Molecular dynamics simulations reveal the determinants of cyclin-dependent kinase 2 inhibition by 5-nitrosopyrimidine derivatives

(Article begins on next page)

18 October 2023

Stefano Guglielmo,*[ab] Daniela Cortese,[ac] Celine Cano[c] and Roberta Fruttero[a],
Molecular dynamics simulations reveal the determinants of cyclin-dependent kinase 2 inhibition by 5-nitrosopyrimidine derivatives,

LETTER TO THE EDITOR

Molecular dynamics simulations reveal the determinants of cyclin-dependent kinase 2 inhibition by 5-nitrosopyrimidine derivatives.

Stefano Guglielmo,*[ab] Daniela Cortese,[ac] Celine Cano[c] and Roberta Fruttero[a]

Correspondence to: Stefano Guglielmo (E-mail: stefano.guglielmo@unito.it)

[a] Dipartimento di Scienza e Tecnologia del Farmaco Universita' degli Studi di Torino Via P. Giuria 9, 10125 Torino, Italy. E-mail: stefano.guglielmo@unito.it.

[b] Scientific Computing Competence Centre (C3S), University of Turin, Italy.

[c] Northern Institute for Cancer Research, School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK.

INTRODUCTION

The cyclin-dependent kinases (CDKs) are members of a family of 26 serine-threonine protein kinases and are key components in maintaining the systematic progression of cells in the cell cycle.

While some isoforms play a direct role, others exert their action through activation of other CDKs, regulation of transcription, differentiation and apoptosis (Whittaker et al., 2017). Among them, CDK2 prompts the transition from G1 to S phase by triggering DNA synthesis (Tadesse et al., 2019). Due to the multi-functional nature of the CDKs, and their aberrant activity being implicated in a range of disease states, a large number of structurally diverse inhibitors have been reported. During recent years, many novel and potent ATP-competitive CDK inhibitors have indeed been developed, providing an extensive array of structurally diverse pharmacophores (Malumbres et al., 2008). Among these a number of purine derivatives, *e.g.* NU6102 (**1**; CDK2 K_i = 6 nM (Davies et al., 2002); Chart 1), proved to be potent and selective inhibitors. Elucidation of the interactions between inhibitors and the CDK2 ATP-binding site has been facilitated by crystallographic studies. Similar to ATP, most of these inhibitors interact with CDK2 through a donor-acceptor motif (Davies et al., 2002; Hardcastle et al., 2004). More recent work has shifted the attention from the purine core pharmacophore to alternative scaffolds, among which one of the most popular is the pyrimidine moiety, as exemplified by the nitrosopyrimidine derivative NU6027 (**2**; CDK2 K_i = 1.3 μM; Chart 1) (Arris et al., 2000). Starting from the nitrosopyrimidine scaffold, in the setting of a collaboration with the authors of the studies previously reported, our research group carried out further structural modulation: in particular, a potent inhibitor (**3**; CDK2 IC$_{50}$ = 3.9 nM; Chart 1) was obtained through formal substitution of the nitroso group with a cyano-*NNO*-azoxy moiety (Boschi et al., 2013).

Chart 1. Structures of CDK2 inhibitors representative of the classes of purines (1), nitrosopyrimidines (2) and cyano-NNO-azoxypyrimidines (3).

1 NU6102   **2** NU6027   **3**
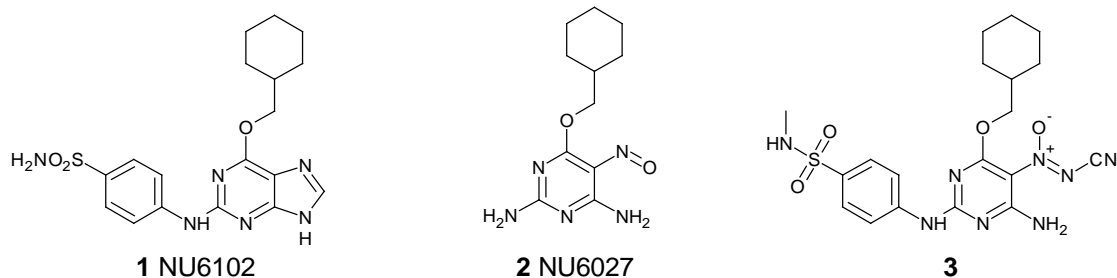
In a subsequent study, a small series of nitroso- and cyano-*NNO*-azoxypyrimidines was prepared (Cortese et al., 2016) to explore the structure-activity relationships (SAR) around the 2-amino group, which was decorated with aliphatic and alicyclic moieties with variable steric hindrance and polarity. Among these compounds, nitrosopyrimidine **4** (CDK2 IC_{50} = 0.16 μM; Chart 2) emerged as the most interesting inhibitor, outperforming the parent compound NU6027 by an order of magnitude in potency. Inhibition of CDK2 in this series proved quite sensitive to structural modification of the moiety linked to 2-amino group: the most sizeable and significant changes in activity were observed with compound **5** (CDK2 IC_{50} = 24.5 μM; Chart 2) and compound **6** (35% of CDK2 inhibition at 100 μM; Chart 2).

Chart 2. Structures of CDK2 inhibitors described in ref. 8 and object of the present study.

4   **5**   **6**
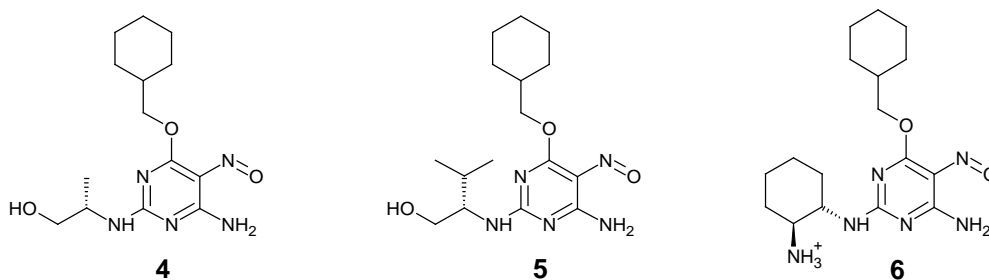
In the present paper a computational study was accomplished with the aim to delineate the molecular bases of the activity data for the three compounds reported above, and clarify SAR for this class of CDK2 inhibitors. The core of this study is constituted by a metadynamics simulation in its well-tempered version (WT-METAD) (Barducci et al., 2008), which allowed mapping the energy landscape of the interaction between the ligands and CDK2. From these simulations, a putative binding mode was identified for each of the three inhibitors and was subjected to an

unbiased molecular dynamics (MD) run, to assess stability and further analyse relevant interactions. Based on these latter simulations, a semi-quantitative determination of free energy of binding was attained through a molecular mechanics/generalized Born solvent area (MM/GBSA) calculation. Finally, a steered molecular dynamics (SMD) simulation was carried out, in order to investigate relevant energetics of the unbinding pathway of the complexes (see Supplemental material for methodological details).

RESULTS AND DISCUSSION

Binding pose of derivatives 4, 5 and 6. WT-METAD simulations were carried out in order to explore the free energy surface (FES) of the system; in order to describe the binding phenomenon to CDK2 two collective variables (CVs) have been defined: the first CV was assigned to the distance (d) between the centre of mass of the pyrimidine ring and the centre of mass of the three hinge region residues of CDK2 which engage the main interactions with pyrimidine inhibitors, namely Glu81, Phe82 and Leu83. The second CV was chosen in order to account for the orientation of the molecules with respect to the same group of residues, and it was defined as the angle χ between the centre of mass of the Glu81-Phe82-Leu83 triplet, the pyrimidine N3 and the oxygen atom at position 6 of the pyrimidine nucleus. A harmonic restraint was applied on d in order to prevent the ligands from getting farther than 20 Å from the binding site. The two-dimensional FESs for the complexes are reported in Fig. 1-3. For each energy basin, a representative structure has been extracted for analysis, through *gmx cluster* module implemented in Gromacs 2016.1, after selecting the frames corresponding to the basin using Wordom (Seeber et al., 2011). As for compound **4**, a free energy basin can be clearly identified in a region corresponding to the interaction with the ATP binding site ($d \approx 5$ Å and $\chi \approx 140°$, Fig. 1A). A representative conformation for this region is reported in Fig. 1B: it features the well-known interaction pattern made by a triplet of hydrogen bonds (2-NH to Leu83 (CO), N3 to Leu83 (NH), 4-NH$_2$ to Glu81 (CO)). Further analysis of the binding pose reveals an additional interaction, stabilizing the complex in the bioactive

conformation: the hydroxyl group of the side chain of the ligand engages a charge-enhanced hydrogen bond interaction with the carboxylate group of Asp86, and a weak electrostatic interaction with the protonated amino group of Lys89.
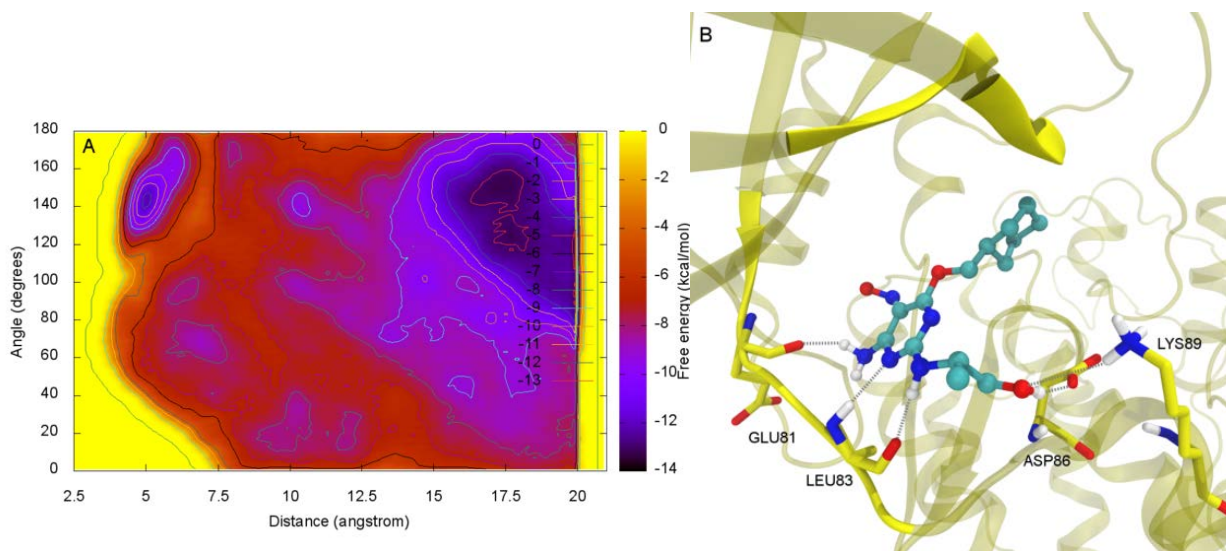

Figure 1A Free energy surface for the complex CDK2-**4**. **B** Representative conformation of the relative free energy minimum.

In the case of complex CDK2-**5**, a less populated basin can be recognized in the FES region of the ATP binding site (Fig. 2A), 4 kcal mol^{-1} above the global minimum: the inspection of a representative frame still reveals the presence of the triplet of hydrogen bonds. On the other hand, a significant difference emerges in the side chain conformation: in this case the bulkier isopropyl moiety points towards Asp86, thus hampering the interaction between the amino acid carboxylate group and the hydroxyl group on the inhibitor (Fig. 2B).
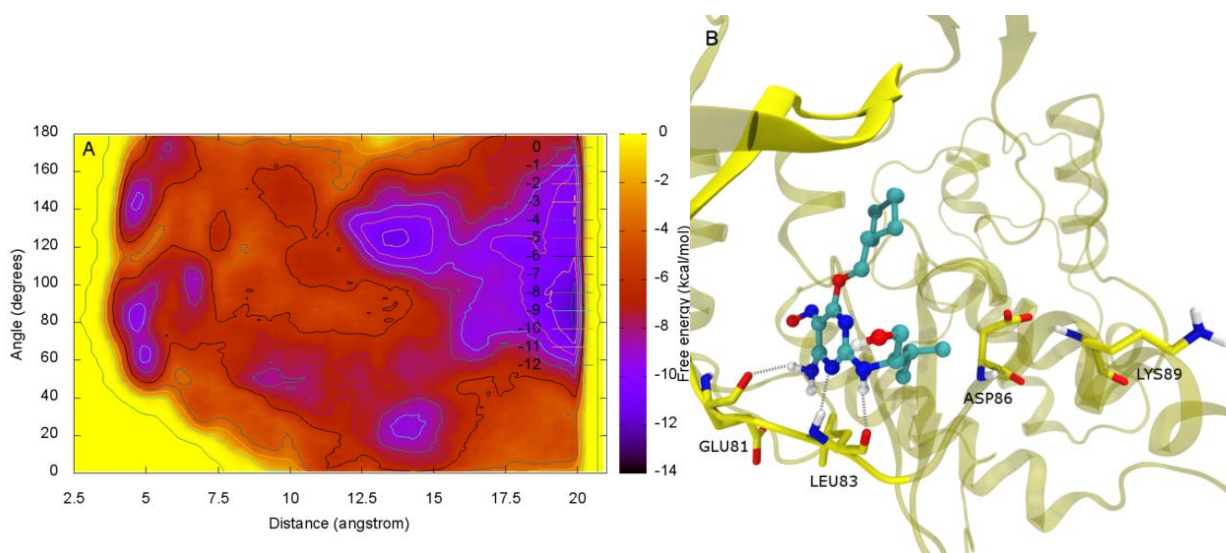
Figure 2A Free energy surface for the complex CDK2-**5**. **B** Representative conformation of the relative free energy minimum.

Coming to the CDK2-**6** complex, the FES lacks the basin identified above for the other two systems. Instead, there is a basin positioned at $d \approx 6$ Å and $\chi \approx 100°$ (Fig. 5A). The corresponding conformation (Fig. 5B) clearly indicates a pose quite different from **4** and **5**: in this case an electrostatic interaction is established between the protonated amino group of the cyclohexyl ring and the carboxylate group of Asp86; this interaction blocks the ligand at the entrance of the ATP binding site, preventing the formation of the characteristic hydrogen bond triplet and allowing only the interaction between the 2-amino group and Leu83.
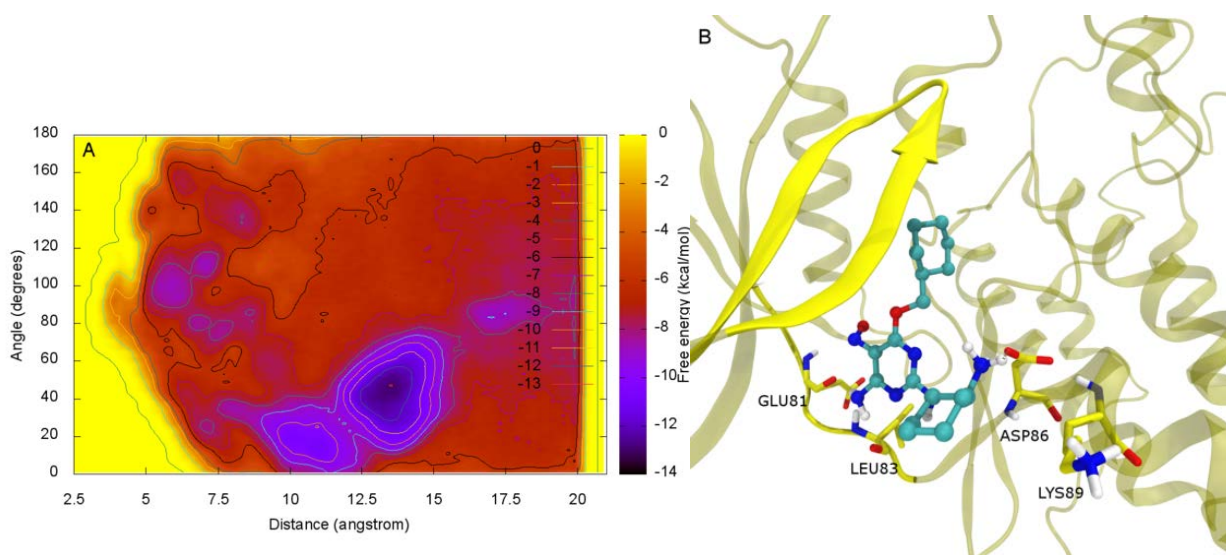
Figure 3A Free energy surface for the complex CDK2-compound **6**. **B** Representative conformation of the relative free energy minimum.

A few other relevant basins characterizing the FESs of the three complexes deserve further consideration. As for the CDK2-**4** complex, a representative snapshot of the basin centred at $d \approx$ 17.5 Å and $\chi \approx 140°$, with a slightly lower free energy compared to that of the bound configuration, has been obtained using the *gmx cluster* module implemented in Gromacs 2016.1 (Fig. S2A). The ligand engages weak polar interactions through the pyrimidine nucleus, but it is almost entirely enclosed in a lipophilic lodge made by side chains of Phe109, Phe286 and Val289. In the case of the other two complexes, the basins located at $d \approx 14$ Å and $\chi \approx 130°$ (for CDK2-**5**, Fig. S2B) and $\chi \approx 50°$ (for CDK2-**6**, Fig. S2C) correspond mainly to the ligands dwelling in close proximity of the glycine-rich loop (G-loop, residues 10 to 20). CDK2-**6** complex is particularly stabilized by the presence of the favourable electrostatic interaction between the protonated -NH_2 group of the ligand and the side chain carboxylate group of Asp86.

In the case of compound **6**, the presence of such a populated conformation at a close distance from the binding site, in addition to the geometric constraints discussed above, represents a significant obstacle to the attainment of a productive binding pose, providing a sound rationale to the very low activity against the target.

Unbiased molecular dynamics: dynamic characterization of the binding pose of 4, 5 and 6 and MM/GBSA free energies. In order to obtain further insight on the binding pose of the three inhibitors, each complex was subjected to a 50 ns molecular dynamics simulation in explicit solvent. From the inspection of salient dynamic parameters, namely root mean square deviation (RMSD) of the protein C_α and of ligand heavy atoms, the system CDK2-**4** proved the most stable; the complex CDK2-**5** showed an initial ligand displacement due to side chain mobility (data not shown), followed by a good stability for the first 20 ns; ligand **6** showed higher values of RMSD throughout the simulation (Fig. S3A, B), likely as a consequence of the absence of hydrogen bond

triplet, nevertheless, for comparison purpose, the trajectory was used for further analysis. Following these considerations, the simulations were run for two additional 20 ns replica: RMSD of the protein C_α and of ligand heavy atoms and root mean square fluctuation (RMSF) of the protein C_α were thus averaged over the three independent runs for each system (Fig. 4A-C). The RMSD plots of the three complexes show some meaningful differences: for complex CDK2-**4** a greater stability is found both for ligand and for C_α, whilst for the remaining two, ligands encounter larger fluctuations and C_α RMSD's display a smooth increase: this can be a confirmation of a larger overall stability of the system CDK2-**4**. The RMSF plot confirms this trend, even though with less pronounced differences. The RMSF trace is in accordance with previous reports on the limited flexibility of the ligand-bound CDK2 binding site and G-loop (Kong et al., 2016).
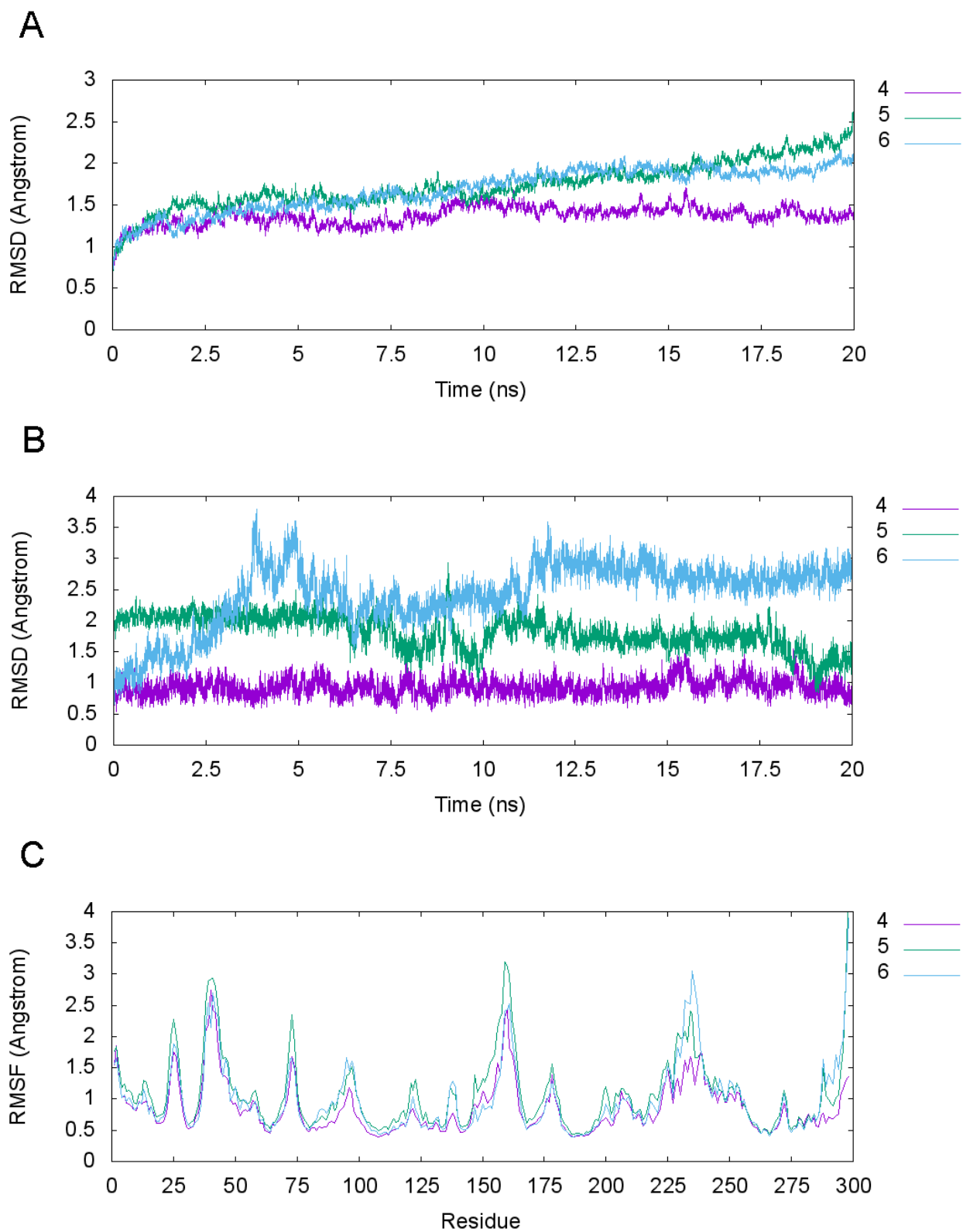
Figure **4A** RMSD plot of the C_α of CDK2 for the three complexes over 20 ns simulation. **B** RMSD

plot of ligands over 20 ns simulation. **C** RMSF plot of the C_α of CDK2 for the three complexes. All

plots represent the average across three independent 20 ns molecular dynamics runs. The dynamic parameters were calculated after alignment of C_α using the first frame as reference structure.

For CDK2-**4** and CDK2-**6** complexes, a specific analysis of ligand interactions with Asp86 was carried out., in view of their prominent role in determining the binding pose of the ligands. In the first case, the distance between the hydroxyl oxygen of ligand **4** and the centre of geometry of the carboxylate of Asp86 was monitored along the three independent trajectories: the hydroxyl group maintains the interaction discussed in the previous section throughout the entire simulation, thus confirming the hypothesis that this interaction has a pivotal role in binding of the compound to CDK2 (Fig. S4A). The same behaviour can be found analysing the interaction between Asp86 carboxylate and the ammonium group of ligand **6**: in this case the interaction is even tighter, with shorter distance and smaller fluctuations (Fig. S4B).

Also MM/GBSA calculated two-end-state binding free energies (ΔG_{bind}), an approximated estimation of free energy that considers only end-point states with an implicit solvent model (Homeyer & Gohlke, 2012),[12] confirm the trend of affinities of the ligands for CDK2 (Table 1). While MM/GBSA calculation is far from reproducing experimental binding affinities, due to its intrinsic approximations and to overestimation of energy as a consequence of neglecting the entropic term, these data still provide useful insight: in particular, the contribution of van der Waals interactions between the cyclohexyl ring and the G-loop of the enzyme stands out as a major component of the overall binding energy. The electrostatic contribution is energetically unfavourable for the three complexes, in particular for CDK2-**4** (4.12 kcal mol^{-1}) and CDK2-**5** (4.83 kcal mol^{-1}). Nonetheless, the analysis of the binding poses highlights the role of electrostatic interactions in shaping the initial orientation of the various ligands in the binding site, thus determining their ability to optimize van der Waals contacts. This effect is particularly evident for the CDK2-**6** complex: in this case the electrostatic energy is less unfavourable (0.40 kcal mol^{-1}), owing to the interaction between the protonated -NH$_2$ group of the ligand and the carboxylate group

of the side chain of Asp86; however, as previously discussed this interaction holds the pyrimidine farther away from the binding site.

Table 1 Binding affinities and individual energy components.

System	ΔG_{bind}[a]	ΔE_{ele}[a]	ΔE_{vdW}[a]
CDK2-**4**	-32.52 ± 0.02	4.12 ± 0.02	-36.65 ± 0.02
CDK2-**5**	-31.04 ± 0.02	4.83 ± 0.01	-35.87 ± 0.02
CDK2-**6**	-29.65 ± 0.03	0.40 ± 0.03	-30.05 ± 0.02

[a] Energies are reported in kcal mol^{-1} and represents the average of three independent simulations ± SE.

Steered molecular dynamics. Steered molecular dynamics (SMD) is a computational technique often used to describe several phenomena taking place during the formation of the ligand-target complex (Patel et al., 2014); in particular, it can give useful information on unbinding events, which in turn can account for the capability of different ligands to interact with biological targets in physiological non-equilibrium conditions. In the present study, simulations were carried out in five independent runs for each complex. The SMD trajectories provide atomistic details about the process of ligand unbinding, and the forces calculated during the simulations can reflect, within a good approximation, the trend of residence times for the different complexes.

Fig. 5 reports the plot of the force profile *vs* simulation time as the average across the five simulations carried out for the three ligands. On the whole, the three force profiles mirror to certain extent the stability of the complexes: the force applied to ligands increases linearly with simulation time and in case of **4** it reaches a value of *ca* 1700 pN required for unbinding; in case of **5** this event requires a force of *ca* 1000 pN, whilst for **6** the force reaches an intermediate value as a consequence of the favourable electrostatic interaction discussed above.
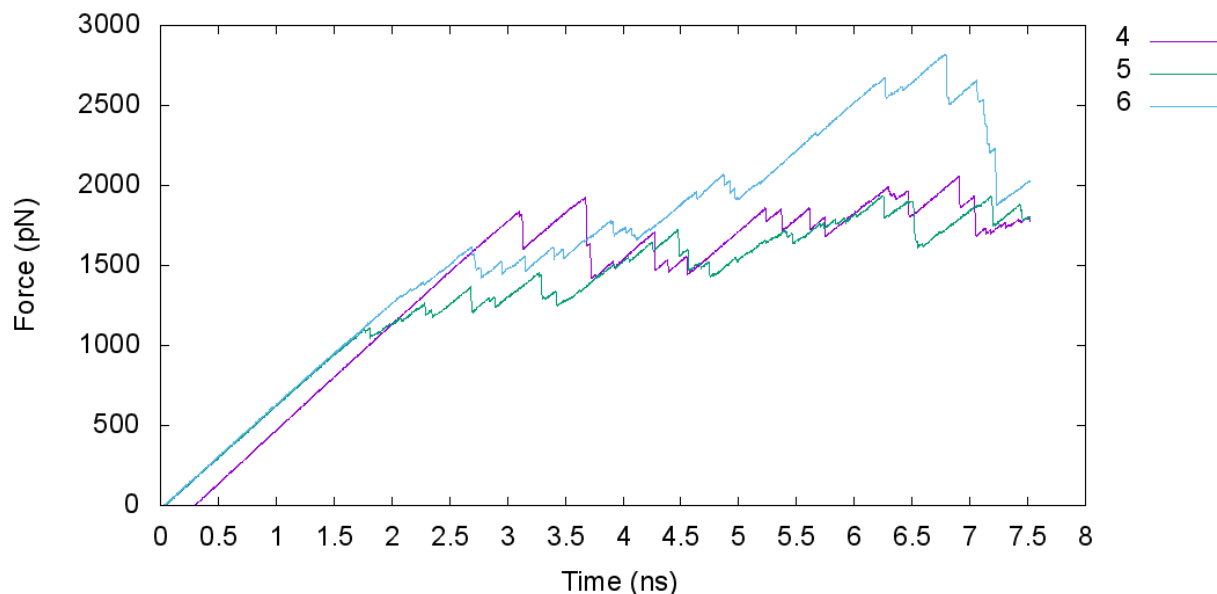
Figure 5. Force curves of the SMD simulations for the three systems. Each curve represents the average of five independent runs.

Snapshots of the unbinding pathway obtained through *gmx cluster*, corresponding to the most relevant force changes for each system, are also reported in Fig. 6-8. In the case of CDK2-**4** complex (Fig. 6), the first frame (A) refers to the bound configuration only; as can be seen from the second snapshot (B), the inhibitor starts to egress from the binding site by "pivoting" on the interaction with Asp86 and projecting the cyclohexyl moiety beyond the G-loop; in this conformation the N3 atom of the pyrimidine ring interacts with the Glu12 of the G-loop through a hydrogen bond. In the following step (C) the ligand has loosened the interaction with the target having lost the hydrogen bond with Asp86, while still maintaining two weak electrostatic bonds with Glu12 and Gly13. The last frame (D) reports the fully unbound ligand in bulk solvent.
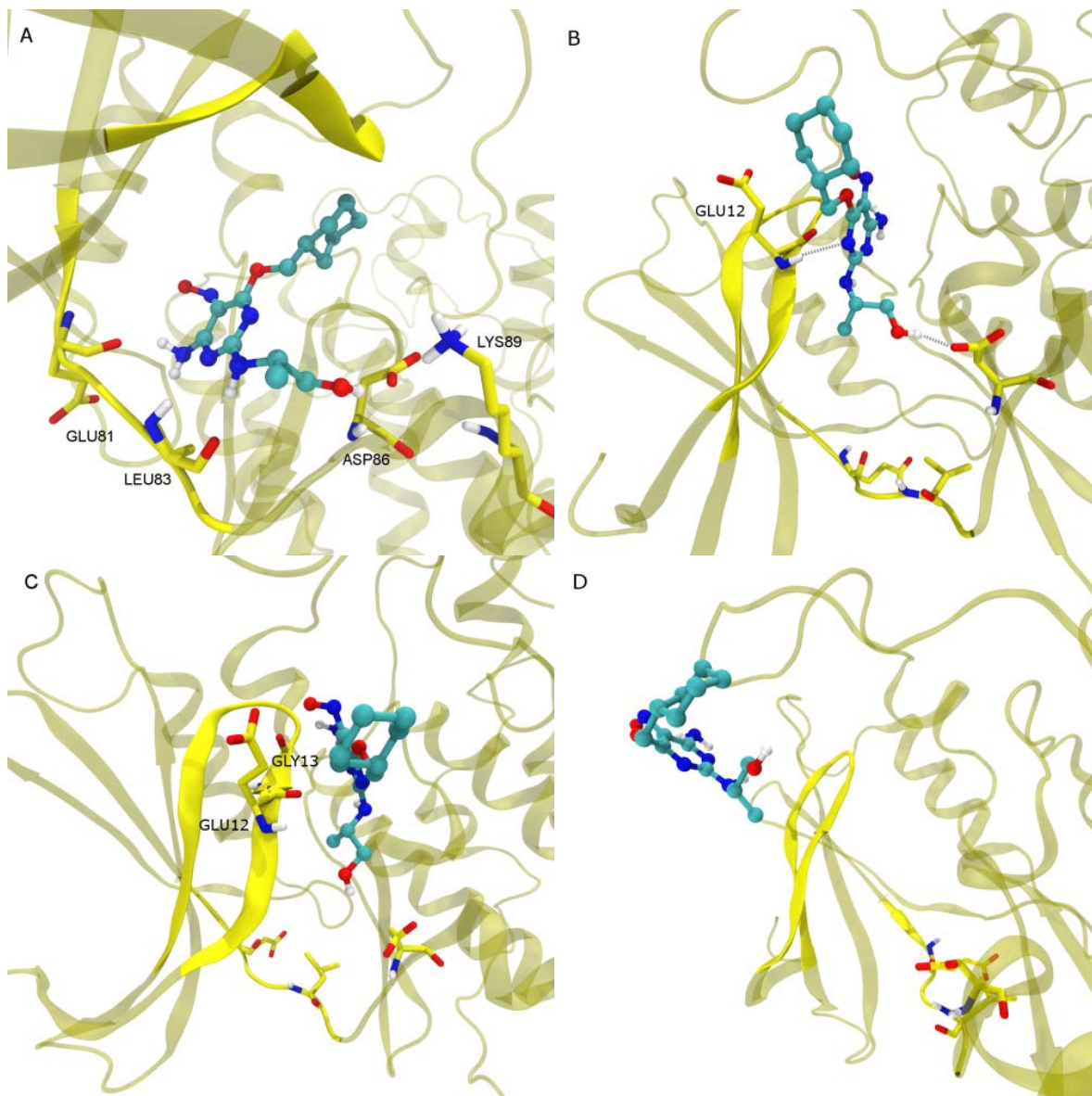
Figure 6. Representative snapshots of the unbinding pathway for system CDK2-**4**.

As to CDK2-**5** complex, Fig. 7A reports the starting pose, whilst in snapshot B the ligand has lost the triplet of hydrogen bonds, replaced by an electrostatic interaction between the oxygen atom of nitroso group and the protonated amino group of Lys33. In the third frame the ligand has got closer to the surface of the enzyme and, before reaching bulk solvent, is still held onto the protein by three interactions, namely the two interactions engaged by the 4-NH$_2$ group with the carboxylate group of Glu162 and with the NH$_2$ of Gln131 and a third weak electrostatic interaction between the oxygen atom of the nitroso group and the protonated amino group of Lys88.
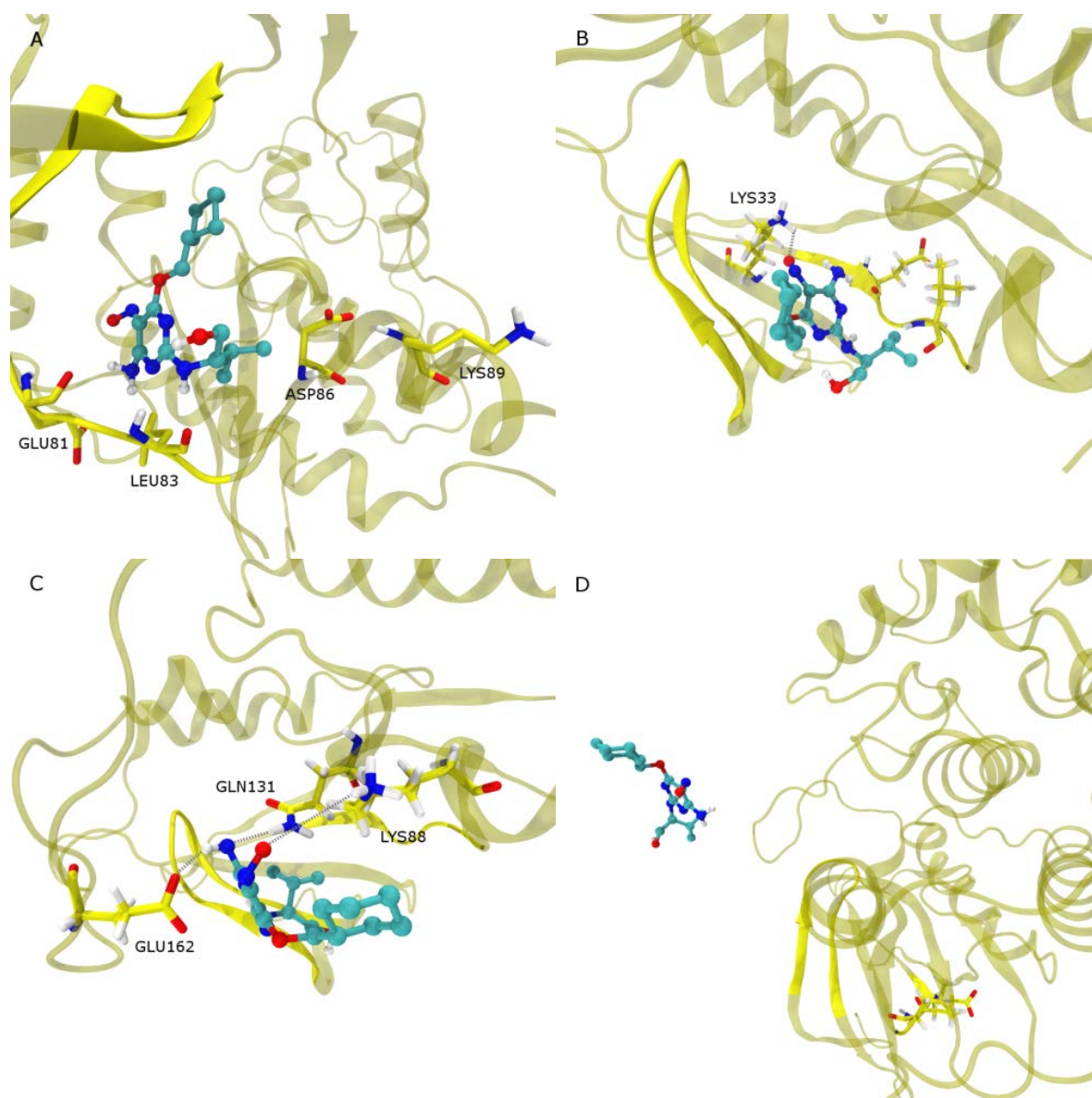
Figure 7. Representative snapshots of the unbinding pathway for system CDK2-**5**.

In Fig. 8 A - E the unbinding pathway of complex CDK2-**6** is depicted. In particular, in frame B the ligand has lost the contact with Leu83, having shifted in a plane almost parallel to the G-loop, thus reaching a conformation stabilized through the weak electrostatic interaction between nitroso group and the protonated amino group of Lys33. In the following step a rotation around the cyclohexylmethoxy group of the ligand has begun, and the cyclic moiety is held in a lipophilic cage delimited by Thr14, Lys 129, Val164 and Thr165. In frame D the rotation has completed: the cyclohexylmethoxy group is still pinned to the lipophilic lodge, whilst the pyrimidine moiety and

the cyclohexylamino group of the side chain are projected towards bulk water. Starting from this conformation, the ligand completes its egress from the enzyme (Fig. 8E).
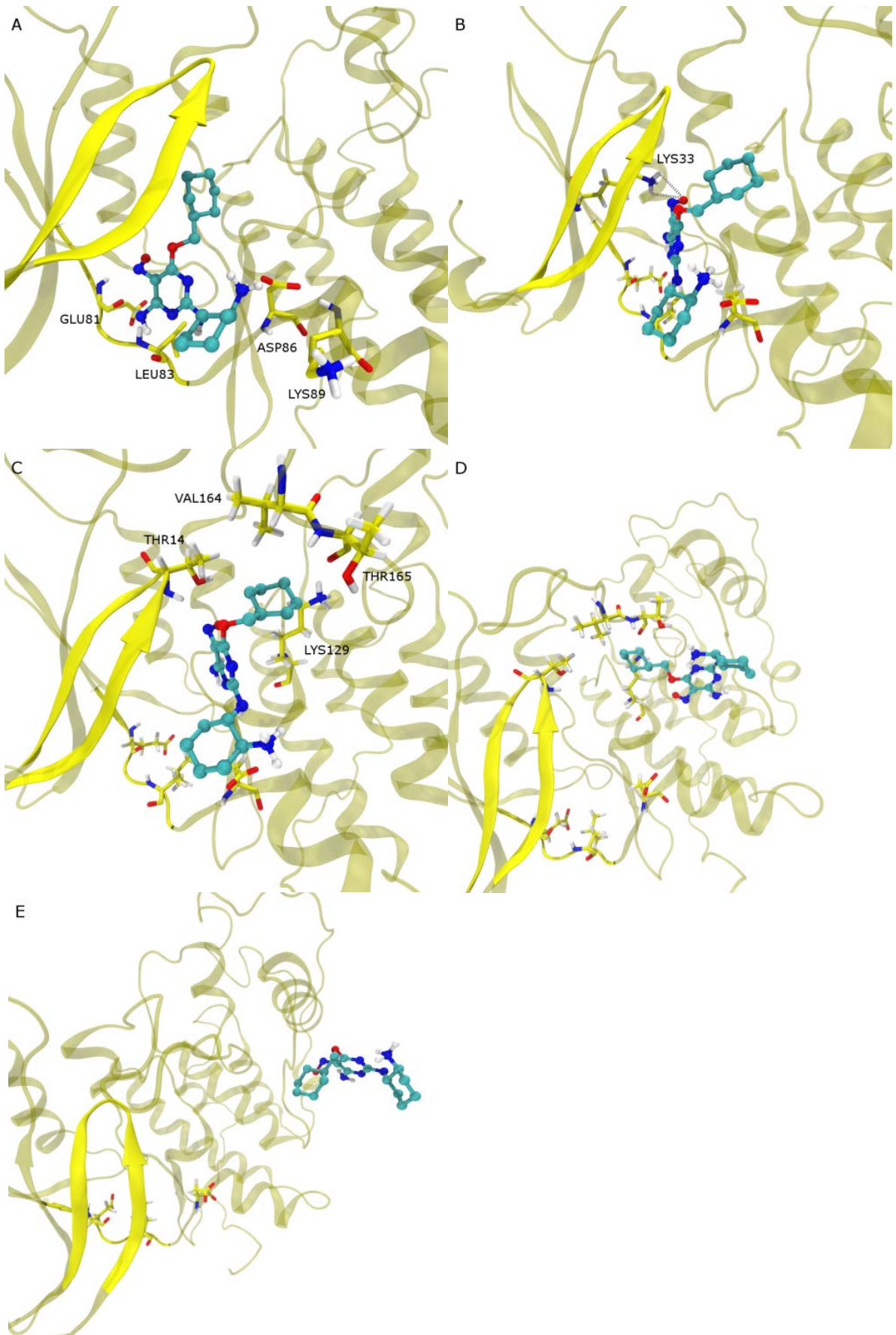
Figure 8. Representative snapshots of the unbinding pathway for system CDK2-**6**.

CONCLUSIONS

A computational study based on the combination of different molecular dynamics simulations has been performed to explain the sizeable differences in binding affinities towards CDK2 of three structurally-related nitrosopyrimidines. The exploration of free energy surface performed through well-tempered metadynamics simulation revealed the characteristic binding pose featuring a triplet of hydrogen bonds for the two more active compounds (**4** and **5**); for ligand **4** an additional interaction between the hydroxyl group on the side chain and Asp86 was also identified, which proved quite stable during unbiased molecular dynamics simulations. Compound **6** behaves differently in that it engages a quite strong electrostatic interaction through its protonated NH_2 group with Asp86, but this event induces a significant displacement of the pyrimidine core from the binding site. Binding free energy calculations through a MM/GBSA approach confirms these findings as well as unbinding forces calculation through SMD. The decomposition of binding energies clearly indicates a prominent role of van der Waals interactions, but highlights the importance of electrostatic interactions to achieve a proper orientation of the ligands within the binding pocket.

Overall, the present investigation provides an explanation for the significant differences in affinity towards CDK2 observed for the three structurally correlated ligands; it also indicates the possibility of optimizing ligand activity exploiting a mainly electrostatic interaction with Asp86. This interaction can be reached through a suitable side chain at N2 position "constructed" in compliance with the geometric requirements which have been elucidated by the comparative analysis of ligands **4** and **6** binding modes.

were carried out with Occam supercomputer (Aldinucci et al., 2017) at Scientific Computing Competence Centre (C3S) of the University of Turin. The authors are thankful to prof. A. Gasco for valuable discussion.

DECLARATION OF INTERESTS

The authors declare no competing interest.

Additional Supporting Information may be found in the online version of this article.

REFERENCES

Aldinucci, M., Bagnasco, S., Lusso, S., Pasteris, P., Rabellino, S., Vallero, S. (2017). OCCAM: a flexible, multi-purpose and extendable HPC cluster. *Journal of Physics: Conference Series*, *898*, 082039. doi: 10.1088/1742-6596/898/8/082039.

Arris, C. E., Boyle, F. T., Calvert, A. H., Curtin, N. J., Endicott, J. A., Garman, E. F., Gibson, A. E., Golding, B. T., Grant, S., Griffin, R. J., Jewsbury, P., Johnson, L. N., Lawrie, A. M., Newell, D. R., Noble, M. E. M., Sausville, E. A., Schultz, R., Yu, W. (2000). Identification of Novel Purine and Pyrimidine Cyclin-Dependent Kinase Inhibitors with Distinct Molecular Interactions and Tumor Cell Growth Inhibition Profiles. *Journal of Medicinal Chemistry*, *43*, 2797-2804. doi: 10.1021/jm990628o.

Barducci, A., Bussi, G., Parrinello, M. (2008). Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Physical Review Letters*, *100*, 020603-1. doi: 10.1103/PhysRevLett.100.020603.

Boschi, D., Tosco, P., Chandra, N., Chaurasia, S., Fruttero, R., Griffin, R. J., Wang, L., Gasco, A. (2013). 6-Cyclohexylmethoxy-5-(cyano-NNO-azoxy)pyrimidine-4-amine: A new scaffold endowed with potent CDK2 inhibitory activity. *European Journal of Medicinal Chemistry*, *68*, 333-338. doi: 10.1016/j.ejmech.2013.07.031.

Cortese, D., Chegaev, K., Guglielmo, S., Wang, L., Golding, B. T., Cano, C., Fruttero, R. (2016). Synthesis and Biological Evaluation of N[2]-Substituted 2,4-Diamino-6-cyclohexylmethoxy-5-nitrosopyrimidines and Related 5-Cyano-NNO-azoxy Derivatives as Cyclin-Dependent Kinase 2 (CDK2) Inhibitors. *ChemMedChem*, *11*, 1705-1708. doi: 10.1002/cmdc.201600108.

Davies, T. G., Bentley, J., Arris, C. E., Boyle, F. T., Curtin, N. J., Endicott, J. A., Gibson, A. E., Golding, B. T., Griffin, R. J., Hardcastle, I. R., Jewsbury, P., Johnson, L. N., Mesguiche, V., Newell, D. R., Noble, M. E., Tucker, J. A., Wang, L., Whitfield, H. J. (2002). Structure-based design of a potent purine-based cyclin-dependent kinase inhibitor. *Nature Structural & Molecular Biology*, *9*, 745-749. doi: 10.1038/nsb842.

Hardcastle, I. R., Arris, C. E., Bentley, J., Boyle, F. T., Chen, Y., Curtin, N. J., Endicott, J. E., Gibson, A. E., Golding, B. T., Griffin, R. J., Jewsbury, P., Menyerol, J., Mesguiche, V., Newell, D. R., Noble, M. E. M., Pratt, D. J., Wang, L., Whitfield, H. J. (2004). N[2]-Substituted *O*[6]-Cyclohexylmethylguanine Derivatives: Potent Inhibitors of Cyclin-Dependent Kinases 1 and 2. *Journal of Medicinal Chemistry*, *47*, 3710-3722. doi: 10.1021/jm0311442.

Homeyer, N., Gohlke, H. (2012). Free Energy Calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area Method. *Molecular Informatics*, *31*, 114-122. doi: 10.1002/minf.201100135.

Kong, X., Sun, H., Pan, P., Tian, S., Li, D., Li, Y., Hou, T. (2016). Molecular principle of the cyclin-dependent kinase selectivity of 4-(thiazol-5-yl)-2-(phenylamino) pyrimidine-5-carbonitrile

derivatives revealed by molecular modeling studies. *Physical Chemistry Chemical Physics*, *18*, 2034-2046. doi: 10.1039/c5cp05622e.

Malumbres, M., Pevarello, P., Barbacid, M., Bischoff, J. R. (2008). CDK inhibitors in cancer therapy: what is next? *Trends in Pharmacological Sciences*, *29*, 16-21. doi: 10.1016/j.tips.2007.10.012.

Patel, J. S., Berteotti, A., Ronsisvalle, S., Rocchia, W., Cavalli, A. (2014). Steered molecular dynamics simulations for studying protein-ligand interaction in cyclin-dependent kinase 5. *Journal of Chemical Information and Modeling*, *54*, 470-480. doi: 10.1021/ci4003574.

Seeber, M., Felline, A., Raimondi, F., Muff, S., Friedman, R., Rao, F., Caflisch, A., Fanelli, F. (2011). Wordom: a user-friendly program for the analysis of molecular structures, trajectories, and free energy surfaces. *Journal of Computational Chemistry*, *32*, 1183-1194. doi: 10.1002/jcc.21688.

Tadesse, S., Caldon, E., C., Tilley, W., Wang, S. (2019). Cyclin-Dependent Kinase 2 Inhibitors in Cancer Therapy: An Update. *Journal of Medicinal Chemistry*, *62*, 4233-4251. doi: 10.1021/acs.jmedchem.8b01469.

Whittaker, S. R., Mallinger, A.; Workman, P., Clarke, P. A. (2017). Inhibitors of cyclin-dependent kinases as cancer therapeutics. *Pharmacology & Therapeutics*, *173*, 83-105. doi: 10.1016/j.pharmthera.2017.02.008.