Molecular Dynamics Simulations of Molecular Glue Degraders: targeting the DDB1-CDK12-Cyclin K complex for Selective Protein Degradation

Final Degree Project in Biochemistry and Molecular Biology

Ferran Jornet Margalef

Academic supervisor: Josep Gómez Alvarez, PhD

Professional supervisor: Jaime Rubio Martinez, PhD

In cooperation with: Modelling of Biological Systems and Drug Design

Tarragona, June 2025

The work presented in this Final Degree Project is based on the results obtained during my internship at the Modelling of Biological Systems and Drug Design group at the University of Barcelona, supervised by Dr. Jaime Rubio Martinez

ABSTRACT

Targeted protein degradation via molecular glue degraders offers a novel therapeutic strategy for modulating previously undruggable proteins such as Cyclin-dependent kinase 12 (CDK12). This thesis investigates whether molecular dynamics (MD) simulations can elucidate the structural and energetic principles underpinning the formation and stability of the DDB1–CDK12–Cyclin K complex mediated by different molecular glues.

Using high-resolution crystallographic structures and structure-based models, a total of eight systems, including seven ligands (CR8, SR-4835, Roscovitine, 919278, HQ461, LL-K12-14, and LL-K12-18) and one ligand-free complex, were subjected to 1 microsecond MD simulations. Structural stability was assessed via RMSD and RMSF analyses, while ligand binding affinities were estimated using MMGBSA and per-residue pairwise decomposition.

The results demonstrate that MD simulations preserve the integrity of the ternary complexes and reveal key residue interactions contributing to binding. Although high binding affinity generally correlated with better complex formation, it did not directly predict Cyclin K degradation efficiency, suggesting that dynamic and positional factors beyond static affinity contribute to molecular glue functionality.

Among the studied ligands, SR-4835 and LL-K12-18 emerged as the most promising candidates due to their favourable binding profiles and previously validated in vitro activity.

Keywords:
Molecular glues degraders, CDK12, DDB1, Cyclin K, Molecular dynamics

Table of contents

1. Introduction

Cancer is a non-transmissible disease characterized by an abrupt proliferation of a specific cell population. Cancer cells are the foundation of the disease, they initiate tumours by carrying the oncogenic and tumour suppressor mutations that define cancer as a genetic disease. The hyperproliferation combined with increased genetic instability create different clonal subpopulations, with different degrees of differentiation, proliferation, vascularity, inflammation and invasiveness[1]. Because of its complex mechanism of action and the impact altering the homeostatic functions of the organism, cancer ranks as a leading cause of death worldwide, just in front of cardiovascular diseases[2].

Various therapeutic approaches have been developed over the years as attempts to counter it. On one hand, the traditional cancer treatments include surgery, radiation therapy and chemotherapy, either individually or in combination. On the other hand, more advanced treatment approaches have been proposed with the objective of being more efficient, accurate and minimally invasive. These emerging therapies include exosome or nanoparticles-based therapy, stem cell therapy and gene therapy[3].

Within all those strategies, cyclin dependent kinases (CDKs) have emerged as promising targets for anticancer therapies. The development of potent inhibitors that selectively target specific members of the CDK family has proven challenging. The whole transcription-associated CDK family rely on ATP as a cofactor, which is attributed to structurally similar active sites[4]. To enhance the selectivity of compounds various strategies have been employed including development of covalent inhibitors[5] and targeted protein degradation approaches such as molecular glue degraders[6].

Targeted protein degradation represents a novel therapeutic paradigm that goes beyond inhibition by making us of the own cell protein disposal, the ubiquitin-proteosome pathway. In this context, drug-induced interactions involving E3 ubiquitin ligases have gained attention, particularly through the use of molecular glues.

1.1 Molecular Glues

Molecular glues (glues) are a class of small-molecule drugs that induce or stabilize interactions between proteins. In the context of ubiquitin ligase, drug-induced interactions can lead to protein degradation. The mechanism behind molecular glue degrades is based on the binding of substrate receptors of E3 ubiquitin ligases and recruitment of target proteins for their ubiquitination. This creates an emerging strategy for inactivation of therapeutic targets that are intractable by conventional pharmacological means[7].

The glues are not dependent on a druggable pocket on the target protein but instead exploit complementary protein-protein interfaces between receptor and target. By reprogramming the selectivity of the ubiquitin ligase, these molecules redirect the ligase to repeatedly catalyse ubiquitination of the target protein. In this way, the constraints of classical inhibitors can be circumvented and the repertoire of druggable proteins can be expanded[7].

Among the promising targets for molecular glue-induced degradation there is the protein complex formed by Damage-specific DNA binding protein 1 (DDB1), Cyclin-dependent kinase 12 (CDK12) and cyclin K, which serve as a test case for expanding the scope of molecular glue degraders.

1.2 DDB1

The Damage-specific DNA binding protein 1 (DDB1) is a multifunctional protein encoded by the *DDB1* gene in humans, located on chromosome 11q112-q13. It was first isolated as part of a heterodimeric complex (together with DDB2) that recognizes DNA lesions induced by ultraviolet light in the nucleotide excision repair (NER) pathway[8]. Beyond its role in DNA repair, DDB1 is a component of the Cul4A-RING ubiquitin E3-ligases (CURL4). It functions as an adaptor protein to link Cullin 4A (Cul4A) and CUL4-associated factors (DCAFs), allowing it to recruit and coordinate a wide variety of substrates for ubiquitination and degradation. This function is crucial for regulating cellular processes such as DNA repair, replication, chromatin remodelling and cell cycle control[9].

As mentioned before, The CUL4-DDB1 complex forms part of a larger E3 ubiquitin ligase assembly, named CRL4 (figure 1). This complex is built around Cul4A/B as a scaffold protein and DDB1 that acts as an adaptor for DCAFs. The DCAFs proteins bind to DDB1 and act as substrates receptors, providing specificity to the complex. The E3 ligase also includes a RING-finger protein (e.g. RBX1) that catalyses the ubiquitin transfer. The ubiquitination process involves the transfer of ubiquitin molecules from an E2 conjugating enzyme to lysine residues on the target protein, a step mediated by the E3 ligase. Consequently, the ubiquitination marks the proteins for degradation by the 26S proteosome, regulating the proteins involved in various pathways[10].

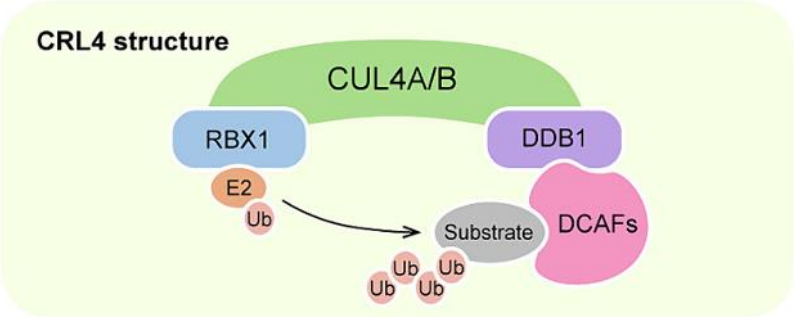


Figure 1. *Schematic representation of the CRL4 (Cullin-RING Ligase 4) E3 ubiquitin ligase complex. Adapted from Cullin-RING Ligase 4 in Cancer: Structure, Functions and Mechanism (Cheng J et al.)*

DDB1 is essential in targeting specific proteins for ubiquitination and subsequent degradation. This specificity makes it a powerful regulatory hub for controlling protein levels involved in essential cellular processes[11]. Importantly, this ability to guide proteins towards degradation has therapeutic potential in strategies that are aimed at eliminating target proteins that are otherwise considered undruggable. Exploiting DDB1 scaffolding function through molecular glues offers a promising approach for modulating the proteome in cancer[12].

1.3 Cyclin K

Cyclin K is a regulatory protein encoded by the *CCNK* gene in humans. It is a member of the transcription cyclin family and plays crucial roles in both cell cycle regulation and transcriptional control. Cyclin K forms complexes with cyclin-dependent kinases, most notably CDK12 and CDK13. These complexes phosphorylate the C-terminal domain (CTD) of RNA polymerase II[13]. It is also involved in the cell cycle regulation, regulating the assembly of the pre-replicative complex (pre-RC) during the G1 phase, crucial step for DNA replication and cell proliferation[14].

1.4 CDK12

CDK12 is a serine/threonine protein kinase that belongs to CDK family, participating in diversified biological functions such as transcription, post-transcriptional modification, translation and cell cycle progression. In humans, it is encoded by the *CDK12* gene and produces a protein comprising 1490 amino acids[15]. The protein consists of three major domains: an N-terminal domain rich in arginine-serine (RS) repeats, a catalytic kinase domain and a C-terminal domain. CDK12 forms an active complex with cyclin K, and upon phosphorylation by CDK-activating kinase at threonine 893 it becomes functionally active. This complex primarily targets the C-terminal domain (CTD) of RNA polymerase II (RNAPII), phosphorylating multiple serine residues to regulate transcription elongation. It is essential for the productive elongation of long, intron-rich genes that contain intronic polyadenylation sites (IPAs)[15,16].

A key role of CDK12 is the transcriptional regulation of genes involved in the DNA damage response (DDR), many of which are essential for maintaining genomic integrity. Through phosphorylation of RNAPII, CDK12 ensures the proper expression of DDR genes such as *BRAC1*, *RAD51* and *ATR*. Depletion or loss of function of CDK12 has been shown to downregulate these genes, leading to malfunction of DNA repair mechanism[15].

Another substantial role of CDK12 is controlling the G1 phase on cell cycle, where its expression peaks to prepare for DNA replication. It transcriptionally regulates critical replication genes such as *CDC6*, *CCNE1* and *CDT1*[17].

The expression of DDR genes which are tightly regulated by CDK12, is critical for the homologous recombination (HR) repair pathway. HR is a high-fidelity mechanism to resolve double-strand DNA breaks. When CDK12 is mutated or inactivated, as observed in certain cancers like high-grade serous ovarian cancer (HGSC) and BRAC-mutant triple-negative breast cancer (TNBC), the HR pathway becomes deficient. CDK12-deficient cells show increased use of IPAs resulting in truncated, non-functional mRNAs. This HR deficiency leads to genomic instability and increased sensitivity to DNA-damaging agents[18,19].

Therapeutically, HR-deficient tumours are susceptible to PARP inhibitors, which exploit their impaired repair capability. However, resistance to PARP inhibitors can emerge through compensatory pathways. In this context, a combination of CDK12 inhibitors with PARP inhibitors is being investigated as a strategy to overcome both primary and acquired resistance[20,21].

Selective inhibition of CDK12 has demonstrated therapeutic potential across various cancer types by interfering with transcriptional regulation of oncogenes, DNA repair genes,

and replication machinery. In HER2-positive breast cancer and papillary thyroid cancer, CDK12 is often co-amplified with *ERBB2*, driving activation of the Wnt/β-catenin pathway and promoting metastasis, effects that are reversed by CDK12 knockdown. In MYC-driven cancers, CDK12 inhibition results in synthetic lethality due to its role in MYC transcription regulation. Similarly, in Ewing's sarcoma, CDK12 inhibitors show synthetic lethality with the EWS/FLI fusion protein, especially when combined with PARP inhibitors. Even in CDK12 wild-type tumours, pharmacological inhibition impairs homologous recombination and induces cell cycle arrest, sensitizing cells to DNA-damaging therapies. These mechanisms position CDK12 as a versatile and powerful target in both monotherapy and combination strategies[15,22,23].

CDK12 was historically considered an undruggable target due to structural and functional similarities with other CDK, particularly the conserved ATP-binding pocket. This high degree of homology made achieving selective inhibition of CDK12 extremely challenging. Molecular glue degraders circumvent these limitations by redirecting E3 ligase activity to degrade cyclin K rather than inhibit the kinase function specifically. The degradation of the cyclin K leads to functional inactivation of CDK12 offering a selective therapeutic strategy[24].

1.5 Discovery of molecular glues for DDB1-CDK12-Cyclin K complex

M. Słabicki and colleagues confirmed the action of the CDK inhibitor (R)-CR8 (CR8) as a molecular glue degrader after correlating the data of clinical and preclinical drugs tested for cancers against the mRNA levels of 499 E3 ligase components[7]. CR8 cytotoxicity correlated with mRNA levels of the protein DDB1. As this DDB1-dependent cytotoxicity of CR8 suggested ubiquitin ligase-mediated degradation, quantitative proteome-wide mass spectrometry was performed after treating cells with CR8. From the total of eight thousand quantified proteins, only cyclin K showed a consistently decrease in abundance. They confirmed that CR8 triggers proteasomal degradation of cyclin K through the activity of DDB1-containing cullin-RING ubiquitin ligase[7].

Their results also provided genetic evidence for the involvement of a functional CUL4-RBX1-DDB1 ubiquitin ligase complex in mediating CR8 cytotoxicity. As no DCAF was identified in the (R)-CR8 resistance screens, they confirmed that the CR8-engaged CDK12-cyclin K complex binds directly to one of the CUL4-RBX1-DDB1 ligase components in the absence of a substrate receptor. The DDB1 domains involved in DCAF binding β-propeller domain A (BPA) and β-propeller domain C (BPC) were sufficient for drug-induced recruitment of CDK12-cyclin K. The DDB1 β-propeller domain B (BPB) that binds CUL4 was dispensable for the interaction[7].

Although CR8 stimulated binding between CDK12-Cyclin K and DDB1, a week interaction between CDK12-Cyclin K and DDB1 was still detectable in vitro in the absence of the drug. CR8 strengthened complex formation by 500- to 1000-fold estimated by isothermal titration calorimetry (ITC)[7].

All the data indicated that CR8-engaged CDK12-Cyclin K is recruited to the CUL4-RXB1-DDB1 ligase core through DDB1, and CR8 tightened the complex sufficiently to drive glue-induced degradation of cyclin K in the absence of a canonical DCAF substrate[7].

CR8 (figure 2) is a pleiotropic CDK inhibitor that is reported to bind CDK1, CDK2, CDK3, CDK5, CDK7, CDK9 and CDK12, yet in cells there is observed a selective destabilization of

cyclin K in the presence of the drug. Cyclin K is reported to associate with CDK9, CDK12 and CDK13, but only CDK13 is also recruited to DDB1 in presence of CR8, albeit with lower binding affinity. CDK13 has a sequence identity of 90,8% with CDK12 whereas CDK9 has a sequence identity of 45,5% where the key difference is in the C-terminal extension. Mutations in the C-terminal abolished basal binding between CDK12 and DDB1, whereas complex formation could still be facilitated by CR8[7,24].

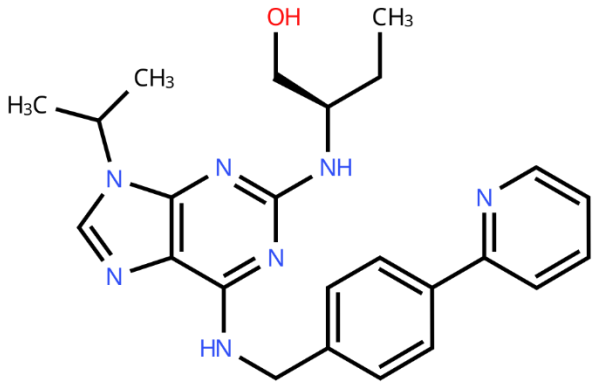


***Figure 2**: Chemical structure of SR-4835 ligand*

These results demonstrate that the selective CDK inhibitor CR8 induces specific protein-protein interactions between CDK12 or CDK13 and DDB1 and suggest that the C-terminal is not essential for drug-dependent kinase recruitment.

1.6 Design principles for molecular glue degraders

After confirming the mechanism of action of CR8 as a cyclin K degrader, other ligands were tested to find the key principles to design new degraders in a logical manner or *de novo*.

CR8 occupies the ATP-binding pocket which is the active site of CDK12 and forms discrete contacts with residues from the BPC domain of DDB1. From the explored purine-based compounds with diverse substituents (figure 3), there was a selective inhibitor that also showed gain-of-function in molecular glue activity, **SR-4835** (figure 4)[25].

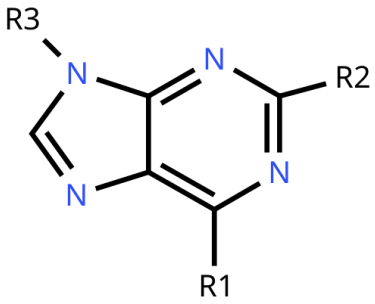


***Figure 3.** General purine-based scaffold used in the design of molecular glues. The core bicyclic purine structure supports various substituents at positions R1, R2 and R3.*

SR-4835 potently recruited DDB1 to CDK12-cyclin K (EC_{50} = 16 ± 1 nM) and consistently degraded cyclin K. The imidazole ring of the gluing moiety interacts with Arg928 (DDB1) and is additionally positioned to hydrogen bond with Tyr815 and Asp819 of CDK12. The high selectivity of this inhibitor compared to CR8 is explained because Tyr815 is specific to CDK12 and CDK13 only[25].

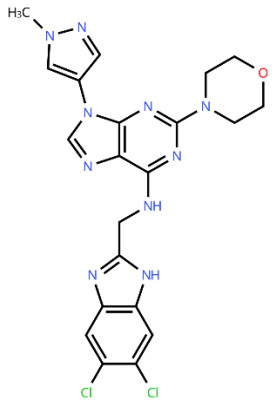


Figure 4. *Chemical structure of SR-4835 ligand*

Pompom Ghosh and colleagues confirmed that SR-4835 promotes cyclin degradation through the ubiquitin-proteosome system, same as CR8. SR-4835 is dependent on CUL4-RBX1-DDB1 ubiquitin ligase, and CDK12 is required for cyclin degradation[26].

SR-4835 also remains one of the few disclosed inhibitors with a set of pharmacokinetics properties suitable for in *vivo studies*. SR-4835 sensitizes triple-negative breast cancer cells to PARP inhibitors and DNA-damaging chemotherapies by reducing the expression of DNA damage response pathway, while also stimulating rapid tumour regression in patient-derived xenograft mouse models[26].

Roscovitine (figure 5), the parent compound of CR8, which lacks the 2-pyridil substituent, was also tested for the ability to drive complex formation. Although it had a binding affinity three times lower than CR8 and there was a degree of cyclin K degradation *in vitro*, this ligand was not capable of inducing cyclin K degradation in cells. The presence and correct orientation of the 2-pyridil moiety on the surface of CDK12 confers the gain-of-function activity that triggers cyclin K degradation[25].

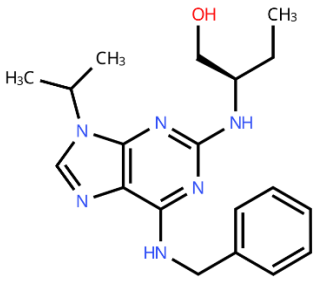


Figure 5. *Chemical structure of roscovitine ligand*

As SR-4835 is a more promising candidate as molecular glue because of its enhanced selectivity and drug-like properties, it was selected as starting point for optimization. Zemin Zhang and colleagues performed Structure-activity relationship (SAR) analysis to elucidate the critical binding site with DDB1 for optimal activity[27].

They performed dynamic simulations to provide insight into an additional targetable pocket to enhance the interaction. LL-K12-4 was formed by replacing the methylpyrazole group by

an ethyl group in R3 position. It successfully degraded cyclin K at a concentration of 5 nm in MDA-MB-231 cells and resulted in EC_{50} = 22.42 in cells[27].

Afterwards they tried to replace the R2 moiety with different linked pyridine rings. **LL-K12-K14** (figure 6) didn't result in increases in activity, although **LL-K12-18** (figure 7) emerged as a remarkably potent degrader of cyclin K, utilizing a diethylglycine derivate as the additional gluing moiety. Fitted kinetic parameters indicated that LL-K12-18 promote a higher affinity interaction between DDB1 and CDK12 compared to SR-4835, largely due to a slower dissociation rate[27].

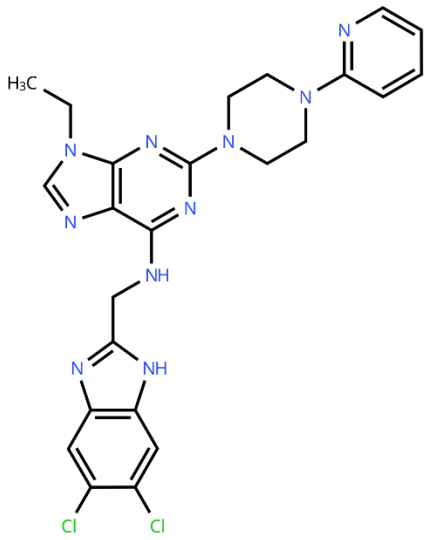


Figure 6. *Chemical structure of LL-K12-14 ligand*

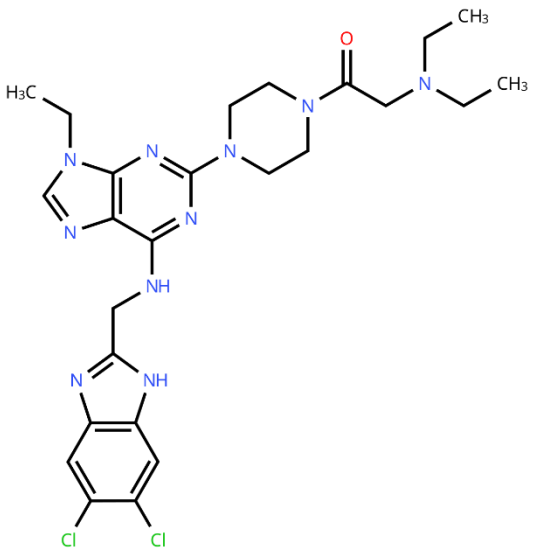


Figure 7. *Chemical structure of LL-K12-18 ligand*

More structurally dissimilar CDK12 inhibitors were investigated, focussing on **919278** ligand (figure 8). This inhibitor showed weak CDK12 binding affinity (5.6 µM) but pronounced downstream transcriptional effects. Despite its low molecular weight and lack of chemical similarity to CR8 or SR-4835, compound 919278 potently recruits DDB1 to CDK12-Cyclin K in vitro (EC_{50} = 38 ± 1 nM)[25].

The compound satisfies the key hinge contacts Glu814 akin to that of CR8 and interacts with CDK12 Tyr815 through hydrogen bonding. The isoindolinone gluing moiety protrudes out into the interface engaging DDB1 Arg928 via π-cation interactions. Hence, the diversity observed among DDB1-CDK12 glues extends beyond typical purine-based kinase inhibitor scaffolds[25].

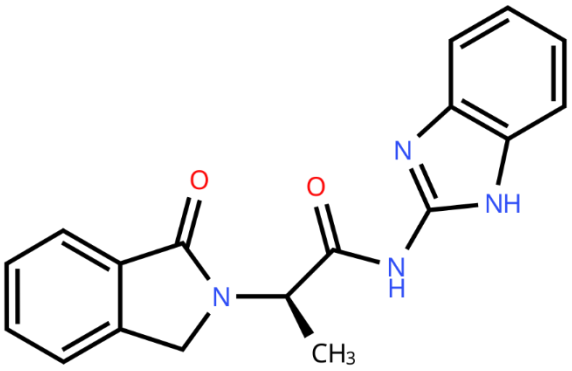


Figure 8. *Chemical structure of 919278 ligand*

Other small cyclin K degraders have been recently published such (e.g. HQ461, Z11, Z7 [28,29]) Especially **HQ461** (figure 9) revealed that all promote the formation of an analogous ternary complex and sustain the hinge hydrogen bonds and π-cation interactions with DDB1 Arg928. It also displays interactions with additional residues, forming a hydrogen-bond with Ile733 of CDK12.

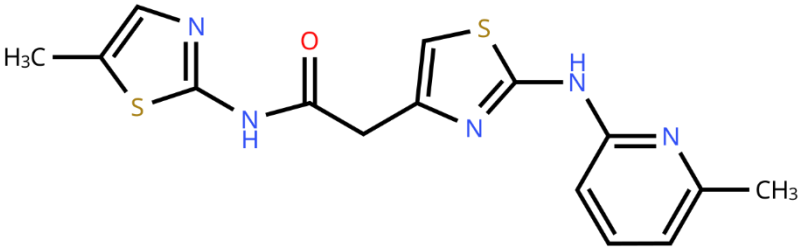


Figure 9. *Chemical structure of HQ461 ligand*

1.7 Crystallographic complex structure

The starting point of this study is the crystallized structures of the protein complex DDB1-CDK12-Cyclin K. A total of five structures from the RSCB Protein Data Bank were selected with different molecular glues bound. The protein structure entries are: 6TD3, 9FMR, 8BU9, 8BUA, 8BUG. All structures follow the same organization. The structures crystalized comprises a truncated version of DDB1 that lacks BPB domain (ΔBPB), CDK12 kinase subunit, cyclin K and the ligand.

In the structures, figure 10 serves as an example, CDK12 forms extensive protein-protein interactions with DDB1. The ligands are positioned in the active site of CDK12. The C-terminal extension of CDK12 binds the cleft between DDB1 domains.

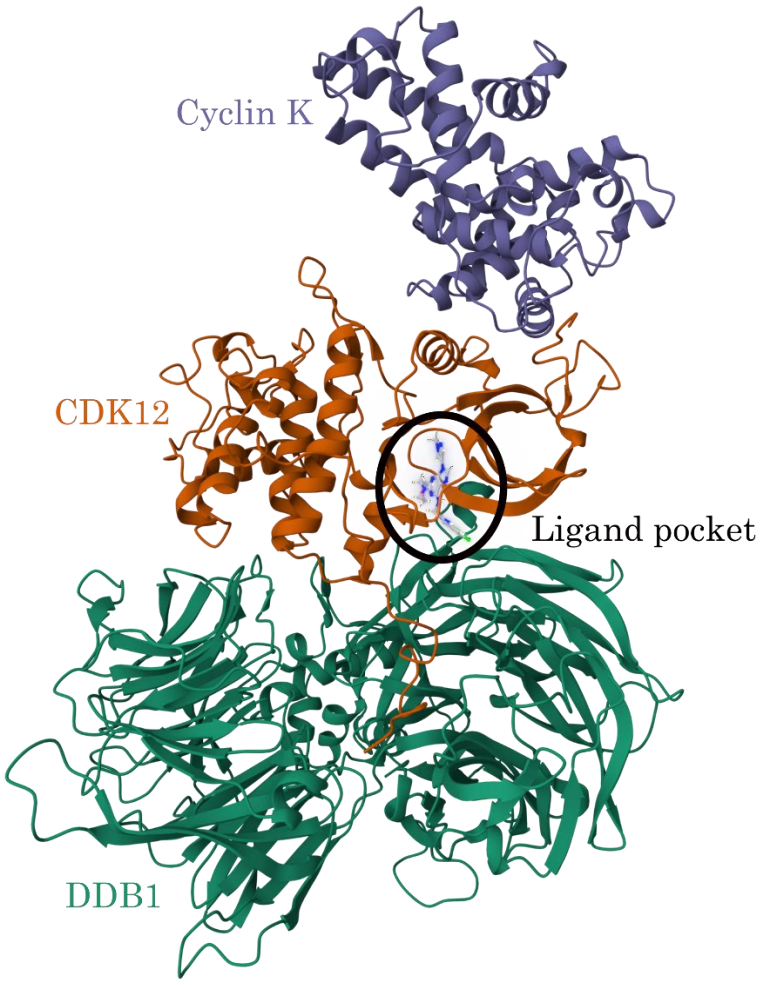


Figure 10. *Crystallographic structure of the complex formed by DDB1 (green), CDK12 (orange), cyclin K (purple) bound to a ligand (highlighted within the ligand pocket). Image generated using Mol 3D Viewer from the RCSB PDB.*

This study aims to perform molecular dynamics (MD) simulations on the DDB1–CDK12–Cyclin K complex bound to a ligand, in order to elucidate the underlying mechanisms of molecular glue-mediated degradation.

2. Hypothesis and Objectives

As introduced in the background, targeted protein degradation through molecular glue degraders has emerged as a promising therapeutic strategy for undruggable targets such as CDK12. Molecular glues function by stabilizing protein-protein interactions with E3 ligases and their substrates, promoting selective protein degradation. Among these, ligands capable of stabilize the protein complex DDB1-CDK12 and degrade cyclin K hold significant therapeutic potential in cancer treatments. Understanding the molecular basis behind the differential degradation efficiency of various ligands is essential for the rational design of more potent and selective degraders.

Hypothesis: molecular glues that form stronger interactions and induce more stable protein complex are the most effective in promoting cyclin K degradation

All this considered, the objectives of this study are:

1) To assess whether molecular dynamics simulations can maintain the structural integrity of the molecular glue degradation system throughout the simulation period.

2) To determine if differences in ligand binding affinity can serve as predictive indicators for their efficiency in both CDK12 inhibition and cyclin K degradation.

3) To discover the main interactions between ligand and the protein complex which constitute an effective molecular glue degrader

3. Methods – Computational Studies

3.1 System Preparation

The crystallographic structures of the complex of proteins DDB1-CDK12-CyclinK with different ligands were the starting point for the present study (table 1). Although the crystallographic structures were a multimeric complex formed from three heterotrimers, only a single heterotrimer was considered for the study. The selection was based on which one had the best-preserved C-terminal of CDK12, because it plays a crucial role mediating the interaction between CDK12 and DDB1. All protein structures were standardized to include the same CDK12 C-terminal length, extending up to residue ARG1169. This was done to ensure a consistent contribution of the CDK12 C-terminus to DDB1 binding across all models. A flexible linker, residues 395 to 404 (GGNGNSGEI), was added to covalently connect two subunits of DDB1 using PRIME function in Maestro. The linker was part of the crystallographic structures sequence, but the amino acids had missing coordinates in the PDB files. Subsequently, the complexes were prepared using the protein preparation wizard within Maestro (Schrödinger) that protonated at pH 7.0, added missing lateral chains and oriented sidechains. Ligands geometries were optimized at the Hartree-Fock (HF) level of theory with the 6-31G** basis set using Gaussian, yielding reliable starting points for molecular-charge determinations.

Next, the protein complex was centred in a cubic simulation box and solvated with rigid, 4-point optimal point charge (OPC) water molecules, ensuring a minimum distance of 15Å between the solute and the box walls. Water molecules closer than 1 Å to any protein were removed. Atoms of Na+ or Cl- were added to neutralize the system using the LEaP module of AMBER22. All calculations were performed using the ff19SB force field for proteins and gaff2 for ligands.

Table 1. *Summary of crystallographic structures used in the study, correlating the PDB access code with the ligand. (ligands marked with an asterisk (*) were derived from an existing crystallographic structure; ligand code marked with a slash (-) is for ligand not included in RCSB protein data base)*

PDB Code	Ligand	Ligand Code	Notes
6TD3	CR8	RC8	
9FMR	SR-4835	RMF	
	LL-K12-14	-	* Created from structure
	LL-K12-18	-	* Created from structure
8BU9	Roscovitine	RRC	
8BUA	919278	RVH	
8BUG	HQ461	RPW	

3.2 Energy minimization

The minimization step is required to eliminate possible steric impediments and to adapt the system to the force field, bringing the system near to a local energy minimum. The multistep minimization procedure took a total of 15000 steps, using the steepest descent method. The first 5000 steps only water molecules, ions and the amino acids without crystallographic coordinates were allowed to relax, meanwhile the rest of the atoms of the proteins were fixed and applying harmonic positional restriction of 5.0 kcal/mol· $Å^{-2}$. The next 5000 steps only the backbone of the protein was being restricted applying the same restriction. Finally, in the last steps all atoms were allowed to move without any restrictions.

3.3 Molecular Dynamics Simulations

After preparing the system for molecular dynamics, the first step was heating gradually the system to 300 K at a rate of 30 K every 20 ps. The atoms of the backbone of the protein were fixed with a harmonic positional restriction of 5.0 kcal/mol· $Å^{-2}$ using the Langevin thermostat algorithm with a collision frequency of 3 ps^{-1} under the NVT ensemble. Subsequently, a second stage of density equilibration was performed using the NPT ensemble to allow the solvent box to adjust to proper density, with a total of 24 ps simulation. Thereafter, 60 ns of simulation was performed reducing the restraint force constant every 20 ns. The harmonic positional restriction was reduced gradually 3.0 kcal/mol·$Å^2$, 1.0 kcal/mol·$Å^2$, 0.5 kcal/mol·$Å^2$. Finally, the production of molecular dynamics was performed allowing the system to have no restraints. A total simulation of 1 μs was conducted for each system under NVT conditions. The molecular dynamics were carried out using the pmemd module from AMBER21.

3.4 Root-Mean-Square Deviation and Root-Mean-Square Fluctuation

Root-mean-square deviation (RMSD) for all the molecular dynamics trajectories was computed using the cpptraj module from AMBER22 to assess the structural stability of the systems along time. RMSD was computed using the first structure of dynamics simulation as a reference[30]. The mathematical definition for the RMSD value at time t is:

$$RMSD(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N} ||r_i(t) - r_i^{ref}||^2}$$

Where N is the total number of atoms, $r_i(t)$ is the position of atom i at time t, and r_i^{ref} is its position in the reference structure.

Root Mean Square Fluctuation (RMSF) for all the molecular dynamics trajectories was calculated using the cpptraj module from AMBER22 as well. RMSF quantifies the flexibility of individual atoms measuring how much each atom fluctuates around its average position throughout the simulation[31]. High RMSF values indicate regions with grater mobility (e.g. protein loops or actives sites). The mathematical definition for atom i is:

$$RMSF_i = \sqrt{\langle ||r_i(t) - \langle r_i \rangle||^2 \rangle}$$

Where $\langle r_i \rangle$ is the time-averaged position for atom i.

3.5 MMGBSA

The molecular mechanics energies combined with generalized Born and surface area continuum solvation (MMGBSA) is a popular approach to predict ligand-binding affinities or the free energy of the binding of small ligands to biological macromolecules[32].

The binding free energy was computed using the MMGBSA procedures, implemented in AMBER22 package. In this method, the binding free energy is calculated according to the following equation:

$$\Delta G_{binding} = \Delta H^{gas} + \Delta G^{solv} - T\Delta S^{gas}$$

where ΔH^{gas} is the gas-phase interaction energy calculated by summing the internal energy, van der Waals (ΔH^{gas}_{vdW}) energy and electrostatic (ΔH^{gas}_{elec}) molecular mechanics energy. The term ΔG^{solv} is computed as the sum of polar (ΔG^{solv}_{polar}) and nonpolar terms ($\Delta G^{solv}_{nonpolar}$). The first term is calculated numerically by the generalized Born (GB) method, while the second one is calculated using the following equation:

$$\Delta G^{solv}_{nonpolar} = \gamma SASA + \beta$$

Where SAS is the solvent-accessible surface area calculated using the Linear Combinations of Pairwise Overlaps (LCPO) method. The values for the γ and β constants were set to 0.0072 kcal/mol·Å^2 and 0 kcal/mol·Å^2, respectively, using the igb = 2 option.

For each MMGBSA calculation the binding free energy was computed using the standard thermodynamic cycle, where the free energy of binding is determined as the difference between the free energy of the complex and the sum of the free energies of the separated components:

$$\Delta G_{binding} = G_{complex} - (G_{receptor} + G_{ligand})$$

The definitions of 'complex', 'receptor' and 'ligand' were adapted for each specific interaction of interest. The different components are denoted as follows: cyclin K (C), CDK12 (K), DDB1 (D) and the ligand (L). A summary of the different MMGBSA calculations and their corresponding components is provided in table x.

Table 2. *Summary of MMGBSA binding free energy calculations for protein-ligand and protein-protein interactions.*

MMGBSA 1	Ligand binding to protein complex ($CKD + L \leftrightarrow CDKL$)
MMGBSA 2	DDB1 binding to Cyclin K-CDK12-Ligand complex ($CKL + D \leftrightarrow CKDL$)
MMGBSA 3	Ligand-DDB1 binding to Cyclin K-CDK12 ($CK + DL \leftrightarrow CKDL$)

3.6 MMGBSA pairwise

The MMGBSA pairwise method is based on the decomposition of the total binding free energy into contributions from specific pairs of residues, atoms or groups within the complex. The decomposition is possible because the Generalized Born model used in MM/GBSA is formulated in a way that allows the calculation of solvation energies and interaction energies on a pairwise basis.

The pairwise decomposition enables the identification of which specific residues or interactions contribute most to the overall binding free energy, providing insight into the molecular determinants of binding affinity.

4. Results

4.1 Root-Mean-Square Deviation - RMSD

The Root Mean Square Deviation (RMSD) is a quantitative measure used to assess the average distance between corresponding atoms of two superimposed molecular structures. It is used in molecular dynamics to evaluate how much the structure deviates from a reference conformation over time.

In this case, RMSD values were calculated for each of the molecular dynamics systems using the initial structure as the reference conformation from which the simulation began.

To assess if the system reached structural stability, the overall trend of the RMSD function is analysed rather than focusing on specific numerical values. In molecular dynamics simulations, structural changes are expected at the beginning as the system adapts to the applied force field. Once this adaptation phase ends, the system may reach a state of convergence, where RMSD values stabilize and show minimal fluctuation over time.

The starting point of the structure coordinates is also important. Structures derived from experimental research, such as crystallographic coordinates, are generally more reliable as starting points for molecular dynamics simulations than those generated through homology modelling or computational design, as they better reflect the true atomic arrangement of the molecule.

In figure 11, the plots show the RMSD values over 1000 ns of molecular dynamics (MD) simulations for each of the ligands. For each system there are two RMSD values calculated, the first iteration was calculated using alpha carbons from all protein amino acids. The second iteration was calculated using the alpha carbons from the amino acids with a low value of RMSF. The second iteration is a better representation to assess whether if the structure reached structural stability because all the flexible parts of the proteins like loops and terminal ends are excluded from the calculations.

The flexible parts of the proteins distort the results by introducing artificial fluctuations that do not necessarily reflect the stability core of the protein-ligand complex. By excluding these regions, the second iteration provides a clearer and more reliable picture of the structural convergence of the system.

As seen in the plots (figure 11), most systems exhibit a noticeable reduction in RMSD fluctuations during the second iteration, indicating that the core structure of the protein remains more stable than what is suggested by the full-protein RMSD.

In almost all the systems, convergence in the RMSD for the second itineration can be observed, indicating that the protein backbone (excluding the flexible regions) reaches a stable conformation over time. A clear example is the system for 919278 ligand, where the RMSD stabilizes with minimal fluctuations after approximately 300 ns. Most other systems follow a similar trend, although some show slight perturbations. For instance, HQ461 displays a minor increase in fluctuation around 600 ns, and CR8 shows mild instability after 800 ns, though overall remain within an acceptable range of deviation, suggesting maintained structural integrity.

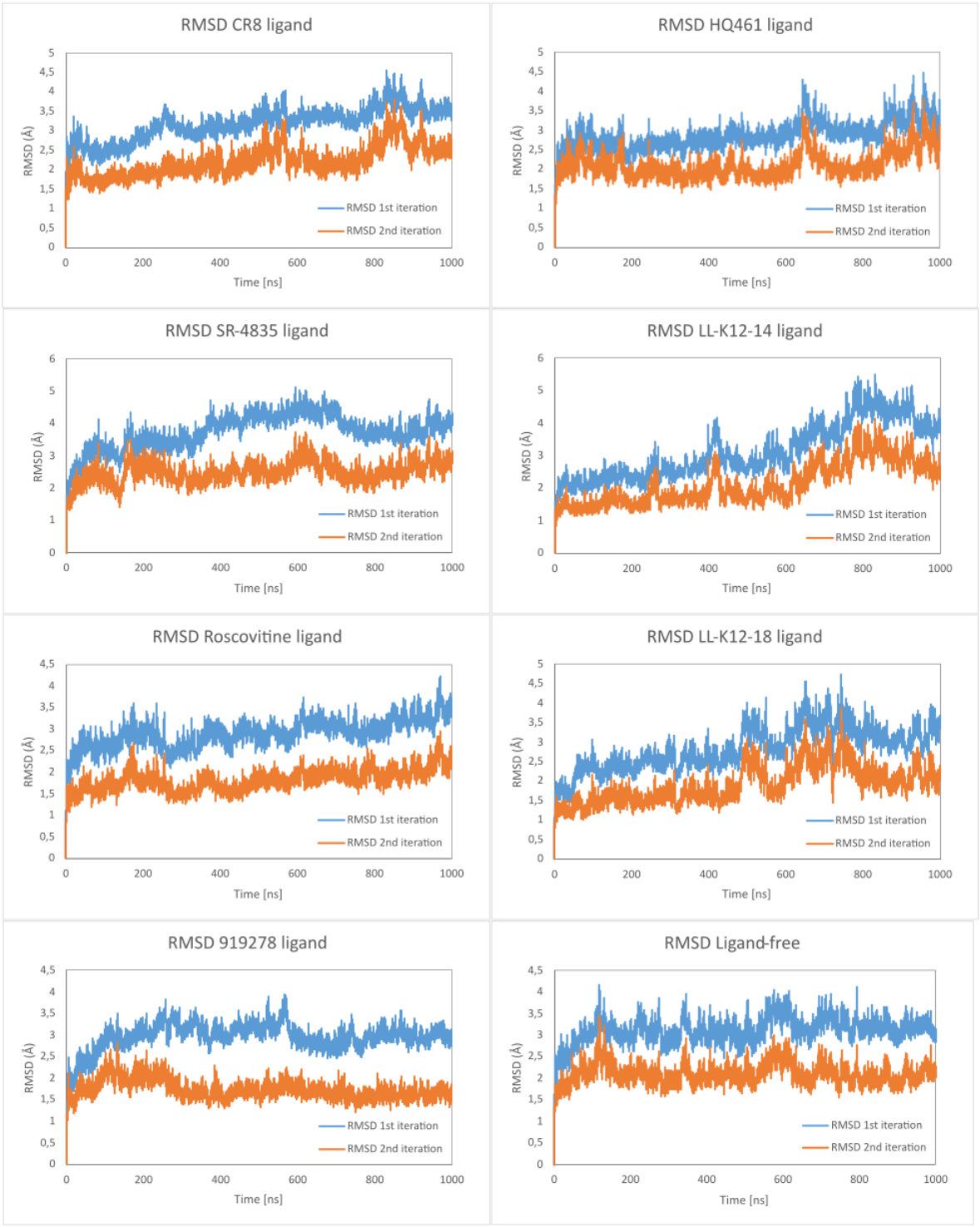


Figure 11. *RMSD plots for eight molecular systems: seven protein–ligand complexes (with ligands (A) CR8, (B) HQ461, (C) SR-4835, (D) LL-K12-14, (E) Roscovitine, (F) LL-K12-18 and (G) 919278) and one protein-only system ((H) ligand-free). Each plot represents the same simulation trajectory over time. The first RMSD calculation was performed using the alpha carbons of all amino acids, while the second used only the alpha carbons of residues with low RMSF values.*

On the other hand, for the systems of the ligands LL-K12-14 and LL-K12-18, it is noticeable that their system needed more time to adapt than the other ligands. This may be explained by the origin from the structure. Both of those ligands were modified from the original structure of the SR-4835 ligand (PDB ID: 9FMR) using MOE. No structural data was available

in the Protein Data Bank (PDB) for LL-K12-14 and LL-K12-18, which is why their atomic coordinate had to be generated computationally. Since the coordinates were not derived from experimental data, the initial binding poses may have been less accurate, leading to longer equilibration period in the MD simulation.

All systems demonstrated structural stability throughout the entire simulation period, with no significant deviations or fluctuations observed in their structural parameters, indicating consistent behaviour under the simulated conditions.

4.2 Root-Mean-Square Fluctuation - RMSF

The Root Mean Square Fluctuation (RMSF) is a measure used in molecular dynamics to assess the flexibility or mobility of individual atoms or residues over time. The RMSF values allow us to determine which parts of the proteins move the most during the simulation. High values of RMSF indicate more flexible regions such as loops or termini, meanwhile lower RMSF values are attributed to more stable and rigid regions.

RMSF plots for the system with CR8 ligand are presented as an example. In Figures 12, the RMSF values for the amino acids of the DDB1 protein are shown (RMSF for the rest of the systems in supplementary material 1 to 6). This protein contains a high number of short loops exposed to the solvent. These solvent-exposed loops are more likely to exhibit movement due to their lack of structural rigidity. Notably, the loop between residues 395 to 405 is highlighted, as this is the linker added between the two domains of the proteins.

In figure 13, the CDK12 protein contains some loops but it is stabilized through interactions with DDB1 and cyclin K. This stabilization limits the fluctuation of its residues. In contrast (figure 14), cyclin K appears to be the most flexible protein within the complex, with both termini forming flexible loops that contribute to higher mobility.

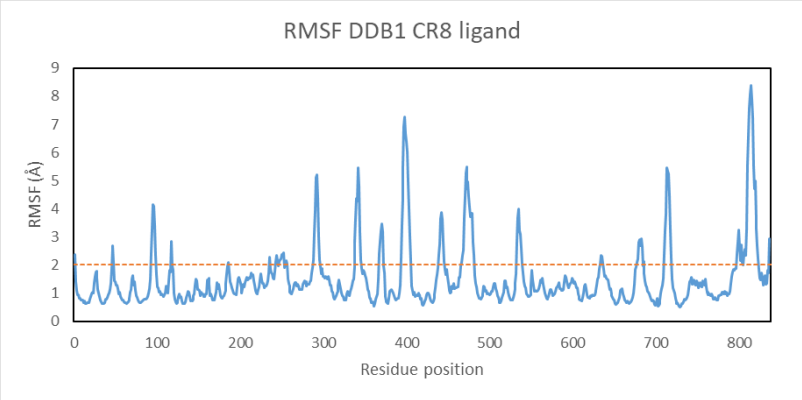


Figure 12. *RMSF of the DDB1 protein for the system with CR8 ligand. The orange dashed line indicates the threshold established for the second iteration calculation of RMSD calculation.*

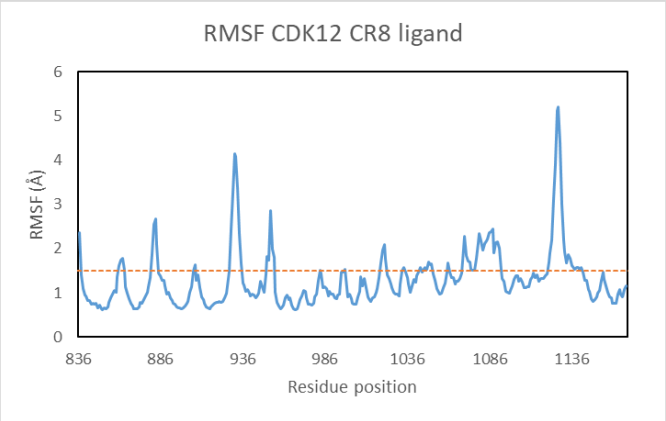


Figure 13. *RMSF of the CDK12 protein for the system with CR8 ligand. The orange dashed line indicates the threshold established for the second iteration calculation of RMSD calculation.*

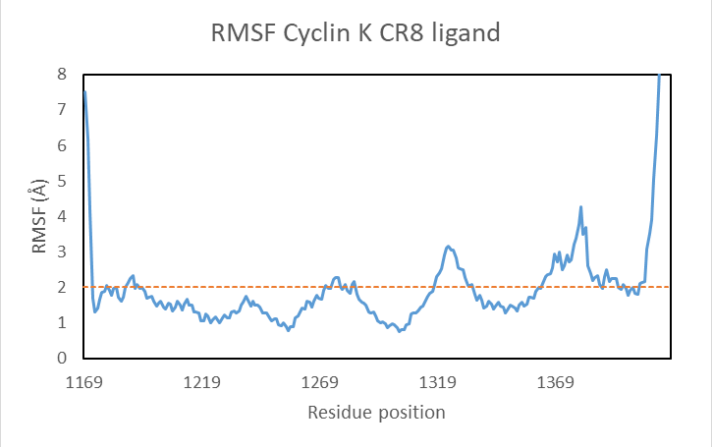


Figure 14. *RMSF of the cyclin K protein for the system with CR8 ligand. The orange dashed line indicates the threshold established for the second iteration calculation of RMSD calculation.*

4.3 MMGBSA

- MMGBSA 1 (DDB1-CDK12-Cyclin K + Ligand)

The MMGBSA is a computational method to estimate the binding free energy between molecules. Figure 15 shows the binding free energy of various ligands interacting with the DDB1-CDK12-Cyclin K protein complex. This method provides relative binding affinities between ligands, rather than absolute value binding free energy values.

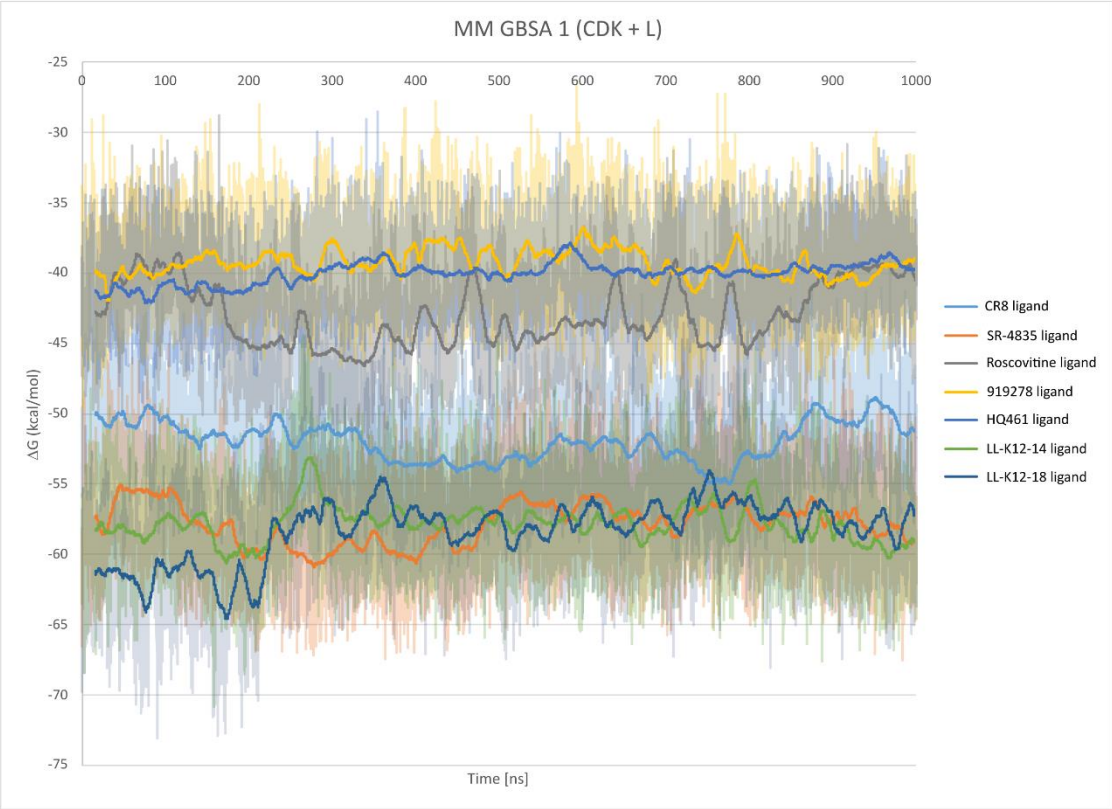


Figure 15. *MMGBSA binding free energy profiles over 1000 ns of molecular dynamics simulation for the DDB1-CDK12-Cyclin K complex with seven different ligands. The solid line represents the mean value of 200 frames (16 ns).*

By analysing the different functions over the time of simulation it is important that each system reaches energetic convergence, where the ligand binding affinity remains relatively stable and does not exhibit large fluctuations. Similar to structural stability, the system undergoes a series of change and adaptation to the force field, which is why significant fluctuations in the ligand binding free energy can be observed during the early stages of the simulation.

When comparing the binding affinity of the different ligands, those with a stronger binding affinity will have a more negative binding free energy. This is the case for three ligands that share a similar binding energy around -58 kcal/mol: SR-4835, LL-K12-14 and LL-K12-18. The next ligand with a more negative binding energy is CR8. On the other hand, the ligands HQ461, 919278, and Roscovitine have a less negative binding energy compared to the others.

To obtain representative binding free energy (ΔG) values from the molecular dynamics (MD) simulations, the mean and standard deviation of the MMGBSA energies were calculated over the last 10 ns, 20 ns, and 50 ns of each trajectory (table x). These time windows were selected to assess whether the systems had reached energetic convergence. By comparing the results across these different time intervals, it was observed that the ΔG values remained consistent within each system, regardless of the length of the averaging window. This stability indicates that the systems had reached convergence, and extending the averaging period did not significantly impact the estimated binding energies. Therefore, the ΔG values obtained can be considered reliable and representative of the final equilibrated state of each ligand-protein complex.

Table 3. *Comparison of the ligands MMGBSA energies over the last 10, 20 and 50 ns of simulation*

	ΔG (last 10 ns)		ΔG (last 20 ns)		ΔG (last 50 ns)	
Ligand	Mean	SD	Mean	SD	Mean	SD
CR8	-51,37	±2,75	-51,20	±2,73	-50,67	±2,80
SR-4835	-58,33	±2,85	-58,99	±2,92	-58,55	±3,06
Roscovitine	-41,11	±2,95	-40,24	±2,75	-40,02	±2,53
919278	-39,01	±2,93	-39,12	±2,79	-39,26	±2,78
HQ461	-39,72	±2,12	-39,59	±2,20	-39,26	±2,30
LL-K12_14	-59,14	±2,70	-59,12	±2,59	-59,57	±2,69
LL-K12-18	-56,95	±3,70	-56,79	±3,32	-57,72	±3,33

The representation in figure 16 confirms the overall tendencies observed in figure 15. There is a substantial difference between the binding free energy of the ligands SR-4835, LL-K12-14 and LL-K12-18 over the other ligands. The ligands that show most favourable binding energies, indicate stronger predicted interactions with DDB1 and CDK12.

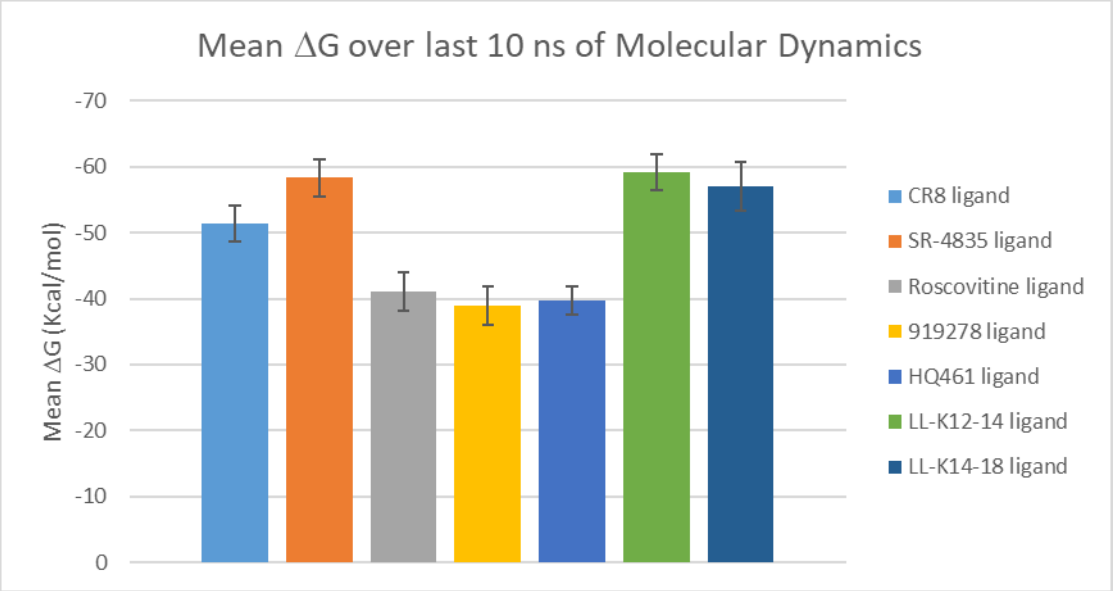


Figure 16. *Representation of the mean binding free energies (ΔG) of seven ligands calculated using the MMGBSA over the final 10 ns of molecular dynamics simulations. Error bars represent standard deviations.*

- MMGBSA 2 (CDK12-Cyclin K-ligand + DDB1) & MMGBSA 3 (CDK12-Cyclin K + Ligand-DDB1)

MMGBSA 2 was calculated to assess if the interaction between DDB1 and the rest of the complex was different depending on the ligand. In the dynamic simulation without any molecular glue the MMGBSA 2 was also calculated to provide information of how the DDB1 interacts with CDK12.

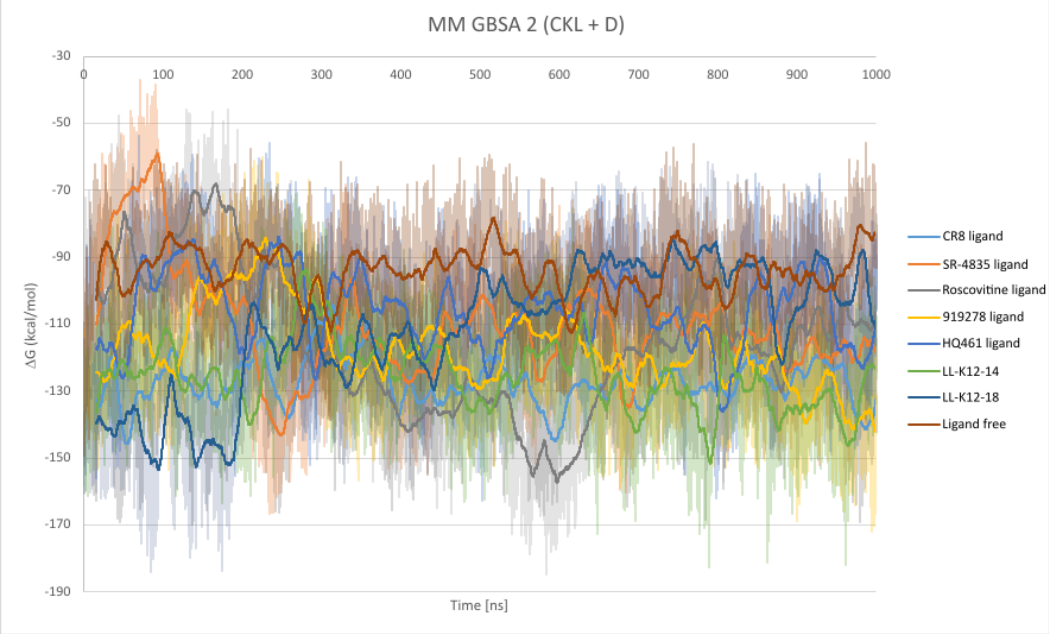


Figure 17. *MMGBSA binding free energy profiles over 1000 ns of molecular dynamics simulation for the CDK12-Cyclin K-ligand complex against the DDB1. The solid line represents the mean value of 200 frames (16 ns).*

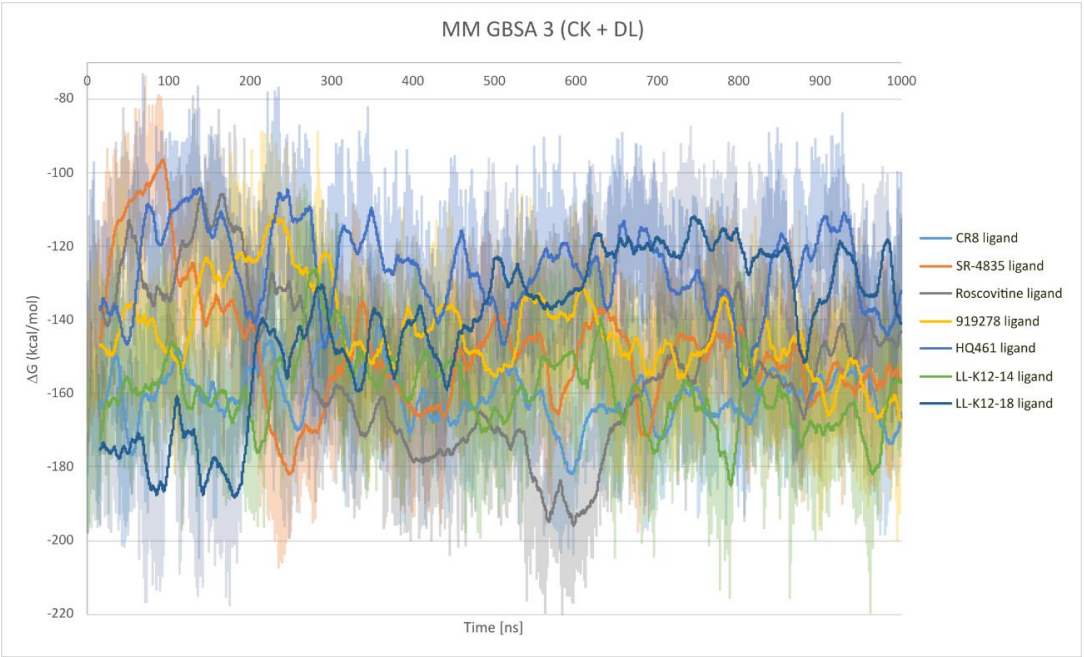


Figure 18. *MMGBSA binding free energy profiles over 1000 ns of molecular dynamics simulation for the CDK12-Cyclin K complex against the DDB1-ligand complex. The solid line represents the mean value of 200 frames (16 ns).*

Figure 17 represents the binding of the DDB1 to the complex CDK12-Cyclin K-ligand, while figure 18 represents the binding of DDB1-ligand to the complex CDK12-cyclin K. In both cases the systems do not reach an energetic convergence, the function of binding energy oscillates. Without reaching energetic convergence it is complicated to extract conclusions of how the different ligands affect the binding energy of the DDB1 protein to the complex. We will need more time of MD simulations to explore how the proteins behave when reaching energetic convergence.

The calculation for the ligand free system in MM GBSA 2 provides information about an interaction between the proteins CDK12 and DDB1 even without ligand. This interaction (around -90 kcal/mol in relative values) confirms some interactions between the kinase and the DDB1 protein. Part of this interaction is mediated by the C-terminal of the CDK12 that bind in the space between the two domains of DDB1.

4.4 MMGBSA pairwise

Per-residue MMGBSA pairwise decomposition of binding free energy contributions (ΔG) for all seven ligands was conducted to compare how individual amino acids contribute to their binding affinities. Figure 19 presents an example of MMGBSA pairwise per-residue decomposition for the CR8 ligand (other ligand MMGBSA pairwise in supplementary material 7 and 8). A threshold was established for all amino acids, and only those with a ΔG contribution lower than -1 kcal/mol were selected.

The calculations of MMGBSA pairwise were calculated for a pair of time frames: 960 to 980 ns and 980 to 1000 ns. Selecting the final part of molecular dynamics assures selecting only frames where both the protein complex and the ligand have reached structural stability and energetic convergence, respectively.

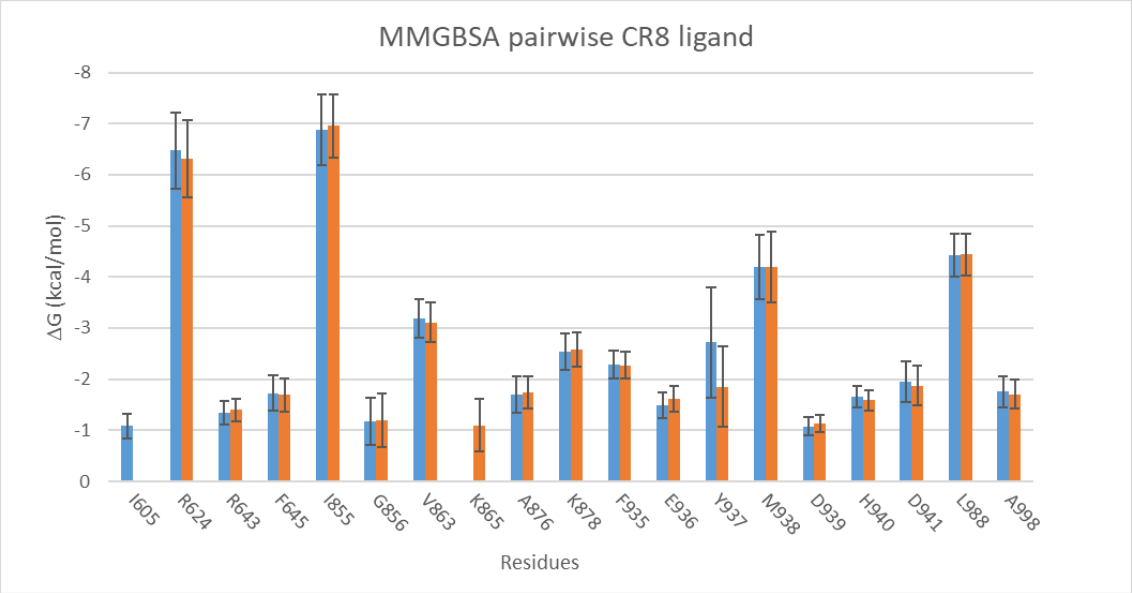


Figure 19. *Per-residue MMGBSA pairwise decomposition of binding free energy contributions (ΔG) for the CR8 ligands. The chart displays the ΔG values (in kcal/mol) for key interacting residues, with error bars representing standard deviations. Blue and orange bars correspond to values from two different time frames (blue 980-1000 ns and orange 960-980 ns) over the MD simulation. Residues 1-836 belong to DDB1, 837-1169 to CDK12.*

A total of 27 amino acids were found to interact with at least one of the ligands. To focus on the most significant interactions, only residues with an energy contribution lower than -2 kcal/mol were considered. These residues represent the strongest contributors to ligand binding. Among them, 15 amino acids showed a strong interaction with at least one of the ligands (figure 20).

The residue numbering corresponds to the concatenated protein complex in the following order: residues 1-836 belong to DDB1, 837-1169 to CDK12 and 1170-1417 to Cyclin K.

The ligands interact with DDB1 in a significant manner with amino acids ARG624, ARG643 and PHE645.Among these, ARG624 contributes the most to ligand binding, with the CR8 ligand showing the strongest interaction at this site.

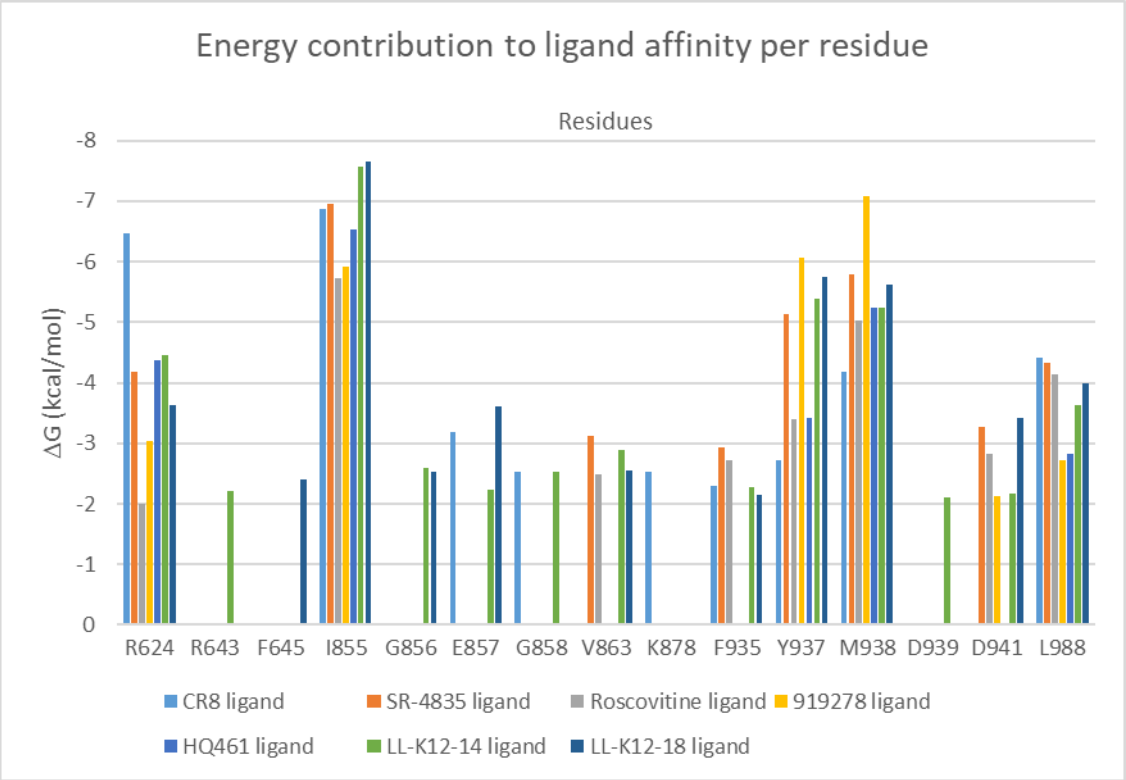


Figure 20. *Per-residue MMGBSA pairwise decomposition of binding free energy contributions (ΔG) for seven ligands. The graph highlights residues with notable contributions to ligand binding affinity (only ΔG values below -2kcal/mol).*

For CDK12 protein, several residues contribute to ligand binding: ILE855, GLY856, GLU857, VAL863, LYS878, PHE935, TYR937, MET938, ASP939, ASP941 and LEU988. Among those, ILE855, TYR937 and MET938 contribute the most to ligand binding.

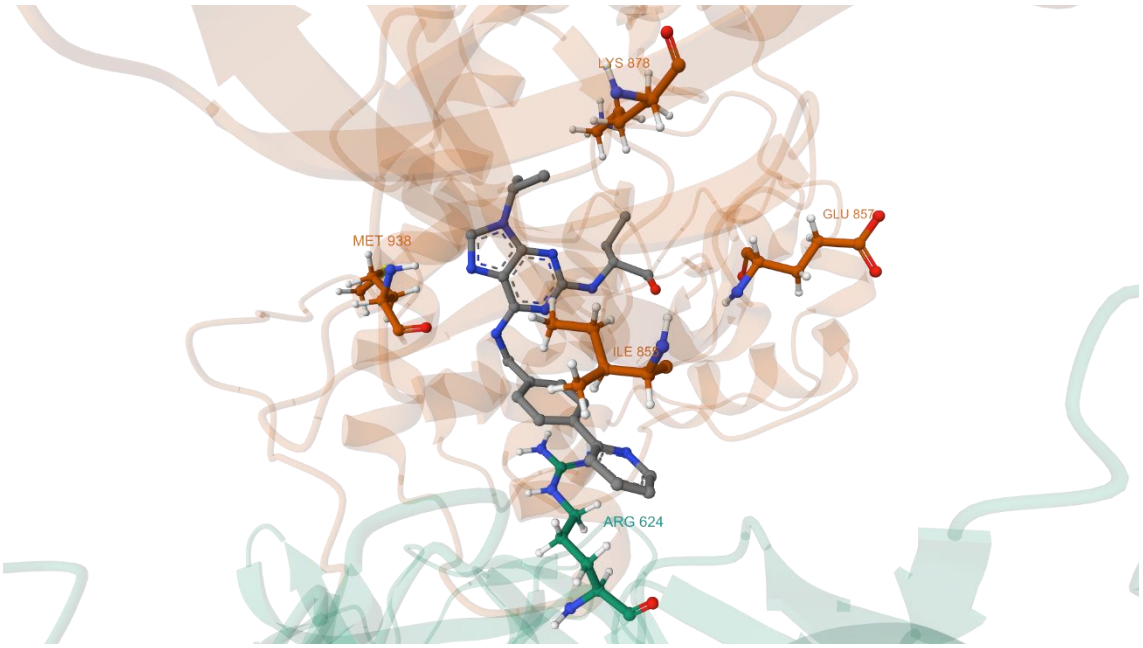


Figure 21. *Representation of the CR8 ligand bound to DDB1-CDK12 protein complex. The ligand is shown in grey, DDB1 is coloured in green and CDK12 in orange. Key amino acids residues contributing to the ligand interaction are highlighted and labelled. Image generated using Mol 3D Viewer from the RCSB PDB.*

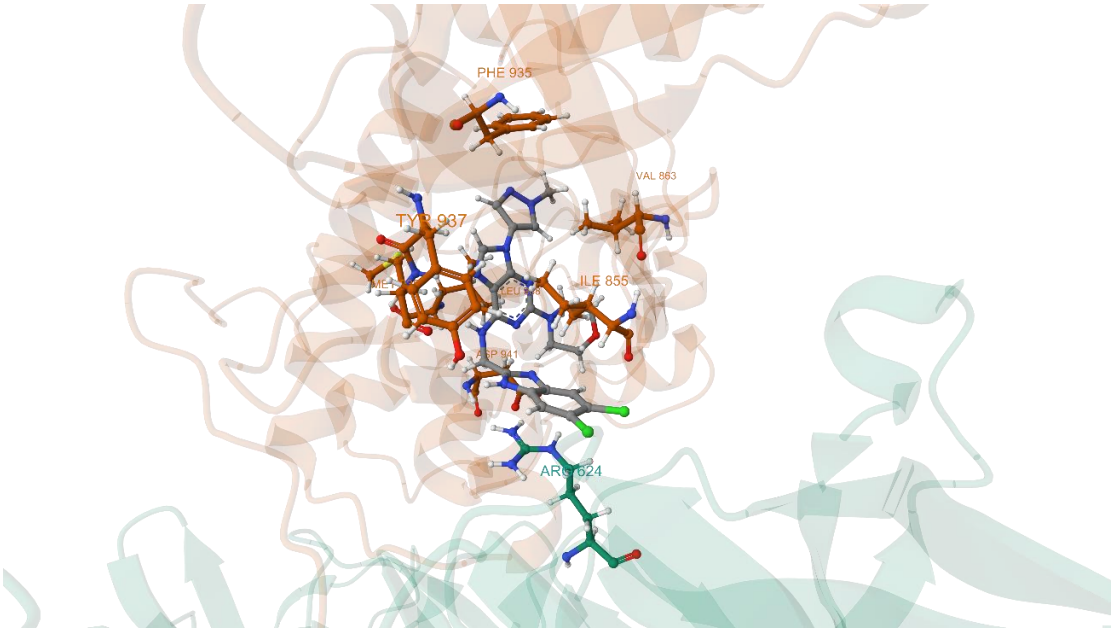


Figure 22. *Representation of the SR-4835 ligand bound to DDB1-CDK12 protein complex. The ligand is shown in grey, DDB1 is coloured in green and CDK12 in orange. Key amino acids residues contributing to the ligand interaction are highlighted and labelled. Image generated using Mol 3D Viewer from the RCSB PDB.*

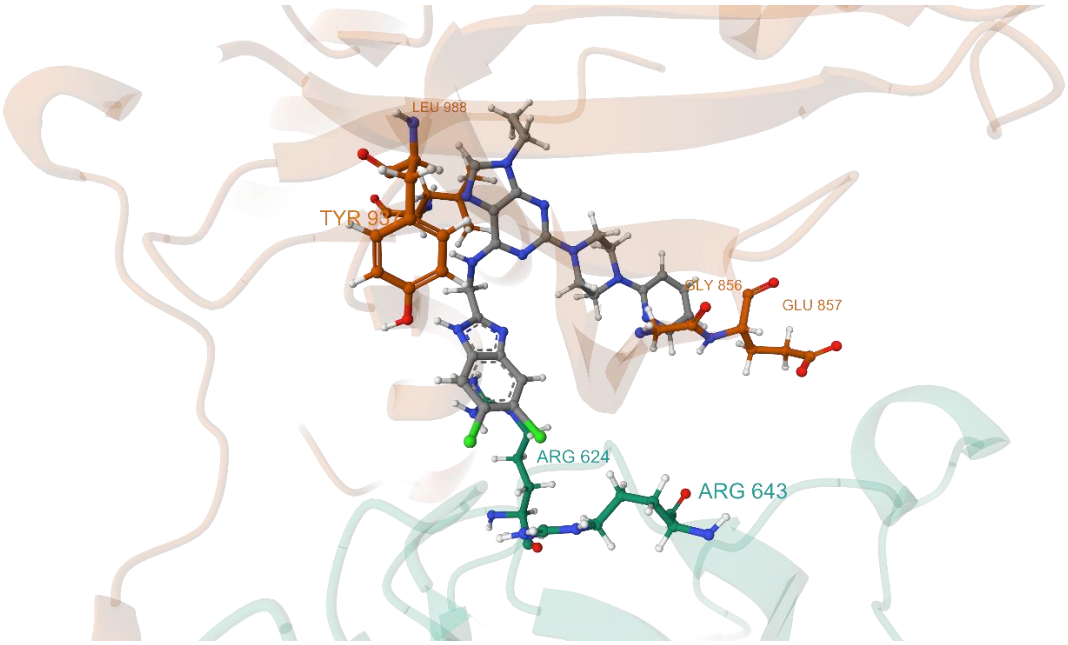


Figure 23. *Representation of the LL-K12-14 ligand bound to DDB1-CDK12 protein complex. The ligand is shown in grey, DDB1 is coloured in green and CDK12 in orange. Key amino acids residues contributing to the ligand interaction are highlighted and labelled. Image generated using Mol 3D Viewer from the RCSB PDB.*

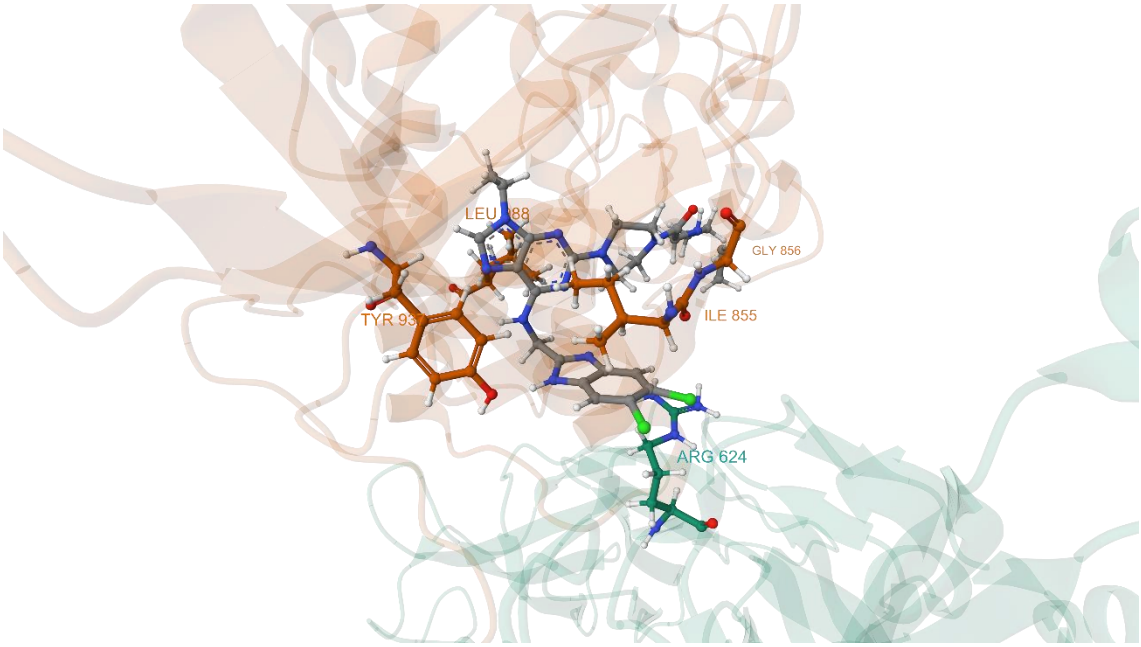


Figure 24. *Representation of the LL-K12-18 ligand bound to DDB1-CDK12 protein complex. The ligand is shown in grey, DDB1 is coloured in green and CDK12 in orange. Key amino acids residues contributing to the ligand interaction are highlighted and labelled. Image generated using Mol 3D Viewer from the RCSB PDB.*

5. Discussion

In this work we performed a computational investigation of seven candidate molecular glue degraders targeting the DDB1-CDK12-Cyclin K complex using molecular dynamics simulations and MMGBSA binding free energy calculations.

The main differences between the ligands are the side chains attached to the purine base they share. These side chains give each ligand characteristics that either benefit or hinder its role as a molecular glue degrader.

Within the group of ligands containing purine bases, two subgroups can be distinguished: CR8-like and SR-4835-like ligands. The first subgroup includes the CR8 ligand and the Roscovitine ligand. Roscovitine has a structure that is almost identical to CR8, except for a difference in the R1 group (figure 2 and 5). In Roscovitine, this functional group responsible for the binding to DDB1 is shorter, as it lacks one of the aromatic rings present in CR8. This structural difference results in a binding energy approximately 10 kcal/mol lower for the Roscovitine complex (figure 15).

Moreover, this gluing moiety plays a key role in anchoring DDB1 to the CDK12/cyclin K complex. It can be observed that the contribution of the main DDB1 amino acid involved in ligand binding, ARG624, is significantly reduced in this case (figure 20). This can lead to a weaker binding strength and a shorter binding time for the complex, potentially preventing the ubiquitination system from properly catalysing cyclin degradation. In conclusion, a bulky functional group at the R1 position improves the binding energy with DDB1, as it allows for better interaction with ARG 624.

Within the SR-4835-like category, the ligands SR-4835, LL-K12-14, and LL-K12-K18 are included. All ligands in this group show high interaction energy with the DDB1-CDK12-Cyclin K complex. However, cyclin K degradation has been reported to be much lower in *in vitro* studies with LL-K12-14 ligand, with only 29,6% of degradation after 24 hours of 50 nM treatment in MDA-MB-231 cells[27]. The structural difference between LL-K12-14 and LL-K12-K18 lies in their R2 moiety (figures 6 and 7). This R2 moiety is also different from that of their parent ligand, SR-4835 (figure 4). These ligands exploit an opening between the CDK12 protein surfaces to insert bulkier functional groups. These functional groups contribute key interactions with residues such as ILE 856, GLU 857, and LEU 998.

Even though there are differences in the energetic contributions of the individual residues between LL-K12-14 and LL-K12-K18, there is no simple explanation for the observed difference in cyclin K degradation levels.

As for ligands 919278 and HQ461, they show more structural differences compared to the others. Ligand 919278 still retains a purine-based functional group that grants it specificity to the CDK12 active site, but it does not follow the same three-substituent (R1, R2, R3) model (figure 8). This structural variation results in different energy contributions, with this ligand showing stronger interactions with residues TYR 937 and MET 938, compared to the other ligands (figure 20). However, the interaction with the main residue of contact of DDB1 with the ligand (ARG 624) is weaker, which could explain the lower formation of the ternary complex (EC_{50}=38$\pm$ 1 nm) observed in *in vitro* studies[25].

As for ligand HQ461 the free energy of binding with the complex has been one of the lowest. This also confirms the reduced complex formation observed in *in vitro* (EC_{50} = 43 $\pm$ 13 nm) assays [25], making it a poor candidate.

From figures 21 to 24 some key interactions can be determined. The lateral chain of ARG 624 establish π-cation interactions with the aromatic R1 moieties of purine-based ligands. Moreover, from highlighted residues from the ligand pocket, it is also observable that hydrophobic interactions are key to ligand affinity. A perfect example is the hydrophobic lateral chain of ILE 855, residue with the highest energy contribution to ligand binding energy.

An analysis of the specific hydrogen bonds involved in ligand binding would be important to assess the flexibility of the ligand. Hydrogen bonds stabilize and anchor the ligand, maintaining its geometry within the active site. Therefore, analysing both the hydrogen bonding and the flexibility of the ligands during the simulation could provide valuable information about differences in their conformations and how these affect their binding to the protein complex.

Among all the candidates, both SR-4835 and LL-K12-18 are the most potent. SR-4835 is a ligand with complex formation similar to CR8 (EC50 = 16 ± 1 nM)[25], but thanks to its pharmacokinetic properties, it is considered a better candidate. LL-K12-18 is the optimized version of SR-4835, which has been modified at the R2 and R3 substituents. Optimization of these substituents has led to improved degradation of cyclin K, reaching 96.4% degradation after 24 hours of treatment at 50 nM in MDA-MB-231 cells, achieving a 50-fold boost compared to SR-4835. An antiproliferative effect has also been observed in the MDA-MB-231 (EC50 = 0.37 nM) and MDA-MB-468 (EC50 = 0.03 nM) cell lines[27]. Based on these observations, the LL-K12-18 ligand emerges as the most promising candidate for molecular glue degrader therapies.

To further advance this line of investigation, future studies should aim to explore the dynamic positioning of cyclin K within the ternary complex and how it may vary depending on the ligand bound. It would be interesting to explore whether any of the ligand induce an allosteric change in CDK12 that propagates to the cyclin K and alters its orientation, thereby influencing its degradation. This could explain why certain ligands with strong binding affinity to the protein complex fail to function effectively as molecular glue degraders. This serves as a hypothesis to explain the currently observed lack of correlation between ligand binding affinity and successful DDB1-CDK12-Cyclin K complex-mediated degradation.

One additional approach to advance the research in this field could be the design of a pharmacophore based on key interactions found in this study, specially those from the most effective molecular glue degraders. By rationally designing new ligands using the structural features of the best performing degraders, the pharmacophore would serve as a blueprint for virtual screening of chemical libraries. The virtual screening would lead to faster identification of potential drug candidates, which could be evaluated *in silico* through molecular docking. The resulting compounds found could improve the binding affinity to the complex, enhance the degradation rate of cyclin K or have better pharmacokinetics properties.

5.1 Limitations of the study

The main limitation of this study lies in the constraints of current computational power. This *in silico* study is based on molecular dynamics which are simulations of the physical motion of atoms within molecules. MD simulate this movement by solving Newton's equation repeatedly. Although these simulations rely on a force field that is a set of mathematical functions and parameters to simulate the molecular behaviour, it allows us to observe intermolecular behaviours that are difficult to observe experimentally.

For each frame of molecular dynamics, the computer must solve the Newton's equation for every atom on the system. The duration of an MD simulation is directly proportional to the size of the system and the length of simulated time. In this study, simulating the complex of proteins DDB1-CDK12-Cyclin K for one microsecond required approximately two weeks of computation time.

As such, the computer power limits how long a system can be simulated. With greater computational resources, it would be possible to simulate the system for longer times, enabling us to observe other behaviours of the system that shorter times simulations cannot capture. For instance, longer simulation could reveal system behaviours such as the reaction equilibrium or if the ligands are not able to maintain the binding of the complex could be observed.

It is also important to note that the DDB1-CDK12-Cyclin K complex studied here represent only a part of the full ubiquitin ligase machinery. If the technological advances allow to perform molecular dynamics with bigger systems, it would be highly beneficial to model the complete CRL4 system. The simulation of the complete assembly could provide deeper insights, not only to understand the movement of the proteins bound to the molecular glue but also understand the special conformation of the whole system. This could help explain how different ligands modulate the positioning of cyclin K, and why in certain cases, the RXB1 protein fails to catalyse its ubiquitination.

6. Conclusion

This thesis set out to explore whether molecular dynamics simulations could help elucidate the underlying mechanisms behind molecular glue-induced degradation. A total of eight structures comprising seven different molecular glues (CR8, SR-4835, Roscovitine, 919278, HQ461, LL-K12-14 and LL-K12-18) and one ligand-free complex were simulated over one microsecond. The simulations provided insight into the overall behaviour of these molecular glue systems but also allowed each of the study's objectives to be addressed in detail:

- The RMSD and RMSF results confirmed that the complexes remained stable throughout the simulation time, with minimal structural drift, only specific loops and protein termini exhibit high fluctuation during the simulation. RMSD overall trend confirmed that the systems reached structural stability.

- The results aligned with known experimental data on complex binding affinity. Ligand such as CR8 and SR-4835-like showed stronger binding energies. These findings did not correlate with cyclin degradation. Although LL-K12-14 binding affinity was one of the strongest, its degradation activity is not sufficient.

  Therefore, the binding affinity of the molecular glue degraders to the protein complex is **not** directly proportional to the levels of degradation of cyclin K.

- The key interactions for an effective molecular glue degrader are an R1 moiety with a voluptuous aromatic group to interact through π-cation with ARG624 and hydrophobic groups through the R2 and R3 moieties to create hydrophobic interactions with lateral chains of amino acids in the binding pocket.

7. References

(1) Hanahan, D.; Weinberg, R. A. Hallmarks of Cancer: The next Generation. *Cell*. March 4, 2011, pp 646–674. https://doi.org/10.1016/j.cell.2011.02.013.

(2) Sung, H.; Ferlay, J.; Siegel, R. L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **2021**, *71* (3), 209–249. https://doi.org/10.3322/caac.21660.

(3) Kaur, R.; Bhardwaj, A.; Gupta, S. Cancer Treatment Therapies: Traditional to Modern Approaches to Combat Cancers. *Molecular Biology Reports*. Springer Science and Business Media B.V. November 1, 2023, pp 9663–9676. https://doi.org/10.1007/s11033-023-08809-3.

(4) Schmitz, M.; Kaltheuner, I. H.; Anand, K.; Düster, R.; Moecking, J.; Monastyrskyi, A.; Duckett, D. R.; Roush, W. R.; Geyer, M. The Reversible Inhibitor SR-4835 Binds Cdk12/Cyclin K in a Noncanonical G-Loop Conformation. *Journal of Biological Chemistry* **2024**, *300* (1), 105501. https://doi.org/10.1016/j.jbc.2023.105501.

(5) Zhang, T.; Kwiatkowski, N.; Olson, C. M.; Dixon-Clarke, S. E.; Abraham, B. J.; Greifenberg, A. K.; Ficarro, S. B.; Elkins, J. M.; Liang, Y.; Hannett, N. M.; Manz, T.; Hao, M.; Bartkowiak, B.; Greenleaf, A. L.; Marto, J. A.; Geyer, M.; Bullock, A. N.; Young, R. A.; Gray, N. S. Covalent Targeting of Remote Cysteine Residues to Develop CDK12 and CDK13 Inhibitors. *Nat Chem Biol* **2016**, *12* (10), 876–884. https://doi.org/10.1038/nchembio.2166.

(6) Toriki, E. S.; Papatzimas, J. W.; Nishikawa, K.; Dovala, D.; Frank, A. O.; Hesse, M. J.; Dankova, D.; Song, J.-G.; Bruce-Smythe, M.; Struble, H.; Garcia, F. J.; Brittain, S. M.; Kile, A. C.; McGregor, L. M.; McKenna, J. M.; Tallarico, J. A.; Schirle, M.; Nomura, D. K. Rational Chemical Design of Molecular Glue Degraders. *ACS Cent Sci* **2023**, *9* (5), 915–926. https://doi.org/10.1021/acscentsci.2c01317.

(7) Słabicki, M.; Kozicka, Z.; Petzold, G.; Li, Y.-D.; Manojkumar, M.; Bunker, R. D.; Donovan, K. A.; Sievers, Q. L.; Koeppel, J.; Suchyta, D.; Sperling, A. S.; Fink, E. C.; Gasser, J. A.; Wang, L. R.; Corsello, S. M.; Sellar, R. S.; Jan, M.; Gillingham, D.; Scholl, C.; Fröhling, S.; Golub, T. R.; Fischer, E. S.; Thomä, N. H.; Ebert, B. L. The CDK Inhibitor CR8 Acts as a Molecular Glue Degrader That Depletes Cyclin K. *Nature* **2020**, *585* (7824), 293–297. https://doi.org/10.1038/s41586-020-2374-x.

(8) Chu, G.; Yang, W. Here Comes the Sun: Recognition of UV-Damaged DNA. *Cell* **2008**, *135* (7), 1172–1174. https://doi.org/10.1016/j.cell.2008.12.015.

(9) Iovine, B.; Iannella, M. L.; Bevilacqua, M. A. Damage-Specific DNA Binding Protein 1 (DDB1): A Protein with a Wide Range of Functions. *International Journal of Biochemistry and Cell Biology*. Elsevier Ltd 2011, pp 1664–1667. https://doi.org/10.1016/j.biocel.2011.09.001.

(10) Cheng, J.; Bin, X.; Tang, Z. Cullin-RING Ligase 4 in Cancer: Structure, Functions, and Mechanisms. *Biochimica et Biophysica Acta - Reviews on Cancer*. Elsevier B.V. September 1, 2024. https://doi.org/10.1016/j.bbcan.2024.189169.

(11) Wu, R.; Wu, X.; Zou, L.; Zhou, L.; Mao, Y. DDB1 Regulates the Activation-Induced Apoptosis of T Cells via Downregulating the Expression of Histone Methyltransferase SETD7. *Medical Oncology* **2023**, *40* (5), 146. https://doi.org/10.1007/s12032-023-02015-8.

(12) Meyers, M.; Cismoski, S.; Panidapu, A.; Chie-Leon, B.; Nomura, D. K. Targeted Protein Degradation through Recruitment of the CUL4 Complex Adaptor Protein DDB1. *ACS Chem Biol* **2024**, *19* (1), 58–68. https://doi.org/10.1021/acschembio.3c00487.

(13) Lim, S.; Kaldis, P. Cdks, Cyclins and CKIs: Roles beyond Cell Cycle Regulation. *Development* **2013**, *140* (15), 3079–3093. https://doi.org/10.1242/dev.091744.

(14) Lei, T.; Zhang, P.; Zhang, X.; Xiao, X.; Zhang, J.; Qiu, T.; Dai, Q.; Zhang, Y.; Min, L.; Li, Q.; Yin, R.; Ding, P.; Li, N.; Qu, Y.; Mu, D.; Qin, J.; Zhu, X.; Xiao, Z.-X.; Li, Q. Cyclin K Regulates Prereplicative Complex Assembly to Promote Mammalian Cell Proliferation. *Nat Commun* **2018**, *9* (1), 1876. https://doi.org/10.1038/s41467-018-04258-w.

(15) Emadi, F.; Teo, T.; Rahaman, M. H.; Wang, S. CDK12: A Potential Therapeutic Target in Cancer. *Drug Discov Today* **2020**, *25* (12), 2257–2267. https://doi.org/10.1016/j.drudis.2020.09.035.

(16) Dubbury, S. J.; Boutz, P. L.; Sharp, P. A. CDK12 Regulates DNA Repair Genes by Suppressing Intronic Polyadenylation. *Nature* **2018**, *564* (7734), 141–145. https://doi.org/10.1038/s41586-018-0758-y.

(17) Chirackal Manavalan, A. P.; Pilarova, K.; Kluge, M.; Bartholomeeusen, K.; Rajecky, M.; Oppelt, J.; Khirsariya, P.; Paruch, K.; Krejci, L.; Friedel, C. C.; Blazek, D. CDK12 Controls G1/S Progression by Regulating RNAPII Processivity at Core DNA Replication Genes. *EMBO Rep* **2019**, *20* (9). https://doi.org/10.15252/embr.201847592.

(18) Krajewska, M.; Dries, R.; Grassetti, A. V.; Dust, S.; Gao, Y.; Huang, H.; Sharma, B.; Day, D. S.; Kwiatkowski, N.; Pomaville, M.; Dodd, O.; Chipumuro, E.; Zhang, T.; Greenleaf, A. L.; Yuan, G.-C.; Gray, N. S.; Young, R. A.; Geyer, M.; Gerber, S. A.; George, R. E. CDK12 Loss in Cancer Cells Affects DNA Damage Response Genes through Premature Cleavage and Polyadenylation. *Nat Commun* **2019**, *10* (1), 1757. https://doi.org/10.1038/s41467-019-09703-y.

(19) Ekumi, K. M.; Paculova, H.; Lenasi, T.; Pospichalova, V.; Bösken, C. A.; Rybarikova, J.; Bryja, V.; Geyer, M.; Blazek, D.; Barboric, M. Ovarian Carcinoma CDK12 Mutations Misregulate Expression of DNA Repair Genes via Deficient Formation and Function of the Cdk12/CycK Complex. *Nucleic Acids Res* **2015**, *43* (5), 2575–2589. https://doi.org/10.1093/nar/gkv101.

(20) Xie, Y.; Xiao, D.; Li, D.; Peng, M.; Peng, W.; Duan, H.; Yang, X. Combined Strategies with PARP Inhibitors for the Treatment of BRCA Wide Type Cancer. *Front Oncol* **2024**, *14*. https://doi.org/10.3389/fonc.2024.1441222.

(21) Johnson, S. F.; Cruz, C.; Greifenberg, A. K.; Dust, S.; Stover, D. G.; Chi, D.; Primack, B.; Cao, S.; Bernhardy, A. J.; Coulson, R.; Lazaro, J.-B.; Kochupurakkal, B.; Sun, H.;

Unitt, C.; Moreau, L. A.; Sarosiek, K. A.; Scaltriti, M.; Juric, D.; Baselga, J.; Richardson, A. L.; Rodig, S. J.; D'Andrea, A. D.; Balmaña, J.; Johnson, N.; Geyer, M.; Serra, V.; Lim, E.; Shapiro, G. I. CDK12 Inhibition Reverses De Novo and Acquired PARP Inhibitor Resistance in BRCA Wild-Type and Mutated Models of Triple-Negative Breast Cancer. *Cell Rep* **2016**, *17* (9), 2367–2381. https://doi.org/10.1016/j.celrep.2016.10.077.

(22) Quereda, V.; Bayle, S.; Vena, F.; Frydman, S. M.; Monastyrskyi, A.; Roush, W. R.; Duckett, D. R. Therapeutic Targeting of CDK12/CDK13 in Triple-Negative Breast Cancer. *Cancer Cell* **2019**, *36* (5), 545-558.e7. https://doi.org/10.1016/j.ccell.2019.09.004.

(23) Choi, S. H.; Kim, S.; Jones, K. A. Gene Expression Regulation by CDK12: A Versatile Kinase in Cancer with Functions beyond CTD Phosphorylation. *Exp Mol Med* **2020**, *52* (5), 762–771. https://doi.org/10.1038/s12276-020-0442-9.

(24) Lv, L.; Chen, P.; Cao, L.; Li, Y.; Zeng, Z.; Cui, Y.; Wu, Q.; Li, J.; Wang, J.-H.; Dong, M.-Q.; Qi, X.; Han, T. Discovery of a Molecular Glue Promoting CDK12-DDB1 Interaction to Trigger Cyclin K Degradation. *Elife* **2020**, *9*. https://doi.org/10.7554/eLife.59994.

(25) Kozicka, Z.; Suchyta, D. J.; Focht, V.; Kempf, G.; Petzold, G.; Jentzsch, M.; Zou, C.; Di Genua, C.; Donovan, K. A.; Coomar, S.; Cigler, M.; Mayor-Ruiz, C.; Schmid-Burgk, J. L.; Häussinger, D.; Winter, G. E.; Fischer, E. S.; Słabicki, M.; Gillingham, D.; Ebert, B. L.; Thomä, N. H. Design Principles for Cyclin K Molecular Glue Degraders. *Nat Chem Biol* **2024**, *20* (1), 93–102. https://doi.org/10.1038/s41589-023-01409-z.

(26) Ghosh, P.; Schmitz, M.; Pandurangan, T.; Zeleke, S. T.; Chan, S. C.; Mosior, J.; Sun, L.; Palve, V.; Grassie, D.; Anand, K.; Frydman, S.; Roush, W. R.; Schönbrunn, E.; Geyer, M.; Duckett, D.; Monastyrskyi, A. Discovery and Design of Molecular Glue Enhancers of CDK12-DDB1 Interactions for Targeted Degradation of Cyclin K. *RSC Chem Biol* **2024**. https://doi.org/10.1039/d4cb00190g.

(27) Zhang, Z.; Li, Y.; Yang, J.; Li, J.; Lin, X.; Liu, T.; Yang, S.; Lin, J.; Xue, S.; Yu, J.; Tang, C.; Li, Z.; Liu, L.; Ye, Z.; Deng, Y.; Li, Z.; Chen, K.; Ding, H.; Luo, C.; Lin, H. Dual-Site Molecular Glues for Enhancing Protein-Protein Interactions of the CDK12-DDB1 Complex. *Nat Commun* **2024**, *15* (1). https://doi.org/10.1038/s41467-024-50642-0.

(28) Kuchukulla, R. R.; Hwang, I.; Kim, S. H.; Kye, Y.; Park, N.; Cha, H.; Moon, S.; Chung, H. W.; Lee, C.; Kong, G.; Hur, W. Identification of a Novel Potent CDK Inhibitor Degrading CyclinK with a Superb Activity to Reverse Trastuzumab-Resistance in HER2-Positive Breast Cancer in Vivo. *Eur J Med Chem* **2024**, *264*, 116014. https://doi.org/10.1016/j.ejmech.2023.116014.

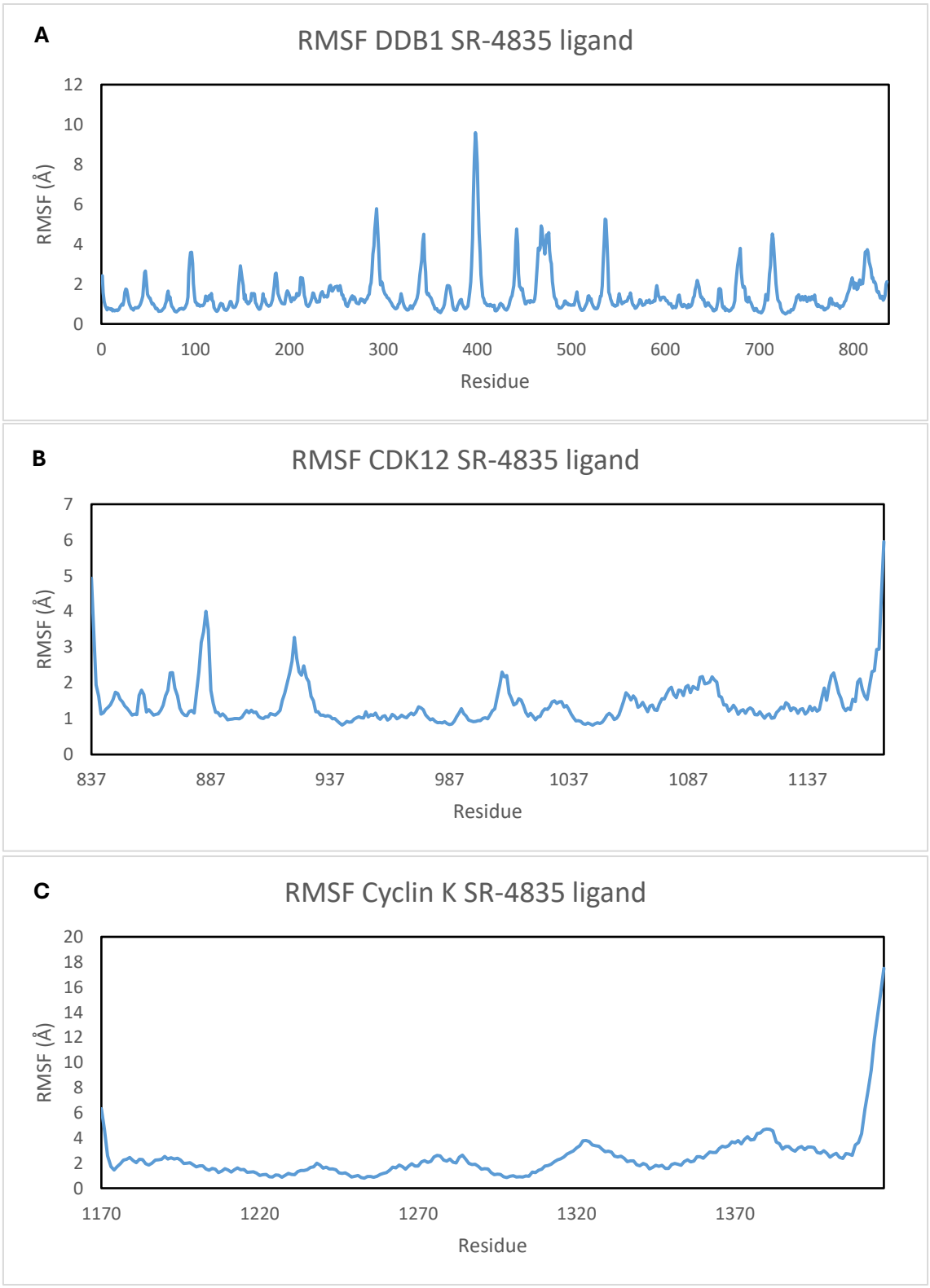
(29) Thomas, K. L.; Bouguenina, H.; Miller, D. S. J.; Sialana, F. J.; Hayhow, T. G.; Choudhary, J. S.; Rossanese, O. W.; Bellenie, B. R. Degradation by Design: New Cyclin K Degraders from Old CDK Inhibitors. *ACS Chem Biol* **2024**, *19* (1), 173–184. https://doi.org/10.1021/acschembio.3c00616.

(30) Pitera, J. W. Expected Distributions of Root-Mean-Square Positional Deviations in Proteins. *J Phys Chem B* **2014**, *118* (24), 6526–6530. https://doi.org/10.1021/jp412776d.

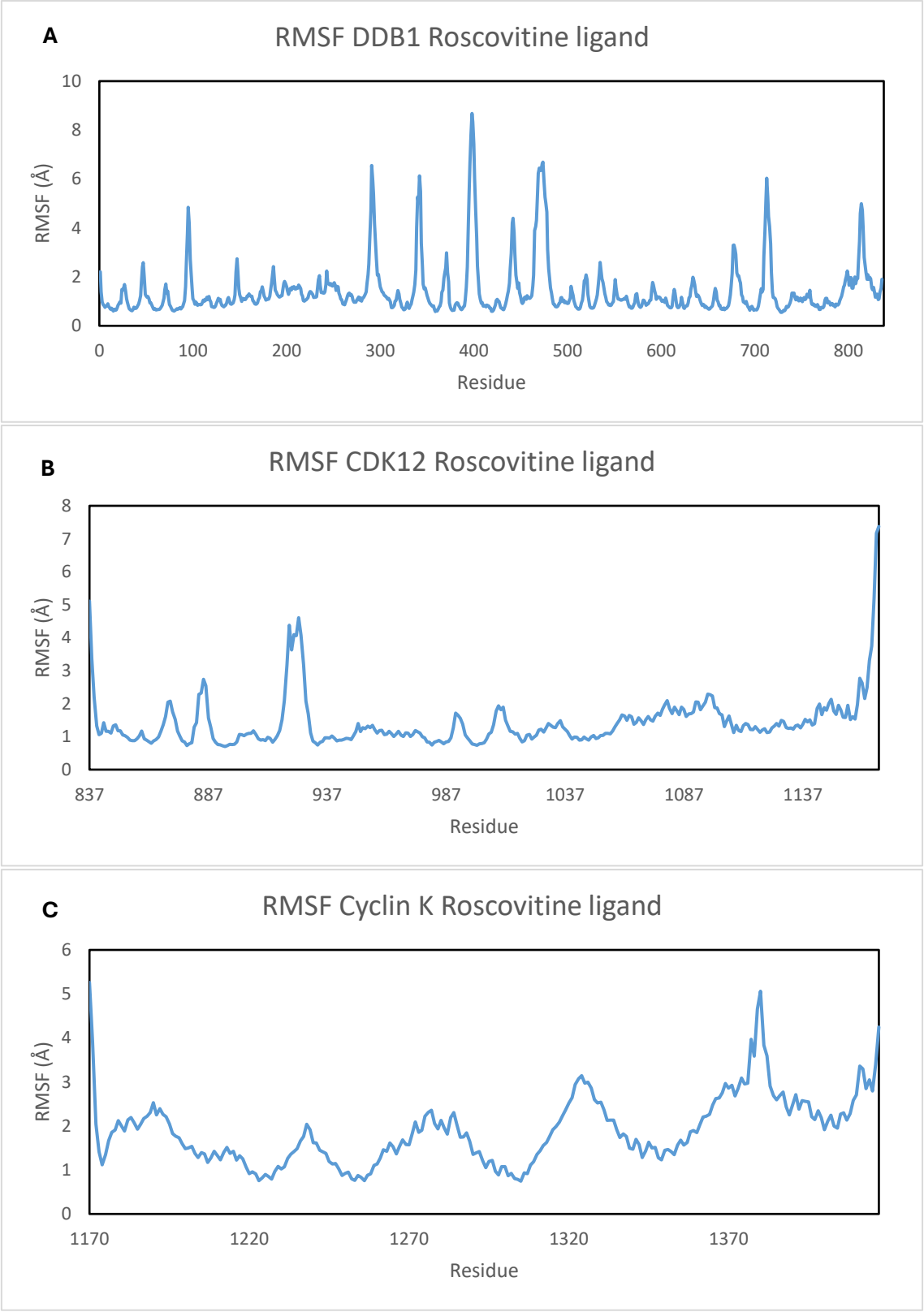
(31) Kuzmanic, A.; Zagrovic, B. Determination of Ensemble-Average Pairwise Root Mean-Square Deviation from Experimental B-Factors. *Biophys J* **2010**, *98* (5), 861–871. https://doi.org/10.1016/j.bpj.2009.11.011.

(32) Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opinion on Drug Discovery*. Informa Healthcare May 1, 2015, pp 449–461. https://doi.org/10.1517/17460441.2015.1032936.

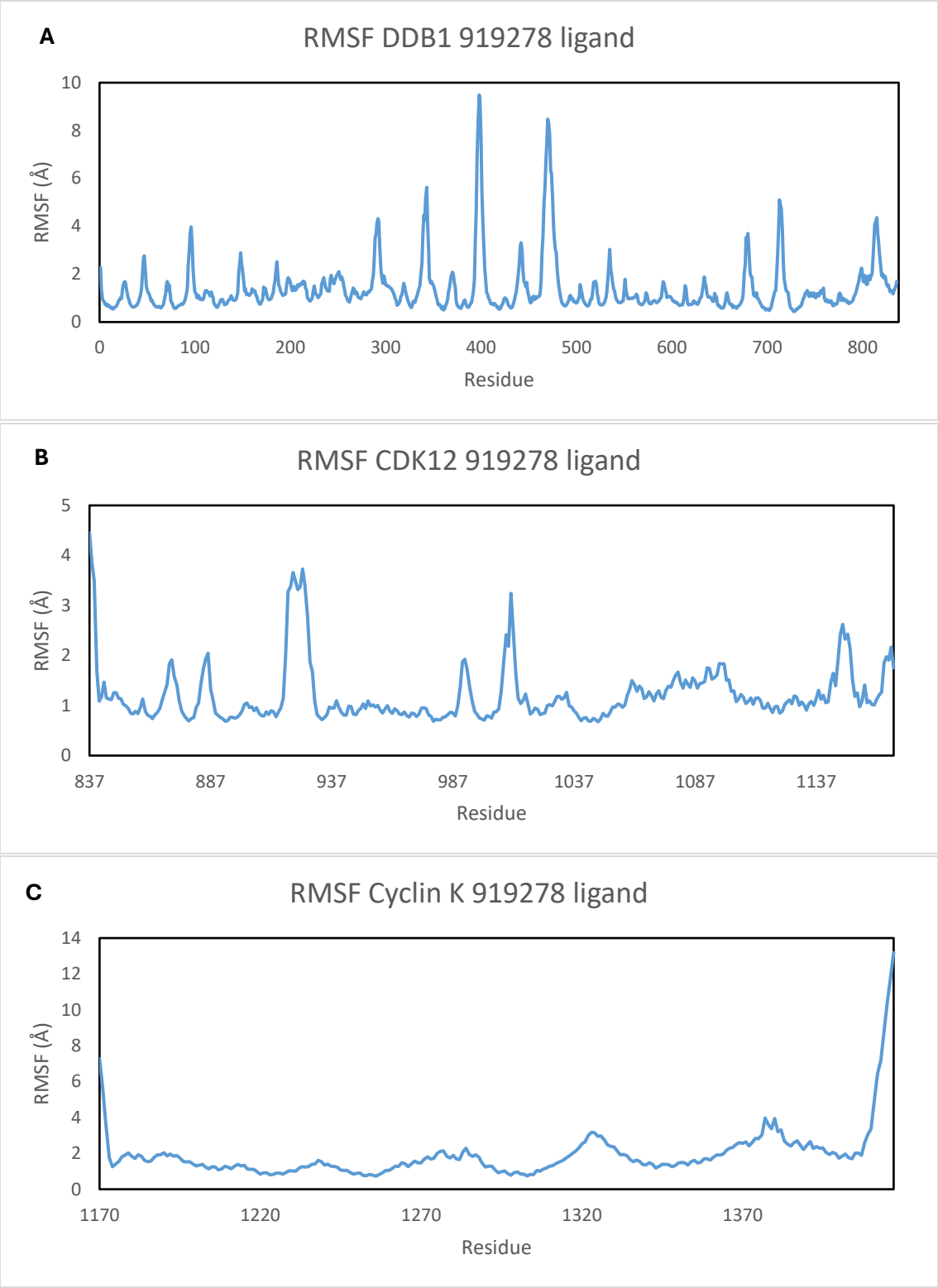
8. Supplementary materials



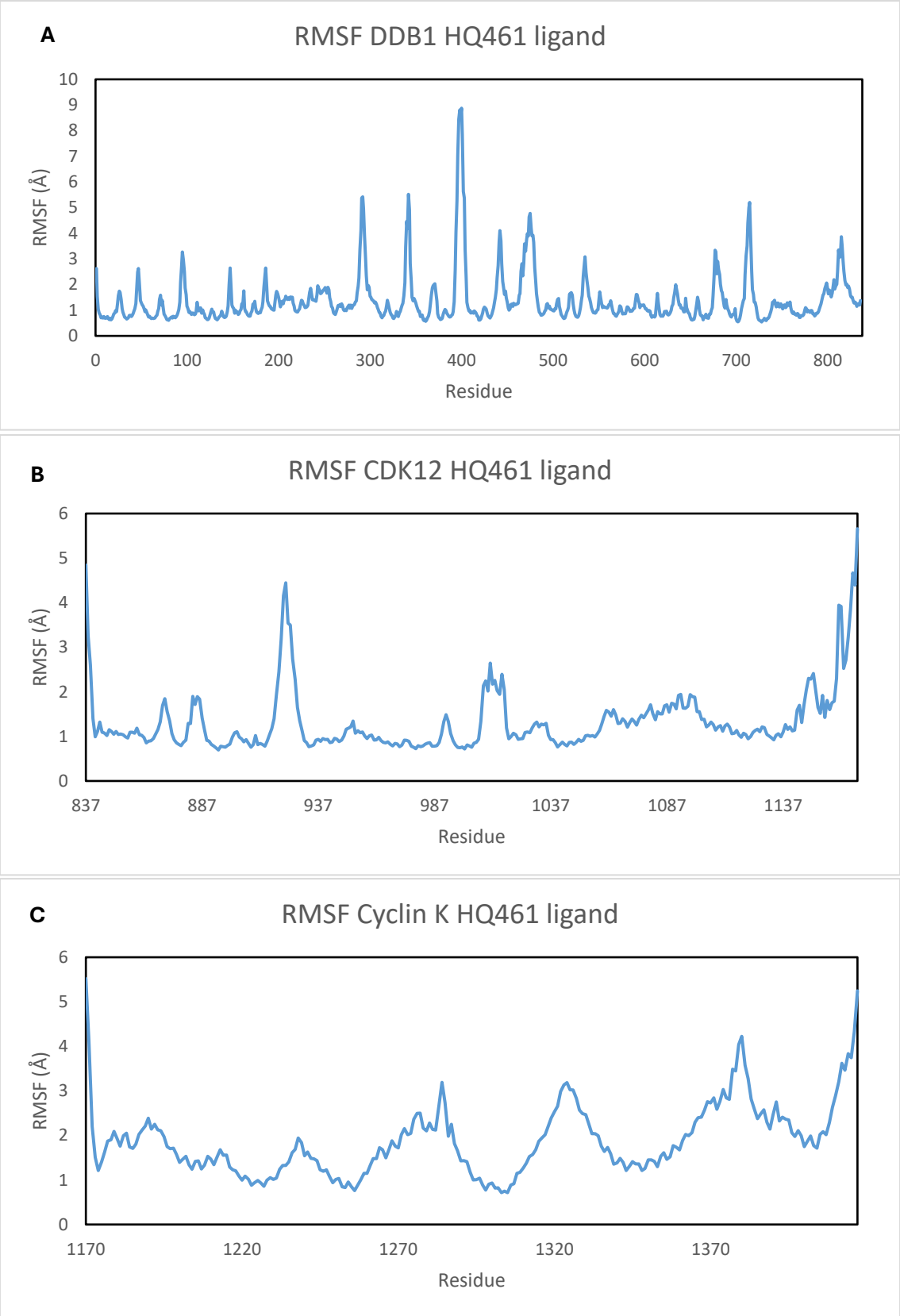
Supplementary Figure 1. *RMSF of the (A) DDB1, (B) CDK12 and (C) Cyclin K protein for the system with SR-4835 ligand.*



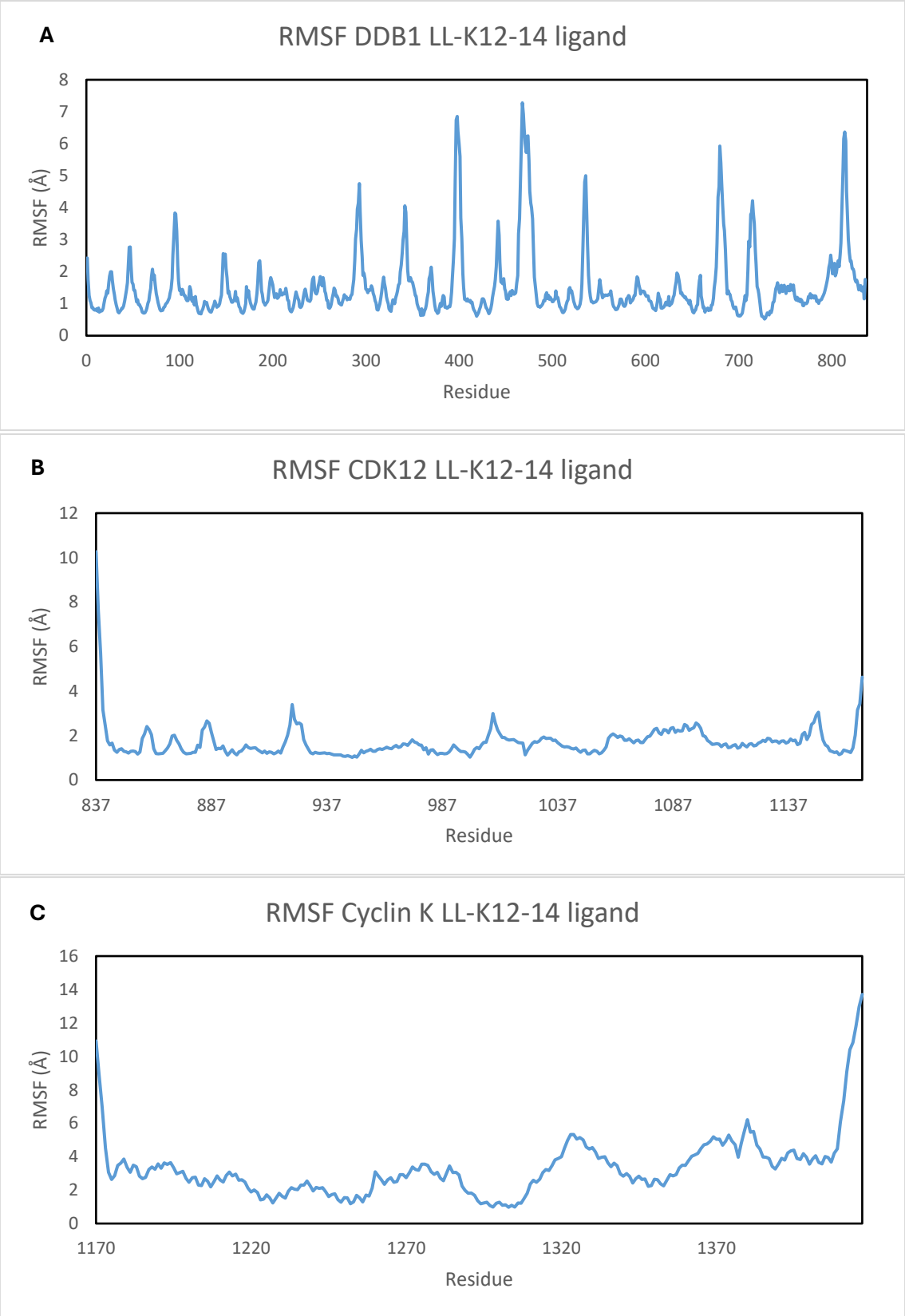
Supplementary Figure 2. *RMSF of the (A) DDB1, (B) CDK12 and (C) Cyclin K protein for the system with Roscovitine ligand.*



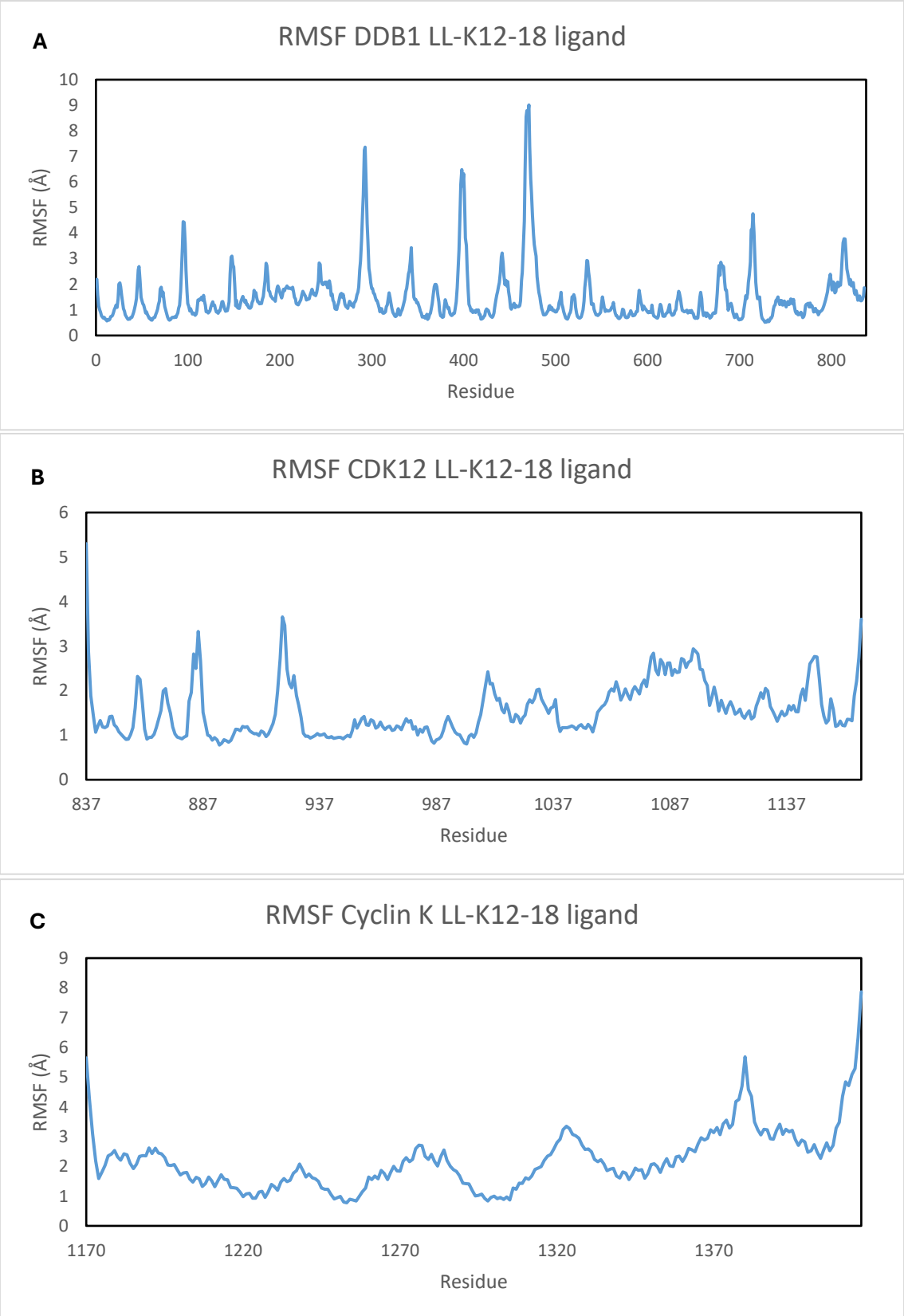
Supplementary Figure 3. *RMSF of the (A) DDB1, (B) CDK12 and (C) Cyclin K protein for the system with 919278 ligand.*



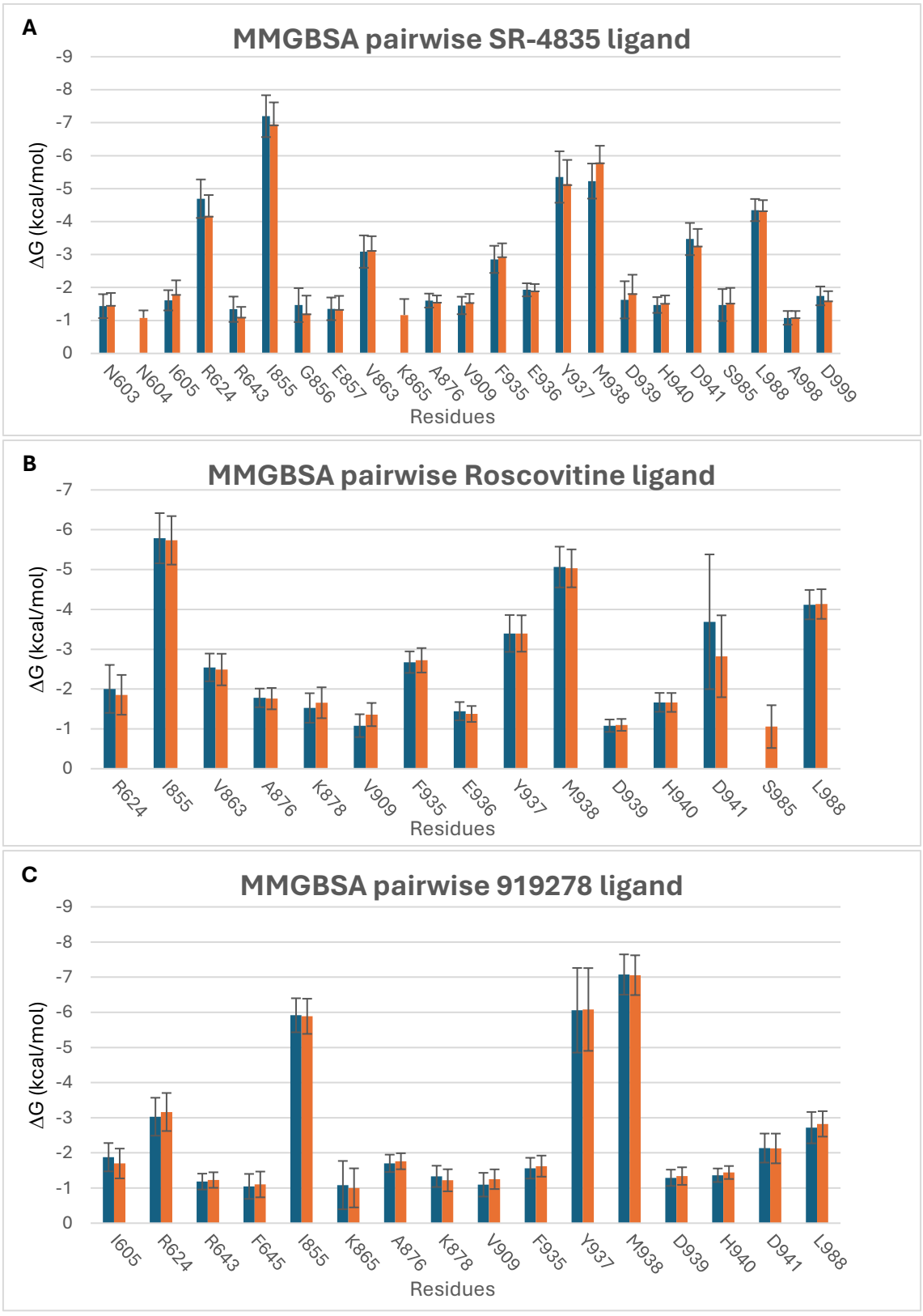
Supplementary Figure 4. *RMSF of the (A) DDB1, (B) CDK12 and (C) Cyclin K protein for the system with HQ461 ligand.*



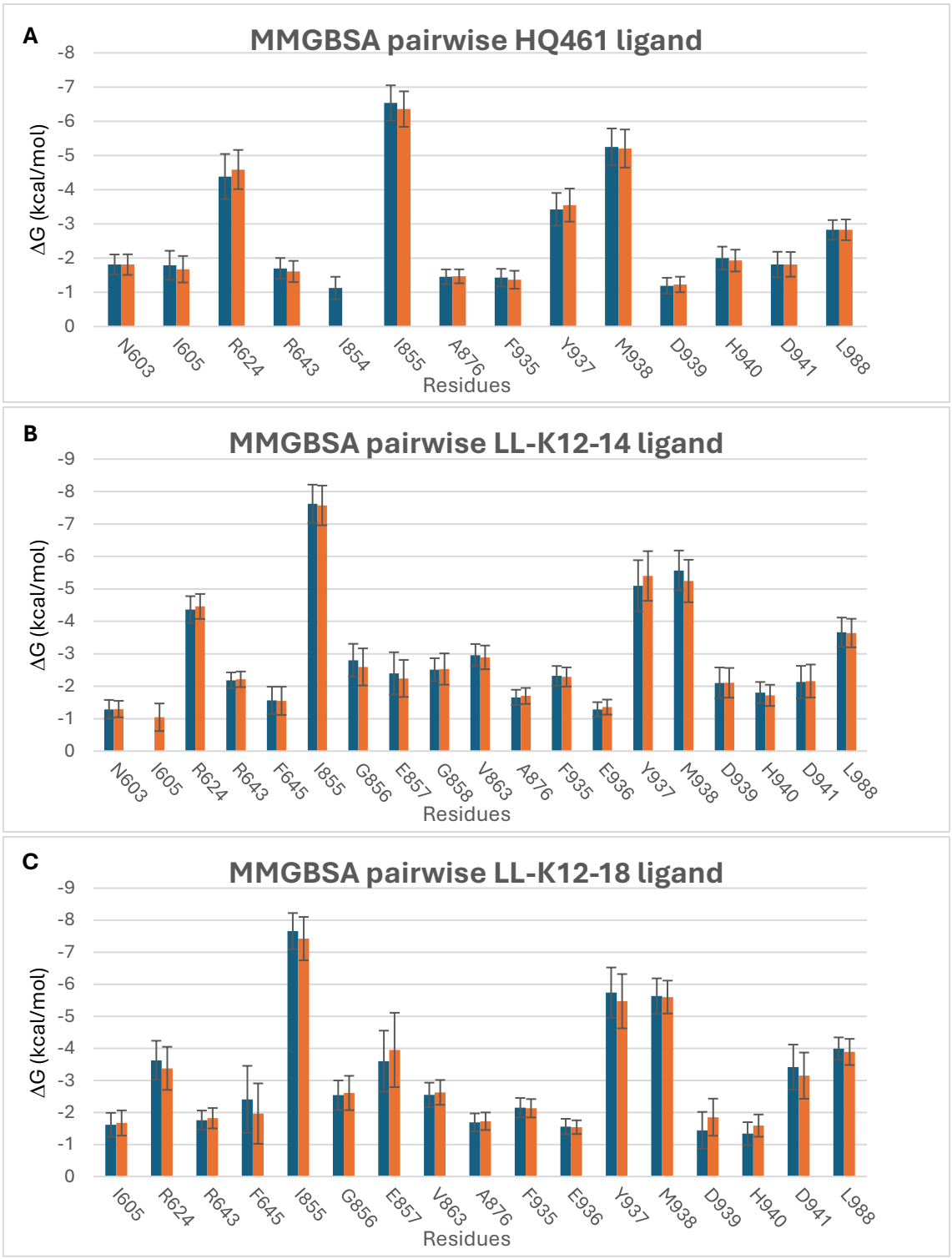
Supplementary Figure 5. *RMSF of the (A) DDB1, (B) CDK12 and (C) Cyclin K protein for the system with LL-K12-14 ligand.*



Supplementary Figure 6. *RMSF of the (A) DDB1, (B) CDK12 and (C) Cyclin K protein for the system with LL-K12-18 ligand.*



Supplementary Figure 7. *Per-residue MMGBSA pairwise decomposition of binding free energy contributions (ΔG) for the (A) SR-4835 (B) Roscovitine and (C) 919278 ligands. The chart displays the ΔG values (in kcal/mol) for key interacting residues, with error bars representing standard deviations. Blue and orange bars correspond to values from two different time frames (blue 980-1000 ns and orange 960-980 ns) over the MD simulation.*



Supplementary Figure 8. *Per-residue MMGBSA pairwise decomposition of binding free energy contributions (ΔG) for the (A) HQ461 (B) LL-K12-14 and (C) LL-K12-18 ligands. The chart displays the ΔG values (in kcal/mol) for key interacting residues, with error bars representing standard deviations. Blue and orange bars correspond to values from two different time frames (blue 980-1000 ns and orange 960-980 ns) over the MD simulation.*