A Mechanistic Model Explaining Ligand Affinity for, and Partial Agonism of, Cannabinoid Receptor 1

Fred Shahbazi[†*1,2], Daniel Meister[†*1,2], Sanam Mohammadzadeh[1], John F. Trant*[1,2,3],

[1] Department of Chemistry and Biochemistry, University of Windsor, 401 Sunset Avenue, Windsor ON, N9B 3P4, Canada

[2] Binary Star Research Services, LaSalle, Ontario N9J 3X8, Canada

[3] We-Spark Health Institute, 401 Sunset Avenue, Windsor ON, N9B 3P4, Canada

† These authors contributed equally to the work.

* Corresponding authors' email: farsheed@uwindsor.ca, j.trant@uwindsor.ca, meister@uwindsor.ca

Abstract

Interaction with cannabinoid receptor 1 (CB1) partially determines the bioactivity of the major phytocannabinoids, tetrahydrocannabinol (THC) and cannabidiol (CBD). However, CB1 may also interact with many of the >120 other minor cannabinoids, and there has also been significant effort in preparing synthetic cannabinoids with either enhanced agonistic or antagonistic activity. The design process of these molecules, and the identification of off-target effects at this receptor for molecules designed to target other proteins, would be aided by a reliable computational model that can quantitatively predict binding; unfortunately, developing this model is challenging as CB1, a G-protein coupled receptor, is highly dynamic. Furthermore, although the mechanism of CB1 agonism is understood, the molecular mechanism of partial agonism is unclear. In this report we provide a highly accurate model predictive for a ligand's affinity to the orthosteric site of CB1, use this model to register the predicted affinity (high and low) of two homologous prophetic cannabinoids, and discuss how a mechanism for THC partial agonism arises natively from the model consistent with experimental data. To be successful, the model accounts not only for molecular interaction, but crucially, the partition of ligands into lipid membranes. This model highlights both the capacity and need to include quantitative physico-chemical properties alongside calculated affinities in predictive tools for protein binding.

Keywords: partition coefficient, GPCR, partial agonist, predictive model, cannabis

Introduction

Cannabinoids, alkyl resorcinol-functionalized diterpenes produced most notably by *Cannabis species* are thought to act primarily through agonism and antagonism of human G-protein-coupled receptors (GPCRs).[1] Δ^9-Tetrahydrocannabinol (**THC**) is a well-established partial agonist of cannabinoid receptor 1 (CB_1) where it occupies the orthosteric site, the primary binding site for endogenous ligands located in the extracellular-facing well that is formed by the transmembrane helices. The majority of drugs that target GPCRs act at the orthosteric site.[2,3] Full agonists sitting in the orthosteric site open the G-protein binding domain on the cytosolic side of the protein, while antagonists prevent G-protein binding by inducing a conformational change that closes the site. Partial agonists have intermediate activity. Reverse agonists can help recruit non-canonical G-proteins, or act through alternative β-arrestin pathways.[4] Since the initial discovery of **THC** and other related cannabinoids,[5,6] numerous modifications and analogs have been synthesized to define the structure–activity relationship (SAR) of **THC** with both CB_1 and CB_2.[7] THC analogues differ primarily at two major sites: the ring system and the alkyl chain (**Figure 1A**). According to Bow

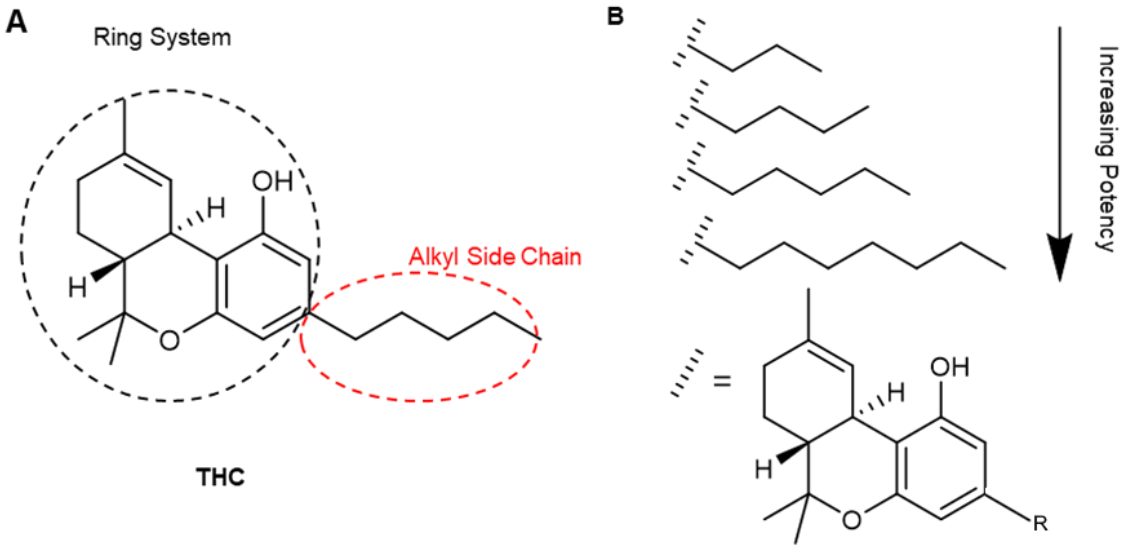

Figure 1 Overview of **THC** and its ring system, highlighting the various side chain lengths discussed in the article.

and co-workers, the length of the alkyl chain is the key parameter for determining CB_1 receptor activity; a minimum of three carbons is necessary to bind the receptor with agonist activity, with activity peaking at eight-carbons and falling off as length increases past that point (**Figure 1B**).[8]

Recently, Citti *et al*. isolated a pair of novel phytocannabinoids: tetrahydrocannabiphorol (**THCP**) with a seven-carbon alkyl chain,[9] and tetrahydrocannabutol (**THCB**) with a four-carbon alkyl chain.[10] Both had higher *in vitro* binding affinity (K_i= 1.2 and 15 nM, respectively) than that reported for five-carbon **THC** (K_i= 40 nM).[8,10] Citti proposes that this differential activity arises because the orthosteric binding site of CB_1 has three hydrophobic pockets[11]: The main hydrophobic pocket (M-pocket) which houses the ring system of **THC** homologs; the long hydrophobic pocket (L-pocket) formed by TMs III, V, and VI, which can accommodate the long heptyl chain of THCP and the pentyl chain of THC; and the hydrophobic sub-pocket (S-pocket) formed by F170, F200, L387, M363, L359, and C386 that lies towards the toggle switch residues needed to activate the receptor. This last is located at the intersection between the M-pocket and the L-pocket (Error! Reference source not found. and **Video S1**). As they are too short to benefit from the hydrophobic L-pocket, the propyl and butyl chains of **THCV** and **THCB** instead sit in the S-pocket. Citti argues that this is the reason for **THCB**'s higher affinity for CB_1 than the longer **THC**.[9,10]

However these new findings contradict the literature: first, **THC** analogues with alkyl chains shorter than 5 carbons—or longer than 8 carbons—have decreased binding affinity compared to those with lengths in that range with affinity peaking at 8 carbons as noted.[8] However, they report that four-carbon **THCB** (K_i=15 nM) has higher binding affinity than five-carbon **THC** (K_i=40 nM). Second, the binding affinity of **THCP** and **THCB** were compared to the **THC** and **THCV** affinity values reported by Bow and Rimoldi which are the least generous available.[8] Several

binding affinities have been reported for **THC** (K_i= 40.7 ± 1.7[8], 35.64 ± 12.4[12], 25.1 ± 5.54[13], 5.05[14] and 2.9 ± 0.3[15]) and **THCV** (K_i = 75.4[8], 46.6[16] and 22 ± 5[17]). Experimental affinity data inherently contains noise, especially when comparing results from different laboratories due to experimental, specific protocol, and even reagent differences.[18] There is a large variation in the experimental data available in the literature which could arise from different testing conditions and protocols and great caution must be taken in overreliance and overinterpretation of small differences that arise from any sort of concentration-dependent non-thermodynamic technique. Access to these minor cannabinoids however, previously restricted due to a combination of synthetic complexity and a low abundance in the plant, is likely to be continually made easier; for example, Magolan recently disclosed a flexible modular synthesis through resorcinol functionalization that will accelerate diversification.[19]

As an aside, it is also important to note the distinction between various pharmacological concepts such as "affinity," "activity," "efficacy," and "potency." These are often interchangeably used in the cannabis literature, and this leads to confusion, and a tendency to attempt to equate incomparable experiments—contributing to the almost two order of magnitude difference in reported THC@CB1 affinities. *Affinity* refers to the strength of binding between a ligand and its receptor, typically measured by dissociation constant (*Ki*) values, although more accurately measured using a biophysical technique such as ITC which can provide ΔG of binding. *Activity* encompasses the ability of a ligand to elicit a biological response upon binding to its target receptor. *Efficacy* specifically relates to the magnitude of the downstream response generated by a ligand-receptor interaction, while *Potency* indicates the concentration of a drug required to produce a specific effect (generally the amount required for a given form of administration, e/g/ oral potency will differ from intravenous or cerebral spinal potency). At the molecular interaction

level, only affinity has any meaning. The other terms (with the possible exception of *activity* should it be defined entirely in terms of a cell-based response) refer to systems-based responses that cannot be measured in a cell-based, or cell-free assay.

For CB1, agonist activity is induced when a ligand occupying the orthosteric pocket forces open the "toggle switch" defined by W356$^{6.48}$ and F200$^{3.36}$; "triggering" this switch induces the conformational change on the cytoplasmic side of the protein allowing for G-protein interaction. An antagonist simply occupies the pocket, preventing any agonist from entering that can trigger the switch.[1] Both agonists and antagonists can have high affinity of course.

The community understands the mechanism of action of agonists, but it is less clear whether there is a consensus mechanism of action for partial agonists. In particular, what makes them *partial* agonists? Lacking any crystallographic data of the receptor with any bound phytocannabinoid, this question remains outstanding. Furthermore, it highlights that we have an imprecise understanding of the experimental binding affinity, even for these well studied major cannabinoids, with reasonable estimates of the K_i varying by well over an order of magnitude. In our program, we primarily focus on modeling the binding affinity of THC derivatives to the CB1 receptor, as reflected by the *Ki* values obtained from experimental studies. These values serve as proxies for ligand-receptor interactions and guide our efforts to understand the structural determinants of binding affinity. By elucidating these distinctions, we wished to develop a model for determining binding affinity for use in screening of new compounds and correct for factors that may not be normally accounted for in methods such as docking. Generally, this is done using an all-atomic molecular modelling study, but this provided inconsistent results: affinity for the receptor was not sufficient, in and of itself, to describe the observed *Ki*. This however can partially be explained by the different mode of entry of ligands into CB_1 compared to many G-protein-coupled receptors: it

enters from the lipid membrane, not the solvent. This is a factor not considered by methods such as docking, which only considers binding to the receptor and assumes the ligand is solvated in water. With this information, and using a library of 21 **THC** homologues with experimental data (**Figure 2**), we propose an empirical correction to predict the affinity of a ligand for CB_1; a conceptual model to determine whether a ligand is likely to be agonist, antagonist, or partial agonist; and a mechanism by which partial agonists function as such. During the preparation of this article, Shukla and colleagues published on the mechanism of action of THC as a partial agonist using complementary techniques to our own, and in close agreement with our proposed mechanism, providing further confidence in the reliability of our hypothesis.[20]

2.0 Methods

Rigid-receptor docking (RRD) and Induced-Fit Docking (IFD) using the Glide module of the Schrödinger suite were conducted to study the binding of THC analogues to CB1, followed by MM-GBSA calculations using Prime MM-GBSA to estimate binding free energies. Molecular dynamics simulations of ligand-bound CB1 complexes were performed using AMBER18 to further investigate protein-ligand interactions and complex stability. The 3D coordinates of CB1 receptor complexes (PDB ID: 6N4B) were obtained from the Protein Data Bank, prepared by removing crystal waters, and subjected to MD simulations for receptor relaxation. Ligands were prepared using Schrödinger software, energy minimized with the OPLS3e force field, and prepared using the Ligprep module. As with all structures in this article, these co-ordinates are available as supporting information in the native file formats from the Borealis Dataverse (see note at the end of the article).

Ligand docking studies were conducted using the Glide module with a grid-based docking protocol, generating an active site using the Receptor Grid Generation module. The IFD extended

sampling protocol was adopted to account for flexibility in both ligand and receptor, followed by Glide docking calculations with default parameters. Prime MM-GBSA was utilized to estimate ligand binding energies and strain energies of complexes, while MD simulations were performed using AMBER18 to observe ligand behavior over time. These computational methods provided comprehensive insights into the binding interactions and dynamics of THC analogues with CB1 receptor, shedding light on ligand-receptor interactions and complex stability.

More information on how proteins and ligands were prepared, and the detailed parameters of the docking, MD and MM-GBSA protocols are available in the supplementary information.

3.0 Results and Discussion

To predict the binding affinity between ligands and receptors as well as to characterize the different binding modes, an *in silico* study was conducted on a total of 21 **THC** analogues with experimentally measured binding affinity towards CB_1, including the antagonist **THCV**,[11] weak agonists **THCA**,[21] partial agonists **THCB** and **THC**,[10] and agonists **THCP**, **AM11542**, **AM841**, **AM12033**, **AM4030**, **HU-210**, ajulemic acid (**AJA**) and **Nabilone**, which have various potency and selectivity (**Figure 2**).

We investigated several parameters to develop a model capable of providing a reliable and accurate correlation between experimental binding affinity and *in silico* docking results. We want to highlight that this is a hard problem: correlation of computational prediction to experiment when the experimental data was all collected in parallel using a single methodology by a single user is still challenging; however, this rich data set does not exist for the CB1 receptor. Instead, we need to compare data collected by multiple research groups using similar (but not identical workflows)

with various ligands. This will introduce variance as experimental *Ki*s measured are highly dependent on protein expression levels and the precise conditions of the data collection. However, although this increases the difficulty in generating the model, it also makes any successful model far more robust, and inherently more useful. To tackle this challenge, we first examined RRD with scaled van der Waals radii of non-polar atoms (1.0, 0.8, and 0.6) to represent some of the flexibility present within the receptor, an approach well precedented to provide good correlation to experiment.[22] Reduced van der Waals radii are crucial for achieving accurate results in some cases where some movement is required from the protein to adopt a binding conformation. Glide employs reduced atomic van der Waals radii to mimic minor protein readjustments, a critical aspect in docking setup. This allows for some flexibility in the rigid-receptor approximation, enhancing ligand binding predictions; however, correct ligand docking may still not be achieved. [23] It generally works best when the initial protein structure best reflects the binding mode of the specific class of ligands, a reasonable expectation seeing the superficial similarity of the ligand library. The docking was followed by further analysis to better determine the free energy of the complex (and consequently the binding energy) using Prime/MM–GBSA calculations.[24] These analyses began with the lowest energy docked conformer in each case, once this pose was visually confirmed to be a reasonable conformer. The MM-GBSA model is a valuable tool for predicting the binding energies between ligands and receptors and is frequently used to evaluate protein ligand interactions, often with good success. However, it simplifies force fields that approximate molecular interactions, often neglecting important electronic effects and quantum mechanical contributions.[25] One of the limitations of MM-GBSA is its neglect of certain physics-based corrections, such as entropic effects, which are crucial for accurately predicting binding affinities. To approximate solvent effects, continuum solvent models like the implicit solvent model are

commonly used. Although computationally efficient, these models may not fully capture the intricate interactions between the solute and solvent molecules.[26] Additionally, MM-GBSA may struggle to accurately capture complex interactions at the ligand-solvent interface, including hydrogen bonding, hydrophobic interactions, and solvent rearrangement effects. The implicit solvent model used in MM-GBSA assumes a uniform dielectric environment, which may not accurately represent the heterogeneous nature of the solvent environment surrounding the ligand.[27] To enhance accuracy, researchers integrate MM-GBSA with techniques such as explicit solvent models[27] and Free Energy Perturbation (FEP) or Thermodynamic Integration (TI). These techniques can account for entropic contributions and provide more accurate predictions of binding free energy.[28,29] The experimental K_i values, rigid docking scores (RRD) and Prime/MM-GBSA predicted binding free energies are listed in **Table 1**.

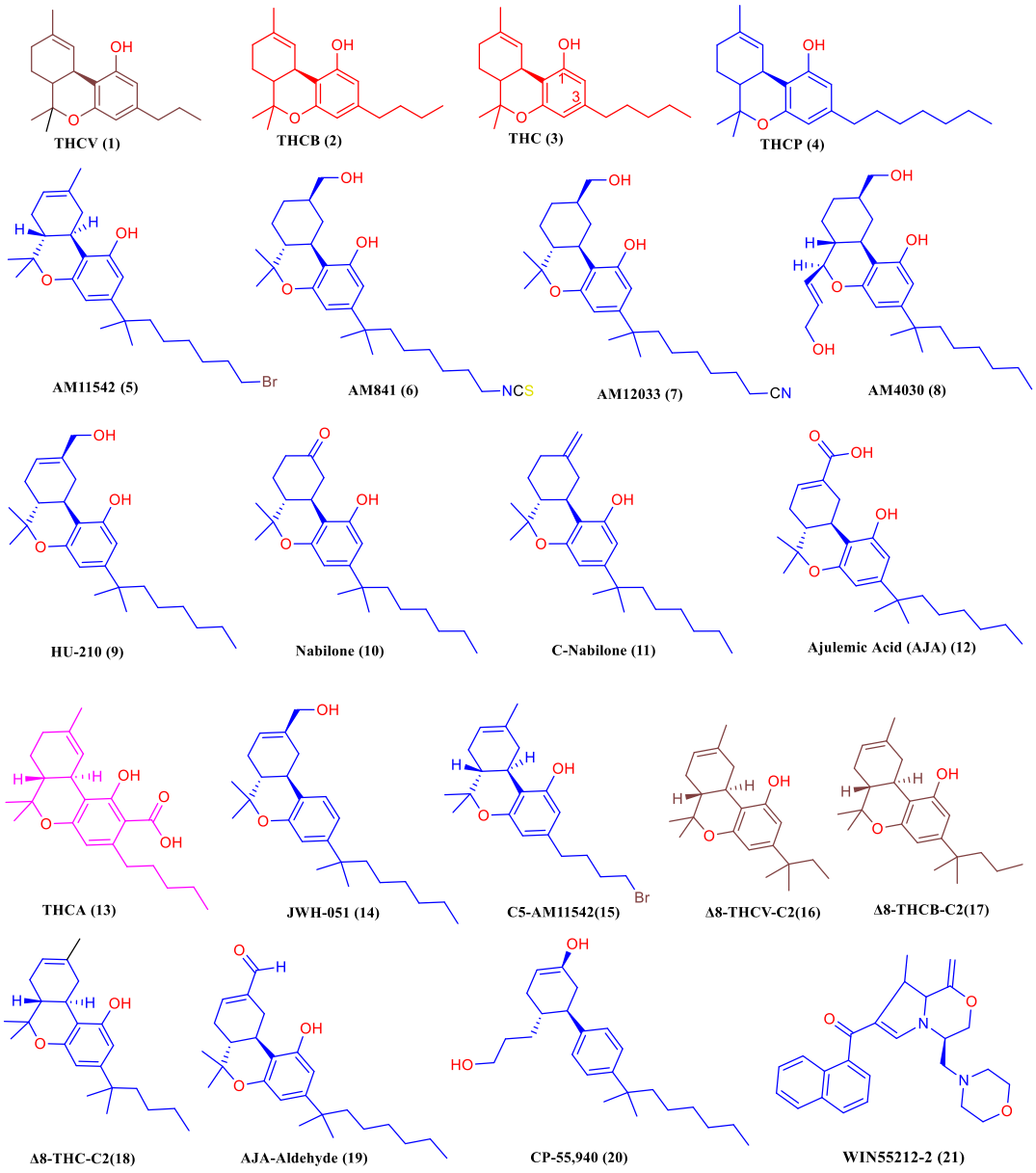

Figure 2 Structures of the **THC** analogues with known experimental binding affinities used in this study. Those in red are partial agonists, in purple weak agonists, in blue potent agonists, in brown antagonists.

Table 1 RRD scores, IFD scores, LogP and predicted binding-free energies (kcal/mol) obtained by Prime/MM–GBSA and Md/MMBGSA of the CB_1 ligands.

Ligand	K_i (nM)[a]	r_w scaling factor: Rigid Docking Score (kcal/mol)			Prime/MM-GBSA ΔG_{bind} (kcal/mol)			IFD docking score (kcal/mol)	LogP	MD/MM-GBSA ΔG_{bind} (kcal/mol)
		1	0.8	0.6	1	0.8	0.6			
THCV (1)	22[17]	-9.12	-8.63	-7.99	-62.09	-54.83	-51.64	-10.81	4.91	-39.09
THCB (2)	15[10]	-9.92	-8.99	-8.24	-62.78	-56.16	-54.25	-10.98	5.3	-43.71
THC (3)	2.9[15,30]	-10.13	-9.07	-8.52	-61.93	-63.9	-55.76	-11.51	5.66	-44.92
THCP (4)	1.2[9]	-5.07	-8.65	-8.02	-57.09	-55.65	-67.06	-11.81	6.44	-47.71
AM11542 (5)	0.11[31]	-8.51	-8.7	-8.05	-55.01	-68.11	-67.9	-12.65	7.58	-57.48
AM841 (6)	1.14[32]	-3.03	-11.36	-8.89	-47.7	-77.74	-73.36	-11.73	5.98	-65.25
AM12033 (7)	0.51[32]	-9.34	-10.03	-9.47	-70.63	-70.08	-73.55	-13.46	4.3	-61.36
AM4030 (8)	0.7[33,34]	-5.44	-8.68	-9.07	-49.09	-64.51	-62.58	-12.28	5.34	-58.92
HU-210	0.73[35]	-8.68	-9.425	-8.25	-81.01	-68.02	-76.49	-12.18	5.83	-63.78
Nabilone (10)	2.19[36]	-9.75	-8.53	-7.78	-70.71	-69.33	-68	-11.91	5.5	-53.71
C-Nabilone (11)	1.82[36]	-5.45	-8.97	-7.51	-64.13	-58.11	-37.19	-11.6	6.72	-63.49
AJA (12)	32.2[37]	---[b]	-9.45	-7.61	---	-55.43	-43.35	-10.62	5.83	-43.94
THCA (13)	23.51[12]	---	-5.98	-5.07	---	-34.99	-34.16	-10.89	5.59	-37.53
JWH-051 (14)	1.2[38]	-9.25	-8.61	-6.92	-68.3	-71.1	-71.67	-12.1	6.56	-63.83
C5-AM11542 (15)	10.8[39]	-9.88	-9.59	-7.85	-61.45	-65.45	-66.44	-10.88	5.9	-49.51
Δ^8-THCV-C2 (16)	14[40]	-8.87	-9.27	-7.9	-62.07	-59.49	-61.29	-10.5	5.99	-40.22
Δ^8-THCB-C2 (17)	10.9[40]	-8.48	-8.79	-8.15	-66.37	-55.06	-65.55	-10.74	5.59	-43.87
Δ^8-THC-C2 (18)	3.9[40]	-8.51	-8.83	-7.8	-63.39	-64.84	-50.54	-11.27	5.79	-52.35
AJA-Aldehyde (19)	2.24[36]	-7.39	-9.27	-8.03	-64.5	-68.4	-64.81	-12.03	5.70	-54.56
CP55940 (20)	0.58[35]	-9.45	-8.71	-7.31	-67.49	-58.55	-74.27	-12.35	5.10	-60.67
Win55212-2 (21)	1.9[41]	---	---	-6.25	---	---	-49.21	-12.35	4.15	-54.35

[a] For ligands where multiple K_i values have been reported in the literature, the lowest reported value was selected; with the differences in reported values ranging to an order of magnitude and dependent on the tool used to measure the value, there is error built into our model. The lone exception is for **HU-210**, where the employed reported value of 0.73 nM is higher than the lowest value, 0.25 nM. This provides better correlation with our model, suggesting that the higher value may prove more correct should the value be redetermined by a third measurement.

[b] --- indicates that the ligand does not dock to the orthosteric binding site of CB_1.

There is only weak correlation between the experimental values (log K_i(nM)) and the RRD score (kcal/mol; **Figure 3A**, **Figure S1**). The Pearson correlation coefficient (R^2) is 0.081, 0.065, and 0.110 for r_w scaling factors 1.0. 0.8 and 0.6, respectively. This is an extremely poor correlation. An MM-GBSA refinement does little to improve the correlations, and although it does become statistically significant with r_w scaling factors of 0.8 or 0.6, this remains a poor tool for predicting

binding affinity (**Figure 3B, Figure S1**). This suggests that there might be more adjustments occurring in the receptor depending on very fine details of the ligand than one would necessarily expect based on their similarity by inspection. This both implies that induced docking might prove more useful, and that mechanism might be dependent on minor adjustments to the binding pocket.

IFD is a far more computationally expensive than RRD, but it allows for considerable flexibility in the binding site residues which works well for systems with moderate flexibility and differences in the binding mode of various ligands.[42] This can be important if the initial pocket in a given conformation is too restrictive or permissive to accommodate a ligand (meaning the RRD will be artificially poor), and both the pocket and ligand must mutually adapt to each other when forming a complex.[43] However, IFD can introduce additional errors in measurement if the pocket is too flexible, and can be less useful for prediction than RRD if the ligand classes are all similar to one another. We have used extended sampling which involves a protocol that automates the process of softening potential and trimming side chains in docking studies. This process aims to enhance sampling efficiency and explore a broader conformational space of ligands within the binding site. IFD generally shows better results in reproducing the native conformations of complexes,[44] and this was used with all 21 ligands (**Table**).

Even by inspection, these results seem to reflect what we know form experimental science: increasing the number of side chain carbon atoms in the series from **THCV** to **THCP** leads to improved docking scores. Overall, the correlation between the experimental values (log K_i(nM)) and IFD (kcal/mol) has dramatically improved compared to the RRD. The Pearson correlation coefficient (R^2) is 0.810 with a ρ-value $<1 \times 10^{-5}$ (**Figure 3C**). However, there are several ligands whose behaviour is not consistent with the model, such as **AM12033** and **AM11542**. This could

be simply that no model is perfect, and that we should be satisfied with a good correlation, or it could be that free energy of binding alone does not model the system correctly.

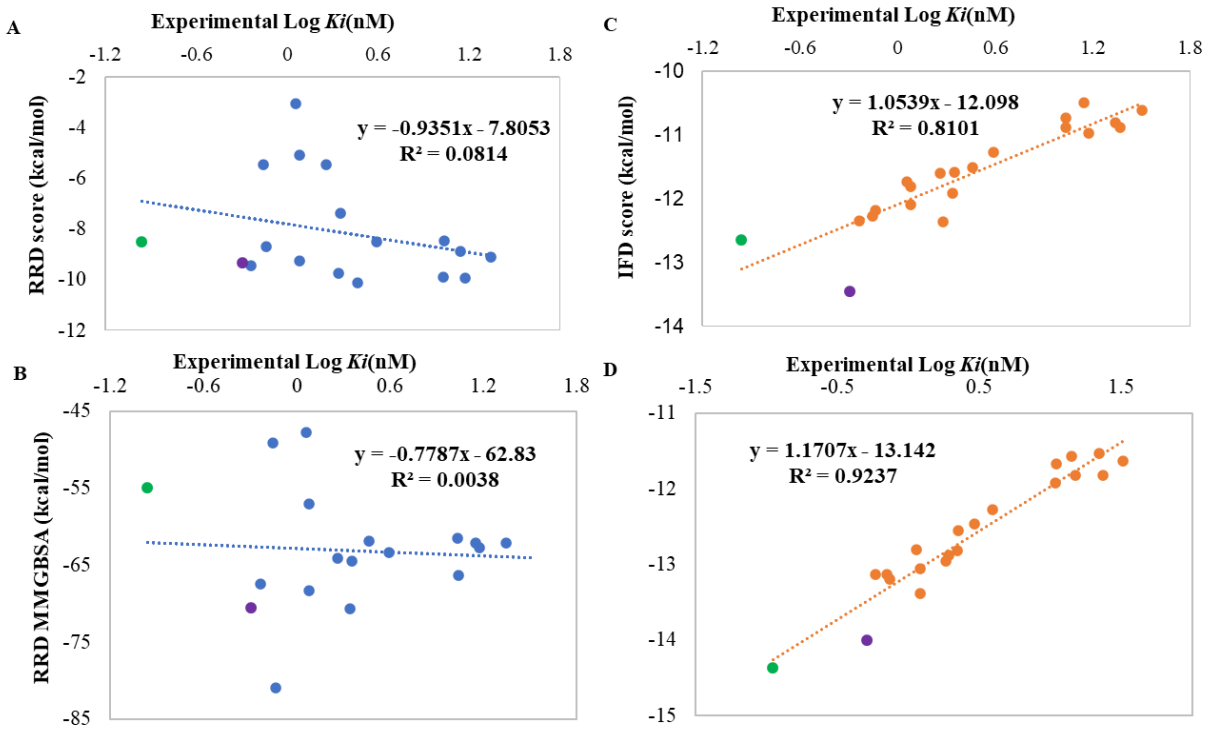

Figure 3 Correlation analysis between experimental values, (A) RRD, (B)Prime/MM-GBSA, (C) IFD scores and (D) IFD scores optimized with Lipophilicity. **AM11542** and **AM12033** are shown in green and purple, respectively.

Let us consider the assumptions of the system. Efficacy depends on several factors beyond simply the affinity of a drug for its target, including the ability of the drug to enter the cell, the stability of the drug over the lifetime of the experiment, and whether it is sequestered through some competing biochemical mechanism. All of these essentially affect the localized concentration of the drug. Generally, for drugs with a similar scaffold, many of these features would be expected to be largely equivalent. Furthermore, most GPCRs, indeed most membrane proteins, interact with their ligand in the bulk extracellular fluid, so many of these mechanisms are not relevant. However, Class A GPCRs can have their orthosteric site opening into the lipid bilayer, and CB_1 is one such protein.[1] Consequently, the relevant concentration is not the concentration of the drug in solution, but rather

the concentration of the drug in the lipid bilayer—these are not the same. The drugs first diffuse into the membrane, and only then do they bind to the receptor (**Figure 4**). The overall observed experimental binding affinity, ΔG_{Obs}, is a combination of ΔG_{Mem} and ΔG_{Bind}. Consider two drugs with the same binding affinity for a type A GPCR like CB_1 that differ only in their water solubility: hydrophobic **A,** and hydrophilic **B**. For the same bulk concentration, **A** would be expected to partition into the lipid bilayer to a greater degree than **B**. This would give **A** a higher localized concentration to bind with the GPCR. Cannabinoids enter the cannabinoid receptors *via* the lipid bilayer.[45-47] Recently, Hurst and *at el.* demonstrated *via* molecular dynamics that ligands access the binding pockets of other class A GPCRs *via* the lipid bilayer.[48] This is consistent with our models where during all MD simulations, the orthosteric site's opening never left the lipid bilayer.

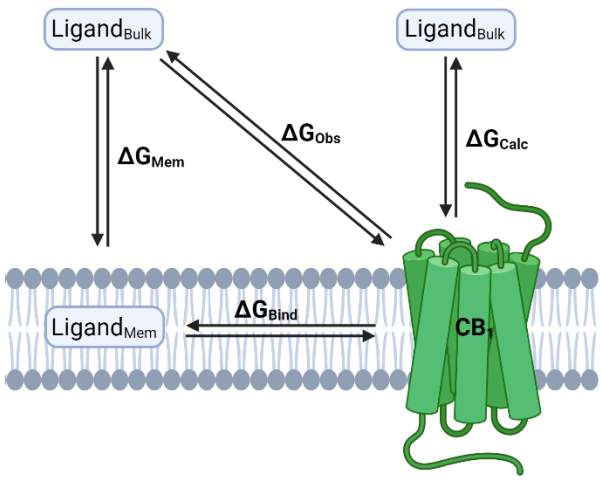

Figure 4 Reaction path of ligand binding to CB_1. The experimentally observed binding affinity (ΔG_{Obs}) is the combination of two steps, the ligand diffusing into the membrane (ΔG_{Mem}) and ligand binding CB_1 (ΔG_{Bind}). Computationally calculated affinities (ΔG_{Calc}) do not account for ΔG_{Mem} and assume the ligand enters from bulk solvent.

This gives rise to a unique challenge when evaluating the binding affinity ligands with methods such as docking or MM-GBSA as these methods assume that the ligand is in the bulk solvent. This gives rise to error when calculating desolvation effects as the bulk solvent is significantly more polar than the membrane and don't accurately reflect the ligand binding pathway (**Figure 4**). The magnitude of the error is expected to be greater with more lipophilic ligands and is proportional to how likely the ligand is to enter the membrane. This can be addressed *via* a lipophilicity correction based on the $\log_p$ of the ligand, the partition coefficient between 1-octanol and water.[49-51] The prediction of this parameter is a key tool in modern drug design.[52] We calculated the $\log_P$ for all 21 ligands using QikProp (**Table 1**).[53] We then employed the imperialist competitive algorithm (ICA), as implemented in MATLAB,[54] to generate a series of best fit equations to the data set with different exponential forms, constants, and relationships between the binding term, derived from the IFD binding, and the hydrophobicity partition term, derived from logP.[55] The best fit equation improved the Pearson correlation coefficient square (R^2) from 0.81 to 0.92 (**Figure 3D**), and correctly shifted the "outlier" ligands towards the trend; partitionability into the lipid bilayer explains the discrepancy between **AM12033** and **AM11542** binding affinity and efficacy. The equation of our fit is as follows:

Optimized Fit: $$K_i = (X + 13.166)/1.1755$$

Where K_i is measured in nM and $X = IFD\ Score - 0.03(\text{logP})^2$. The values of the constants are, of course, empirically derived. To the best of our knowledge, this is the first empirically corrected model for predicting efficacy based on combining binding affinity calculated through all-atomistic modelling, and hydrophobicity, not only for the cannabinoids but for any class of ligand with a membrane protein; we see no reason why this same methodology could not be applied to any other

system where ligands need to partition into compartments, although permeability functions might prove a more useful parameter if the ligand simply needs to passively pass through a bilayer rather than act from the bilayer as in this case. This is proving to be the case with other proteins under study in our lab. An enzymatic stability term could similarly be employed for ligands that enter a cell through the lysosome and must survive processing to engage with their target, although this is admittedly a bit more challenging to estimate.

We then looked at how well the model worked to predict the binding affinities of different analogues that were not included in the training data set. We compared the binding of **Δ^8-THC**, whose binding affinity for CB_1 has been variously reported as 44 nM[8,56] or 47 nM with binding predicted by our model.[57] **Δ^8-THC** differs from **THC** only by the location of the olefin in Ring C meaning we can expect similar lipophilicity and likely a similar binding mode; under this understanding the values do seem rather high compared to that of **THC** (2.9 nM). Docking the ligand using IFD provided a reasonable conformation, and a calculation of the lipophilicity and its use in our equation estimates a K_i of 11.02 nM (**Error! Reference source not found.C** and **5D**). This is lower than **THC** and while within range of values reported in the literature, is on the lower end. This is important as one of the reports for **Δ^8-THC** also estimated the binding affinity of **Δ^9-THC** to be 40 nM,[8] which is an outlier compared to other measurements (see above). Based on our model, we propose that the binding affinity of **Δ^8-THC** has been significantly underestimated in reports to date, and its value might benefit from re-measurement.

A B

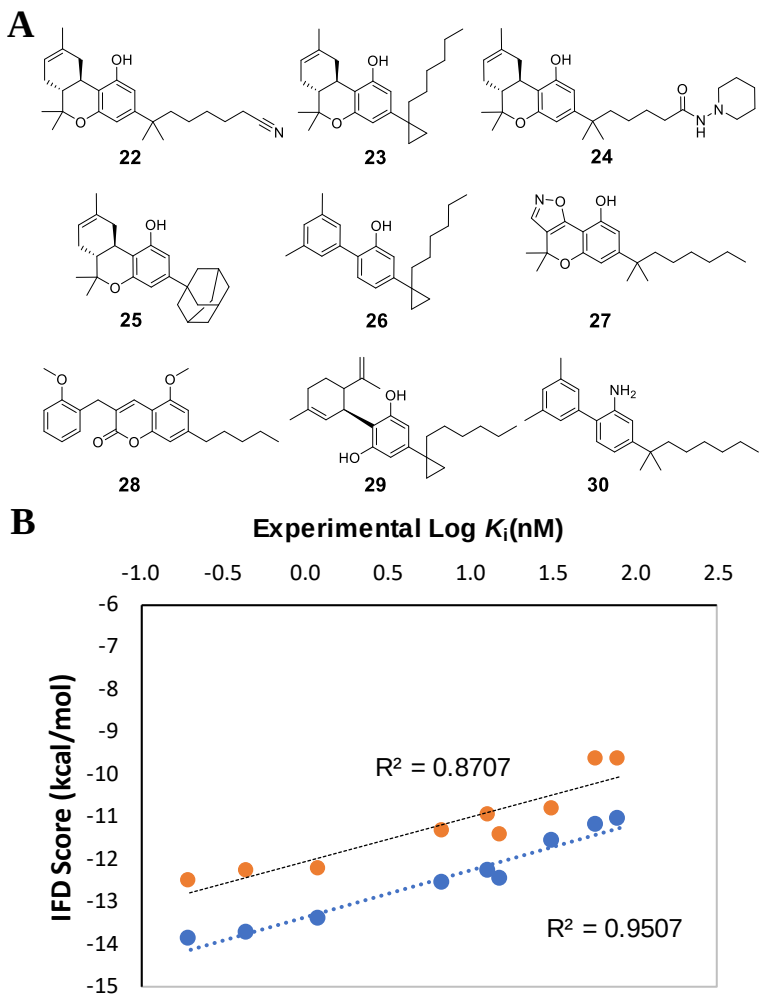

Figure 5 (A) Structures of cannabinoids used in evaluation of the model. (B) Plot of IFD scores of the cannabinoid test set. Uncorrected scores (orange) and lipophilic corrected scores (blue).

We also tested nine other cannabinoids which were more structurally diverse and not part of the original set to see if the correction improved correlation with experimental data (**Figure 5A**, **Table 2**). These spanned a range of K_i values from 0.20 to 80 nM. While docking scores showed good initial correlation (R^2=0.87), this was improved upon inclusion of the lipophilic correction (R^2=0.95) again indicating that lipophilicity of the molecules plays an important role (Figure 5B). Using equation 1 the docking scores of the compounds, the K_i values of the compounds were

calculated. These agreed well with the reported experimental values (**Table 2**). It's important to note that this test data set is highly structurally diverse, the model is appropriate for ligands of CB1, not just THC-like derivatives, many of these molecules lack the resorcinol-terpene core all together.

The establishment of a standardized dataset of binding data enabled the development of a computational molecular docking model capable of accurately categorizing binding affinity.[58] This model effectively distinguished critical structural characteristics of THC derivatives that either enhance or reduce binding affinity. Using our model and our understanding of the structural features responsible for CB_1 binding, we prophesize two new related molecules of which we predict one will prove a very high affinity binder and CB_1 full agonist, while the other will be inactive. **THC** has been the subject of many structure activity relationship studies (Error! Reference source not found.**A**). Gómez-Jeria and coworkers developed a pharmacophore model for classical cannabinoid-CB_1 interactions (Error! Reference source not found.**B**).[59] The C1 phenol group is required for good selectivity for CB_1 over CB_2, and we have already discussed the importance of the alkyl chain. Binding affinity can also be enhanced by hydroxylation of the C11 methyl group as can be seen in the AM-series (Error! Reference source not found.**A**).[57] Using this information, and aiming for synthetic simplicity, we propose two unknown compounds, both simple **THC** homologues, **THCN** with 9 methyl groups and **THCU** with 11 methyl groups. We conducted the IFD and calculated the lipophilicity and then predicted the binding affinity based on our model (Error! Reference source not found.**C**). The alkyl side chain of **THCN** extends perfectly into S-pocket while **THCU** is too long and does not fit into the orthosteric site; it will not be able to fit in the receptor, and we expect it to be largely inactive. THCU exhibited a docking score of −12 kcal/mol, translating to a predicted binding affinity of 0.84 nM when considering logP. This

forecast is unsurprising, given the maintenance of toggle switch movement in THCU, potentially contributing to its superior binding affinity compared to THC., which would make it the best binding phytocannabinoid-like molecule. The synthesis of these new compounds is currently underway for their evaluation, but we wish to register the prediction in the literature in advance.

Table 2 IFD scores, LogP, corrected docking score and predicted *Ki* of the CB_1 ligand test set.

Ligand	K_i (nM)	IFD docking score (kcal/mol)	LogP	Corrected Score	Predicted *Ki* (nM)
22	0.20	-12.50	6.75	-13.86	0.25
23	0.44	-12.25	7.01	-13.73	0.33
24	1.20	-12.21	6.36	-13.42	0.60
25	6.80	-11.32	6.47	-12.57	3.20
26	13.00	-10.94	6.68	-12.28	5.70
27	15.40	-11.44	5.92	-12.49	3.79
28	32.00	-10.79	5.13	-11.58	22.48
29	58.70	-9.65	7.12	-11.17	50.61
30	80.00	-9.66	6.80	-11.05	64.29

It is important to know which one of the reported binding affinities for **THC** and **THCV** correlates best with predicted values. We calculated the correlation of experimental binding affinity and IFD score for all ligands except **THCV** and **THC** (**Figure S3**). The Pearson correlation coefficient (R^2) was 0.797 or 0.918 for IFD (kcal/mol) and IFD scores optimized with lipophilicity respectively. A calculation of the IFD score and the lipophilicity their processing through our equation estimates a K_i for **THCV** and **THC** of 26.93 and 4.11 nM respectively, which is near the expected values.

Although determining a model for predicting binding affinity of designer cannabinoids is critical to our current research program, affinity, as can be clearly seen, does not define the functional role

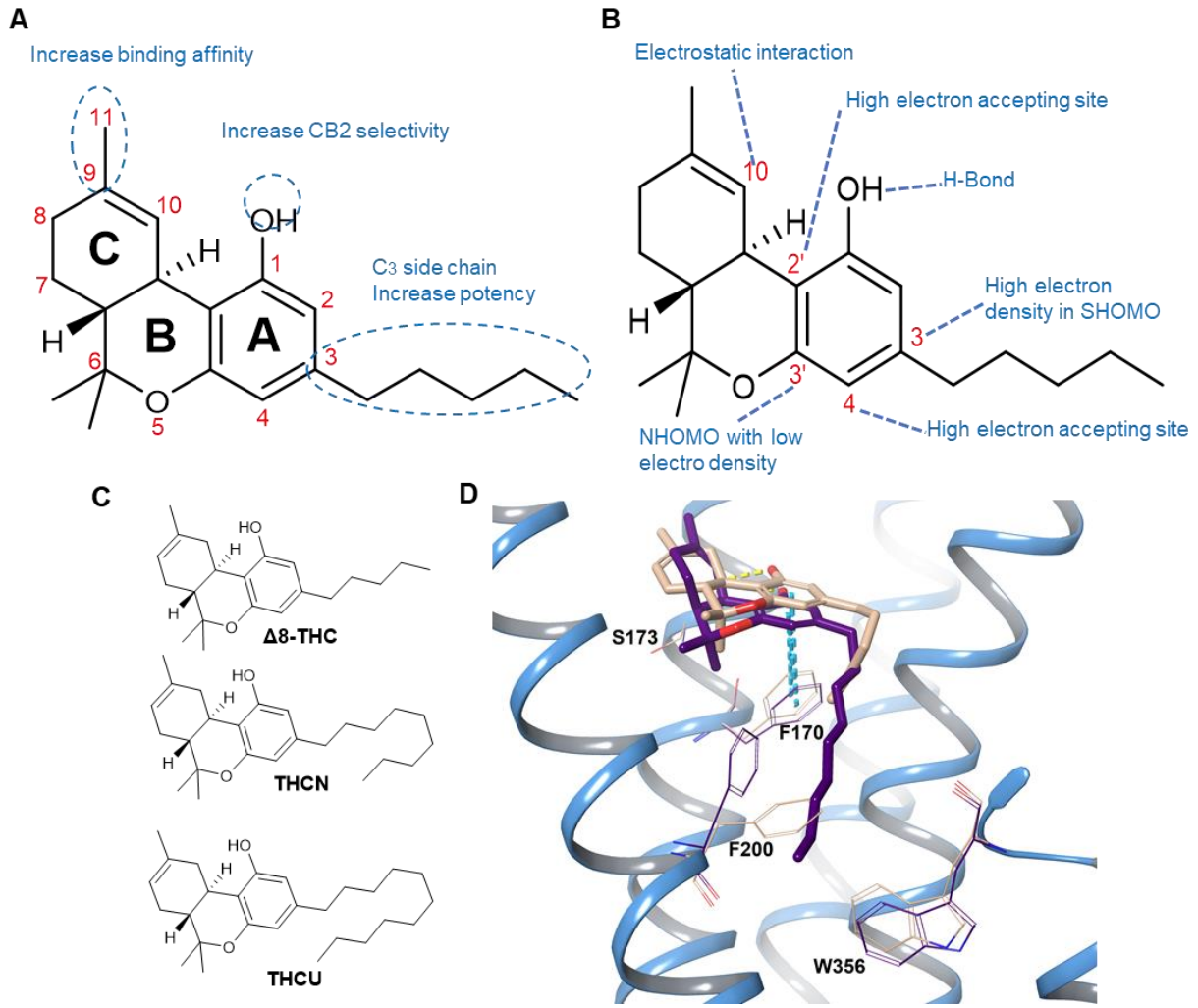

Figure 6 (A) common chemical modifications on **THC** skeleton and (B) Proposed pharmacophore for classical cannabinoids interacting with CB1 receptors. (C) structures of **Δ^8-THC**, **THCN** and **THCU**. (D) binding poses of **THCA** (bisque) and **THCN** (indigo) in complex with CB1. In all figures, oxygen is in red, and nitrogen is in blue. H-bonds are represented by yellow dotted lines, and π-π interactions by blue dotted lines. Key residues related to the ligands are highlighted in the same color as their present ligand.

of the ligand. Tight binders and weak binders can be either antagonists, partial-agonists, reverse agonists, or full agonists. The activity of a ligand arises from the specific of how the ligand interacts with the receptor.[11] We have not identified a clear theoretical literature model that differentiates between these roles. Consequently, we more closely investigated the binding mode of the homologous series of **THCV**, **THCB**, **THC** and **THCP** using IFD as differential receptor response to the ligands likely explains why the first is an antagonist, the middle two partial

agonists, and the latter a full agonist (**Table**). Highly potent agonist **AM11542** was included as a control.

Table 3 IFD scores and predicted lipophilicity of THC homologues ligands.

Ligand	IFD Score (kcal/mol)		Δ(IFDscore)* (kcal/mol)
	S-pocket Pose	L-pocket Pose	
THCV (1)	-10.37	-10.81	0.44
THCB (2)	-10.98	-10.8	-0.18
THC (3)	-11.51	-11.51	-0.07
THCP (4)	-11.81	-10.75	-1.058
AM11542 (5)	-12.65	-10.16	-2.49

*Δ(IFDscore)= $IFDscore_{(S\text{-}pocket)}$-$IFDscore_{(L\text{-}pocket)}$

THCV, **THCB**, **THC** and **THCP** all adopt similar conformations in the orthosteric ligand-binding site. Their ring systems sit in the M-pocket in nearly superimposable geometries: they only differ in that the alkyl side chains of **THCB**, **THC** and **THCP** protrude into the smaller S-pocket towards the receptor-activating toggle switch (formed by F200 and W356), which does not occur for **THCV**, which instead extends into the L-pocket (**Figure 7A**). The phenolic C1-OH of all four cannabinoids forms a hydrogen bond with S173; in the case of **THCV** and **THCB**, it forms an additional H-bond with H178 (**Figure S3A**). The ring systems, excepting that of **THCV**, participate in π-π interactions with the receptor's F170, which sits at the intersection of the three pockets. Hydrophobic interactions help retain the ligands affinity to the rest of the surface, and, as expected, these interactions increase in strength as the surface area increases due to a lengthening alkyl chain with the IFD score rising from −10.81 to −11.81 moving through the series from **THCV** to **THCP** (**Table**).

Interested in mechanism, we focused in on the effects that cannabinoid binding has on the dynamics of the toggle switch formed by F200 and W356, respectively located on transmembrane α-helix 3 (TM3) and TM6. When an alkyl chain pushes between them, it forces open the two helices like chopsticks revealing the G-protein binding site on the cytoplasmic face, allowing for binding, and activating the receptor.[60,61] Their different positions are best described by comparing their form in the presence of **THC** and the highly potent inverse agonist Taranabant (**TNB**, **Figure 7C**; PDB ID: *5U09*)[62]. **TNB@CB$_1$** is akin to the empty inactivated receptor; but reduces its flexibility (hence inverse agonism), locking the two aromatic residues that make up the switch parallel to one another. This holds the transmembrane helices together. The ligand sits in the M-pocket, extending its side chain down the L-pocket with high affinity to prevent other ligands from binding. **THC** on the other hand, extending its tail into the S-pocket pushes the residues open, activating the receptor.

THC, along with its shorter homologues **THCV** and **THCB**, all have similar effects on the toggle switch with the key residues adopting the same conformation in the activated form (**Figure 7B** and **C**). **THCP** extends deeper into this pocket, forcing the residues even further apart, further opening up the G-protein binding site, facilitating activity, and helping to explain its full agonist role (**Figure 7D**). However, this does not explain why the shorter analogues are only partial agonists, or why **THCV** is an antagonist, as they interact the same way. The true story is more complicated.

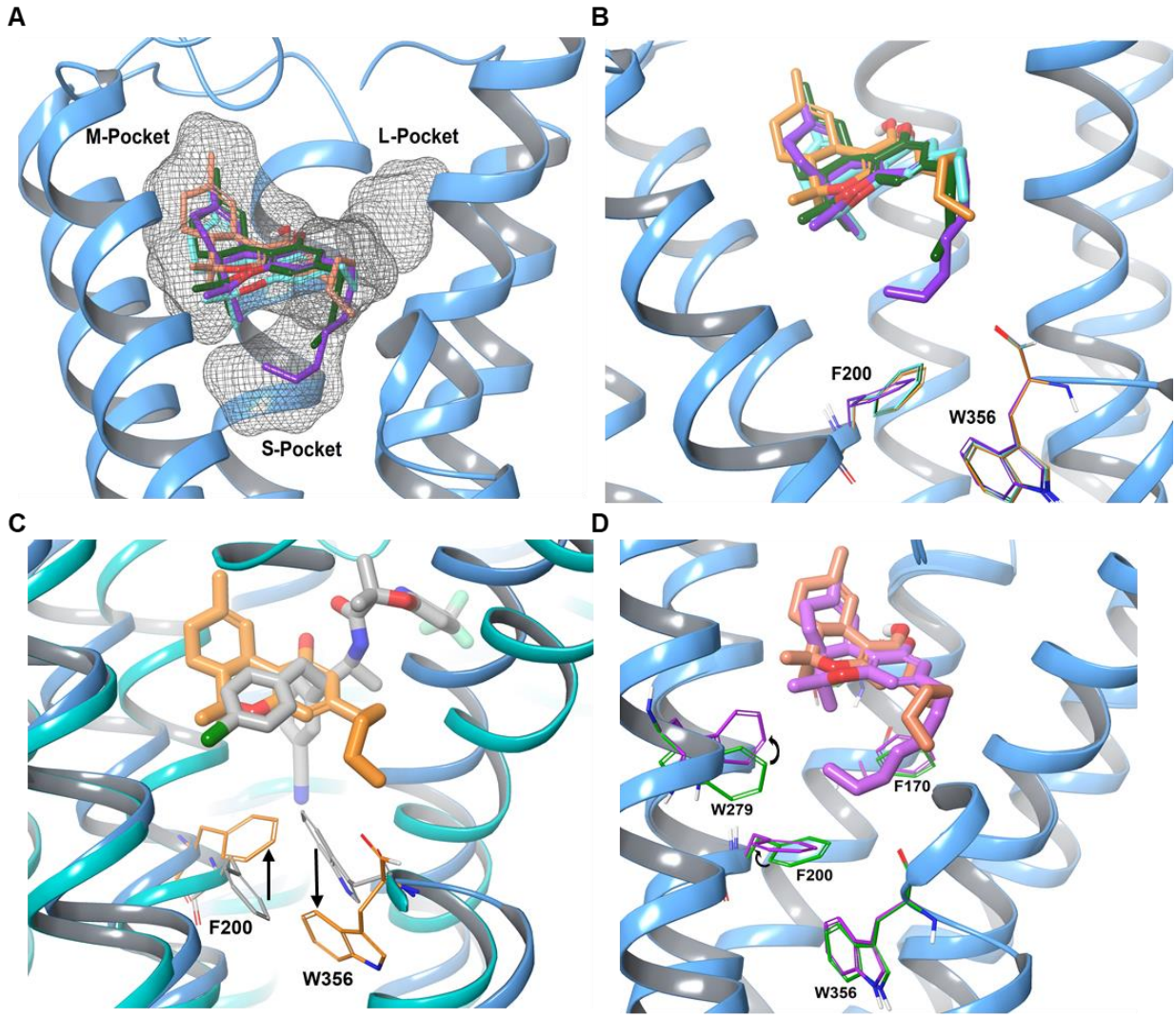

Figure 7(A) binding poses of **THCV** (Cyan), **THCB** (dark green), **THC** (orange) and **THCP** (purple) in complex with CB1 (PDB ID: 6N4B). (B) binding poses of **THCV**, **THCB**, **THC** in complex with CB1. (C) superimposition of **THC@CB1** and **TNB@CB1** (gray) ligand-binding pockets; (D) binding poses of **THCP** and **THC** in complex with CB1. The oxygen atoms are in red, nitrogen in blue and sulfur in yellow, H-bonds in yellow dotted lines, π-π interactions in blue dotter lines and hydrophobic pocket is bordered in dash dark gray mesh. Key residues related to ligands have the same colors.

We then turned to the very potent AM-series analogues. Consistent with the literature and published crystal structures,[1] our model places the C ring system of all ligands into the M-pocket. Most of them extend their alkyl chain into the S-pocket, but that of **C5-AM11542** folds back over itself to extend into the L-pocket (**Figure S4A**). They are all however, highly effective at forcing

open the toggle switch, with the distance between F200 and W356 starting higher than for **THC**, and increasing in the order of **C5-AM11452**, **AM4030**, **AM11542**, **AM12033** and **AM841** (**Figure S4A**). **AM11542**, **C5-AM11542**, **AM841**, **AM12033** each have one π-π interaction with F170. **AM841**, a covalent inhibitor in its final form, has an extra H-bond with S173 and π-π interaction with F268 when it sits non-covalently in the pocket. The phenolic hydroxyl at C1 of **AM12033** forms a H-bond with H178 and the aliphatic OH group at C11 forms two H-bonds with D176 and S173. **AM4030** forms an extra π-π with F268 and the OH of the 6β-((E)-3-hydroxyprop-1-enyl) group form H-bond with F268.

Other agonists studied, like **HU-210**, **Nabilone**, **JWH-051**, **Δ^8-THCV-C2**, **Δ^8-THCB-C2**, and **Δ^8-THC-C2** behave very similarly (**Figure S4B**). In all cases the S-pocket-occupying side chain forces open the toggle switch. Exceptions are **THCA** and **AJA**, which adopt a different conformation in the orthosteric site. The carboxyl group of **THCA** forms two H-bonds, one with S383, the other with H178, which induces a repositioning of the ring system, and consequently the alkyl chain remains in the atrium between the S- and L-pockets entering neither (**Figure S4C**). Unusually, the ring system of **AJA** rotates 180˚C compared to all other ligands. The carboxyl and phenolic OH- groups form strong H-bonds with K192 and S173, respectively locking this unusual conformation (**Figure S4C**, found for the lowest 10 energy docking poses) however this may be an outlier resulting from docking returning an incorrect pose and experiments would be required to determine if this pose is truly how it binds. All poses of **C-nabilone** show far different binding, with the ring sitting at the intersection of the S- and L-pockets and the alkyl chains sticking up into the M-pocket (**Figure S4D**). The third lowest energy pose of **AJA-Aldehyde** is identical to the activation state, and its alkyl chain extends to the S-pocket. Its docking score is also consistent

with the equation, and it marginally improves the R-Squared value when added to the training set (≈0.03) (**Figure S4D**).

These lowest energy docked conformers, however, fail to capture the complexity of the dynamic binding of the phytocannabinoids. **THC** is a CB_1 partial agonist, meaning that upon binding, it does not completely induce the conformational change associated with agonists. There are multiple mechanisms by which this could occur. One would be that in the docked conformation, **THC** simply does not induce enough pressure on the toggle switch to open the G-protein site. This is not supported by our model which predicts that it forces a similar conformation onto the protein as full agonists like **THCP** do. Alternatively, **THC** might drift away from the core of the orthosteric site and occupy a position higher in the cavity as **CBD** is predicted to do in the presence of **THC**. A third possibility is that the alkyl chain can flip from the S-pocket to the L-pocket. Shao *et al.* computationally docked **THC** to a relaxed receptor derived from the antagonist TNB-bound structure (PDB: *5TGZ*).[62] They predicted that the alkyl side chain of **THC** extends just towards toggle residue W356, and would likely activate it as an agonist. Similarly, when Hua and colleagues docked **THC** to the full-agonist bound structure (PDB: *5XRA*)[31], they predicted that **THC** would behave similarly to **AM11542** and that its alkyl side chain of **THC** extend towards F200 and W356. However, in the docking study accompanying their Cryo-EM structure of CB_1, Kumar *et al.*[63] proposed that **THC**'s alkyl chain is more flexible and potentially able occupy either the **L** or S-pockets. This has been further supported by Dutta and colleagues who, like us, proposed that this "switch hitting" behaviour explains the partial agonism of **THC**.[20] Evidence appears to support that **THCV** and **THCB** protrude into the S-pocket towards the toggle switch,[9-11] while **THCP** behaves similarly to **THC**, and occupies the L-pocket.[9]

As this might help mechanistically explain partial agonism, we analyzed the behavior of the alkyl chains of **THCV**, **THCB**, **THC**, **THCP** and **AM11542** in the orthosteric site. We employed IFD and MM-GBSA refinements of conformations of these ligands occupying both the S- and L-pockets and calculated the difference in preference for the two pockets (ΔIFDscore$_{S/L}$, **Table** , **Figure 8**). Among these ligands only **THCV** has a positive ΔIFDscore$_{S/L}$, meaning that it prefers to occupy the non-triggering L-pocket. This explains why it is an antagonist. However, the side chain is very short, and does not extend far into either pocket: even when it does insert into the S-pocket, it does not disrupt the toggle switch residues (**Figure 8A**). **THCV** forms similar hydrophobic interactions in both conformations, interacting with S383, C382, F379, I362, L359 and F170. For the slightly longer **THCB** and **THC**, the ΔIFDscore$_{S/L}$ are −0.18 and −0.07, respectively. This is essentially 0, meaning that in both cases we would predict that the ligand fluctuates rapidly between occupying the two pockets. Unlike for **THCV**, the toggle switch residues do significantly change orientation depending on the location of the alkyl chain (**Figure 8B**). This arises because although both ligands form the same core interactions at the M-pocket with C382, F379, I362 and F170 (and, for **THC**, with M363, S383, L359) regardless of the orientation; they differ in their additional interactions when the alkyl chain enters one or the other pocket (**Figure 8C**). However, for **THCP** (**Figure 8D**) and **AM11542**, ΔIFDscore$_{S/L}$ is high as their alkyl chains are effectively unable to occupy the L-pocket if the ring system is in any reasonable position within the M-pocket. This means they are locked into a conformation that forces open the toggle switch. They cannot move their alkyl chain into the non-activating L-pocket. Consequently, when bound, they must activate the toggle switch, explaining why they are full agonists. While we were working on this project, Dutta and co-workers employed MD simulations to show that THC's alkyl side chain plays a crucial role in determining its partial agonism.[20] Their

research revealed that this side chain is essential for stabilizing the ligand in both agonist and antagonist-like conformations within the receptor binding pocket.[20] Like us, they also showed that it can also fluctuate between the two pockets.

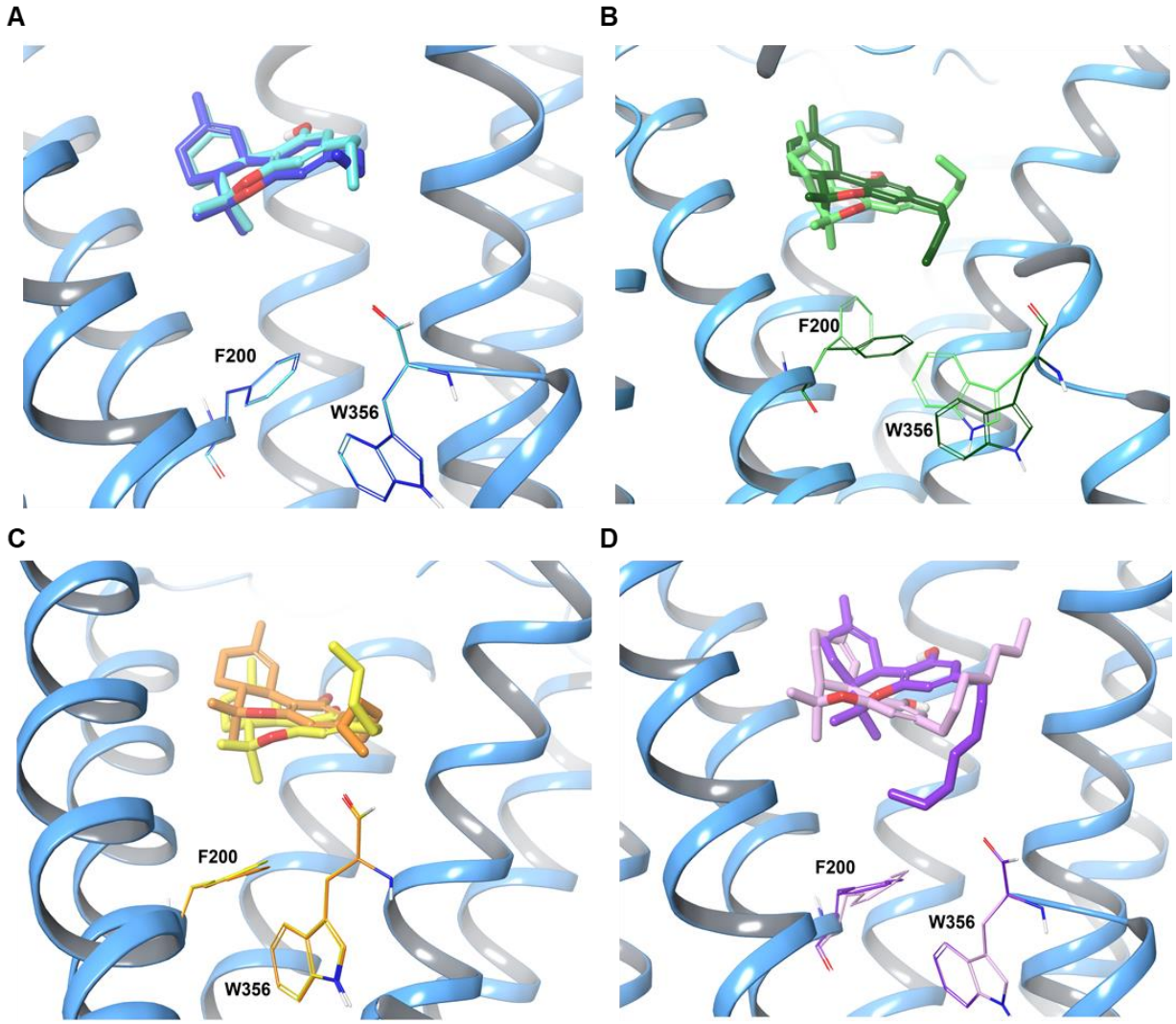

Figure 8 Superimposed docking poses of ligands (A) **THCV** in the **L-pocket** (cyan) and **S-pocket** (blue), (B) **THCB** in the **L-pocket** (dark green) and **S-pocket** (light green), (C) **THC** in the **L-pocke**t (yellow) and **S-pocket** (orange), (D) **THCP** in the **L-pocket** (purple) and **S-pocket** (rosepink). Ribbons are shown in light blue color and residues are colored as same as their related ligands.

We conducted molecular dynamics (MD) simulations for a diverse set of 21 ligands, embedding the CB1 receptor within a phosphatidylcholine (POPC) membrane and solvated with water and NaCl ions to achieve a physiological concentration of 0.15 M. Extending over a 200 ns simulation period, our analysis employed MM–GBSA calculations to quantitatively estimate binding free energies within the orthosteric site. Notably, our findings revealed a strong correlation (R^2 = 0.66, **Figure S5A**) between calculated MM-GBSA values and experimentally determined *Ki* values, affirming the reliability of our computational approach for predicting ligand binding affinities. However, the addition of the hydrophobicity partition term to the training set did not significantly enhance the R-Squared value (0.03), suggesting its limited influence on binding free energy predictions in this context (**Figure S5B**). MD simulations of **THCV**, **THCB**, **THC** and **THCP** and **AM11542** in the L- or S-pocket were performed to investigate their effects on CB_1 activation *via* their interaction with toggle residues and conformational changes in the CB_1 transmembrane helices. CB_1 activation is characterized by the outward movement of TM5, TM6 and TM7 after the ligand interacts with toggle residues F200 and W356, which opens the G-protein binding pocket. This provides better correlation between MM-GBSA and *Ki* value (R^2= 0.94, **Figure S5C**) showing the importance of the alkyl chain towards right pocket. GPCR activation and conformational change can take a long time but occurred rather quickly in our simulations with changes observable within the first 200 ns of simulations. MD simulations were also extended up to 1500 ns, however this only resulted in the eventual movement of ligands out of the long hydrophobic pocket and once ligands (agonists or partial agonists) were no longer interacting with the toggle residues the receptor quickly converted to the inactive conformation, thus the analysis focused on the time frame were ligands remained within the pockets and interacting with the toggle residues to compare differences in receptor activation in these states.

In the case of **AM11542**, an agonist, clear activation and helix movement is observed when compared to the inactive receptor (**Figure 9A**). For **THCV**, very little movement is observed in the helices (**Figure 9B**). This agrees with the experimental observations that it is an antagonist as receptor activation is not observed. In the cases of **THCB**, **THC** and **THCP**, some helical movement is observed and a partial opening of the G-protein binding site (**Figure 9C**, **D**, and **E**). The most notable change was observed in **THCP** which showed the largest movement of TM5 and TM6, though changes were not as pronounced as in **AM11542**, as **THCP** had begun migrating out of the binding pocket. In all cases besides **THCV**, the ligands bound with the side chain in the L-pocket exhibited greater movement in the helices than those in the S-pocket. This partial opening of the G-protein binding site could be the reason that some ligands act as partial agonists or antagonists. It could open just enough for the G-protein to be able to bind, however as it is not fully open, G-protein binding affinity is decreased and overall, a lower response is observed.

An examination of the distances between the helices, specifically TM5 and TM6 shows an interesting trend (**Table 4**). These helices move the most during receptor activation to open the G-protein binding site. **AM11542**, a strong agonist showed the greatest movement of the helices, consistent with full activation. In nearly all cases the ligand with the alkyl chain in the L-pocket resulted in greater receptor activation than when placed in the S-pocket. The exception being **THCV**, however both L and S conformations showed minimal movement and both conformations are consistent with an antagonist.

Taking a closer look at the toggle switches following MD simulations, the reason for the partial loop movement can be observed. The alkyl chain of **AM11542** extends deep into the CB_1 pocket, hitting both F200 and W356 of the toggle switch—significant movement is observed for both residues (**Figure A**). In the case of **THCV**, little movement is observed in the toggle residues, with

a slight shift in F200 but not enough to trigger activation (**Figure B**). **THCB** sits deeper in the pocket and as a result in addition to this slight shift in F200, W356 also experience a slight shift downwards (**Figure 10C**). **THC** interacts effectively with F200, and a significant rotation is observed (**Figure 10D**). Lastly, **THCP** sits significantly deeper in the pocket and can interact with both toggle residues in a manner like **AM11542** (**Figure E**). In all cases, the ligands in the L-pocket resulted in a more significant movement of toggle residues compared to those in the S-pocket.

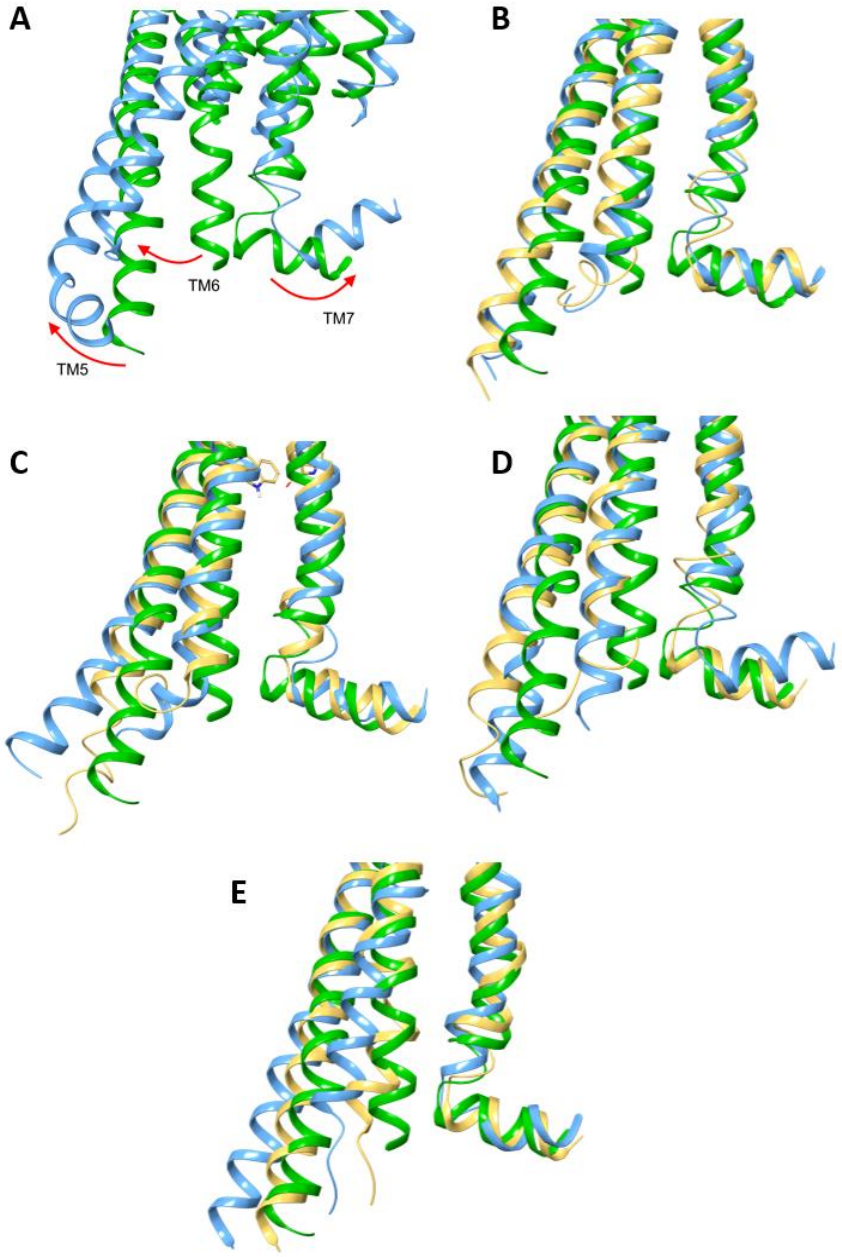

Figure 9 Superimposed structures of CB1 post MD simulation showing helix movement and receptor activation. Receptor with no ligand is represented in green, receptor with the ligand L-pocket in blue and receptor with the ligand originating in the **S-pocket** in orange. (A) **AM11542** (B) **THCV** (C) **THB** (D) **THC** (E) **THCP**.

Table 4 Distances between TM6 and TM5 from TM1. Used as a surrogate measurement for how open the G-protein binding site is; the greater the value, the more open the G-protein binding site.

	Distance between intracellular ends of TM6 and TM1	Difference from inactive receptor	Distance between intracellular ends of TM5 and TM1	Difference from inactive receptor2
Receptor	23.15	0.00	22.72	0.00
AM11542_L	29.94	6.78	29.34	6.63
AM12033_L	26.51	3.35	26.01	3.29
AM841_L	27.04	3.89	25.76	3.05
THCV_L	23.83	0.67	26.24	3.52
THCV_S	25.63	2.47	24.51	1.80
THCB_L	27.72	4.57	28.31	5.59
THCB_S	23.67	0.51	25.38	2.66
THC_L	26.85	3.70	26.98	4.27
THC_S	25.14	1.99	26.69	3.97
THCP_L	26.67	3.52	28.35	5.63
THCP_S	25.01	1.86	24.67	1.95

This indicates that the ability of the ligands to interact with these toggle residues is key to receptor activation and that smaller ligands with shorter chains fail to induce the structural changes required for full activation. Instead, what occurs is a partial activation, characterized by partial movement of TM5, TM6, and TM7, which correlates to the degree of how well the ligands can interact with either F200, W256 or both. This explains why some ligands such as **THC** behave as partial agonists, despite their high binding affinities and provides insight into the mechanism of partial agonists. This also highlights the importance of looking beyond the binding affinity when designing new ligands for receptors. The method through which they enter the binding pocket, in this case through the lipid membrane, is a key factor, along with the exact binding mode and residues that ligand interacts with. Depending on the active site residues that are interacted with, vastly different biological effects can be observed.

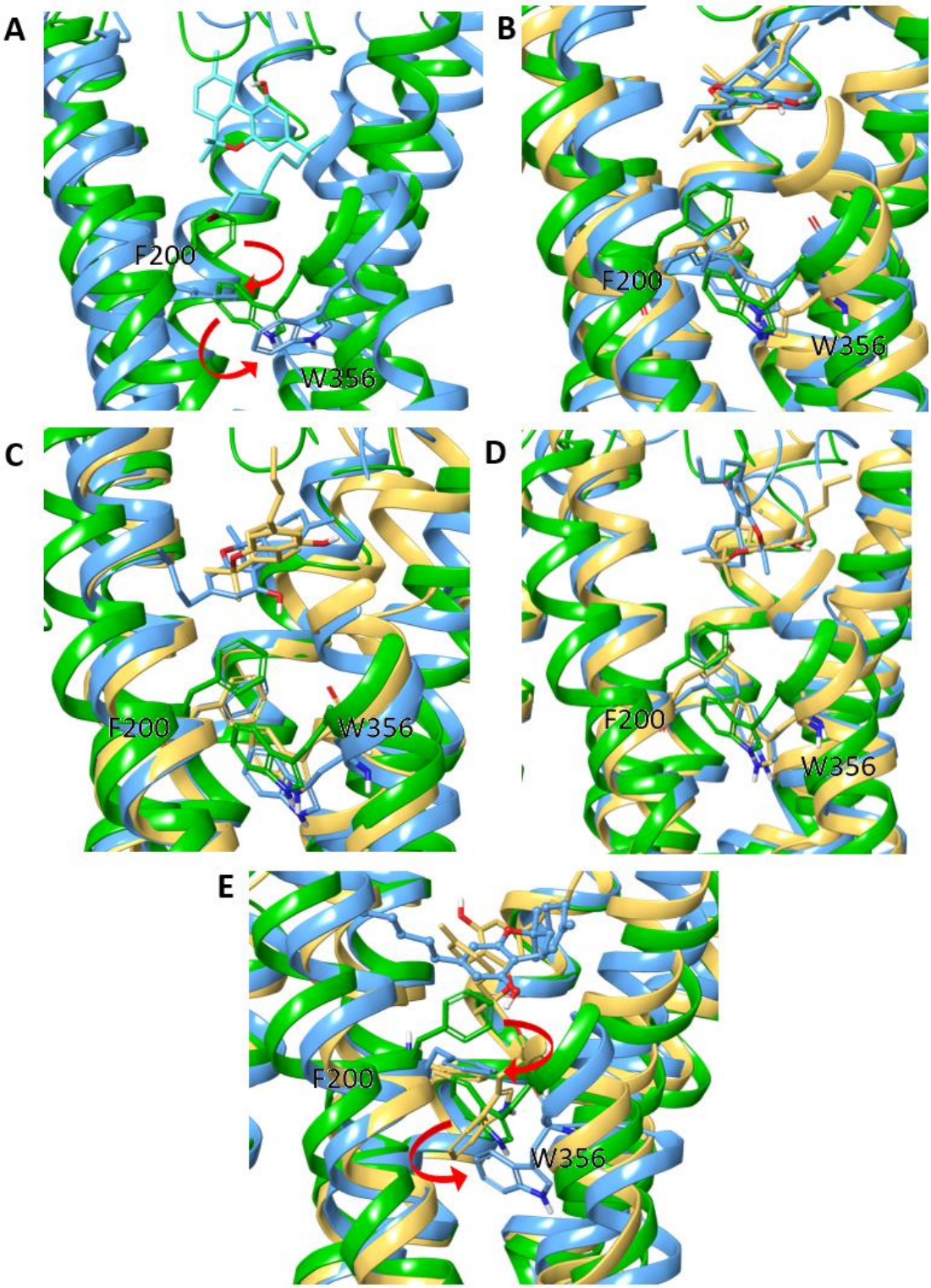

Figure 10 Superimposed structures of CB1 post MD simulation showing positions of toggle switches F200 and W356. Receptor with no ligand is represented in green, receptor with the ligand originating in the L-pocket in blue and receptor with the ligand originating in the S-pocket in orange. (A) **AM11542** (B) **THCV** (C) **THCB** (D) **THC** (E) **THCP**.

4.0 Conclusion

The hydrophobicity of the ligands was found to be essential for modelling and predicting binding affinity as the ligands enter CB_1 through the membrane. We developed a model for predicting binding affinity and activity of cannabinoids which can be used for further drug design efforts in the design of new cannabinoid-based ligands. We also determined that the binding pocket which the alkyl chain of cannabinoids occupy in the orthosteric site has a significant impact on their ability to activate the receptor and whether the ligands will act as agonists or antagonists. The ligands have to be able to interact with the toggle residues P200 and W356. How well they interact with the toggle residues also determines the degree of structural change in the receptor. Full agonists induct a larger conformation change in the toggle residues and subsequently helices 5, 6, and 7, move outwards to open the G-protein binding site. Partial agonists and antagonists were found to adopt an intermediate structure, where the binding site was neither fully open nor fully closed, which could be the cause of reduced activity, despite high binding affinity of ligands. This explanation is likely extendable to other GPCRs with partial agonist activity and a toggle switch. Combined this gives a more thorough understanding of how ligands interact with CB_1 and receptor activation, which in turn can be used to design and evaluate new cannabinoids.

Data Availability

All input and output files for the computational analyses can be obtained from the deposited data available from the Borealis Dataverse at http://doi.org/10.5683/SP3/3KJKR8.

Supporting Information

Full computational methodological details, additional tables and figures showing the binding mode of various ligands discussed are also provided. A video showing the comparison of the agonistic and non-agonistic binding modes of THC is also provided.

Competing Interests

JFT, DM and FS are all associated with Binary Star Research Services (BSRM). This generates an apparent conflict of interest. BSRM has no commercial interests in the subject of this manuscript and holds no intellectual property related to this manuscript. The interests of BSRM had no input into the methodology, research goals, or conclusions of this manuscript, and the company does not benefit from the publication of this manuscript nor did any of the authors receive any benefit from BSRM from its preparation or publication. BSRM provided no funding for this project.

Acknowledgements

The authors would like to thank the University of Windsor Faculty of Science Research Chair (to JFT) and the Natural Sciences and Engineering Research Council of Canada (2018-06338 to JFT) for providing funding. All authors wish to recognize that this work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca) and Compute/Calculation Canada, now the Digital Research Alliance of Canada (https://alliancecan.ca/en, qkh-310 to JFT). The other authors wish to recognize SM for preparing the table of contents graphic.

Author Contributions

Conceptualization, JFT and FS; Funding acquisition JFT; Investigation, DM, FS, SM; Methodology, All authors ; Visualization, DM, FS; Project administration, JFT; Graphical

abstract, SM; Supervision, JFT; Writing original draft, FS & DM; Writing – review and editing, All authors.

5.0 Reference List

1 Shahbazi, F., Grandi, V., Banerjee, A. & Trant, J. F. Cannabinoids and cannabinoid receptors: The story so far. *iScience* **23**, 101301, (2020).

2 Piscitelli, F. *et al.* Indole-2-carboxamides as allosteric modulators of the cannabinoid CB_1 receptor. *J. Med. Chem.* **55**, 5627-5631, (2012).

3 Shao, Z. *et al.* Structure of an allosteric modulator bound to the CB1 cannabinoid receptor. *Nat. Chem. Biol.* **15**, 1199-1205, (2019).

4 Kee, T. R. *et al.* The multifaceted functions of β-arrestins and their therapeutic potential in neurodegenerative diseases. *Experimental & molecular medicine* **56**, 129-141, (2024).

5 Adams, R., Hunt, M. & Clark, J. H. Structure of Cannabidiol, a Product Isolated from the Marihuana Extract of Minnesota Wild Hemp. I. *Journal of the American Chemical Society* **62**, 196-200, (1940).

6 Wollner, H. J., Matchett, J. R., Levine, J. & Loewe, S. Isolation of a physiologically active tetrahydrocannabinol from *Cannabis sativa* resin. *J. Am. Chem. Soc.* **64**, 26-29, (1942).

7 Banerjee, A. R., Hayward, J. J. & Trant, J. F. "Breaking bud": the effect of direct chemical modifications of phytocannabinoids on their bioavailability, physiological effects, and therapeutic potential. *Organic & Biomolecular Chemistry*, (2023).

8 Bow, E. W. & Rimoldi, J. M. The structure-function relationships of classical cannabinoids: CB1/CB2 modulation. *Perspect. Med. Chem.* **8**, 17-39, (2016).

9 Citti, C. *et al.* A novel phytocannabinoid isolated from *Cannabis sativa L*. with an *in vivo* cannabimimetic activity higher than Δ9-tetrahydrocannabinol: Δ9-Tetrahydrocannabiphorol. *Sci. Rep.* **9**, 20335, (2019).

10 Linciano, P. *et al.* Isolation of a high-affinity cannabinoid for the human CB1 receptor from a medicinal *Cannabis sativa* variety: Δ9-tetrahydrocannabutol, the butyl homologue of Δ9-tetrahydrocannabinol. *J. Nat. Prod.* **83**, 88-98, (2020).

11 Jung, S. W., Cho, A. E. & Yu, W. Exploring the ligand efficacy of cannabinoid receptor 1 (CB1) using molecular dynamics simulations. *Sci. Rep.* **8**, 13787, (2018).

12 Rosenthaler, S. *et al.* Differences in receptor binding affinity of several phytocannabinoids do not explain their effects on neural cell cultures. *Neurotoxicol. Teratol.* **46**, 49-56, (2014).

13 McPartland, J. M., Glass, M. & Pertwee, R. G. Meta-analysis of cannabinoid ligand binding affinity and receptor distribution: Interspecies differences. *Br. J. Pharmacol.* **152**, 583-593, (2007).

14 Iwamura, H., Suzuki, H., Ueda, Y., Kaya, T. & Inaba, T. *In vitro* and *in vivo* pharmacological characterization of JTE-907, a novel selective ligand for cannabinoid CB2 receptor. *J. Pharmacol. Exp. Ther.* **296**, 420-425, (2001).

15 Pagé, D. *et al.* Novel benzimidazole derivatives as selective CB_2 agonists. *Bioorg. Med. Chem. Lett.* **18**, 3695-3700, (2008).

16 Pertwee, R. G. The diverse CB1 and CB2 receptor pharmacology of three plant cannabinoids: delta9-tetrahydrocannabinol, cannabidiol and delta9-tetrahydrocannabivarin. *Br. J. Pharmacol.* **153**, 199-215, (2008).

17 Husni, A. S. *et al.* Evaluation of phytocannabinoids from high-potency *Cannabis sativa* using in vitro bioassays to determine structure-activity relationships for cannabinoid receptor 1 and cannabinoid receptor 2. *Med. Chem. Res.* **23**, 4295-4300, (2014).

18 Landrum, G. A. & Riniker, S. Combining IC50 or Ki Values from Different Sources Is a Source of Significant Noise. *Journal of chemical information and modeling*, (2024).

19 Magolan, J. *et al.* Processes for the preparation of ortho-allylated hydroxy aryl compounds. USA patent US20240150269A1 (2024).

20 Dutta, S., Selvam, B., Das, A. & Shukla, D. Mechanistic origin of partial agonism of tetrahydrocannabinol for cannabinoid receptors. *J. Biol. Chem.* **298**, 101764, (2022).

21 McPartland, J. M. *et al.* Affinity and efficacy studies of tetrahydrocannabinolic acid A at cannabinoid receptor types one and two. *Cannabis Cannabinoid Res.* **2**, 87-95, (2017).

22 Halgren, T. A. *et al.* Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **47**, 1750-1759, (2004).

23 Friesner, R. A. *et al.* Extra precision GLIDE: Docking and scoring incorporating a model of hydrophobic enclosure for protein− ligand complexes. *J. Med. Chem.* **49**, 6177-6196, (2006).

24 Lyne, P. D., Lamb, M. L. & Saeh, J. C. Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *J. Med. Chem.* **49**, 4805-4808, (2006).

25 Huang, S.-Y., Grinter, S. Z. & Zou, X. Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions. *Phys. Chem. Chem. Phys.* **12**, 12899-12908, (2010).

26 Knight, J. L. & Brooks Iii, C. L. Surveying implicit solvent models for estimating small molecule absolute hydration free energies. *J. Comput. Chem.* **32**, 2909-2923, (2011).

27 Mikulskis, P., Genheden, S. & Ryde, U. Effect of explicit water molecules on ligand-binding affinities calculated with the MM/GBSA approach. *J. Mol. Moldel.* **20**, 2273, (2014).

28 Mulakala, C. & Viswanadhan, V. N. Could MM-GBSA be accurate enough for calculation of absolute protein/ligand binding free energies? *J. Mol. Graphics Modell.* **46**, 41-51, (2013).

29 Forouzesh, N. & Mishra, N. An effective MM/GBSA protocol for absolute binding free energy calculations: A case study on SARS-CoV-2 spike protein and the human ACE2 receptor. *Molecules* **26**, (2021).

30 Ji, B. *et al.* Prediction of the binding affinities and selectivity for CB1 and CB2 ligands using homology modeling, molecular docking, molecular dynamics simulations, and MM-PBSA binding free energy calculations. *ACS Chem. Neurosci.* **11**, 1139-1158, (2020).

31 Hua, T. *et al.* Crystal structures of agonist-bound human cannabinoid receptor CB1. *Nature* **547**, 468-471, (2017).

32 Hua, T. *et al.* Activation and signaling mechanism revealed by cannabinoid receptor-G_i complex structures. *Cell* **180**, 1-11, (2020).

33 Drake, D. J. *et al.* Classical/nonclassical hybrid cannabinoids: Southern aliphatic chain-functionalized C-6β methyl, ethyl, and propyl analogues. *J. Med. Chem.* **41**, 3596-3608, (1998).

34 Tius, M. A. *et al.* Classical/non-classical cannabinoid hybrids; Stereochemical requirements for the southern hydroxyalkyl chain. *Life Sci.* **56**, 2007-2012, (1995).

35 Showalter, V. M., Compton, D. R., Martin, B. R. & Abood, M. E. Evaluation of binding in a transfected cell line expressing a peripheral cannabinoid receptor (CB2): identification of cannabinoid receptor subtype selective ligands. *J. Pharmacol. Exp. Ther.* **278**, 989-999, (1996).

36 Gareau, Y. *et al.* Structure activity relationships of tetrahydrocannabinol analogues on human cannabinoid receptors. *Bioorg. Med. Chem. Lett.* **6**, 189-194, (1996).

37 Rhee, M.-H. *et al.* Cannabinol derivatives: Binding to cannabinoid receptors and inhibition of adenylylcyclase. *J. Med. Chem.* **40**, 3228-3233, (1997).

38 Huffman, J. W. *et al.* Synthesis and pharmacology of a very potent cannabinoid lacking a phenolic hydroxyl with high affinity for the CB2 receptor. *J. Med. Chem.* **39**, 3875-3877, (1996).

39 Charalambous, A. *et al.* Pharmacological evaluation of halogenated Δ8-THC analogs. *Pharmacol. Biochem. Behav.* **40**, 509-512, (1991).

40 Huffman, J. W. *et al.* Structure-activity relationships for 1',1'-dimethylalkyl-Delta8-tetrahydrocannabinols. *Bioorg. Med. Chem.* **11**, 1397-1410, (2003).

41 Wiley, J. L., Marusich, J. A. & Huffman, J. W. Moving around the molecule: Relationship between chemical structure and *in vivo* activity of synthetic cannabinoids. *Life Sci.* **97**, 55-63, (2014).

42 Induced Fit Docking protocol, Glide, Prime (Schrödinger, LLC, New York, NY, , 2021).

43 Sotriffer, C. A. Accounting for induced-fit effects in docking: What is possible and what is not? *Curr. Top. Med. Chem.* **11**, 179-191, (2011).

44 Zhong, H., Tran, L. M. & Stang, J. L. Induced-fit docking studies of the active and inactive states of protein tyrosine kinases. *J. Mol. Graphics Modell.* **28**, 336-346, (2009).

45 Pei, Y. *et al.* Ligand-binding architecture of human CB2 cannabinoid receptor: evidence for receptor subtype-specific binding motif and modeling GPCR activation. *Chem. Biol.* **15**, 1207-1219, (2008).

46 Barnett-Norris, J. *et al.* Agonist alkyl tail interaction with cannabinoid CB1 receptor V6.43/I6.46 groove induces a helix 6 active conformation. *Int. J. Quantum Chem.* **88**, 76-86, (2002).

47 Reggio, P. H. Endocannabinoid binding to the cannabinoid receptors: what is known and what remains unknown. *Curr. Med. Chem.* **17**, 1468-1486, (2010).

48 Hurst, D. P. *et al.* A lipid pathway for ligand binding is necessary for a cannabinoid G protein-coupled receptor. *J. Biol. Chem.* **285**, 17954-17964, (2010).

49 Henchoz, Y. *et al.* Analytical tools for the physicochemical profiling of drug candidates to predict absorption/distribution. *Anal. Bioanal. Chem.* **394**, 707-729, (2009).

50 Waring, M. J. Lipophilicity in drug discovery. *Exp. Opin. Drug Discov.* **5**, 235-248, (2010).

51 Pallicer, J. *et al.* Evaluation of logPo/w values of drugs from some molecular structure calculation software. *ADMET & DMPK* **2**, (2014).

52 Mannhold, R., Poda, G. I., Ostermann, C. & Tetko, I. V. Calculation of molecular lipophilicity: State-of-the-art and comparison of log P methods on more than 96,000 compounds. *J. Pharm. Sci.* **98**, 861-893, (2009).

53 QikProp - 20109- Schrödinger, LLC: New York, NY (2019).

54 Pirhadi, S., Maghooli, K., Moteghaed, N. Y., Garshasbi, M. & Mousavirad, S. J. Biomarker Discovery by Imperialist Competitive Algorithm in Mass Spectrometry Data for Ovarian Cancer Prediction. *J Med Signals Sens* **11**, 108-119, (2021).

55 MATLAB and Statistics Toolbox Release (2012b,).

56 Huffman, J. W. *et al.* 3-(1',1'-Dimethylbutyl)-1-deoxy-delta8-THC and related compounds: Synthesis of selective ligands for the CB2 receptor. *Bioorg. Med. Chem.* **7**, 2905-2914, (1999).

57 Prandi, C., Blangetti, M., Namdar, D. & Koltai, H. Structure-activity relationship of cannabis derived compounds for the treatment of neuronal activity-related diseases. *Molecules* **23**, (2018).

58 Ellis, C. R., Kruhlak, N. L., Kim, M. T., Hawkins, E. G. & Stavitskaya, L. Predicting opioid receptor binding affinity of pharmacologically unclassified designer substances using molecular docking. *PLoS ONE* **13**, e0197734, (2018).

59 Gomez-Jeria, J. S., Soto-Morales, F., Rivas, J. & Sotomayor, A. A theoretical structure-affinity relationship study of some cannabinoid derivatives. *J. Chil. Chem. Soc.* **53**, 1382-1388, (2008).

60 McAllister, S. D. *et al.* Structural mimicry in class A G protein-coupled receptor rotamer toggle switches: The importance of the F3.36(201)/W6.48(357) interaction in cannabinoid CB1 receptor activation. *J. Biol. Chem.* **279**, 48024-48037, (2004).

61 Latek, D. *et al.* Modeling of ligand binding to G protein coupled receptors: cannabinoid CB1, CB2 and adrenergic beta 2 AR. *J. Mol. Model.* **17**, 2353-2366, (2011).

62 Shao, Z. *et al.* High-resolution crystal structure of the human CB1 cannabinoid receptor. *Nature* **540**, 602+, (2016).

63 Krishna Kumar, K. *et al.* Structure of a signaling cannabinoid receptor 1-G protein complex. *Cell* **176**, 448-458.e412, (2019).