresham, H.D., Haynes, M.K., Hjelle, B., Hromas, R., Hudson, L., Mackenzie, D.A., Muller, C.Y., Reed, J.C., Simons, P.C., Smagley, Y., Strouse, J., Surviladze, Z., Thompson, T., Ursu, O., Waller, A., Wandinger-Ness, A., Winter, S.S., Wu, Y., Young, S.M., Larson, R.S., Willman, C., Sklar, L.A., 2011. Drug repurposing from an academic perspective. *Drug Discov. Today Ther. Strateg.*, *Drug repurposing* 8, 61–69. <https://doi.org/10.1016/j.ddstr.2011.10.002>

- PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC, n.d.
- Rivers-Auty, J., Daniels, M.J.D., Colliver, I., Robertson, D.L., Brough, D., 2018. Redefining the ancestral origins of the interleukin-1 superfamily. *Nat. Commun.* 9, 1156.
<https://doi.org/10.1038/s41467-018-03362-1>
- Rueda, M., Bottegoni, G., Abagyan, R., 2010. Recipes for the Selection of Experimental Protein Conformations for Virtual Screening. *J. Chem. Inf. Model.* 50, 186–193.
<https://doi.org/10.1021/ci9003943>
- Satopaa, V., Albrecht, J., Irwin, D., Raghavan, B., 2011. Finding a “Kneedle” in a Haystack: Detecting Knee Points in System Behavior, in: 2011 31st International Conference on Distributed Computing Systems Workshops. Presented at the 2011 31st International Conference on Distributed Computing Systems Workshops, pp. 166–171.
<https://doi.org/10.1109/ICDCSW.2011.20>
- Scott, D.E., Bayly, A.R., Abell, C., Skidmore, J., 2016. Small molecules, big targets: drug discovery faces the protein–protein interaction challenge. *Nat. Rev. Drug Discov.* 15, 533–550. <https://doi.org/10.1038/nrd.2016.29>
- Shaftel, S.S., Griffin, W.S.T., O’Banion, M.K., 2008. The role of interleukin-1 in neuroinflammation and Alzheimer disease: an evolving perspective. *J. Neuroinflammation* 5, 7. <https://doi.org/10.1186/1742-2094-5-7>
- Vigers, G.P.A., Anderson, L.J., Caffes, P., Brandhuber, B.J., 1997. Crystal structure of the type-I interleukin-1 receptor complexed with interleukin-1 β . *Nature* 386, 190–194.
<https://doi.org/10.1038/386190a0>
- Wermuth, C.G., Ganellin, C.R., Lindberg, P., Mitscher, L.A., 1998. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure Appl. Chem.* 70, 1129–1143. <https://doi.org/10.1351/pac199870051129>

Appendix A: Python Script for Pairwise ligand RMSD calculation

```
1  #!/usr/bin/env python3
2  # -*- coding: utf-8 -*-
3  """
4  Created on Mon Jun 13 02:58:50 2022
5
6  @author: neba
7
8  PairWise_RMSD_smallLigs.py
9
10 This code does pair-wise RMSD calculations for two ligands in .pdb format at a time.
11
12 """
13
14 # PairWise_RMSD_smallLigs.py
15 #For large throughput alignments, do not put script in the same directory as pdb files
16 #Let directory of input files contain only the pdb files
17 #Use only "underscore" to name directories with >1 word. Do not use "stroks"
18
19 import __main__ #Needed to run Top-level code" It's "top-level" because it imports
20 #all other modules that the program needs.
21
22 __main__.pymol_argv = [ 'pymol', '-qc'] # Quiet and no GUI
23
24 import re
25 import pymol
26 from pymol import cmd
27 import glob
28 import pandas as pd
29
30
31 pymol.finish_launching()
32
33 directory = "/Users/neba/Desktop/RUN_3/test_runP2/*.pdb" #This is a sample file path containing the pdb files
#directory2 = "/Users/neba/Desktop/CyclicaVsFDA/Results/test1/*.pdb"
34
35
36 # Read User Input
37 structurePath = glob.glob(directory)
38 #structurePath2 = glob.glob(directory2)
39
40 structurePath.sort()
#structurePath2.sort()
41 extension = re.compile( '^(.*[\\/]|\.\pdb)$' )
42
43 object_list = []
44 for filename in structurePath:
45     object_list.append(extension.sub ('',filename))
46
47 #for filename in structurePath2:
#object_list.append(extension.sub ('',filename))
48
49 rmsd = {}
50 rmsd_list = []
51 for i in range (len(structurePath)):
52     for j in range(i+1, len(structurePath)):
53         obj_name1 = extension.sub('',structurePath[i].split("/")[-1].split(".")[0])
54         cmd.load(structurePath[i],obj_name1)
55         obj_name2 = extension.sub('',structurePath[j].split("/")[-1].split(".")[0])
56         cmd.load(structurePath[j],obj_name2)
57
58         rmsd_val = cmd.align(obj_name1, obj_name2, cycles=0, cutoff=5,
59             quiet=0, transform=0)#good ligand alignment when transform=0
#transform=0/1: do superposition {default: 1}
60
61         rmsd.setdefault(obj_name1,{})[obj_name2] = rmsd_val[0],rmsd_val[1],rmsd_val[2],
62         rmsd_val[3], rmsd_val[4],rmsd_val[5], rmsd_val[6]
63
64         rmsd_list.append((round(rmsd_val[3],2), round(rmsd_val[0],2), round(rmsd_val[4],2),
65             round(rmsd_val[5],2), obj_name1, obj_name2)) #indexing to creat
#columns in df
66
67
68         #Delete object and replace with new items when
69         #loop loops over directory, the delete commands must
#rightly indent with loops for good results
70
71         cmd.delete(obj_name2)
72         cmd.delete(obj_name1)
73
74 #Re-order table
75 rmsd_list.sort()
76
77 #testing
78 print (rmsd_list)
79
80 #Create tables with output pairwise rmsds
81 rmsd_df = pd.DataFrame(rmsd_list)
82
83
84 #Renaming columns to match pymol's "align" column names
85 rmsd_df.columns = ['RMSD_BR', 'RMSD_AR', '#atomsAligned',
86                   'Alignment_score','Ref_Ligand',
87                   'Aligned_Ligand']
88
89 #save file to directory
90 rmsd_df.to_csv("rmsd115fd_P2Cyclica.csv", index=False) #A .csv output can be generated and
# saved in working directory as in this sample
91
92 rmsd_df.head(50)
93
94 #The "align" tool of pymol returns a list with 7 items:
95 #RMSD after refinement
96 #Number of aligned atoms after refinement
97 #Number of refinement cycles
98 #RMSD before refinement
99 #Number of aligned atoms before refinement
100 #Raw alignment score
101 #Number of residues aligned
102
103 #jupyter notebook --NotebookApp.iopub_Data_Rate_Limit=1e12
104
105
106
107
108
109
110
111
```

Supplementary Table 1: Nineteen (19) FDA drugs shortlisted by both kneedle algorithm and percentile filtering.

ZINC000203757351	-7.70465	128	12	1.7636331	S(=O)(=O)(NC(=O)[C@]12NC(=O)[C@H]3N(C(=O)[C@@H](NC(=O)c4ncc(C)nc4)CCCC/C=C\ C@H]1C2)C[C@H](Oc1nc2c(c4c1cccc4)cccc2)C3)C1CC1
ZINC000049841054	-7.66765	137	6	3.0212128	O=C(N[C@H](C(=O)N[C@H](C(=O)[C@]1(C)OC1)CC(C)Cc1cccc1)[C@@H](NC(=O)[C@@H](NC(=O)CN1CCOCC1)CCc1cccc1)CC(C)C
ZINC000169677008	-7.66315	138	9	2.3188729	O(CCO)[C@H]1[C@H](OC)C[C@H](C[C@@H](C)[C@H]2OC(=O)[C@H]3N(C(=O)C(=O)[C@@]4(O)[C@H](C)CC[C@H](O4)C[C@H](OC)/C(/C)=C/C=C/C/[C@@H](C)C[C@@H](C)C(=O)[C@H](OC)[C@H](O)/C(/C)=C/[C@@H](C)C(=O)C2)CCCC3)CC1
ZINC000256097213	-7.55126	203	6	2.6107681	O(C[C@@H](O)C)C[C@H]1[C@H](OC[C@@H](O)C)[C@@H](OC[C@@H](O)C)[C@H](OC[C@@H](O)C)[C@@H](OC[C@@H](O)C)[C@@H](OC[C@@H](O)C)[C@@H](OC[C@@H](O)C)[C@H](OC[C@@H](O)C)O[C@H]2CO[C@H](O)C)O1

Supplementary Table 2: Fifteen (15) Cyclica hits shortlisted based on Kneedle algorithm threshold

Original IDs	IDs	Energy (Kcal/mol)	Pose Rank /1927	Rank Compound /223	% Score	SMILES
il1b_Pocket1_Compound 1740.pdb	p1_1740	- 6.6526709	1	1	Top 1%	O=C(N[C@@H](C[C@@H](NC(=O)C(CC=C)(C)C)C)C(=O)NC1CC1
il1b_Pocket1_Compound 7351.pdb	p1_7351	- 6.6050763	3	2	Top 1%	S(=O)(=O)(NC[C@@H]1OCCC1)c1ccc(CNC(=O)c2c(Cc3cccc3)[nH]nc2)cc1
il1b_Pocket1_PV-006428509081.pdb	p1_006428509081	- 6.5701141	4	3	Top 1%	Fc1c(C(=O)NCC[C@@H](NC(=O)C(=O)NC2CC2)C)cc2c(NC(=O)CC2)c1
il1b_Pocket1_PV-005701626593.pdb	p1_005701626593	- 6.4124036	10	4	Top 1%	FC1(C(=O)N2[C@H]3CC(CNC(=O)C(=O)NCC)C[C@@H]2CC3)CC2(C1)CCC2
il1b_Pocket1_PV-004497748307.pdb	p1_004497748307	- 6.3736086	14	5	Top 1%	Fc1c(C)[nH]nc1C(=O)NC1(CNC(=O)C(C)C)C2CCC(O)CC2)CCC1
il1b_Pocket1_Compound 4810.pdb	p1_4810	- 6.3471098	16	6	Top 1%	O=C(NC[C@@H]1[NH2+]C(C)(C)OC1)C(=O)N[C@H]1CC[C@H](C(C)(C)C)CCC1
il1b_Pocket1_PV-005187765264.pdb	p1_005187765264	- 6.3128824	18	7	Top 1%	Fc1c(F)cc2[nH]cc(C[C@H](NC(=O)NCc3oc(C4CC4)nn3)C)c2c1
il1b_Pocket1_PV-006689806556.pdb	p1_006689806556	- 6.22283564	26	8	Top 5%	O=C(NC1CC2(C1)CC(NC(=O)c1c(C)n(C(C)C)cc1)C2)C(=O)NC1CC1
il1b_Pocket1_PV-005170713925.pdb	p1_005170713925	- 6.2190900	27	9	Top 5%	FC(F)(CC(=O)N[C@H]1[C@H](O)C[C@@H](CNC(=O)[C@@H]2Cc3[nH]ncc3CC2)C1)C
il1b_Pocket1_Compound 1299.pdb	p1_1299	- 6.1347499	34	10	Top 5%	O=C(NC[C@@](O)(CC1CCCC1)C)C(=O)Nc1cc(-c2ocnc2)ccc1
il1b_Pocket1_PV-004510776337.pdb	p1_004510776337	- 6.1003475	40	11	Top 5%	Clc1c(F)c2oc(C(=O)N[C@](C(C)C)(CNC(=O)[C@H](O)C)C)cc2cc1
il1b_Pocket1_PV-002583140598.pdb	p1_002583140598	- 6.0759358	42	12	Top 5%	O=C(NCC1CCC(C(C)(C)CC1)Nc1c(C(=O)NC2CC2)nn[nH]1
il1b_Pocket1_Compound 2195.pdb	p1_2195	- 6.0286584	47	13	Top 5%	S(=O)(=O)(CC(=O)Nc1[nH]nc(SC(F)(F)F)n1)CCCC1cccc1
il1b_Pocket1_Compound 1255.pdb	p1_1255	- 5.9748898	52	14	Top 5%	O=C(NC[C@H]1[NH2+]C(C)(C)OC1)C(=O)NC1CCN(c2ncc(C#N)cc2)CC1
il1b_Pocket1_Compound 4756.pdb	p1_4756	- 5.9354911	58	15	Top 5%	FC(F)CC(C(=O)N[C@H]1[C@H](O)C[C@@H](CNC(=O)c2nnn(C)c2)C1)(C)C