PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Conformational flexibility of the complete catalytic domain of Cdc25B phosphatases

Raphael S. R. Sayegh, Fabio K. Tamaki, Sandro R. Marana,[†] Roberto K. Salinas, and Guilherme M. Arantes[†]*

Departamento De Bioquímica, Instituto De Química, Universidade De São Paulo, Av. Prof. Lineu Prestes 748, São Paulo, Brazil

ABSTRACT

Cdc25B phosphatases are involved in cell cycle checkpoints and have become a possible target for developing new anticancer drugs. A more rational design of Cdc25B ligands would benefit from detailed knowledge of its tertiary structure. The conformational flexibility of the C-terminal region of the Cdc25B catalytic domain has been debated recently and suggested to play an important structural role. Here, a combination of experimental NMR measurements and molecular dynamics simulations for the complete catalytic domain of the Cdc25B phosphatase is presented. The stability of the C-terminal α-helix is confirmed, but the last 20 residues in the complete catalytic domain are very flexible, partially occlude the active site and may establish transient contacts with the protein core. This flexibility in the C-terminal tail may modulate the molecular recognition of natural substrates and competitive inhibitors by Cdc25B.

Proteins 2016; 84:1567–1575.

Key words: NMR spectroscopy; protein dynamics; molecular dynamics; computer simulation; protein phosphatase.

INTRODUCTION

Progression in the cell cycle is controlled by molecular checkpoints where activation of cyclin-dependent kinases (Cdk) by Cdc25 phosphatases is an essential step.[1] Three Cdc25 homologous forms (A, B, and C) are encoded in the human genome. For instance, Cdc25A has been linked to the G1/S transition of the cell cycle, while Cdc25B acts in the G2/M transition.[2] Not surprisingly, the three forms have been found to be over-expressed in several types of human tumors,[3] and Cdc25 phosphatases have become active targets for the development of new anticancer drugs.[4,5]

Cdc25 phosphatases are considered members of the protein tyrosine phosphatase (PTP) family,[6] but they share very low sequence and structural similarity with other PTPs besides the conserved active site motif Cys–$(Xxx)_5$–Arg, known as the phosphate binding loop (P-loop). This low similarity may be useful for the design of selective inhibitors. Cdc25s also show substrate promiscuity as they dephosphorylate both pThr and pTyr residues found in Cdk via the same cysteine side-chain nucleophile.[1,7,8]

The similarity between the amino-acid sequences of the three human Cdc25 homologous forms is approximately 70% for the complete catalytic domain but drops to less than 45% in the C-terminus (last 33 residues). Crystal structures are available for Cdc25A and Cdc25B catalytic domains.[9,10] The most notable structural difference between the two forms is a C-terminal α-helix found only in Cdc25B, whereas a coiled coil is found in the same region of Cdc25A. Additional segments of 27 (Cdc25A) and 16 (Cdc25B) amino-acids were present in the C-terminus of protein constructs used for crystallography, but structures for these terminal regions were not provided in both crystallographic models probably due to an undefined electronic density.[9,10]

Previous computer simulations suggested that the C-terminal α-helix observed in Cdc25B could be (partially) unfolded in solution.[11,12] But, recent nuclear

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

Grant sponsor: CNPq and FAPESP projects; Grant numbers: 12/00543–1 (R.S.R.S.), 08/55914–9, and 14/19439–5 (S.R.M.), 13/17883–2 (R.K.S.), and 14/21900–2 (G.M.A.).
S.R.M. and G.M.A. are CNPq research fellows.
The authors declare no conflict of interests.

*Correspondence to: Guilherme M. Arantes; Departamento De Bioquímica, Instituto De Química, Universidade De São Paulo, Av. Prof. Lineu Prestes 748, São Paulo, Brazil. E-mail: garantes@iq.usp.br

Received 27 April 2016; Revised 17 June 2016; Accepted 1 July 2016
Published online 13 July 2016 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.25100

magnetic resonance (NMR) experiments refuted this conclusion and provided evidence that the C-terminal helix has little flexibility and is well folded in the ps–ms time scale.[13] Unfortunately, the protein construct used for that NMR study lacked the last 15 residues present in the full-length C-terminal tail of wild-type (WT) Cdc25B.

Notably, the C-terminus is involved in Cdc25 substrate recognition and discrimination between Cdk/Cyclin complexes.[14,15] Removal of residues 556R and 562R in the Cdc25B C-terminal tail decreases more than 10-fold the phosphatase activity toward the bis-phosphorylated Cdk2-pYpT/Cyclin-A natural substrate.[14] Although the homologous Cdc25C has a 100-fold lower activity toward Cdk2-pYpT/Cyclin-A, a chimeric construct with the last 18 residues from the Cdc25B C-terminal tail fused to the Cdc25C catalytic core shows a gain-of-function with a 50-fold increase in activity.[14] Given the spatial proximity of the C-terminus and the P-loop in the Cdc25B structure,[10] the terminal region could play important roles in complexation processes such as: release and egress of the product inorganic phosphate,[7,8] binding and stability of bulky artificial phosphatase substrates such as O-methyl fluorescein phosphate,[5] and complexation of competitive inhibitors.[11]

Here, the conformational flexibility of the *complete* catalytic domain of Cdc25B was probed with experimental NMR and molecular dynamics simulations. The last 20 residues in the Cdc25B C-terminus were found to be highly flexible, partially occlude the active site and establish contacts with the protein core.

METHODS

Cloning, expression, and purification of Cdc25B

The complete Cdc25B catalytic domain (residues 373S–566Q) and a C-terminal truncated construct (residues 373S–534E) were amplified from a plasmid purchased from Origene (code: NM_021873) with the cDNA sequence of human Cdc25B and cloned into pET-28a vector using NdeI/BamHI restriction enzymes.

Expression of Cdc25B fused to a His_6-tag followed by thrombin cleavage site at the N-terminal end was done in *Escherichia coli* BL21 (DE3-Gold). Cells were grown in M9 medium labeled with $^{15}NH_4Cl$ and ^{13}C-glucose at 37°C until they reach log phase of growth (OD_{600nm} = 0.6–0.8). At this point, the temperature was set to 20°C and the protein expression was induced with 0.8 m*M* IPTG for 18 h before collection of cells by centrifugation. NMR samples used for backbone assignment were expressed in M9 medium prepared with ^2H_2O to allow random and partial deuteration of the protein.

Cells were lysed by sonication in buffer containing 20 m*M* Tris pH 7.4, 500 m*M* NaCl, 10% glycerol, 5 m*M* β-mercaptoethanol, and 1 m*M* phenylmethylsulfonyl fluoride (PMSF). After centrifugation, the supernatant was incubated with Ni^{2+} resin and washed with phosphate buffer (20 m*M* NaPi, pH 6.7, 500 m*M* NaCl, 5 m*M* β-mercaptoethanol, and 20 m*M* imidazole) before elution of the protein with phosphate buffer containing 500 m*M* imidazole.

The His_6-tag was cleaved by incubation with thrombin–agarose resin (Sigma) under gentle shaking for 2 h at room temperature before submitting to size exclusion chromatography on a Hiload 26/600 Superdex 75 (GE Healthcare) column equilibrated with buffer containing 20 m*M* NaPi (pH 6.7), 50 m*M* NaCl, 2 m*M* β-mercaptoethanol, and 2 m*M* DTT. The homogeneity of eluted fractions was evaluated by SDS polyacrylamide gel electrophoresis. Homogeneous fractions were pooled and concentrated for NMR spectroscopy experiments using an Amicon Ultra device (cut-off 10 kDa, Millipore).

NMR experiments and assignment of the complete catalytic domain

Samples of approximately 0.2 m*M* $^{13}C/^{15}N$-labeled and partially deuterated Cdc25B in buffer containing 20 m*M* NaPi (pH 6.7), 50 m*M* NaCl, 5% ^2H_2O, 2 m*M* β-mercaptoethanol, and 2 m*M* DTT were used in NMR experiments.

Backbone resonance assignments were obtained from the analysis of a standard set of three-dimensional experiments with deuterium decoupling: HNCA, HN(CO)CA, HN(CA)CB, HN(COCA)CB, HNCO, HN(CO)CA, using traditional pulse sequences.[16] All spectra were acquired at 298 K on a 800 MHz Bruker Avance™ III spectrometer equipped with a cryogenic probe and processed with NMRPipe package.[17] Spectra visualization and assignment were done in CCPNMR Analysis 2.4 software.[18]

A chemical shift perturbation (CSP) index of NH amide groups in the backbone was calculated as:

$$\mathrm{CSP} = \sqrt{(\Delta\delta_{1H})^2 + 0.15(\Delta\delta_{15N})^2}, \qquad (1)$$

where $\Delta\delta_X$ is the difference between the chemical shifts measured for X= ^1H or ^{15}N nuclei.

Relaxation rates, heteronuclear NOE, and model-free analysis

Spin-lattice relaxation (R_1) and rotating frame relaxation ($R_{1\rho}$) rates of backbone amide resonances from uniformly labeled ^{15}N Cdc25B were measured by inverse detected 2D NMR experiments.[19,20] R_1 and $R_{1\rho}$ relaxation rates were calculated using 12 different relaxation delays for the acquisition of pseudo-3D planes. Delay times used for R_1 were: 0.1, 0.2, 0.35, 0.5, 0.6, 0.75, 0.9, 1.2, 1.4, 1.6, 1.8, and 2.0 s. Delay times used for $R_{1\rho}$ were: 0.01, 0.024, 0.04, 0.054 (2×), 0.08, 0.096, 0.112,

0.124, 0.14, 0.16, and 0.18 s. Exponential decay of peak intensities were fitted in MATLAB®R2009b software. Spectra were acquired with 16 scans and a recycle delay of 3 s.

Spin–spin relaxation rates (R_2) were calculated from $R_{1\rho}$ rates using the following relation:[20]

$$R_{1\rho} = R_2\cos^2\theta + R_1\sin^2\theta$$

where $\theta = \tan^{-1}(2\pi\Delta\nu/\gamma_N B_1)$, $\Delta\nu$ is the resonance offset and $\gamma_N B_1$ is the strength of the spin-lock field. A spin-lock field-strength of 1 kHz was used in the $R_{1\rho}$ experiments.

Heteronuclear NOE of ^{15}N spins was obtained from the ratios of the peak intensities on 2D ^{1}H–^{15}N spectra recorded in the presence or in the absence of proton saturation. Saturation time of 3 s and recycle delay of 6 s were used.

Order parameters of ^{15}N spins were calculated with ModelFree[21] and FAST-Modelfree[22] programs. The central structure of the most populated cluster from a MD simulation of the Cdc25B complete catalytic domain was used to obtain the initial parameters of an axially symmetric diffusion tensor using PDBinertia[23] to align the protein reference axis with its moments of inertia and R2R1_diffusion program:[24] $\tau_c = 12.6$ ns, $D_{\parallel/\perp} = 1.26$, $\theta = [0]^\circ$ and $\phi = 0^\circ$.

Parameters were adjusted with ^{15}N chemical shift anisotropy and NH bond length values of -160 ppm and 1.02 Å, respectively. The axially symmetric rotational diffusion tensor was optimized together with the motional models for the ^{15}N spins in a global grid search that converged after 5 iteractions and 600 Monte Carlo simulation steps were used to estimate the precision of fitted parameters.

Residual dipolar couplings

Partial alignment was achieved in liquid crystalline media prepared with PEG C12E5/n-hexanol.[25] Scalar couplings of backbone amides ($^{1}J_{NH}$) were measured in isotropic media using anti-phase in-phase experiments.[26] Residual dipolar couplings of backbone amides (D_{NH}) were calculated by subtracting $^{1}J_{NH}$ from NH splittings measured in the presence of anisotropic media.

A comparison between calculated and measured D_{NH} was performed with the DC program included in NMRPipe package.[17] The crystal structure of the Cdc25B catalytic domain (PDB: 1QB0)[10] was used to calculate D_{NH}. Hydrogens were added to the PDB structure with pdb2gmx tool, available in GROMACS suite of programs,[27] followed by an energy minimization procedure with heavy atoms restrained to their initial coordinates.

Molecular dynamics simulation

Simulations of the Cdc25B complete catalytic domain were started with coordinates from the crystal structure deposited in PDB (code 1QB0). Additional N-terminal residues (GSHMEFQ) resulting from the adapter sequence used for cloning and the C-terminal residues (552G–566Q) absent in the crystal structure were manually added in an extended conformation. Crystallographic waters, ions, and β-mercaptoethanol were removed, hydrogens were added to the structure and an inorganic phosphate was modeled in the P-loop in the dianionic form. A MD simulation in implicit solvent with coordinate restrains for all atoms present in the crystal structure was initially carried out for 12 ns. The final structure was solvated in a dodecahedron water box and NaCl was added to neutralize the system up to a 10 m*M* salt concentration.

The Amber99SB-ILDN force field[28–30] was used for protein and ions and the TIP3P model[31] was used for water molecules. A 2 μs-long MD simulations was performed at 300 K and 1 bar pressure with a 2 fs integration step in the NPT ensemble, using the structure described in the last paragraph as a starting geometry. The PME method[32] with 0.9 nm real space cutoff and a 0.15 nm Fourier spacing in the mesh was used to treat long-range electrostatics. The Berendsen barostat[33] and the Bussi thermostat[34] were both used with period 0.5 ps. Construction of the models, MD simulations and analysis of results were performed with the GROMACS 4.6.1 suite of programs.[27]

RESULTS AND DISCUSSION

Assignment of the Cdc25B complete catalytic domain

Backbone atoms of the complete catalytic domain of WT Cdc25B (segment 373S–566Q) were assigned by conventional NMR triple resonance experiments. The ^{1}H–^{15}N HSQC spectrum [Fig. 1(A)] showed relatively good peak dispersion, except for the region around $\delta^{1H} = 8.2$ ppm and $\delta^{15N} = 120$ ppm which suggests the presence of conformational disorder in the polypeptide chain.[35] Out of 187 expected backbone amide resonances in the HSQC spectrum, 173 were observed and 143 were assigned (~82%). All backbone amide resonances of the C-terminal segment 552G–566Q absent in the crystal structure and in the Cdc25B construct used by Lund and Cierpicki[13] were located in regions of the HSQC spectrum expected for residues in random coil conformation (red dots in Fig. 1).[35]

The assigned backbone resonances (^{1}H and ^{15}N) were in close agreement with the chemical shifts provided by Lund and Cierpicki, except for residues 434G, 509K, 511G, and 529R. These differences are probably due to a shorter spectral window used by Lund and Cierpicki for the acquisition of NMR spectra in the ^{15}N dimension.[13] The complete assignment obtained here was deposited in the BMRB databank (accession number 26779).

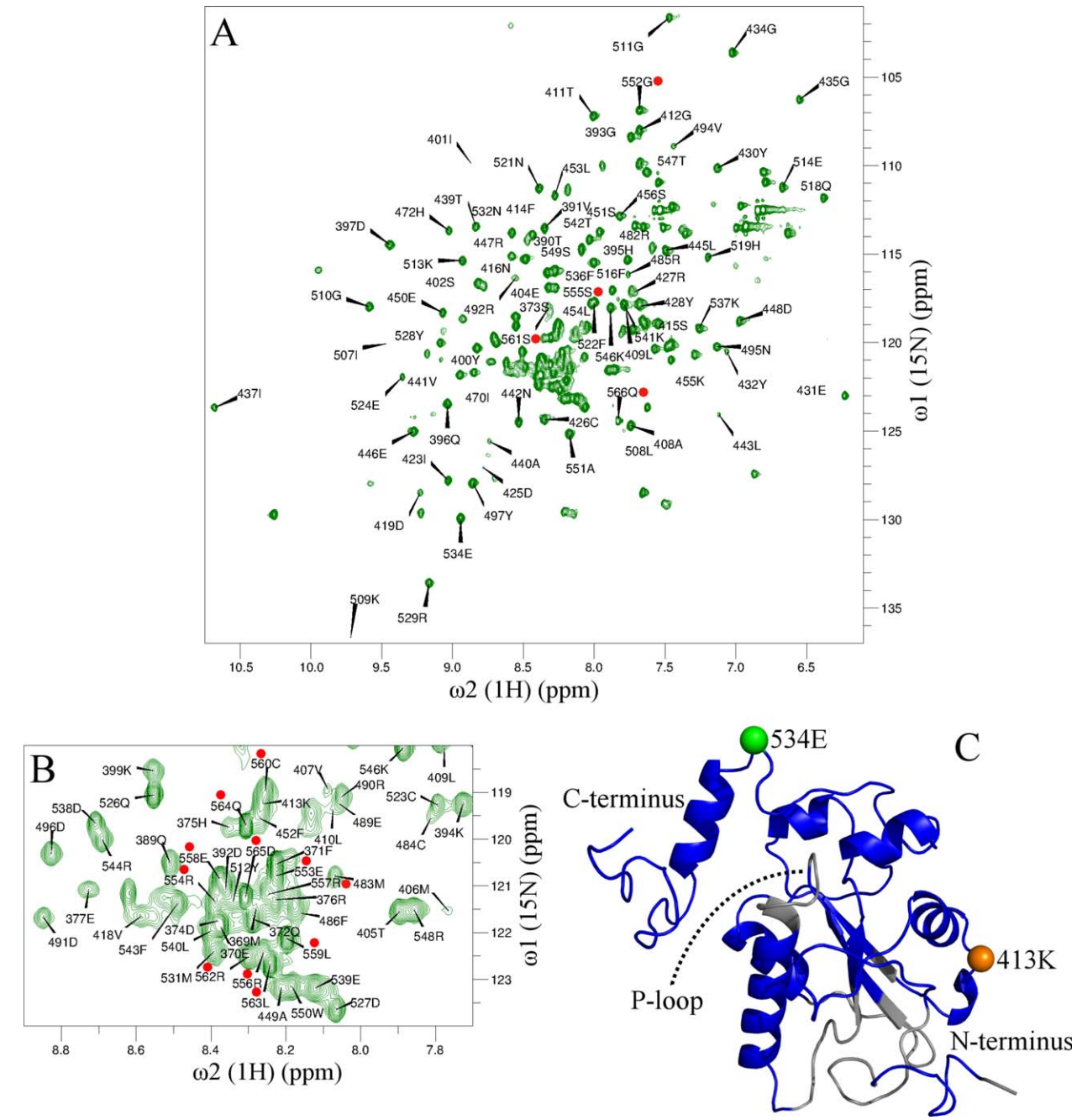

Figure 1

Assignment of backbone amide resonances. Panel **A** shows the ^{1}H–^{15}N HSQC spectrum of the Cdc25B complete catalytic domain where residues of the C-terminal segment (552G–566Q) are marked with red dots. Panel **B** is a magnification of the central region in the spectrum. Panel **C** shows a snapshot structure where Cdc25B regions with assigned NMR chemical shift are colored in blue and residues 534E and 413K are represented by balls. [Color figure can be viewed at wileyonlinelibrary.com]

Most of the unassigned or unobserved resonances correspond to segments of the protein located in exposed loop regions [378L–388L, 458I–467V, and 499S–502Y, Fig. 1(C)] where conformational exchange should broaden the NMR signal. Although relatively rigid, the active site P-loop (474E–480G) was also not assigned due to signal broadening probably caused by intermediate rate exchange of $HOPO_3^{2-}$ ions with the buffer, as also observed in a low molecular weight protein tyrosine phosphatase.[36] Part of this P-loop (475F–480G) was assigned by Lund and Cierpicki,[13] but using a truncated protein construct and with a point mutation at the

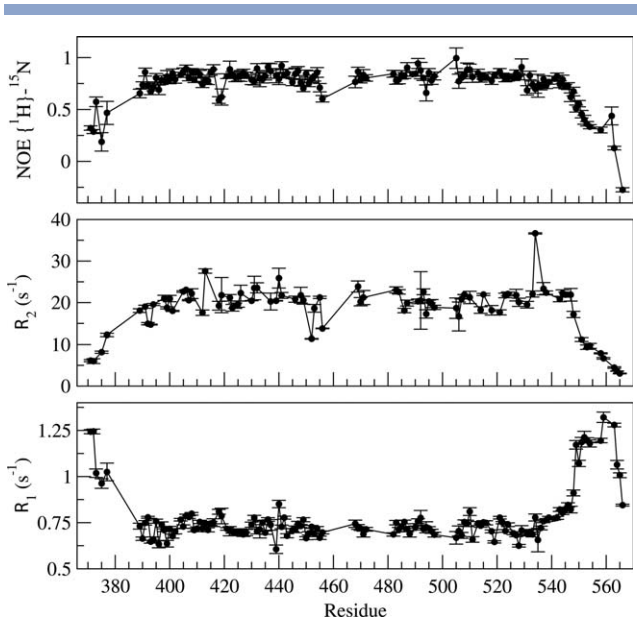

Figure 2

R_1 and R_2 relaxation rates and $\{^1H\}$–^{15}N NOE of backbone ^{15}N spins measured for the Cdc25B complete catalytic domain.

catalytic cysteine (C473S) that could alter the binding and dynamic proprieties of the active site.

Residues 373S–377E, 394K, 396Q, 416N, 439T, 446E, 495N–497Y, 510G, and 547T–550W were assigned here but not by Lund and Cierpicki. On the other hand, they assigned 387L, 388L, 420K–422V, 433E, 438K, 459A, 468I, 488R, 505M, 506Y, and 535A, which was not done here. These differences may be attributed to the partial deuteration of the protein conducted here and to different buffers used in protein samples for acquisition of NMR spectra.

Fast dynamics in the Cdc25B terminals

Values of R_1 and $R_2^{15}N$ relaxation rates and heteronuclear $\{^1H\}$-^{15}N NOE of the N-terminal 373S–377E and C-terminal 547T–566Q segments shown in Figure 2 are clearly different from the rest of the protein. Higher R_1 and lower R_2 and NOE values found in the Cdc25B terminals in comparison to the rest of the protein chain are related to the presence of considerable motion in the ps–ns time scale in these terminal regions.

The amplitude of motions on fast (ps–ns) time scales was quantified with order parameters (S^2) adjusted from the ^{15}N relaxation and NOE measurements. Lower S^2 values also pointed to increased conformational flexibility of the terminal regions and S^2 values close to unity indicate a relative rigidity of the protein core [Fig. 3(A)]. A significant contribution of conformational exchange (R_{ex}) for residues 413K and 534E [Fig. 1(C)] was determined here, indicating motion in the μs–ms time scale for these two residues [Fig. 3(B)]. Lund and Cierpicki have not observed a large R_{ex} contribution for 534E, but their R_2 value for the same residue in the WT Cdc25B is considerably higher than in the C473S mutant, suggesting that this point mutation has altered the conformational flexibility around the active site. Correlation times obtained from adjustment of relaxation measurements are presented as Supporting Information (Fig. S1).

The last 20 residues in Cdc25B C-terminus are disordered

Experimental residual dipolar couplings (D_{NH}) of the Cdc25B complete catalytic domain measured in neutral PEG/hexanol alignment media agree well with D_{NH} calculated for the residues present in the crystal structure with a $Q = 0.26$ quality factor as shown in Figure 4(A). Thus, the protein core and the C-terminal α-helix are well folded and populate conformations in solution similar to the crystal structure.

However, the observed D_{NH}^{obs} for residues in the C-terminus (549S–566Q) decay to values closer to 0 Hz when compared with the rest of polypeptide chain [Fig. 4(B)], suggesting again conformational disorder in this region. The same is also observed for the N-terminus. Given that the Cdc25B C-terminus is enriched with charged residues, partial alignment in a neutral media as performed here should perturb the C-terminus conformational distribution less than alignment in a charged media.

The global flexibility of Cdc25B has been debated in the literature.[11–13] Previous simulations employing either molecular dynamics[11] with the OPLS-AA force field[37] or a library of rotamers sampled by Monte Carlo[12] suggested that the Cdc25B C-terminal α-helix may (partially) unfold in solution. NMR experiments

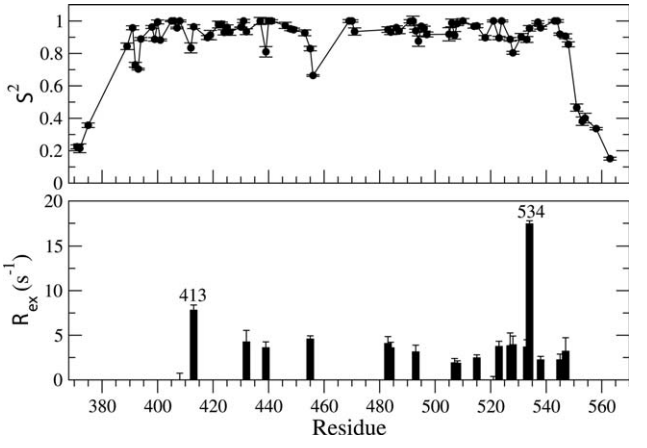

Figure 3

Order parameters (S^2) and rates of conformational exchange (R_{ex}) obtained after the adjustment of relaxation measurements with Lipari–Szabo formalism.

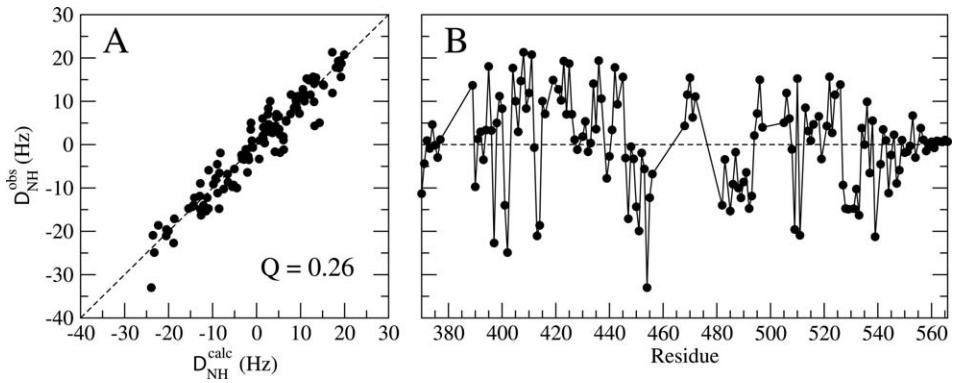

Figure 4

Residual dipolar couplings of backbone amides of the Cdc25B complete catalytic domain. Panel **A** shows a comparison between D_{NH} calculated from crystal structure (PDB: 1QB0) and measured in neutral PEG/n-hexanol alignment media. Panel **B** shows the measured D_{NH} for each residue.

presented previously for a truncated construct of the Cdc25B catalytic domain[13] and obtained here for the complete catalytic domain show this C-terminal α-helix is folded and has little flexibility. Thus, the energy functions used in previous simulations[11,12] artificially destabilized helical conformations. This supports the notion that force fields should be carefully tested and further developed, and simulations should be compared and validated with experimental data.[38,39]

Although NMR measurements are excellent probes of protein dynamics in solution, the broad time range of small to large scale motions expected in disordered protein regions results in highly averaged NMR observables that often cannot be used to resolve conformational substates.[40] Thus, in order to investigate the conformations visited by the C-terminus and its possible contacts with the protein core, a long MD simulation was carried out with the Amber99SB-ILDN force field, which has a balanced helix–coil description and has been used successfully to simulate disordered protein regions.[38,41]

Figure 5 shows that the protein core (residues 373–546) and the C-terminal helix (534–546) remained stable and well folded during the whole MD trajectory, in agreement with the NMR observations. The C-terminal tail has a large deviation from the initial structure. The tail shows periods of high flexibility (between 450–750 ns and 1500–1900 ns) and periods of relative stability (between 800 and 1500 ns) when it forms transient contacts with the protein core. The Cα root mean-square fluctuations [RMSF, Fig. 5(B)] characterize large flexibility and local disorder in the Cdc25B terminals. The simulated RMSF profile is similar to the fluctuations derived from random coil index (RCI) analysis of referenced[42] backbone (CA, CB, C, and N) chemical shifts.[43] This is particularly true for the protein C-terminal and suggests that the simulation approximately samples the amplitude of movements for this region. Backbone order parameters (Supporting Information Fig. S2) and residual dipolar couplings (Supporting Information Fig. S3) back-calculated from the MD simulation trajectory are also in good agreement with the measured values.[44,45]

Structures obtained from the MD trajectory [Fig. 5(C)] show that the C-terminus fluctuates widely and occupies a large volume close to the active site. Thus, it may directly interact with possible competitive inhibitors binding to Cdc25B.[11] Note that the side-chains of residues 556R and particularly 562R, involved in Cdc25 substrate recognition, are mostly exposed to solvent during the MD trajectory. However, some residual structure should still be found around residues 553E–557R which display smaller RMSF, a plateau in the S^2 values (Fig. 3) and D_{NH}^{obs} different from zero (Fig. 4).

The flexible C-terminus may establish contacts with the protein core

The molecular dynamics simulation also suggests that the flexible C-terminus may establish transient contacts with exposed side-chains in helical regions on the protein core. Notably, there is a large fraction of arginine residues (13 out of the 17 found in the Cdc25B catalytic domain) exposed to solvent in the surface of the P-loop helix, the C-terminal helix and the disordered C-terminus. Although these arginine side-chains form transient salt-bridges with acidic residues (e.g., between 482R and 553E), more stable contacts with half-life of 100–300 ns are cation–π stacking interactions[46,47] formed between the indole of 550W and the guanidinium groups of 447R, 482R, 485R, and 544R [Fig. 6(A); Supporting Information Fig. S4).

A multiple sequence alignment of the human Cdc25B with sequences from other species indicates that residues 482R, 485R, 544R, and 550W are strongly conserved (Supporting Information Fig. S5)[48] and suggests a possible functional role for these residues and their transient contacts. On the other hand, 447R is not conserved and

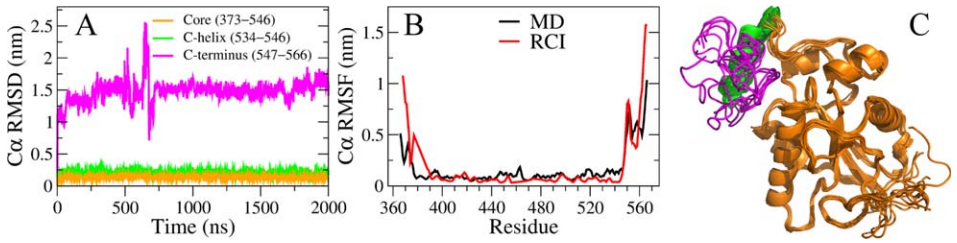

Figure 5

Conformational flexibility of the Cdc25B complete catalytic domain probed by a molecular dynamics simulation. Panel **A** shows Cα root mean-square deviations (RMSD) of the protein core (orange), C-terminal helix (green) and C-terminal tail (magenta) with the respective primary sequences given in brackets. Panel **B** shows root mean-square fluctuations (RMSF) of the trajectory (black) and derived from RCI analysis (red). Panel **C** shows a superposition of eight representative snapshots taken from the trajectory.

found only in a few sequences other than the human one.

In order to probe experimentally the stability of some of these transient contacts, the ^{1}H–^{15}N HSQC spectra measured for the complete Cdc25B catalytic domain (segment 373S–566Q) was compared with the spectra of a truncated construct (segment 373S–534E) which lacks the last 32 residues containing the C-terminal α-helix and disordered region, including 550 W found to establish transient cation–π contacts in the MD simulation [Fig. 6(B)]. The highest chemical shift pertubations are observed in the C-terminus of the truncated construct [residues 528Y–534E, Fig. 6(C)] and in regions in close contact to it such as the α-helix segments 426C–435G and 510G–513K [red and blue segments in Fig. 6(A,C)]. The increased flexibility and different conformational distribution of the new terminal region in the truncated construct should cause a large perturbation on its NMR spectra and of the nearby protein regions.

Smaller but still significant perturbations ($0.1 < \mathrm{CSP} < 0.05$ ppm) are observed for segments 443L–449A and 482R–487I [Fig. 6(A,C)]. These regions are spatially

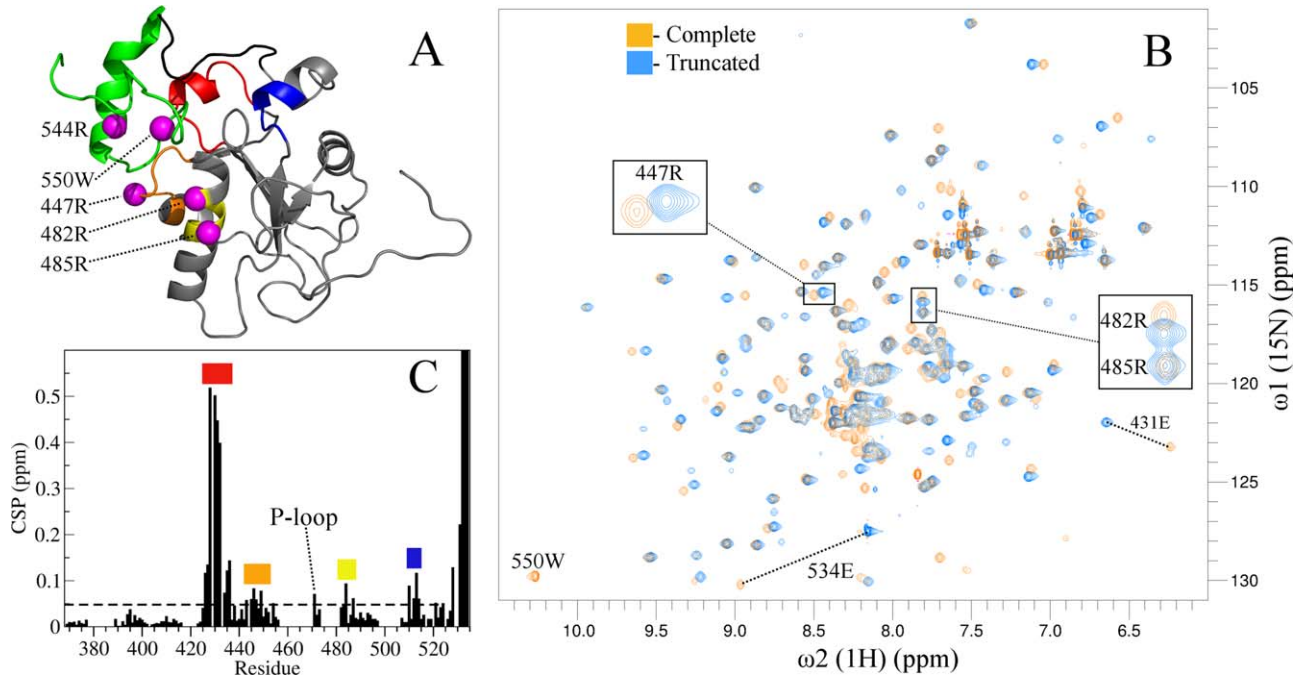

Figure 6

Possible transient contacts between the disorded C-terminal and the stable Cdc25B protein core. Panel **A** shows residues forming transient cation–π contacts with 550 W in the MD simulations. The region colored in green corresponds to the removed segment in the truncated construct. Regions in red, orange, yellow, blue, and black correspond to residues with CSP greater than 0.05 ppm as shown in Panel **C**. Panel **B** shows a superposition of the ^{1}H–^{15}N HSQC spectra for the complete (orange) and truncated (blue) constructs. Panel **C** shows the CSP index for each residue calculated from the perturbations measured in Panel **B**. [Color figure can be viewed at wileyonlinelibrary.com]

distant from the terminal residues of the truncated construct. But, accordingly to the molecular dynamics simulation presented above, these two segments make contacts with the disordered C-terminus of the complete construct, particularly transient cation–π interactions of 550W with 447R, 482R, and 485R. Thus, in the truncated construct the observed CSP on the two segments 443L–449A and 482R–487I, that contain the arginine residues just mentioned, might be caused by disruption of contacts with the disordered region found in the complete Cdc25B construct.

CONCLUSIONS

The conformational flexibility of the Cdc25B complete catalytic domain was evaluated with experimental NMR measurements and MD simulations. The protein construct used here contains all residues at the C-terminus that were not present in studies previously published.[11–13]

Our results show that the last 20 C-terminal residues (547T–566G) show large fluctuations when compared with the rest of the chain. Conformational disorder for the C-terminal tail is supported by a range of observations such as NMR chemical shifts compatible with random coil conformation, ^{15}N relaxation rates in accord with fast backbone motions, dynamic averaging of residual dipolar couplings, and large mean fluctuations in the long time MD simulation.

The C-terminal helix is well folded during the whole MD trajectory presented here (Fig. 5), in disagreement with previous simulations.[11,12] This is due to the previous use of the OPLS-AA force field which has been shown to give an unbalanced description of helix–coil conformations.[38,41] Additionally, the inclusion of the last 15 residues in the C-terminal tail might stabilize the terminal helix in the simulations.

The large fluctuations in position observed here in the MD simulations are validated by the RCI indexes derived from the NMR chemical-shifts. The volume occupied by the terminal residues partially occludes the entrance of the shallow Cdc25B active site. Transient contacts between the C-terminus and the protein core were also observed.

The combination of experimental NMR spectroscopy and computer simulation is a powerful method to describe in atomic detail the conformational flexibility of proteins in solution. We are now applying this combination to determine the role of the C-terminal tail in the recognition of small-molecule ligands by Cdc25B.

REFERENCES

1. Boutros R, Dozier C, Ducommun B. The when and wheres of CDC25 phosphatases. Curr Opin Cell Biol 2006;18:185–191.
2. Rudolph J. Cdc25 phosphatases: structure, specificity, and mechanism. Biochemistry 2007;46:3595–3604.
3. Boutros R, Lobjois V, Ducommun B. CDC25 phosphatases in cancer cells: key players? Good targets? Nat Rev Cancer 2007;7:495–507.
4. Lavecchia A, Di Giovanni C, Novellino E. CDC25A and B dual-specificity phosphatase inhibitors: potential agents for cancer therapy. Curr Med Chem 2009;16:1831–1849.
5. Lazo JS, Wipf P. Is Cdc25 a druggable target? Anti-Cancer Agents Med Chem 2008;8:837–842.
6. Zhang ZY. Protein-tyrosine phosphatases: biological function, structural characteristics, and mechanism of catalysis. Crit Rev Biochem Mol Biol 1998;33:1–52.
7. Arantes GM. Free-energy profiles for catalysis by dual-specificity phosphatases. Biochem J 2006;399:343–350.
8. Arantes GM. The catalytic acid in the dephosphorylation of the Cdk2-pTpY/CycA protein complex by Cdc25B phosphatase. J Phys Chem B 2008;112:15244–15247.
9. Fauman EB, Cogswell JP, Lovejoy B, Rocque WJ, Holmes W, Montana VG, Piwnica-Worms H, Rink MJ, Saper MA. Crystal structure of the catalytic domain of the human cell cycle control phosphatase, Cdc25A. Cell 1998;93:617–625.
10. Reynolds RA, Yem AW, Wolfe CL, Deibel MR, Chidester CG, Watenpaugh KD. Crystal structure of the catalytic subunit of Cdc25B required for G2/M phase transition of the cell cycle. J Mol Biol 1999;293:559–568.
11. Arantes GM. Flexibility and inhibitor binding in Cdc25 phosphatases. Proteins 2010;78:3017–3032.
12. Mamonov AB, Bhatt D, Cashman DJ, Ding Y, Zuckerman DM. General library-based monte carlo technique enables equilibrium sampling of semi-atomistic protein models. J Phys Chem B 2009; 113:10891–10904.
13. Lund G, Cierpicki T. Solution NMR studies reveal no global flexibility in the catalytic domain of CDC25B. Proteins 2014;82:2889–2895.
14. Wilborn M, Free S, Ban A, Rudolph J. The C-terminal tail of the dual-specificity Cdc25B phosphatase mediates modular substrate recognition. Biochemistry 2001;40:14200–14206.
15. Sohn J, Parks JM, Buhrman G, Brown P, Kristjansdottir K, Safi A, Edelsbrunner H, Yang W, Rudolph J. Experimental validation of the docking orientation of Cdc25 with its Cdk2-CycA protein substrate. Biochemistry 2005;44:16563–16573.
16. Sattler M. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog Nucl Magn Reson Spectrosc 1999;34:93–158.
17. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 1995;6:277–293.
18. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 2005;59:687–696.
19. Kay LE, Torchia DA, Bax A. Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. Biochemistry 1989;28: 8972–8979.
20. Massi F, Johnson E, Wang C, Rance M, Palmer AG. NMR $R_{1\rho}$ rotating-frame relaxation with weak radio frequency fields. J Am Chem Soc 2004;126:2247–2256.
21. Palmer AG. ModelFree 4.2. http://www.palmer.hs.columbia.edu/software/modelfree.html; 2014a.
22. Cole R, Loria JP. FAST-Modelfree: a program for rapid automated analysis of solution NMR spin-relaxation data. J Biomol NMR 2003;26:203–213.
23. Palmer AG. PDBinertia 1.11. http://www.palmer.hs.columbia.edu/software/pdbinertia.html, 2014b.
24. Palmer AG. R2R1_diffusion 1.11. http://www.palmer.hs.columbia.edu/software/r2r1_diffusion.html, 2014c.
25. Rückert M, Otting G. Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. J Am Chem Soc 2000;122:7793–7797.

26. Cordier F, Dingley AJ, Grzesiek S. A doublet-separated sensitivity-enhanced HSQC for the determination of scalar and dipolar one-bond J-couplings. J Biomol NMR 1999;13:175–180.
27. Hess B, Kutzner C, Van Der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 2008;4:435–447.
28. Wang J, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J Comput Chem 2000; 21:1049–1074.
29. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 2006;65:712–725.
30. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 2010;78:1950–1958.
31. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926
32. Darden T, York D, Pedersen L. Particle mesh Ewald: an N log(N) method for Ewald sums in large systems. J Chem Phys 1993;98:10089–10092.
33. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
34. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys 2007;126:014101
35. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD. ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. J Biomol NMR 1995;5:67–81.
36. Stehle T, Sreeramulu S, Löhr F, Richter C, Saxena K, Jonker HRA, Schwalbe H. The apo-structure of the low molecular weight protein-tyrosine phosphatase A (MptpA) from *Mycobacterium tuberculosis* allows for better target-specific drug development. J Biol Chem 2012;287:34569–34582.
37. Jorgensen WL, Jorgensen WL, Maxwell DS, Tirado-Rives J, Tirado-Rives J. Development and testing of the olps all-atom force field on conformational energetics and properties of organic liquids. J Am Chem Soc 1996;118:11225–11236.
38. Rauscher S, Gapsys V, Gajda MJ, Zweckstetter M, de Groot BL, Grubmüller H. Structural ensembles of intrinsically disordered proteins depend strongly on force field: a comparison to experiment. J Chem Theory Comput 2015;11:5513–5524.
39. Best RB, Zhu X, Shim J, Lopes PE, Mittal J, Feig M, MacKerell AD. Jr Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain $\chi 1$ and $\chi 2$ dihedral angles. J Chem Theory Comput 2012;8:3257–3273.
40. Brookes DH, Head-Gordon T. Experimental inferential structure determination of ensembles for intrinsically disordered proteins. J Am Chem Soc 2016;138:4530–4538.
41. Palazzesi F, Prakash MK, Bonomi M, Barducci A. Accuracy of current all-atom force-fields in modeling protein disordered states. J Chem Theory Comput 2015;11:2–7.
42. Zhang H, Neal S, Wishart DS. RefDB: a database of uniformly referenced protein chemical shifts. J Biomol NMR 2003;25:173–195.
43. Berjanskii MV, Wishart DS. A simple method to predict protein flexibility using secondary chemical shifts. J Am Chem Soc 2005; 127:14970–14971.
44. Showalter SA, Bruschweiler R. Validation of molecular dynamics simulations of biomolecules using nmr spin relaxation as benchmarks: application to the amber99sb force field. J Chem Theory Comput 2007;3:961–975.
45. Maragakis P, Lindorff-Larsen K, Eastwood MP, Dror RO, Klepeis JL, Arkin IT, Jensen MO, Xu H, Trbovic N, Friesner RA, Palmer AG, Shaw DE. Microsecond molecular dynamics simulation shows effect of slow loop dynamics on backbone amide order parameters of proteins. J Phys Chem B 2008;112:6155–6158.
46. Ma JC, Dougherty DA. The cation-π interaction. Chem Rev 1997; 97:1303–1324.
47. Gallivan JP, Dougherty DA. A computational study of cation-π interactions vs salt bridges in aqueous media: implications for protein engineering. J Am Chem Soc 2000;122:870–874.
48. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. Bioinformatics 2007;23:2947–2948.