Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 27 (2008) 1–12
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Molecular modeling study of CodX reveals importance of N-terminal and C-terminal domain in the CodWX complex structure of *Bacillus subtilis*

Navaneethakrishnan Krishnamoorthy [a], Poornima Gajendrarao [a], Soo Hyun Eom [b], Yong Jung Kwon [c], Gang-Won Cheong [a], Keun Woo Lee [a,*]

[a] *Department of Biochemistry, Division of Applied Life Sciences (BK21 Program), Environmental Biotechnology National Core Research Center (EBNCRC), Gyeongsang National University, Jinju 660-701, Republic of Korea*
[b] *Department of Life Science, Kwangju Institute of Science and Technology, Kwangju 500-712, Republic of Korea*
[c] *Department of Chemical Engineering, Kangwon National University, Chunchon 200-701, Republic of Korea*

Received 12 November 2007; received in revised form 17 January 2008; accepted 27 January 2008
Available online 23 February 2008

Abstract

In *Bacillus subtilis*, CodW peptidase and CodX ATPase function together as a distinctive ATP-dependent protease called CodWX, which participates in protein degradation and regulates cell division. The molecular structure of CodX and the assembly structure of CodW–CodX have not yet been resolved. Here we present the first three-dimensional structure of CodX N-terminal (N) and C-terminal (C) domain including possible structure of intermediate (I) domain based on the crystal structure of homologous *Escherichia coli* HslU ATPase. Moreover, the biologically relevant CodWX ($W_6W_6X_6$) octadecamer complex structure was constructed using the recently identified CodW–HslU hybrid crystal structure. Molecular dynamics (MD) simulation shows a reasonably stable structure of modeled CodWX and explicit behavior of key segments in CodX N and C domain: nucleotide binding residues, GYVG pore motif and CodW–CodX interface. Predicted structure of the possible I domain is flexible in nature with highly coiled hydrophobic region (M153–M206) that could favor substrate binding and entry. Electrostatic surface potential observation unveiled charge complementarity based CodW–CodX interaction pattern could be a possible native interaction pattern in the interface of CodWX. CodX GYVG pore motif structural features, flexible nature of glycine (G92 and G95) residues and aromatic ring conformation preserved Y93 indicated that it may follow the similar mode during the proteolysis mechanism as in the HslU closed state. This molecular modeling study uncovers the significance of CodX N and C domain in CodWX complex and provides possible explanations which would be helpful to understand the CodWX-dependent proteolysis mechanism of *B. subtilis*.

Keywords: ATPase; ATP-dependent protease; CodX; Electrostatic potential calculation; Homology modeling; HslU; MD simulation

1. Introduction

In all living cells, ATP-dependent proteases are major contributors for the protein degradation which is vital for eliminating abnormal proteins and regulation of many cellular processes [1]. The 26S proteasome, a macromolecular machine contains a barrel shaped proteolytic core complex (20S proteasome) and a 19S regulatory complex that utilizes ATP to degrade ubiquitilated proteins in eukaryotic cells [2–5]. The heat shock locus HslVU (ClpQY) protease of *E. coli* is homologous with the eukaryotic 26S proteasome [6–8]. In *E. coli* the *hslVU* operon encodes two heat shock proteins: HslV 19 kDa peptidase and HslU 50 kDa ATPase which can function together as a two-component protease [6–9]. HslU, a member of the Clp/Hsp100 family, itself has ATPase activity and also acts as a chaperone [10–12]. Members of this ATPase family participate in protein unfolding machinery and direct them into the inner proteolytic chamber of their associated peptidase [13,14].

The structure determination of HslU by X-ray crystallographic studies proved that it exists as a hexamer, each

* Corresponding author. Tel.: +82 55 751 6276; fax: +82 55 752 7062.
E-mail address: kwlee@gnu.ac.kr (K.W. Lee).

1093-3263/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2008.01.009

monomer folded into three distinct domains: the N (S2–K109 and I244–L332), I (M110–A243) and C (Q333–L443) domains [15]. In the HslVU complex, HslU contacts HslV through the N and C domain and the activation of HslV requires its binding to the C domain tail region of HslU located opposite to I domain [16]. Moreover, nucleotide is essential for the biologically active HslU hexameric assembly formation [13,17].

Pores are the channeling components that control and direct the protein substrates into the molecular machines. In HslU, structural features of GYVG pore motif are essential for the unfolding-coupled translocation mechanism. Three HslU conformational states were described previously on the basis of HslU primary pore motif Y91 side chain conformation, an essential amino acid that played a vital role in the translocation of the protein substrates [13]. Also, the flexible glycine residues (G90 and G93) at the central pore are significant for the degradation of natively folded proteins [14].

CodWX protease in *B. subtilis* and HslVU protease in *E. coli* exhibit 52% identity in the amino acid sequence of their corresponding peptidase and ATPase [18]. EzrA, a protein that functions in the regulation of *B. subtilis* cell division by inhibiting FtsZ ring formation [19], is degraded by ATP-dependent CodWX protease [20]. The sequence of *B. subtilis codVWXY* operon was identified by Slack et al. [18]. The *cod* operon encodes CodW peptidase and CodX ATPase together as a distinctive ATP-dependent CodWX protease which has N-terminal serine active site [21].

Among the other known ATP-dependent proteases, CodWX has a unique molecular architecture and proteolytic mechanism [20]. CodX, a single ring consisting hexameric or heptameric form was not found under all tested experimental conditions, but found as a dimer of two hexameric or heptameric rings forming a 'spool-like' structure, in which I domains are in contact with each other. CodW exists as a dodecamer that has two stacked dyad hexameric rings. Asymmetric WWXX or symmetric WWXXWW and elongated WWXXWWXX configurations are few other double ring CodWX complexes. One possibility of the CodWX complex formation is that a single ring form of CodX might be predominant under physiological conditions that could interact with CodW to form the WWX or XWWX. Another possibility is CodX double ring into single ring dissociation that could happen by the binding of protein substrates to the lateral side of the I domains, would result in the complex formation of the WWX or XWWX structure with CodW. In addition, experimental results suggest that ADP or ATP analogs are sufficient for CodX oligomerization.

In HslU, N and C domain play a major role in nucleotide binding, substrate unfolding and transformation, and interact with HslV to form the HslVU complex. But in CodWX, the 3D structure of CodX ATPase has not yet been resolved. Therefore, these domains roles are unknown. To unravel their importance we carried out this molecular modeling study. We built the monomer CodX model based on the homologous HslU crystal structure (1HT2_H). This model expressed the well-constructed structure of conserved N and C domain with one of the possible structures of I domain that lack structural template. There are no structural templates presently available to complete the model of I domain. Thus, we believe the structure prediction of part of this domain, in this study is only based on amino acid sequences and one possible structure out of many possibilities. However, we allow this domain to remain in our model to analyze its adaptation during MD simulation. The biologically irrelevant hybrid protein crystal, which consists of *E. coli* HslU and *B. subtilis* CodW was recently identified [22]. We used this X-ray structure of CodW–HslU hybrid and our monomer CodX model to build the biologically functional CodW–CodX (WWX) complex structure (Fig. 3B). To the best of our knowledge, this is the first giant CodWX octadecamer (18 subunits) protein (48,474 atoms) in water box (444,789 atoms) analyzed using MD simulation at neutral pH.

The main aim of this study is to obtain the biologically functional WWX complex structure and to reveal the importance of CodX N and C domain including key segments using dynamical observation. From the obtained WWX functional oligomer, we focused on the modeled CodX and its interaction with CodW and nucleotide, as CodW crystal structure on hand [22]. Though many studies have been previously reported on HslVU structures [15,13,17] and CodWX architecture [20], our dynamical elucidation is a novel approach for this multiprotein complex. Moreover, this study reveals: flexibility of I domain which could possibly disturb its crystallization, the surface charge complementarity which contributes significantly to the WX complex formation, conformational difference between HslU and CodX GYVG pore motifs and apical hydrophobic site of I domain that may favor substrate binding and unfolding.

2. Materials and methods

2.1. Homology modeling for CodX monomer structure

CodX monomer homology model was constructed using its amino acid sequence (accession number P39778) retrieved from UNI-PROT [23] protein sequence database. The homology modeling server SWISS-MODEL [24] First Approach mode was used to generate the model. We allowed the SWISS-MODEL server to select the template for the model construction. Protein Data Bank (http://www.rcsb.org) monomeric form of *E. coli* HslU crystal structure (1HT2_H) was used as template using SWISS-MODEL for the CodX monomer structure building. The stereochemical quality of the homology modeled CodX monomer was validated using Ramachandran plot generated using PROCHECK [25].

2.2. CodX hexamer and CodWX octadecamer complex construction

INSIGHTII [26] HOMOLOGY and BIOPOLYMER modules were used to obtain the hexameric assembly of CodX by superimposition of the modeled monomer with each subunit of HslU hexamer of CodW–HslU hybrid complex [22]. To optimize the CodX hexameric model in its one of the native configurations: the same hybrid crystal was used to produce the biologically applicable CodWX (WWX) octadecamer complex and the complex was subsequently refined using GROMACS

v3.3 [27] as described below. Refined model stereochemical quality was also validated using PROCHECK. This reliable model was further used for MD simulation study.

2.3. Molecular dynamics simulation of CodWX octadecamer complex

The CodWX octadecamer complex MD simulation was performed using GROMACS v.3.3 with GROMOS87 force field that describes the inter-atomic interactions. The molecular topology file and force field parameters for the dADP were generated using the PRODRG server [28]. Protonation states for protein ionizable groups were set to pH 7. CodWX complex was solvated by generating a triclinic box (15.18 nm × 16.39 nm × 20.72 nm) of Simple Point Charge (SPC) water model [29], with 1.8nm as the minimum distance between the protein and the edge of the box. To neutralize the system, water molecules at the most favorable electrostatic potential were replaced with sodium ions, which is essential to apply the Particle Mesh Ewald (PME) method for calculation of coulomb-interactions [30]. The giant MD simulation system composed of protein 48,474 (WWX), 444,789 water, 186 nucleotides (6dADP) and 96 Na^+, in total 493,545 atoms. Non-bonded interactions were evaluated by using twin range cut-off: 0.8 nm for electrostatic interactions and 1.4 nm for Lennard–Jones interactions, and the neighboring pair list was updated every fifth integration step. LINCS algorithm [31] was used to constrain the bond lengths within proteins and SETTLE algorithm [32] for constraining the geometry of water molecules. Each component of the system was coupled separately to the external bath at 300 K by Berendesen thermostat with a coupling constant $\tau_T = 0.1$ ps, a constant pressure of 1 bar applied with a coupling constant of $\tau_p = 1$ ps, and compressibility of $4.5e^{-5}$ bar^{-1} [33]. Energy minimization was done with steepest descent algorithm to improve the model quality; it was performed up to convergence on maximum force lower than 2000 kJ/mol. Subsequently, to relax the solvent

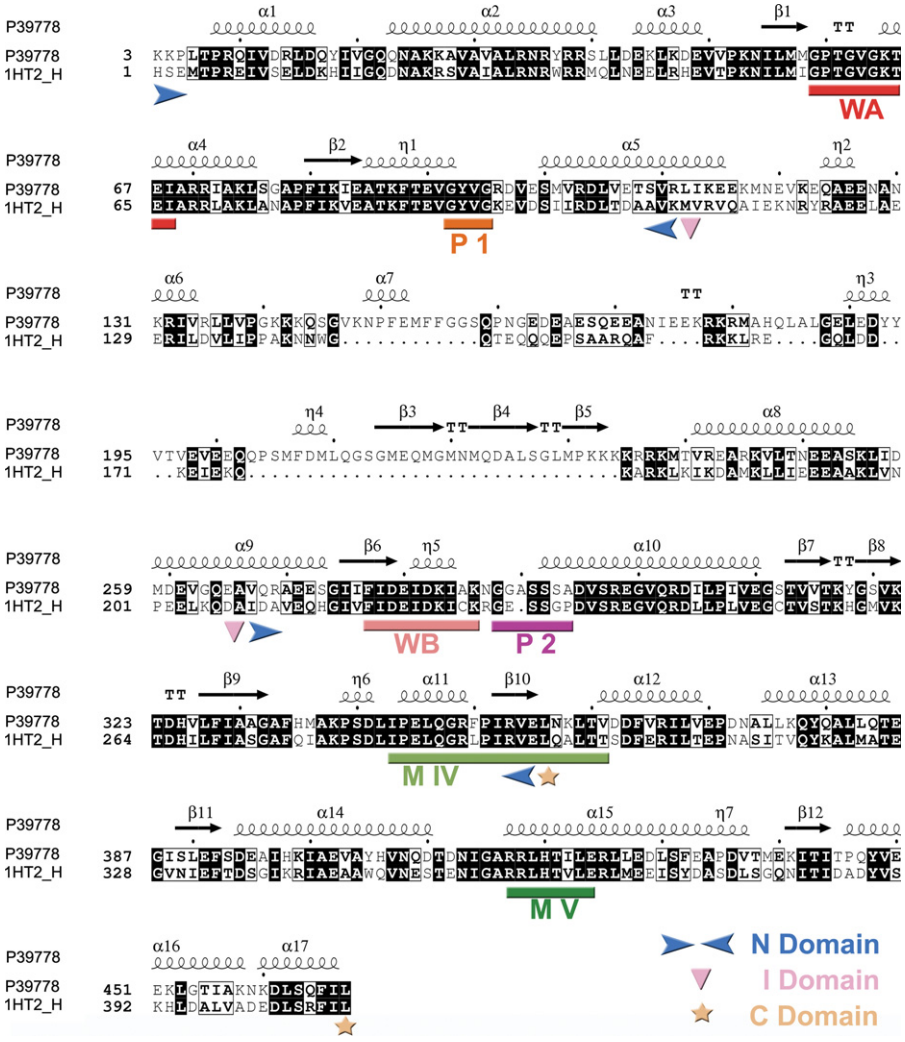

Fig. 1. Sequence alignment: HslU and CodX. SWISS-MODEL and ESPript produced sequence alignment of template (HslU: 1HT2h) with target (CodX: P39778) secondary structure details represented (helices: squiggles, beta strands: arrows, turns: TT letters). Identical residues (black box) are mostly present in N (start: ➤ end: ◀) and C (start and end: ★) domain including building blocks of NBP: Walker box A (WA), Walker box B (WB), motif IV (M IV) and V (M V) of Hsp100/ClpP ATPase, underlined with red, salmon, pale green, and dark green color bars, respectively. And conserved residues (open box) are very less in I domain (start and end: ▼). The primary (P1) and secondary (P2) pore forming residues in CodX hexamer are underlined by orange and magenta bars, respectively.

molecules of the system, 100 ps of position restrained MD was performed as an equilibration run. This pre-equilibrated system was used for 1.5 ns MD simulation with a time step of 2 fs without any position restraints and conformations were collected at every 1 ps for further analysis. Simulation trajectory analyses were done using analysis tools in GROMACS. The hardware resources used in this work were multi-node parallel cluster computer in Linux operating system.

2.4. Electrostatic potential calculation

The DELPHI [34] module of INSIGHTII was used to calculate the electrostatic potential with Consistent Valence Force Field (CVFF) [35] charges for each atom of the energetically optimized CodWX by solving the linear Poisson–Boltzmann equation. Dielectric constant of 80 and 2 were used for solvent and solute, respectively.

3. Results

3.1. Sequence conservation in CodX N and C domain

To determine how amino acid sequences of CodW and HslU are related, SWISS-MODEL produced structure-based sequence alignment used in ESPript [36], web interface to generate the alignment with target secondary structure information (Fig. 1). From the sequence alignment, it was observed that the nucleotide binding residues of Walker box A or P loop motif and Walker box B were highly conserved in CodX. The Hsp100/ClpP ATPases motif IV and motif V were also conserved in CodX, in which motif IV connects most conserved N and C domain. A conserved GYVG sequence located next to the Walker box A that forms the innermost pore and responsible for translocation of substrates in HslU [13,14], was also found and fully conserved in CodX. Earlier study has reported the possibility of substrate binding and entry at I domain [20]. In our model I domain was not conserved well and vastly predicted as coils with a few β-sheets and α-helices. This structure-based sequence alignment indicates that CodX N and C domain residues are mostly conserved including essential nucleotide binding pocket (NBP) and central pore.

3.2. Validation of homology modeled structures

The constructed models, HslU crystal structure based monomer CodX and hybrid crystal based hexamer, were examined for validation purpose using different criteria. The Root Mean Square Deviation (RMSD) analysis of the monomer model returned from the SWISS-MODEL was used to evaluate by means of deviation from its template. The N and C domain Cα atoms RMSD for the monomeric model and the template crystal structure was 0.12 Å, but more deviation was observed from structurally unconserved I domain. The stereochemical quality of the model was evaluated using PROCHECK. The Ramachandran plot of the Φ/Ψ distribution of backbone conformational angles for each residue of raw model obtained from PROCHECK revealed 90.9% of N and C domain and 72.3% of I domain residues are in the most favored regions and only 1.5% I domain residues in the disallowed regions of CodX monomer. This high percentage of N and C domain residues in the most favored regions indicates that the N and C domain of the model is well constructed and more reasonable. Thus, SWISS-MODEL generated monomer and HslU hexamer of hybrid crystal CodW–HslU [22] were used to build CodX hexamer. Furthermore, the same hybrid crystal (WWU) was used to construct the biologically functional WWX assembly. The constructed CodWX (WWX) octadecamer assembly is comprised of CodX hexamer and CodW dodecamer (Fig. 3B). This setup was refined using GROMACS v3.3 software (Section 2) energy minimization to relieve short contacts, which might have been produced while octodecamer generation.

The optimized CodX hexamer model quality validated again by Ramachandran plot using PROCHECK, revealed that 92.1% of the Φ/Ψ distribution of N and C domain residues are in the allowed regions (Fig. 2A). Only eleven residues (1.3%) found in the disallowed regions, all from I domain of CodX, were mostly constructed without template structural information (Fig. 2B). Quality of our optimized model shows that it can be

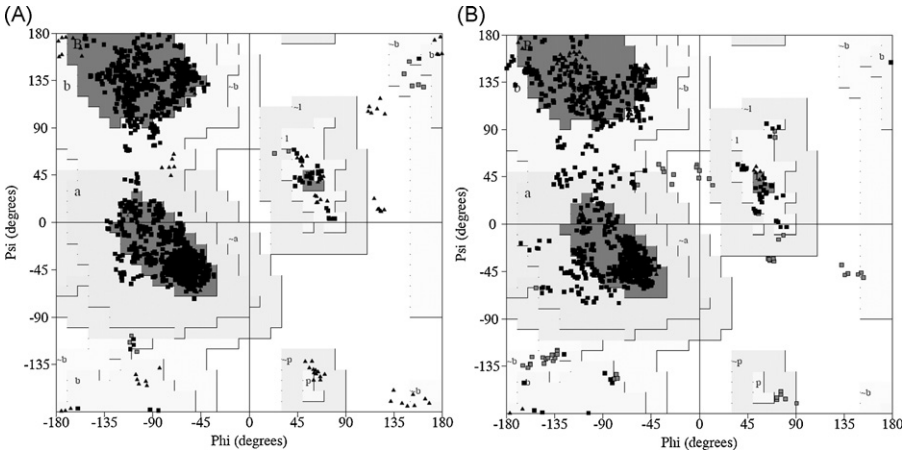

Fig. 2. CodX model structure validation by Ramachandran plot. In refined CodX hexamer, most of the N and C domain (A) residues (1525, 92.1%) were located in favored region (dark gray) and (B) eleven residues (1.3%) of I domain in disallowed area (white).

securely used in the subsequent MD simulation to reveal the structural behavior.

3.3. Structural features of the CodX model

In order to understand the structural features of CodX and its relationship with HslU, structural alignment was done for the target with template (Fig. 3A). The CodX structural domains are proposed based on the sequence conservation and structural alignment (Figs. 1 and 3A). The well-defined HslU domains [15] are also present in CodX and shows similar arrangement, namely the N domain, I domain and the C domain (Fig. 3C).

3.3.1. N domain

A long domain comprising residues K3–R111 and V267–L356, mostly constructed with α-helices and β-sheets. This domain consists of many important key segments required for nucleotide binding which are highly conserved including Walker box A, Walker box B and motif IV (Figs. 1 and 3C–E). In addition, the translocation pore consisting of a primary motif and a secondary motif forming residues are present in N domain analogous to HslU [13]. In CodX hexamer (Fig. 3E and D), the primary pore motif (GYVG) which forms the inner most pore, is also conserved among all protease-associated ATPases and the secondary motif forms the outer layer next to the primary pore.

3.3.2. I domain

L112–A266 referred as I domain in CodX. Further, this domain being constructed mostly as coils with a few short helices and turns in between β-sheets is neither structurally nor sequentially well conserved (Figs. 1 and 3C and A). The longest coil was found in I domain from residues M153 to M206 without any sheets or helices. This region may be more flexible during MD simulation. In CodX hexameric form, I domains are not in contact with each other and these gap regions in between I domains on the apical side (Fig. 3B) could provide site for substrate binding and entry as expected [20].

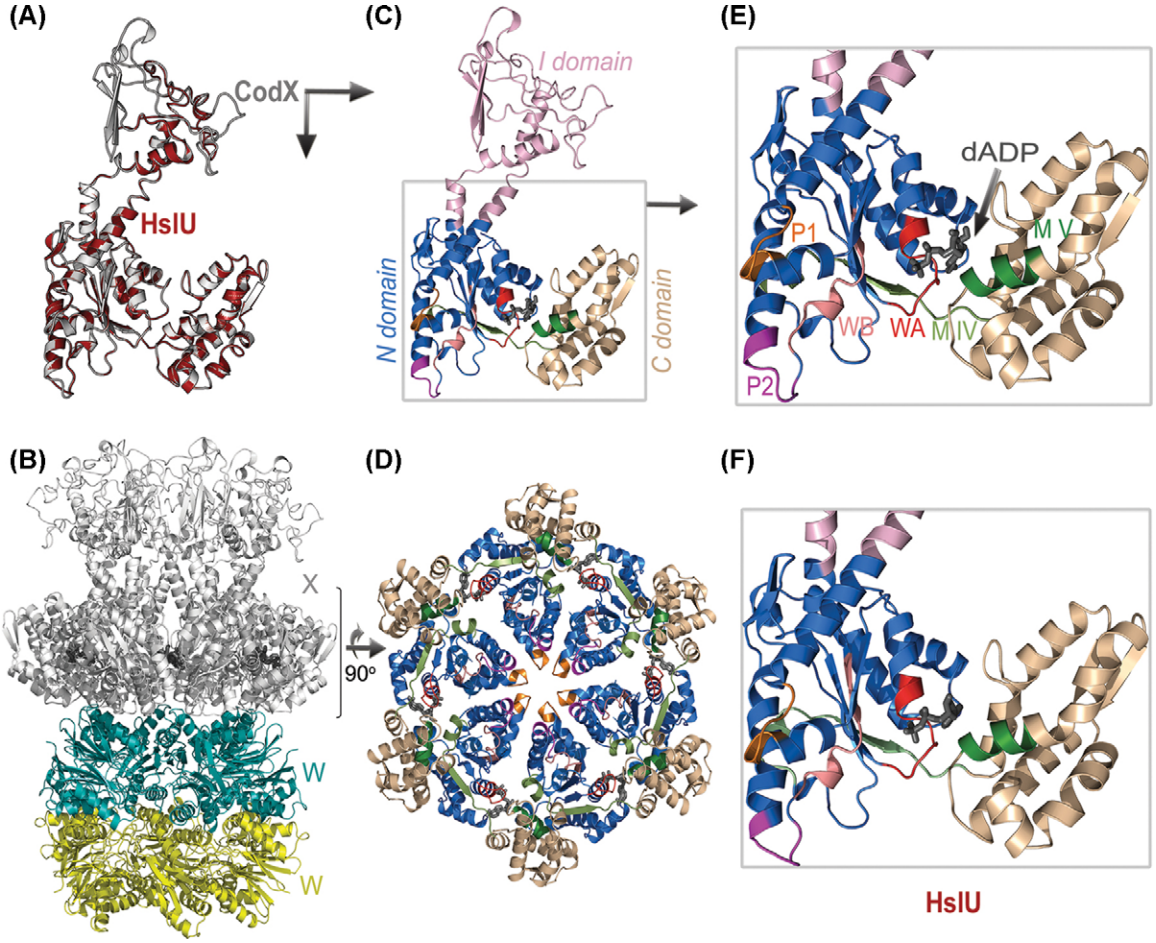

Fig. 3. Structural arrangement of functional domains in modeled CodX and CodWX (WWX) octadecamer assembly. (A) Superimposition of CodX (gray) monomer model with homologous HslU (1HT2 chain H) crystal structure (firebrick) shows the structural relationship between two ATPases. (B) Constructed CodWX octadecamer assembly model using CodW–HslU hybrid, each hexamer of CodWX (WWX) colored in yellow, teal and gray from bottom to top. (C) Structural domains of CodX: N, I, and C are represented by marine, light pink and wheat colors, respectively. (D) A view from CodW shows that the arrangement of structural domains including key segments (residues of NBP and pores) in N and C domain of CodX hexamer, I domain is removed for clarity. (E and F) Comparative extended view illustrates the CodX and HslU [13] key segments in N and C domain. The bound dADP (dark gray) positioned in between the N and C domain of each monomer of hexamer. The same colors and labels in Fig. 1 are used for the key segments in CodX and correspondent HslU. This figure was prepared using PYMOL (http://pymol.org).

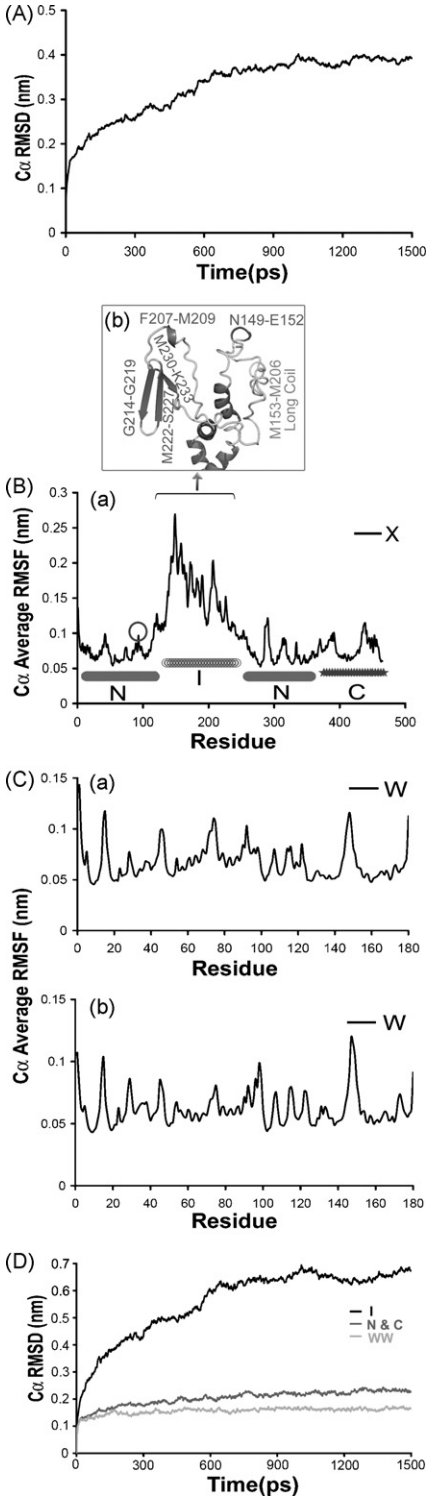

Fig. 4. MD simulation results: structural stability of CodWX complex and fluctuation of I domain. (A) The RMSD of overall CodWX complex Cα atoms, illustrates a stable behavior of CodWX after 1000 ps (temporal equilibration period). (Ba) According to the RMSD analysis, CodX Cα atoms RMSF average of final 500 ps with respect to the residue number for the monomer, in which structural domains are indicated as: solid bar, N; open circle bar, I; and C, star bar. In CodX, GYVG pore motif residues (G92–G95) fluctuations are showed by circle on the plot. (Bb) This extended cartoon representation depicts CodX I domain with hydrophobic channel forming β-sheets (G214–G219, M222–S227, M230–K233), and highly flexible long coil (M153–M206) in between two short helices (F207–M209, N149–E152). (Ca and Cb) Cα average RMSF of CodW plots express the very stable nature of CodW in assembly. (D) Detailed RMSD results of CodX domains and WW elucidate the I domain motion effect on the overall deviation in MD.

3.3.3. C domain

Motif V of the Hsp100/ClpP ATPase is located in C domain which is also essential to control the N and C domain motion on nucleotide binding [13]. Small C domain (N357–L467) is mainly made up of α-helices and a few β-sheets (Figs. 1 and 3C–E). This C domain location in our initial model (Fig. 3B and D) appears that it may play a role in CodWX (WWX) complex formation and CodW activation as described earlier [16]. dADP (1YYF) [22] is positioned in between N and C domain, in other words, between two subunits of the CodX hexamer (Fig. 3B–E). These structural features of CodX and 3D model structure alignment with template illustrate that this ATPase is not only in sequence but also structurally conserved including functionally important domains except I.

3.4. Reasonably stable CodWX octadecamer complex with flexible I domain

Evaluations of several geometric properties and energetic terms of protein will define the stability and dynamic behavior. To evaluate these properties for the overall CodWX octadecamer complex and functionally significant CodX domains, we examined trajectory of CodWX MD simulation using GROMACS analysis tools. Comparison of the global structural drift of the homology model during the simulation from their starting structure provides relative conformational stability over the timescale. To visualize this, the positional RMSD of Cα atoms were plotted with respect to the starting ($t = 0$) structure. This RMSD plot evidently explains the temporal equilibration period (first 1000 ps) followed by stable state of the CodWX complex during the course of 1500 ps simulation (Fig. 4A). Based on the RMSD observation, the entire system appears to be stable after 1000 ps. Thus, the subsequent Root Mean Square Fluctuation (RMSF) analysis of the residues Cα atoms was done after gradual equilibration period for CodW dodecamer (WW) and CodX hexamer (X), and the calculated hexameric residues Cα RMSF was further averaged for monomer and represented (Fig. 4Ba and C). Average Cα RMSF exhibited very low fluctuation of CodW dodecamer residues with a value of less than 0.1 nm throughout the simulation (Fig. 4Ca and Cb), structure of which was already determined by X-ray crystallography [22]. Results of elaborate RMSD analyses (Fig. 4D) for CodX domains and WW clearly describe how the unstable structure of I domain motion causes effect on the overall complex deviation.

In Hsp100 proteins, I domain is the most intricate part of X-ray crystallographic studies. For instance, HslU I domain structure has not yet been resolved completely [15,13]. In our WWX complex, CodX plays a major role in increasing the overall RMSD by nearly 0.4 nm. According to average RMSF result, this deviation mainly emerged from I domain (Fig. 4Ba, Bb and D). This unconserved and a possible structure of I domain contains many coiled regions and solvent exposed in

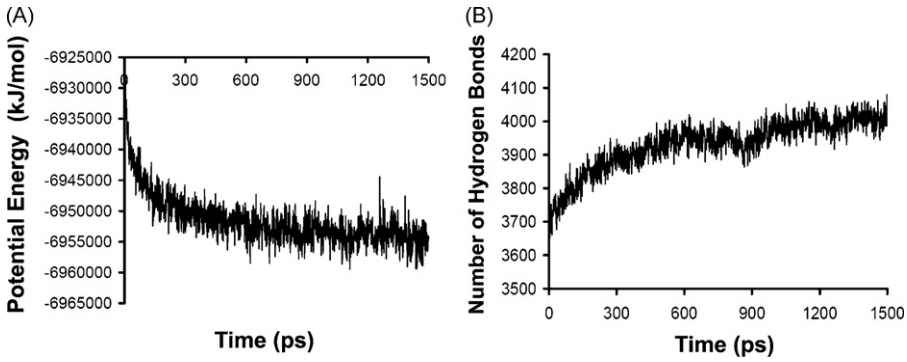

Fig. 5. MD simulation results: potential energy and hydrogen-bonds. (A) The CodWX complex potential energy plot indicates decreasing trend ended after the first 1000 ps and fluctuating around 6,955,000 kcal/mol. It is representing a temporal equilibration period of the system approximately starting from 1000 ps of the 1500 ps simulation. (B) The increase in the total number of hydrogen bonds in our system throughout the simulation directly indicates the stabilization by intra-atomic interactions.

WWX configuration. We observed more fluctuation mainly from the long coil (M153–M206) that is located between two short helices (N149–E152, F207–M209). However, the most significant result is that CodX N and C domain residues, which are the building blocks for key segments, show a tendency to be less mobile.

In our CodX model, I domain (Fig. 4Bb) has three anti parallel β-sheets (G214–G219, M222–S227, M230–K232) connected through turns. These β-sheets firm secondary structural nature indicate that it could form a hydrophobic channel like structure in CodX hexamer (Fig. 4Bb). This channel like structure may facilitate substrate entry to CodX central cavity. The other side of this domain (M153–M206) is adopted with low stable secondary structural elements of coils, bends and turns which are the major sources for its fluctuation.

Additional analyses also confirmed the over all system stability through out the simulation period (Fig. 5). The CodWX complex potential energy gradual decreasing trend coincides with RMSD result (Fig. 5A). The protein stability during our simulation verified with total number of hydrogen bond calculation (Fig. 5B), also shows positive correlation. These structural behavior analyses of CodX suggest that N and C domain are stable in nature and predicted possible I domain is flexible than others because of the highly coiled and solvent exposed hydrophobic end, which corresponds with previous crystallographic studies of other ATP-dependent proteases [15,17].

3.5. Adaptation of bound dADP in CodX

In CodX model, we placed HslU (1YYF) dADP [22] in order to obtain the knowledge about dADP conformational preservation in between N and C domain and its dynamic interaction with protein. We noticed the nucleotide–protein interaction and its dynamical relationship (Fig. 6). CodX sequence and structural alignment previously indicated the highly conserved nature of the NBP including Walker box A, in which residues mostly interacted with phosphate of the nucleotide as HslU [13,17]. Comparison of residues interaction with dADP in CodX and previously studied HslU demonstrated the similar arrangement of protein–nucleotide interactions between these two ATPases, especially Walker box A residues. Excluding Walker box A residues, there were few other residues (V20, R417) from CodX NBP, which contact dADP through hydrogen bonds (Fig. 6A). In CodX, positioned dADP maintained its *anti* conformational state by polar contacts with V20, the pattern that is incredibly

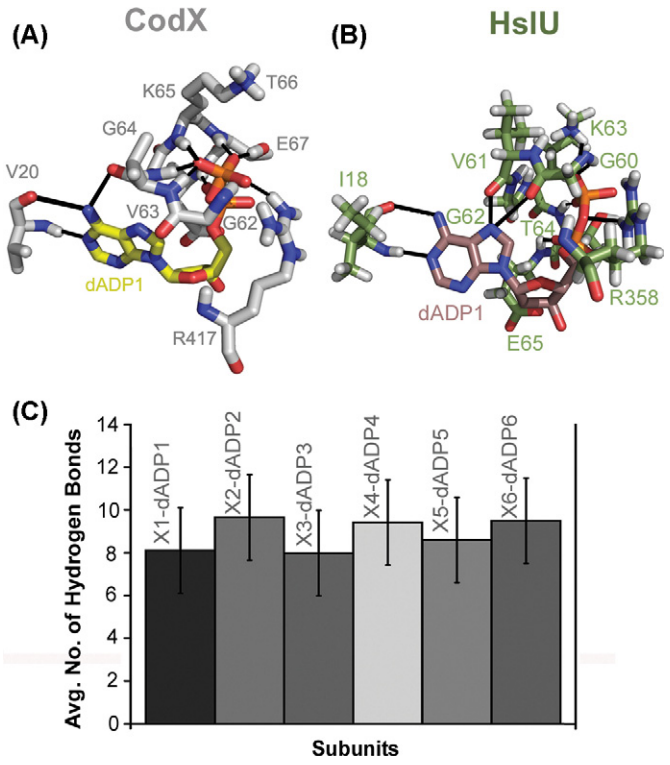

Fig. 6. Protein–dADP interactions: CodX vs. HslU. (A and B) Comparative view of protein dADP interactions in CodX and HslU: CodX-dADP, most of the hydrogen bonds (black lines) are from Walker box A residues (G62–E67) and a few from other residues (V20, R417) which is similar to HslU–dADP, interactions (G60–E65, I18, R358). These residues are identical except V20 in CodX-dADP (X1-dADP1) instead of I18 in HslU–dADP (U1-dADP1), the noteworthy pattern that explains HslU as an ATPase [7,13]. Stick representation of residues atom colors are as follows: carbon in CodX, gray; carbon in HslU, light green; in both: nitrogen, blue; oxygen, red; hydrogen, white; phosphate, orange. (C) During the course of the MD simulation, interaction between each monomer of CodX hexamer and anchored cognate dADP (dADP1–6) is illustrated with the average of 8 hydrogen bonds. Figs. 7–9 prepared using PYMOL and InsightII (Accelrys Inc., San Diego, CA).

analogous to the bound dADP conformation in HslU [13]; detection of this pattern in HslU described its significance as an ATPase [7]. In HslU (Fig. 6B), I18 contacted the adenine of the *anti* conformation dADP but in CodX, instead of I18, V20 amide donated its hydrogen to adenine N1 and carbonyl accepted hydrogen from adenine N6 (Fig. 6A). The time dependent protein–dADP interactions were also observed by calculating hydrogen bonds for the period of the simulation (Fig. 6C). This analysis proved that the Walker box A residues and also other NBP residues actively interacted with dADP all through the simulation. These results of protein–dADP interactions probably suggest that the nucleotide is anchored by protein residues in the center of the N and C domain and also dynamically preserve their contacts.

3.6. The surface charge complementarity contributes to CodWX complex formation

In an attempt to elucidate electrostatic nature of interactions between CodW and CodX, electrostatic potential was calculated using DELPHI module of INSIGHTII for the optimized model (Fig. 7C, A, B and E). This result explains the charge distribution on the surface of CodW that contacts neighbouring CodX has a preponderance of negative charge due to the major occurrence of glutamic acids (Fig. 7C, B and E). CodX also has predominant negative charge on the surface but when compared to CodW, it has few positively charged patches due to the charged histidines (Fig. 7C, A and E). Slight charge complementarities at the CodWX interface shows that CodW and CodX could form a complex, as HslV and HslU [13]. Net charges of CodX hexamer and CodW dodecamer at pH7 were −54.0 and −24.0 respectively. Polar contact observation between the charged residues on the surfaces of optimized model showed the possibility of hydrogen bonds from CodX histidines to CodW glutamic acids, which highly contributes to the structural interaction of CodWX complex (Fig. 7E). A few other polar contacts between arginines (CodW) and aspartic acids (CodX) on the interface were also observed.

To verify the CodWX structural association throughout the MD simulation, the total number of hydrogen bonds between CodW and CodX were measured over the period of simulation (Fig. 7D). The time dependent hydrogen bonding pattern growth curve apparently showed that the two binding partners maintain dynamically and attain their close contact in a polar environment. Surprisingly this result displayed that the two surfaces are approaching to obtain a native or closer molecular interaction from their initial modeled position based on the hybrid (HslU–CodW) structure. This result also suggested an excess of hydrogen bonds between ATPase and peptidase surfaces during the MD simulation. To precisely map these additional key residues, we further examined hydrogen bond possibilities in the final conformation of 1.5 ns MD simulation (Fig. 7F). It is found that not only histidines but also few lysines and aspartic acids played a key role in the intermolecular

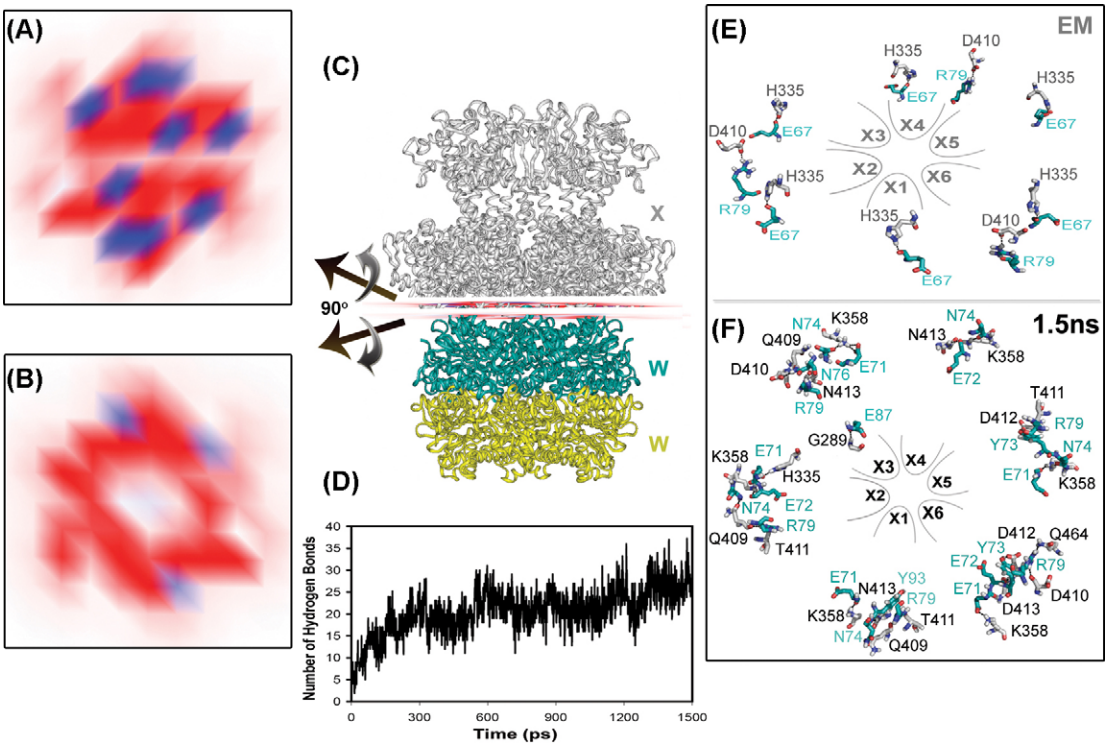

Fig. 7. Electrostatic nature of CodWX complex interface. (A) Sliced CodX surface charge distribution view from CodW illustrates a few positive (blue) patches in between predominant negative (red). (B) CodW surface sliced view depicts extremely negative charge allocation. (C) CodWX complex electrostatic potential calculated the interface is sliced with 2 planes, one at CodW (teal ribbon) surface and other at CodX (gray ribbon) surface. (D) Rising total number of hydrogen bonds between CodW and CodX also demonstrate that the surfaces are approach over the time progress. (E) Residues (sticks) in energy minimized structure contribute to the CodW (teal) and CodX (gray) complex formation at interface by hydrogen bonds (black dashes), nitrogen (blue) and oxygen (red) atoms. (F) Interface after 1.5 ns MD simulation, exploring view illustrates those additional residues contributions when the surfaces are closer than the initial model.

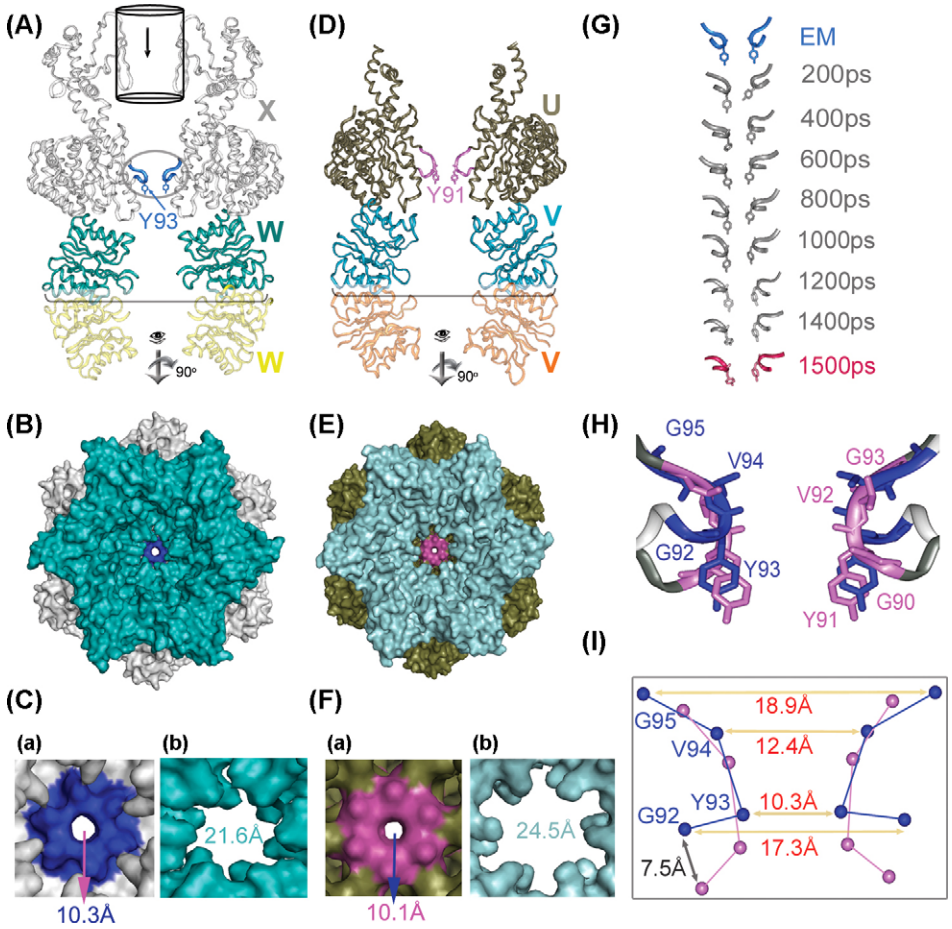

Fig. 8. Structural comparison of GYVG pore motif: HslU vs. CodX. (A and D) CodX and HslU closed states, represented as ribbon in dimeric form of complexes with tyrosine as sticks at primary pore motif, facing toward CodW (in CodWX) and HslV (in HslVU), and colored as follows: WWX in yellow, teal and gray from bottom to top; VVU in orange, aquamarine and smudge from bottom to top; CodX inner pore with Y93, blue; HslU inner pore with Y91, magenta. In CodWX dimer, a view from apical side of I domain shows that the hydrophobic channel forming β-sheets in I domain (cylinder). (B) A view from CodW in CodWX complex depicts the surface of CodX and CodW central pores at the same time. (C) Extended view of CodWX central pores, CodX (Ca) and CodW (Cb). (E) Surface representation of HslV and HslU, and their extended (Fa and Fb) view. (G) Dynamic nature of closed state Y93 at CodX pore motif demonstrated by snap shots collected at regular interval of 1.5 ns MD simulation (1.5 ns snap: dark magenta). (H) Superimposition of HslU and CodX GYVG pore motif, residues are illustrated as sticks. (J) The same superimposed view but Cα-atoms of CodX and HslU central pore residues are displayed as ball and stick model.

contact stabilization with the help of few other polar (N, T, Q) residues from CodX surface. Whereas in CodW, glutamic acids played a major role together with few aspargine and arginine residues for the interface maintenance. CodW and CodX preservation of charge complementarity-based interaction and interface residues prediction provided an insight into the electrostatic nature of their intermolecular interaction pattern and subsequent residues that made favorable contribution to complex stability in a polar environment.

3.7. Structural comparison of HslU and CodX GYVG pore motif

In order to compare the structure of CodX and HslU pores, we observed the generated surface models and superimposed structures. In addition, many conformational snapshots during the simulation were also analyzed to examine the dynamic behavior of CodX GYVG pore motif (Fig. 8). We found Y93 at the CodX primary pore motif, the position that corresponded to Y91 of HslU. It is believed that the aromatic ring structure in Y91 is essential for the substrates translocation in HslU [14]. Closed state of HslU was illustrated in the earlier study [13] where the aromatic ring of Y91 in the primary pore motif pointed toward HslV (Fig. 8D). In our modeled complex, the aromatic ring of CodX Y93 facing towards CodW suggests that our CodX in WWX complex structure is in closed conformational state (Fig. 8A).

Surface models clearly show that the size of CodX pore (Fig. 8B and Ca) is slightly bigger than that of HslU central pore (Fig. 8E and Fa) and CodW pore opening to the proteolytic core facing towards the CodX central pore (Fig. 8A and B). In this complex, CodW and CodX pores are adjacently aligned (Fig. 8A and B and D and E). The size of CodX pore is smaller than that of CodW pore (Fig. 8B). In addition, a view from the apical side of I domain in CodX exhibited the long hydrophobic coil region at the apical end and the following β-sheets (G214–G219, M222–S227, M230–K233), which internally formed a hydrophobic channel like structure (Figs. 8A and 4Bb). This

hydrophobic channel like structure may act as substrate entrance to the CodX central cavity (Fig. 8A). Moreover, CodW and HslV pore size observation (Fig. 8Cb and Fb) explained that the CodW pore (21.6 Å) is smaller than the opening in HslV (24.5 Å).

Structural superimposition of HslU and CodX GYVG pore motifs shows the conformational difference between the two central pores especially in glycine residue (G92) of CodX (Fig. 8H and I). The G92 is the key residue in CodX which deviates 7.5 Å from the backbone of the corresponding G90 in HslU. It creates a tight curvature in CodX central pore near the G92. This can be the key difference between the two pores. The measured distance between the Cα atoms of the key residues in the CodX primary pore motif (Fig. 8I) clearly shows the wider opening in both sides of the CodX central pore (G92 and G95). Furthermore, it elucidates that the CodX opening is larger than the opening in HslU.

During the MD simulation, dynamic behavior of the CodX GYVG pore motif was analyzed by observing conformational snapshots and residues average RMSF (Figs. 8G and 4Ba). Eight different conformational snapshots extracted during the 1.5 ns MD simulation time for every 200 ps showed that the aromatic ring direction of Y93 retained its state towards CodW with slight fluctuation (Fig. 8G). This result also indicates that our WWX complex is in the closed conformation. The two glycine residues (G92 and G95) are more fluctuating during the simulation time. The MD simulation results reveal the high fluctuation nature of the glycine residues and preservation of the Y93 conformational state in CodX central pore, which are the features required for the protein unfolding and translocation [13,14]. Structural comparison of the CodX and HslU central pores and dynamic behavior analysis for the CodX GYVG pore motif provides the difference and similarity of the key residues structural features in the central pore between the two ATPases.

4. Discussion

This molecular modeling study of CodX offers the 3D model of CodX N and C domain and its hexameric form including one of the possible structures of flexible I domain. We also obtain its reasonably stable quaternary structure and the information of surface charge complementarity-based association with CodW dodecamer, completely as a biologically relevant CodWX (WWX) assembly. Additionally, the structural relationship between CodX and HslU (Figs. 3, 6 and 8) revealed that the arrangement of structural domains are identical, particularly the conservation in their N and C domain structure including key segments. Hence, we propose that functionally critical domains of CodX have the same arrangement as HslU, but it differs in the mode of association with another CodX [20].

In spite of many limitations behind the oligomeric (18 subunits) protein simulation in a polar environment, we considered the accuracy of this protein native behavior and thus we simulated in one of its *in vivo* configurations. The reasonably stable nature and compactness of the overall CodWX complex in a polar environment at neutral pH indicated that it may exist predominantly in WWX form under physiological conditions, as anticipated [20]. Simulation results suggested the stable nature of N and C domain in CodX model and their significance in CodWX complex structure, also a mobile I domain. The structure of I domain of CodX that mostly predicted without template was observed as one possible structure out of many possibilities. In our model, structurally unconserved CodX I domain is flexible in nature because of the congregate of the coil region at the edge of this domain, hydrophobicity and also solvent exposed in WWX configuration. These results propose that it is perhaps the most intricate part for the crystallographic study like HslU [15,13]. Structural arrangement of hydrophobic region at the apical end of I domain suggests that it could promote substrate binding and entry as HslU [15,13,37].

HslU based CodX model with structurally conserved NBP probably played a vital role in the overall conformational stability of CodX that could also affect the CodW conformational nature since the nucleotide dependent conformational changes in HslU induce changes in HslV conformation [13,17]. On the other hand, in our CodX model positioned dADP (1YYF) [22] between N and C domain adopted well with Walker box A residues (Fig. 7A) and preserved its *anti* conformation by interacting with V20 (I18 in HslU), suggesting a similar pattern which described HslU as an ATPase [7,13]. However, our current dADP-bound model alone is not sufficient to understand the complete nucleotide-dependent conformational changes in CodWX. In future, molecular modeling study with ATP and empty state comparison with our dADP-bound conformation may be employed to uncover the nucleotide-binding induced conformational changes in CodWX complex elaborately (under investigation) and extend to crystallographic studies as well.

The electrostatic findings of CodWX demonstrated this model as two negatively net charged units (CodW and CodX) but they interact through the slight charge complementary at the surfaces (Fig. 8). In the CodW–CodX complex formation, hydrophobic effect is not playing a significant role compared to surface charge property. The charge-based complex interface pattern of CodWX is consistent with HslVU interface [13]. In MD simulation analysis, the hydrogen bonding pattern increasing trend between the CodW and CodX surfaces suggests that the native distance of CodW–CodX could be closer than that of the template CodW–HslU hybrid. Even though additional number of residues are involved in the CodWX interface formation at the end of the MD simulation, CodX and CodW surfaces preserved their predominance in positive and negative charges, respectively. A similar pattern of this charge complementarity in the initial and final conformations of 1.5 ns MD simulation expresses that it could be a possible native interaction pattern in the interface of CodWX complex.

CodX GYVG pore motif MD simulation results, flexible glycine (G92 and G95) residues and Y93 aromatic ring conformational state retained towards CodW are coincide with previously studied experimental results of structural features of GYVG pore motif in HslU [13,14]. Thus, it indicates that these glycine and tyrosine residues may also play essential role in the CodX protein substrates unfolding-coupled translocation.

Although sequentially the GYVG pore motif of CodX and HslU are conserved, structural comparison displayed the conformational variation especially by glycine at the starting position of the central pores. The conformational difference at the CodX primary pore motif suggests that the structure of this ATPase central pore may also differ from other ATPases, as it differs in the molecular architecture [20].

The pores of CodW and CodX are adjacently aligned in CodWX complex. This pattern proposes that the unfolded substrates appeared to be threaded from CodX pore to CodW pore with the help of hydrophobic paddle (Y93) as proposed in HslVU [13]. Moreover, the apical edge of the CodX hydrophobic site (highly coiled I domain apical end) appears to be a possible site for binding and entry of the substrates, as anticipated [20]. Further extension of this hydrophobic region resembles a channel that could act as a substrate opening to the CodX central cavity. CodX structural relationship with HslU including key segments especially conformation of Y93 aromatic ring could possibly elucidate that it may follow the similar mode during the mechanism as in the HslU closed conformation [13]. To reveal the substrate unfolding-coupled translocation mechanism of this multiprotein complex in polar environment by MD simulation will be followed in near future.

5. Conclusions

ATPases play essential roles in the ATP-dependent proteolysis mechanism. Structural insights into CodX describes the significance of N and C domain, which are essential for the CodWX complex formation, translocation of the protein substrates and nucleotide binding, as they are in HslU [13]. In agreement with our predictions, the experimental reports proved the importance of CodX C domain tail in the activation of CodW peptidase [16]. The newly modeled CodX structure primarily reveals the structural features of ATPase in *B. subtilis* including a possible structure of flexible I domain which contains highly coiled hydrophobic end and may participate in substrate binding and entry. Structural information of this molecular modeling study will be helpful to design further experiments to access CodX structural data at high resolution. Our computer modeled WWX, in particular, position of possible I domain and its substrate binding and entry hypotheses are in agreement with the previous CodX electron microscopic study [20]. The present study uncovers the structural relationship between the ATPases of *B. subtilis* and *E. coli*, two significant bacterial model systems widely used in laboratory for mechanistic studies.

MD simulation results proposed a CodX model consisting of stable N and C domain and a mobile I domain. In addition, it unveils the reasonably stable assembly structure of biologically relevant CodWX (WWX) octadecamer and explicit behavior of various functionally important regions. CodWX surface charge distribution actively facilitates the CodW and CodX interaction and shows a possible native complex interaction pattern. HslU and CodX GYVG pore motif structural features resemblance indicate that the polypeptide-threading related chaperon activities [13,14] could also be applicable to CodWX complex. Our CodW–CodX assembly model could be used as a template for the structure determination of the biologically functional CodWX complex construction using molecular replacement method, once the crystal data is available for CodX. Even though considerable progress has been made on this friendly microbe, widening our knowledge based on the CodWX structure would enable us to acquire the insight into the essential CodWX-dependent proteolysis mechanism and its comprehensive role in the regulation of cell division in *B. subtilis*.

Acknowledgements

We are indebted to Dr. S. Balasubramanian (TANUVAS, India), Mr. Nagakumar Bharatham (GNU, Korea) and Mrs. Kavitha Bharatham (GNU, Korea) for critical suggestions during the manuscript preparation. This work was supported by grants from the MOST/KOSEF for EBNCRC (grant #: R15-2003-012-02001-0) and for the Basic Research Program (grant #: R01-2005-000-10373-0). N. Krishnamoorthy and P. Gajendrarao were supported by the scholarship from the BK21 program, Korea.

References

[1] A.L. Goldberg, The mechanism and function of ATP-dependent protease in bacterial and animal cells, Eur. J. Biochem. 203 (1992) 9–23.
[2] W. Seufert, S. Jentsch, In vivo function of the proteasome in the ubiquitin pathway, EMBO J. 11 (1992) 3077–3080.
[3] O. Coux, K. Tanaka, A.L. Goldberg, Structure and functions of the 20S and 26S proteasomes, Annu . Rev. Biochem. 65 (1996) 801–847.
[4] M. Hochstrasser, Ubiquitin-dependent protein degradation, Annu. Rev. Genet. 30 (1996) 405–439.
[5] W. Baumeister, J. Walz, F. Zuhl, E. Seemuller, The proteasome: paradigm of a self-compartmentalizing protease, Cell 92 (1998) 367–380.
[6] D. Missiakas, F. Schwager, J.M. Betton, C. Georgopoulos, S. Raina, Identification and characterization of HsIV HsIU (ClpQ ClpY) proteins involved in overall proteolysis of misfolded proteins in *Escherichia coli*, EMBO J. 15 (1996) 6899–6909.
[7] M. Rohrwild, O. Coux, H.C. Huang, R.P. Moerschell, S.J. Yoo, J.H. Seol, C.H. Chung, A.L. Goldberg, HslV-HslU: a novel ATP-dependent protease complex in *Escherichia coli* related to the eukaryotic proteasome, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 5808–5813.
[8] S.J. Yoo, J.H. Seol, D.H. Shin, M. Rohrwild, M.S. Kang, K. Tanaka, A.L. Goldberg, C.H. Chung, Purification and characterization of the heat shock proteins HslV and HslU that form a new ATP-dependent protease in *Escherichia coli*, J. Biol. Chem. 271 (1996) 14035–14040.
[9] C.H. Chung, Proteases in *Escherichia coli*, Science. 262 (1993) 372–374.
[10] J.H. Seol, S.J. Yoo, D.H. Shin, Y.K. Shim, M.S. Kang, A.L. Goldberg, C.H. Chung, The heat-shock protein HslVU from *Escherichia coli* is a protein-activated ATPase as well as an ATPdependent proteinase, Eur. J. Biochem. 247 (1997) 1143–1150.
[11] A.F. Neuwald, L. Aravind, J.L. Spouge, E.V. Koonin, AAA+: a class of chaperone-like ATPases associated with the assembly, operation and disassembly of protein complexes, Genome Res. 9 (1999) 27–43.
[12] I.S. Seong, J.Y. Oh, J.W. Lee, K. Tanaka, C.H. Chung, The HslU ATPase acts as a molecular chaperone in prevention of aggregation of SulA, an inhibitor of cell division in *Escherichia coli*, FEBS Lett. 477 (2000) 224–228.
[13] J. Wang, J.J. Song, M.C. Franklin, S. Kamtekar, Y.J. Im, S.H. Rho, I.S. Seong, C.S. Lee, C.H. Chung, S.H. Eom, 2001a, Crystal structures of the HslVU peptidase-ATPase complex reveal an ATP-dependent proteolysis mechanism, Structure 9 (2000) 177–184.

[14] E. Park, Y.M. Rho, O.J. Koh, S.W. Ahn, I.S. Seong, J.J. Song, O. Bang, J.H. Seol, J. Wang, S.H. Eom, C.H. Chung, Role of the GYVG pore motif of HslU ATPase in protein unfolding and translocation for degradation by HslV peptidase, J. Biol. Chem. 280 (2005) 22892–22898.
[15] M. Bochtler, C. Hartmann, H.K. Song, G.P. Bourenkov, H.D. Bartunik, R. Huber, The structures of HsIU and the ATP-dependent protease HsIU-HsIV, Nature 403 (2000) 800–805.
[16] I.S. Seong, M.S. Kang, M.K. Choi, J.W. Lee, O.H. Koh, J. Wang, S.H. Eom, C.H. Chung, The C-terminal tails of HslU ATPase act as a molecular switch for activation of HslV peptidase, J. Biol. Chem. 277 (2002) 25976–25982.
[17] J. Wang, J.J. Song, I.S. Seong, M.C. Franklin, S. Kamtekar, S.H. Eom, C.H. Chung, Nucleotide-dependent conformational changes in a protease-associated ATPase HslU, Structure 9 (2001) 1107–1116.
[18] F.J. Slack, P. Serror, E. Joyce, A.L. Sonenshein, A gene required for nutritional repression of the *Bacillus subtilis* dipeptide permease operon, Mol. Microbiol. 15 (1995) 689–702.
[19] P.A. Levin, I.G. Kurtser, A.D. Grossman, Identification and characterization of a negative regulator of FtsZ ring formation in *Bacillus subtilis*, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 9642–9647.
[20] M.S. Kang, S.R. Kim, P. Kwack, B.K. Lim, S.W. Ahn, Y.M. Rho, I.S. Seong, S. Park, S.H. Eom, G. Cheong, C.H. Chung, Molecular architecture of the ATP-dependent CodWX protease having an N-terminal serine active site, EMBO J. 22 (2003) 2893–2902.
[21] M.S. Kang, B.K. Lim, I.S. Seong, J.H. Seol, N. Tanahashi, K. Tanaka, C.H. Chung, The ATP-dependent CodWX (HslVU) protease in *Bacillus subtilis* is an N-terminal serine protease, EMBO J. 20 (2001) 734–742.
[22] J. Wang, S.H. Rho, H.H. Park, S.H. Eom, Correction of X-ray intensities from an HslV-HslU co-crystal containing lattice-translocation defects, Acta Crystallogr. D 61 (2005) 932–941.
[23] A. Barioch, R. Apweiler, The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000, Nucl. Acids Res. 28 (2000) 45–48.
[24] T. Schwede, J. Kopp, N. Guex, M.C. Peitsch, SWISS-MODEL: an automated protein homology-modeling server, Nucl. Acids Res. 31 (2003) 3381–3385.
[25] R.A. Laskowski, M.W. Macarthur, D.S. Moss, J.M. Thornton, Procheck-a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[26] INSIGHTII, version 2005.3L, Accelrys Inc., San Diego, 2005 http://www.accelrys.com.
[27] D. Van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: fast, flexible and free, J. Comp. Chem. 26 (2005) 1701–1718.
[28] A.W. Schuttelkopfa, D.M.F. van Aalten, Prodrg: a tool for high-throughput crystallography of protein–ligand complexes, Acta Crystallogr. D 60 (2004) 1355–1363.
[29] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, Intermolecular forces, in: B. Pullman (Ed.), Interaction Models For Water In Relation To Protein Hydration, D. Reidel Publishing Company, Dordrecht, The Netherlands, 1981, pp. 331–342.
[30] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[31] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[32] S. Miyamoto, P.A. Kollman, SETTLE: an analytical version of the SHAKE and RATTLE algorithms for rigid water models, J. Comp. Chem. 13 (1992) 952–962.
[33] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.K. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3689.
[34] M.K. Gilson, B. Honig, Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies and conformational analysis, Proteins 4 (1988) 7–18.
[35] P. Dauber-Osguthorpe, V.A. Roberts, D.J. Osguthorpe, J. Wolff, M. Genest, A.T. Hagler, Structure and energetic of ligand binding to proteins: *E. coli* dihydrofolate reductase-trimethoprim, a drug–receptor system, Proteins 4 (1988) 31–47.
[36] P. Gouet, E. Courcelle, D.I. Stuart, F. Metoz, ESPript: analysis of multiple sequence alignments in PostScript, Bioinformatics 15 (1999) 305–308.
[37] M.C. Sousa, C.B. Trame, H. Tsuruta, S.M. Wilbanks, V.S. Reddy, D.B. Mckay, Crystal and solution structures of an HslUV protease–chaperone complex, Cell 103 (2000) 633–643.