Effects of mutants in bHLH region on structure stability and protein-DNA binding energy in DECs

Yi Kong\textsuperscript{a}, Zhen Wang\textsuperscript{b}, Yanfei Jia\textsuperscript{a}, Ping Li\textsuperscript{a}, Shuhua Hao\textsuperscript{a} and Yunshan Wang\textsuperscript{a}\textsuperscript{*}

\textsuperscript{a}Central Laboratory, Jinan Central Hospital Affiliated to Shandong University, No. 105 Jiefang Road, Jinan 250013, Shandong Province, P.R. China; \textsuperscript{b}Department of Oncology, Zhang Qiu People Hospital, No. 1920 Huiquan Road, Zhangqiu 250200, Shandong Province, P.R. China

Communicated by Ramaswamy H. Sarma.

(Received 24 February 2016; accepted 29 May 2016)

The human DEC subfamily contains two highly conserved members belonging to basic helix-loop-helix (bHLH) transcription factors. This conserved family is spread widely among various species with the function of regulating various crucial molecular signaling pathways. Due to the significance of DECs for important biological processes, their relationship with diseases and the lack of experimentally proven structures, we have implemented a comparative modeling for the bHLH region of DECs as homodimers with themselves and heterodimers with HES-1. Three mutants with predicted roles in reducing intramolecular binding (H57A, R65A, and LL7879AA in DEC1 and LL7071AA in DEC2) were investigated on DEC monomers. Molecular dynamics (MD) simulations were also employed to evaluate the behavior of the mutant molecules in aqueous solution. The monomer was divided into subregions for accurate investigation. The fluctuation in the basic region of mutants was higher than that of wild-type molecules. The binding energy value between protein and DNA obviously increased in the homodimer harboring R65A mutants, which led to more unstable status between protein and DNA. Thus, the mutant R65A interfered DNA-binding affinity. A study on the spatial structures of wild-type and mutant DECs may facilitate functional prediction for mutation effects and dynamic behavior under various conditions and may ultimately help in targeted drug design.

Keywords: DNA-binding affinity; DEC1/2; protein comparative modeling; molecular dynamics simulation

Introduction

Molecules of the differentiated embryonic chondrocyte-expressed gene (DEC) subfamily play a crucial role in embryological development, differentiation and circadian rhythm. This family contains two members, DEC1 and DEC2. The conserved family is spread widely among various species with an orthologous gene, clockwork orange, in \textit{Drosophila}. In vertebrates, the function of DEC1 and DEC2 is mainly focused on circadian rhythm (Li et al., 2013), hypoxia response (Jia et al., 2013), DNA mismatch repair (Inaguma et al., 2013), and tumor genesis (Liu et al., 2013). DEC proteins, members of the basic helix-loop-helix (bHLH) transcription factor family, are composed of a basic DNA-binding region and a dimerization region with two amphipathic \(\alpha\)-helices bridged by a loop domain. DECs form homodimers or heterodimers with other bHLH proteins to specifically bind to the DNA sequence CANNTG (E-box) located in the promoter regions of the target genes (Fujimoto et al., 2007).

The dimerization process plays a crucial role in DEC regulation activity. Homodimers of DECs suppress target gene expression through direct binding of E-box elements in the target gene promoter, leading to down-regulation of MLH1 and MyoD (Fujimoto et al., 2007; Nakamura et al., 2008). Heterodimers consist of DECs and HIF1\(\alpha\) or BAML1, which are identified through pull-down studies (Ma, Shi, Lu, Wu, & Wang, 2013; Sato et al., 2004). DECs preferentially form homodimers, while heterodimers also regulate target gene expression. DECs can also bind to HIFs to promote HIF proteasomal degradation (Montagner et al., 2012). The heterodimers formed by DEC2 and MyoD integrate through the basic and HLH regions, inhibiting the transcriptional activity of MyoD (Fujimoto et al., 2007). Before now, there have been no reports comparing DNA-binding affinity between homodimers and heterodimers formed by DECs and other partners.

DECs possess dual functions as repressors and coactivators of mammalian clock gene expression in a tissue-specific manner (Rossner et al., 2008; Tsang et al., 2012). Mutated DECs often lead to disorders of cellular rhythms. A DEC1 mutant prevents checkpoint recovery by inhibiting the down-regulation of p53, which is a key molecule in controlling DNA damage response (Kim et al., 2014). A DEC2 mutant phenotype modifies the

*Corresponding author. Email: sjinwys@sohu.com

© 2016 Informa UK Limited, trading as Taylor & Francis Group
normal inhibition of BHLHE41 of CLOCK/BMAL1 transactivation, which is associated with a resistance to sleep deprivation. (He et al., 2009; Pellegrino et al., 2014). A circadian defect is often considered a cancer risk factor, and this potential association is supported by pathological studies on hepatocellular and lung carcinomas (Liu et al., 2013; Shi et al., 2011) and experimental results in vivo (Ma et al., 2013).

Over 70 single nucleotide polymorphisms (SNPs) in the DEC1 and DEC2 gene code sequences have been reported in the NCBI database (DEC1:71; DEC2:70). Sixty percent of these DEC1 SNPs and 50% of DEC2 SNPs are missense mutants that change the amino acids. Three missense mutations in DEC1 were reported by Sato (Sato et al., 2004), which may reduce their DNA-binding ability or abolish retinoid X receptor alpha (RXRA) repression. A mutant in DEC2 was reported by Cho (Cho et al., 2009), and it abolishes RXRA repression. These mutants are considered to be important because they reduce either DNA binding or dimerization, which often act in a dominant-negative manner.

The three-dimensional (3D) structures of the DEC proteins have not been identified experimentally. Because of the innate correlation between protein structure and function, the elucidation of the 3D structure of DECs could facilitate function prediction, investigation of mutation effects and dynamic behavior under various conditions, and targeted drug design. Theoretical methods would provide a convenient and reliable way for resolving these issues due to the lack of published protein 3D structures. Comparative modeling of the bHLH region in DECs for homodimers and heterodimers with HES-1 is performed.

Hes Family BHLH Transcription Factor 1 (HES-1) belongs to the bHLH family of transcription factors and acts as a transcriptional repressor. It is related in transcription factor activity, sequence-specific DNA binding and sequence-specific DNA binding. Diseases associated with HES1 include hepatic lipase deficiency and diabetes mellitus, noninsulin-dependent. Among its related pathways are Signaling by GPCR and downstream signaling system.

In this report, we investigate the behavior of several important mutants in aqueous solution through molecular dynamics (MD) simulations. Various mutations that are predicted to reduce DNA binding (H57A, R65A, and LL7879AA in DEC1 and LL7071AA in DEC2) are also studied. Investigation on these mutants would provide new aspect for specific targeted drug design.

Method

Establishment of DEC1/DEC2 dimer models

The amino acid sequences of DEC1 and DEC2 were acquired from the Uniprot database. The conserved domain and secondary structure were evaluated by, respectively, the eukaryotic linear motif resource and DISOPRED, a simple approach based on a running sum of the propensity for amino acids (Jones & Ward, 2003; Puntervoll et al., 2003). DISOPRED is based on a neural network. Disordered regions are flexible, partially dynamic, or completely distributed in solution. The propensity for a given amino acid was calculated to be in a ‘random coil’ and regular ‘secondary structure’ as defined in DSSP. A homologous search for DEC1/DEC2 protein sequences was implemented with the NCBI BLASTp program (Basic Local Alignment Search Tool for proteins (Altschul, 1990)) against protein structure sequences in the RCSB Protein Data Bank (Bernstein et al., 1977). The best template was chosen according to the highest similarity combined with coverage regions. Alignment between the selected template and DEC1/DEC2 was implemented using the CLUSTALW2 program (Larkin et al., 2007; McWilliam et al., 2013) with default parameters.

The tridimensional spatial structures of the DEC1/DEC1 and DEC2/DEC2 homodimers, the DEC1/HES-1 and DEC2/HES-1 heterodimers and the monomeric DEC1 and DEC2 mutated models H57A, R65A, LL7879AA (DEC1), and LL7071AA (DEC2) were constructed using the MODELLER 9v14 package (Sali & Blundell, 1993). We also modeled a protein–DNA interaction following a previous article (Sokkar, Sathis, & Ramachandran, 2012). One hundred models were randomly built from the template structure for each model, including the wild type and mutants. The model with the lowest Z-scores was sorted and evaluated by MODELLER scripts through the DOPE method. The model with the best quality was employed in a sequential root mean square deviation (RMSD) analysis and compared to the template considered as a reference. Comparative modeling in the MODELLER program was performed on a workstation computer with an Intel Xeon Quad-Core CPU (2.13 GHz) running the CentOS Linux operating system.

Model assessment

The models predicted with MODELLER were submitted to tools that were available at the SwissModel Workspace for structural assessment after detailed calculation (Arnold, Bordoli, Kopp, & Schwede, 2006). Global model quality was calculated with QMEAN6 (Benkert, Tosatto, & Schwarzbach, 2008), and then, a pseudoenergy value was provided for ranking alternative models of the same target. The best model was found by sorting according to predicted energy (from lowest to highest). The stereochemical quality of a protein was evaluated with PROCHECK by appraising Ramachandran plots (Laskowski, MacArthur, Moss, & Thornton, 1993). A
reference state, which constructed a residue-specific all-atom potential of mean force, was established through DFIRE. The lowest energy value implies the highest model stability, closer to the ‘natural status’ of the modeled protein (Zhou & Zhou, 2002). The QMEAN Z-score was used to measure the absolute quality of a model, reflecting the concordance between modeled structural features and native conformation. Strongly negative Z-scores usually indicate models with poor quality (Benkert, Biasini, & Schwede, 2011).

MDs simulation

GROMACS package v. 5.0.2 (Van Der Spoel et al., 2005) was employed for MDs simulation. First, a solvated orthorhombic box with a volume of 70 Å × 70 Å × 70 Å was created, with the best quality of modeled protein locating. Second, water molecules were used padding following the TIP4P model (Jorgensen & Chandrasekhar, 1983). Then, the system was neutralized by the random addition of negatively charged Cl− counter ions. An OPLS-AA (optimized potentials for liquid simulations, including every atom explicitly) force field was chosen for liquid simulations (Kaminski, Friesner, Tirado-Rives, & Jorgensen, 2001). Periodic boundary conditions were accepted in all directions. The minutiae of the configuration are described in Supplementary Table S1. Third, a process of energy minimization with 100 steps was run to abolish any solid geometry conflict in the established system. The LINCS algorithm was selected to restrict all of the bond lengths. In the pre-MD (molecular dynamic) stage, the entire system was slacked via the energy minimization process to avoid solid geometry conflicts. Temperature, pressure, and density were balanced with 100 ps. All of the MD simulations were executed at 300 K and 1 atm according to the modified Berendsen algorithm. During the pre-MD equilibration stage of 1 ns, all atomic protein positions were restrained. Finally, the sequential simulation was conducted for 50 ns with a 2-fs time step applied without the trammel of protein atoms. (Van Der Spoel et al., 2005). All calculations were performed on a computer with an Intel(R) Xeon (R) X5650 six-core CPU (2.67 GHz) and an Nvidia GTX 750 Ti graphics card with CUDA acceleration.

The analysis toolkit package in GROMACS was employed to evaluate trajectories for both homo- and heterodimers in aqueous solution. DSSP was employed to predict the second structure of the modeled dimers. Structural and dynamic properties were described by several essential indexes, such as the interaction potential energy between the monomers of the dimers, the RMSD, the root mean square fluctuation of the residues (RMSF), the radius of gyration (Rg) of the monomer backbone and, for each region, the minimum distance between the centers of mass and the collective motions. The variation of the solvent accessible surface area (ΔSASA) was obtained by subtracting the sum of the SASAs of the sole monomers from the SASA of their respective dimers. We also simulated the system containing protein–DNA complex. Binding energy between protein and DNA was also calculated with gnmmbpsa program (Kumari, Kumar, & Lynn, 2014). Simulations for the same initial model were implemented three times independently.

Results

Prediction of DEC1 and DEC2 models

The human DEC1 and DEC2 amino acid sequences were obtained from the Uniprot database (Entry: O14503; Q9C09J). The human DEC1 consists of four regions: the N-terminal (residues 1–51), the bHLH domain (residues 52–107), the Orange domain (residues 142–175), and the C-terminal domain (residues 176–412), while DEC2 exhibits an analogous structure with N-terminal (residues 1–43), bHLH (44–99), Orange (131–166), and C-terminal (residue 167–489) domains. DEC1 does not have a WRPW domain, which facilities protein degradation.
through a proteinases pathway. The DEC2 protein contains an alanine/glycine-rich region in the C-terminus. Moreover, there is a proline-rich region followed by an HDAC-interacting region in the C-terminus of DEC2 (He et al., 2009). Using the DISOPRED program, we identified highly disordered regions at the N-terminus, both in DEC1 and DEC2 (Figure 1). The disordered region at the C-terminus of DEC2 is longer than that in DEC1 (Supplementary Figure S1). In comparison with DEC1, there is an alanine/glycine-rich region and a proline-rich region followed by an HDAC-interacting region in the C-terminus of DEC2 (He et al., 2009). The bHLH domain was the prior segment of the DEC1 and DEC2 sequences according to Blastp results against known 3D structures in the RCSB Protein Data Bank. Results of the homologous blast are listed in Supplementary Table S2. The coordinates of the bHLH region in the HES-1 crystal (PDB code 2mh3 (Popovic et al., 2014)) were selected for subsequent modeling.

Some sites possess a high degree of conservation, as shown in the alignment of the bHLH region among different species of DECs proteins and the HES-1 protein in Figure 2. Conservational sites appeared in the consensus sequence, while mutant sites were marked by a colored box. Figure 2 illustrated the pair-wise alignments between the bHLH region of DEC1/DEC2 and the HES-1 monomer sequence. The dimers of DEC1 or DEC2 consisted of two monomers with two α-helices separated by a loop region. The DOPE value was calculated to assess qualities of the wild-type and mutant models. The best in quality model exhibits the lowest DOPE value. The scheme of the DEC1 and DEC2 dimer complex with DNA and models with homo-dimers and hetero-dimers are shown in Figure 3.

The predicted structural alterations caused by the H57A, R65A, LL7879AA, and LL7071AA mutations in the DEC1 and DEC2 monomers were assessed (Figure 3(F)-(I), respectively). The H57A and R65A substitutions were predicted to lead to polarity alterations because of changes in the positive electric charge at these sites. Both these substitutions led to an alteration in the AA charge (positive to neutral), with a fluctuation of the solvent accessible surface area (ΔSASA) of the mutated amino acid (Figure 4). However, the substitution LL7879AA in DEC1 and LL7071AA in DEC2 hardly resulted in change of polarity. Therefore, mutated residues brought different variation of protein structures.

**Model quality assessment**

The PROCHECK program was employed to sequentially evaluate the stereo chemical structural validation of models to the model construction using the MODELLER program. The results from the Ramachandran plot demonstrated that the S74 residue in Chain B was in the generous allowed region, and trivial conflicts at this site may amplify and transmit conflicts into the 3D structure with the HES-1 protein as a template. The modeled structures are shown in Figure 3.

![Figure 2. Schematic structure of the human DEC1/DEC2 protein.](image)

Notes: Multiple sequence alignments for the bHLH domain of human DEC1/DEC2 and sequences belonging to different species are shown. The bHLH region of HES-1 are choosen as model template. Conserved residues among different species are highlighted in black blocks. The modeled mutation positions H57A, R65A and LL7879AA are indicated by colored boxes.
structures had improved values compared to the template structure, presenting 87.1% of the residues in the allowed regions (Table 1). DFIRE and QMEAN 6 analyses are typically used to analyze non-bonded atomic interactions and global model quality. All of the models have higher scores in the above analyses than the basic template (2mh3: −153.92 and 0.678), as presented in Table 1. The Z-scores in QMEAN 6 emerged as analogous trends that the scores of the modeled protein were higher than those of the template (−0.885). The heterodimer, which presented lower Z-scores closer to the template, may derive from the influence of the template structure. Modeled proteins with the best qualities were chosen to subsequent analysis.

MD simulations of wild-type and mutant proteins

The simulation time was set to 50 ns, of which only the last 30 ns were employed to complete the analysis due to protein stability. Results are reproducible in the trajectories for the same initial model. The interaction potential energy between monomers remained steady during the simulation time for the wild-type structures (Figure 4). Considering all atoms, time-course analysis was executed in the GROMAC software. In this analysis, both Ca and the backbone atoms were evaluated to seek significant fluctuations of residues. The backbone was chosen to investigate the structural variation without the interference of side-chains.
Figure 4. Structural stability assessment during the MD simulations. 

Notes: The images in the left column present the wild-type and mutant homodimers. The images in the middle column exhibit the wild-type and mutant heterodimers, and the images in the last column depict the DEC2 homodimer and DEC2/HES-1 heterodimers. The RMSD analysis (first and second rows) for all of the monomers was calculated separately as a function of the backbone structure according to the monomers (Figure 4(A)–(C)). The third row represents the variation of the solvent accessible surface area (SASA) for the dimers and was calculated by subtracting the sum of the SASA of the individual monomers from the SASA of their respective dimers (Figure 4(D)). Å – angstrom ($10^{-10}$ m); kJ – kilojoule ($10^3$ joules); mol – 6.02 x $10^{23}$ particles; ns – nanoseconds ($10^{-9}$ s).
Table 1. Summary of the SWISSMODEL evaluation package pseudo-energies.

|                | HES1 | DEC1_A/DEC1_B wt | DEC1_A/DEC1_BH57A | DEC1_A/DEC1_BR65A | DEC1_A/DEC1.BLL7897AA | DEC1/DEC1_HES1 wt | DEC1/DEC1_HES1_H57A | DEC1/DEC1_HES1_R65A | DEC1/DEC1_HES1_H87879AA | DEC2_A/DEC2_BL7701AA | DEC2_A/DEC2_HES1 wt | DEC2/DEC2_HES1_H57A | DEC2/DEC2_HES1_R65A | DEC2/DEC2_HES1_H87879AA |
|----------------|------|------------------|-------------------|-------------------|------------------------|-------------------|---------------------|---------------------|-------------------------|------------------------|---------------------|---------------------|---------------------|-------------------------|
| ENERGY Analysis |      |                  |                   |                   |                        |                   |                     |                     |                         |                        |                     |                     |                     |                         |
| DFIRE          | -153.92 | -140.40 | -138.59 | -136.80 | -130.26 | -148.52 | -146.42 | -144.75 | -144.07 | -146.98 | -136.20 | -152.26 | -145.32 |                         |
| QMEAN6         | 0.678 | 0.704 | 0.763 | 0.777 | 0.761 | 0.733 | 0.730 | 0.743 | 0.74 | 0.718 | 0.753 | 0.728 | 0.754 |                         |
| Z-score        | -0.885 | 0.301 | -0.026 | 0.116 | -0.048 | -0.326 | -0.359 | -0.227 | -0.256 | -0.46 | -0.105 | -0.392 | -0.132 |                         |
| RAMACHANDRAN PLOT |      |                   |                   |                   |                        |                   |                     |                     |                         |                        |                     |                     |                     |                         |
| Most favored   | 87.1% | 93.8% | 90.2% | 92.0% | 92.9% | 91.0% | 90.2% | 91.0% | 90.7% | 91.5% | 88.0% | 89.6% |                       |
| Allowed        | 2.1% | 5.4% | 8.9% | 8.0 | 7.1% | 7.4% | 8.2% | 7.4% | 6.5% | 9.3% | 7.6% | 11.2% | 9.6% |                       |
| Generously allowed | 0.8% | 0.9% | 0.9% | 0.0% | 0.0% | 0.0% | 0.8% | 0.8% | 0.8% | 0.0% | 0.0% | 0.0% | 0% |                       |
| Disallowed     | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 1.6% H42/A | 0.8% H42/A | 0.8% ST4/B | 0.8% ST4/B | 0.0% | 0.0% | 0.0% | 0.8% ST4/B |                       |
|                | ST4/B | ST4/B | ST4/B | ST4/B |

Note: The bold scores are in bold, and the template scores are in bold and italicized. HES-1 hairy and enhancer of split 1; wt – wild-type.
In the equilibrating configuration stage, the potential energy was minimized while the temperature, pressure, and density were equilibrated as a reference. The RMSD and the $R_g$ were used to analyze the protein structure in the solvent. According to the RMSD value, the wild-type (wt) dimers demonstrated an analogous deviation pattern over time, while $R_g$ analysis hardly showed obvious variations higher than 2 Å, which means that there were no principal distortions on the protein. (Supplementary Figure S2).

According to the results of RMSD analysis, all modeled dimers reached convergence status. The monomers of DEC1 homodimer exhibited lower values than dimers harbored mutated residues and their counterpart heterodimers. The solvent accessible surface area variation was evaluated. Consequently, the value of $\Delta$SASA, which was obtained by subtracting the sum of the SASA of the individual monomers from the SASA of their intact dimers, was introduced as an index to predict binding free energy difference ($\Delta\Delta G_{binding}$) upon mutations. Among the dimers, the wt type exhibited lower values than mutated type, implying a more cohesive structure of monomers.

The hydrophobic/hydrophilic SASA were introduced to investigate protein behavior during the simulation (Supplementary Figure S3). The ratio between the hydrophobic and the hydrophilic atoms was calculated, and then, the values more than 1 were considered as increasing of the parameters. Mutants of H57A and R65A bring more intensely inference on SASA ratio. For the mutant LL7879AA in DEC1 and LL7071AA in DEC2, the SASA ratio showed less variation between the wt and the mutated residues in dimers. Compared with the mutated models, the heterodimers showed lower divergence in hydrophobic/hydrophilic SASA ratios (Supplementary Figure S3). Fluctuation on this index implied protein structural variations in solvent, which affected molecular binding. The fluctuation hydrophobic/hydrophilic SASA was accompanied with an undulation of hydrogen bonds. Mutants of H57A and R65A observably decreased hydrogen bonds between monomers and water rather than that within monomers (Supplementary Figure S4). Hydrogen bonds played a crucial role in maintaining protein secondary structure.

The DSSP program (http://www.cmbi.ru.nl/dssp.html) was employed to investigate the secondary structure of the bHLH region in homodimers and heterodimers; nevertheless, the results of DSSP analysis hardly showed significant variations (data not shown). This means that most protein structures remained folded without various residue exposure.

The RMSF was calculated and plotted to reflect the mobility of each atom (Figure 5(A)). Fluctuations up to 1.5 Å were observed in both the basic region and minor peaks in the loop region of all monomers, implying higher mobility due to complete exposure to the solvent. The basic domains of homodimers and heterodimers showed more fluctuations than any other regions, and the mutated proteins demonstrated more variations than the wt counterparts. The higher fluctuation at the N-terminals and C-terminals may be due to the absence of the remnant for stabilizing the 3D structure. This trend was confirmed by calculating the b-factor parameter (Figure 5(B)), and the high fluctuation of the DEC1/2 monomer basic domain of the dimers was marked with a colored tube. Higher RMSF are related to structural flexibility in N-terminal and C-terminal.

Each monomer can be spliced into four regions: basic, helix I, loop and helix II. An RMSD curve was plotted to reflect the time-course fluctuation for each region. The backbone RMSD indicated fluctuations that ranged up to 1.06 Å for the basic region, 0.36 Å for the helix-I region, 0.35 Å for the loop region and 0.34 Å for the helix-II region. The region with the highest RMSD value implied the highest deviation from the reference equilibrated structure (Supplementary Figure S5). All of the simulated systems demanded authentic time for organization and structural stabilization. There was no evidence supporting redistribution of atomic position because of constant steadfast $R_g$ values (Supplementary Figure S2). The RMSD values of mutated proteins were higher than those of their wt counterparts.

**Tracking the region motion by essential dynamic analysis**

Principal component analysis (PCA) was employed to visualize the basic domain and clarify the integrated motive pattern. In a biomolecular system, PCA is found on the presumption that major fluctuant collective modes govern the functional dynamics. The trajectory data in the simulation was engaged to construct the mass-weighted covariance matrix of the atomic coordinates, where the eigenvectors provide the motive direction and determine the associated motive extent (RMSF of the collective motion) (Amadei, Linssen, & Berendsen, 1993). The results of the PCA analysis in Figure 6 present the percentage of cumulative eigenvalues and the movements projected along the original eigenvectors for all wt and mutants. The cones present the direction and amplitude that the atoms moved with. The images in Figure 6(A) demonstrate the contribution of each eigenvector to the protein fluctuations, and 97% of the motions of wt dimers resulted from the first five eigenvectors. Analogously, this ratio in H57A, R65A, and LL7879AA mutated homodimers was 99, 98, and 95%, respectively (Figure 6). The five most representative collective motions for the mutated dimers reinforce the observation that the basic domain demonstrated an intense motion, which brought more flexibility on molecular binding.
Figure 5. Comparison of the atomic fluctuations per residue during the simulation. The RMSF values for each monomer of the homodimer and heterodimer are displayed in (A). The b-factors represented in cartoon form for the homo- and heterodimers (the wt and mutants) are displayed in (B). The most fluctuating residues are colored in blue (terminal residues), and the remainder of the dimer is in red (the most stable), in accordance with the observed fluctuations.

Note: Å – angstrom \( (10^{-10} \text{ m}) \); ns – nanoseconds \( (10^{-9} \text{ s}) \).
Figure 6. Collective motion analysis. (A) represents the percentage of cumulative eigenvalues as a function of eigenvector indices for the homodimers and heterodimers. (B) demonstrates the porcupine plot of the first eigenvector of the homodimer and heterodimer. Note: The cones point in the direction of the indicated mode of atomic movement, and the amplitude of the motion is represented by the length of the cone.

Table 2. Results from decomposition of the binding energy calculation of DEC1 homodimers through MD simulation.

|                      | DEC1_A/DEC1_B | DEC1_A/DEC1_B-H57A | DEC1_A/DEC1_B-R65A | DEC1_A/DEC1_B-LL7879AA |
|----------------------|---------------|---------------------|---------------------|-------------------------|
| van der Waals energy | −677.966 ± 1.803 | −610.527 ± 2.608 | −533.923 ± 3.920 | −626.749 ± 3.383       |
| Electrostatic energy | −14,835.381 ± 18.547 | −14,430.246 ± 27.483 | −11,969.127 ± 23.703 | −14,333.233 ± 30.910   |
| Polar solvation energy | 4248.679 ± 14.090 | 3762.398 ± 24.247 | 3246.786 ± 19.603 | 3716.928 ± 24.695      |
| SASA energy          | −77.119 ± 0.204 | −71.046 ± 0.346 | −66.715 ± 0.324 | −71.076 ± 0.388        |
| Binding energy       | −11,341.869 ± 9.308 | −11,349.376 ± 14.823 | −9322.980 ± 15.527 | −11,314.171 ± 15.314   |
Binding energy between protein and DNA

Binding energy between protein and DNA was calculated based on MM/PBSA method. Homodimer-harboring R65A mutant exhibited highest binding energy values, while the other homodimer did not obviously diverge. The binding range was from $-11,349.376 \pm 14.82$ to $-9322.980 \pm 15.527$ kJ/mol (Table 2). Mutated homodimers brought various interferences on binding energy. Contribution of residues to the binding energy was also calculated to assess all mutants (Figure 7).

Discussion

At present, there is still no integrated DECs spatial structure reported. Predicting these structures with important mutations in the different regions through homology modeling facilitates investigation on the behavior of protein structures in aqueous solution. Lacking of whole protein structure results from wide distribution of intrinsic disordered (ID) region. ID segments widely spread in human transcript factors, which permit interaction between multiple proteins. ID proteins and their transcripts are present in relatively low levels and for short periods of time compared to structured proteins due to a higher proportion of predicted ubiquitination sites and miRNA target sites (Tsveltov, Reuven, Prives, & Shaul, 2009). Dimerization of ID containing bHLH proteins can be interfered by binding small molecular. (Follis, Hammoudeh, Wang, Prochownik, & Metallo, 2008). The large disordered region in DEC1N-terminal is essential for its interaction with ARNTL/BMAL1, E-box binding, and repressor activity against the CLOCK-ARNTL/
BMAL1 heterodimer. The C-terminal region of DEC1, which is highly conserved in vertebrates, contains three α-helices (334–341; 357–363; 394–399). It plays a critical role in repressive transcript activity (Sun & Taneja, 2000). In addition, DEC2 also possesses a similar region intercalated by an additional proline-rich region. Lacking the C-terminal WRPW domain of other bHLH families, DEC1 and DEC2 demonstrated distinct repressive machineries through modulation of the basal transcriptional machinery and histone deacetylase (Chakrabarti et al., 2004).

Our previous studies demonstrated that DEC1 played a crucial role in tumor progression in human gastric and hepatic cancer (Jia et al., 2013; Ma et al., 2013; Zheng et al., 2012). DEC1s usually bind to E-box elements with bHLH domains. Thus, mutants in the domains may change spatial structures, which alter the DNA-binding affinity and produce derangement events. The bHLH domains of the DECs attract more attention due to the reported mutants located in these regions related to transcription factor functionality. Human H57, R65, LL7879 in DEC1 and LL7071 in DEC2 present a high level of conservation among different species. DEC1, DEC2, and HES-1 have been reported to be associated with gastric and breast cancers and myelomonocytic leukemia (Jia et al., 2013; Klinakis et al., 2011; Montagner et al., 2012). The expression pattern of DEC1 is positively associated with the differentiation status of patients with hepatic carcinoma (Shi et al., 2011). Dimerization plays an important role in keeping biological functions of the bHLH domain containing transcription factors. Mutations at these sites may cause interference on DNA-binding capacity dominating the transcription factor functionality. The mutants H57A and R65A in DEC1 alter the charge and polarity of the residue side chain, in contrast to mutants LL7879AA in DEC1 and LL7071AA in DEC2. Different mutants bring distinct effects on protein structure and behavior in aqua solution. The mutant R65A in DECs may play a crucial role in affecting the protein behavior. The acutest raise of DNA-binding energy in homodimer of DEC1 harboring the R65A mutant implies an obvious fluctuation on protein-DNA affinity (Figure 7).

Moreover, the loop region may recruit hydrogen bonds, bridging helices I and II to determine the stability of the HES-I dimer. The basic domain possesses higher mobility, which facilitates binding to different target genes. The helix-I region, which is longer than the helix-II region, is restricted by the geometry of dsDNA binding, while the helix–II region defines the dimer interface. Mutants R65A change the DNA-binding energy obviously, resulting from the altering polarity of this residue, which change the spatial structure of helix-I region. Although mutants H57A also change the polarity of residue, they does not altering the DNA-binding energy due to the high flexibility of basic region that cover the effects of mutated residues. DNA binding of DEC1 and DEC2 was dependent on E-box recognition and interaction, which may alter normal circadian rhythms and contribute significantly to the pathogenesis of many diseases, including cancer (Fujimoto et al., 2007; Li et al., 2004). Our predicted model is consistent with the hypothesis that mutants in the bHLH region extenuate DNA-binding ability (El-Ghouzzi et al., 2001). Previous studies on the structure of DEC1 and DEC2 mainly focused on the gene structure or primary amino acid structure without 3D structure modeling and MDs. Our results may extend the research on the relationship between function and spatial structure, which may bring new insights for regulation mechanisms.

All models of the bHLH region of the DEC1 and DEC2 homodimers and the DEC1/HES-1 and DEC2/ HES-1 heterodimers showed stable secondary structural conformation. However, only binding free energy in DEC1 homodimer with mutated R65A deviated from the wt and other mutated homodimers. Mutants interfering protein–DNA affinity in this kind of transcript factors could be potential targeted drug affecting sites.

Author contributions
YK carried out the protein modeling and MDs simulation and drafted the manuscript. ZW carried out the model validation assay. YK, YJ, PL, SH, and ZW participated in the design of the study and performed the statistical analysis. YW conceived of the study, and participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This work was supported by the National Natural Science Foundation of China (NSFC) [81000869], [81272588]; Project 973 grants [2012CB966503], [2012CB966504].

Supplementary material
The supplementary material for this paper is available online at http://dx.doi.org/10.1080/07391102.2016.1196463.

References
Altschul, S. F. (1990). Basic local alignment search tool. Journal of Molecular Biology, 215, 403–410.
Arnaud, A., Linssen, A. B., & Berendsen, H. J. (1993). Essential dynamics of proteins. Proteins: Structure, Function, and Genetics, 17, 412–425.
Puntervoll, P., Linding, R., Gemünd, C., Chabanis-Davidson, S., Matingsdal, M., Cameron, S., ... Gibson, T. J. (2003). ELM server: A new resource for investigating short functional sites in modular eukaryotic proteins. Nucleic Acids Research, 31, 3625–3630.

Rossner, M. J., Oster, H., Wichert, S. P., Reinecke, L., Wehr, M. C., Reinecke, J., ... Nave, K. A. (2008). Disturbed clockwork resetting in Sharp-1 and Sharp-2 single and double mutant mice. PLoS One, 3, e2762.

Sali, A., & Blundell, T. (1993). Comparative protein modelling by satisfaction of spatial restraints. Journal of Molecular Biology, 234, 779–815.

Sato, F., Kawamoto, T., Fujimoto, K., Noshiro, M., Honda, K. K., Homma, S., ... Kato, Y. (2004). Functional analysis of the basic helix-loop-helix transcription factor DEC1 in circadian regulation. Interaction with BMAL1. European Journal of Biochemistry, 271, 4409–4419.

Shi, X.-H., Zheng, Y., Sun, Q., Cui, J., Liu, Q.-H., Qi, F., & Wang, Y.-S. (2011). DEC1 nuclear expression: A marker of differentiation grade in hepatocellular carcinoma. World Journal of Gastroenterology, 17, 2037–2043.

Sokkar, P., Sathis, V., & Ramachandran, M. (2012). Computational modeling on the recognition of the HRE motif by HIF-1: Molecular docking and molecular dynamics studies. Journal of Molecular Modeling, 18, 1691–1700.

Sun, H., & Taneja, R. (2000). Stra13 expression is associated with growth arrest and represses transcription through histone deacetylase (HDAC)-dependent and HDAC-independent mechanisms. Proceedings of the National Academy of Sciences, 97, 4058–4063.

Tsang, A. H., Sanchez-Moreno, C., Bode, B., Rossner, M. J., Garaulet, M., & Oster, H. (2012). Tissue-specific interaction of Per1/2 and Dec2 in the regulation of fibroblast circadian rhythms. Journal of Biological Rhythms, 27, 478–489.

Tsvetkov, P., Reuven, N., Prives, C., & Shaul, Y. (2009). Susceptibility of p53 unstructured N terminus to 20 S proteasomal degradation programs the stress response. Journal of Biological Chemistry, 284, 26234–26242.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. C. (2005). GROMACS: Fast, flexible, and free. Journal of Computational Chemistry, 26, 1701–1718.

Zheng, Y., Shi, X., Wang, M., Jia, Y., Li, B., Zhang, Y., ... Wang, Y. (2012). The increased expression of DEC1 gene is related to HIF-1α protein in gastric cancer cell lines. Molecular Biology Reports, 39, 4229–4236.

Zhou, H., & Zhou, Y. (2002). Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. Protein Science: A Publication of the Protein Society, 11, 2714–2726.