CD247 can bind SHC1, no matter if CD247 is phosphorylated

Tao Liu\textsuperscript{a,1}, Fangjin Chen\textsuperscript{a,1}, Ning Tang\textsuperscript{a,1}, Jiannan Feng\textsuperscript{b}, Dongsheng Zhao\textsuperscript{c}, Kaihua Wei\textsuperscript{d}, Yunping Zhu\textsuperscript{e}, Fuchu He\textsuperscript{e,f}, and Shaojun Liu\textsuperscript{a,*}

INTRODUCTION

On T cell receptor (TCR) stimulation, src homology 2 domain-containing transforming protein C1 (SHC1) had been found to bind the tyrosine-phosphorylated CD247 chain of the receptor via its src homology 2 (SH2) domain, delivering signals that control T cell development and activation. However, how the phosphorylation of CD247 led to the instant binding has not been characterized clearly. To study the binding process in detail, we simulated and compared the interaction processes of the SH2 domain with CD247 and phosphorylated CD247, respectively. Unexpectedly, the simulation revealed that SHC1 can also bind the nonphosphorylated CD247 peptide, which was further validated to be a weak binding by affinity pull-down experiment. The molecular dynamics (MD) simulation also revealed that the CD247 peptide formed a folding conformation with its Leu209 inserted into the hydrophobic binding pocket in SHC1. And on phosphorylation, it was the electrostatic attraction between the CD247 Tyr(P)206 and the SHC1 Tyr(P)-binding pocket that destroyed the folding conformation of the nonphosphorylated CD247 and, aided by the electrostatic attraction between SHC1 and the Asp203 of CD247, led to the extended conformation of the phosphorylated CD247 binding to SHC1 strongly. The results suggest that nonphosphorylated CD247 can recruit SHC1 in advance to prepare for the instant needs for SHC1 on TCR stimulation. In view of the ubiquity of phosphorylation in protein interaction regulation, we think this study also exemplified the usefulness of MD in more interactome research involving phosphorylation. Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: molecular dynamics simulation; interaction; T cell receptor; pull-down experiment; MALDI-TOF; phosphorylation; CD247; SHC1

*Correspondence to: Prof. S. Liu, State Key Laboratory of Proteomics, Department of Neurobiology, Institute of Basic Medical Sciences, 27 Taiping Road, Beijing 100850, P.R. China. E-mail: liusj@nic.bmi.ac.cn

\textsuperscript{a} T. Liu, F. Chen, N. Tang, S. Liu State Key Laboratory of Proteomics, Department of Neurobiology, Institute of Basic Medical Sciences, Beijing 100850, P.R. China

\textsuperscript{b} J. Feng Department of Immunology, Institute of Basic Medical Sciences, Beijing 100850, P.R. China

\textsuperscript{c} D. Zhao Super Biomedical Computing Center, Academy of Military Medical Sciences, Beijing 100850, P.R. China

\textsuperscript{d} K. Wei Laboratory of Mass Spectrometry, National Center of Biomedical Analysis, Beijing 100850, P.R. China

\textsuperscript{e} Y. Zhu, F. He State Key Laboratory of Proteomics, Biology, Institute of Radiation Medicine, Beijing 100850, P.R. China

\textsuperscript{f} F. He Institutes of Biomedical Sciences, Fudan University, Shanghai, 200032, P.R. China

Abbreviations used: APBS, adaptive Poisson–Boltzmann solver; GST, glutathione S-transferase; ITAM, immunoglobulin receptor tyrosine-based activation motifs; MALDI-TOF, matrix assisted laser desorption/ionization time-of-flight; MD, molecular dynamics; MM/PBSA, molecular mechanics/Monte Carlo Poisson-Boltzmann/surface area method; MS, mass spectrometry; RMSD, root mean square deviation; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; SH2, src homology 2; SHC1, src homology 2 domain-containing transforming protein C1; TCR, T cell receptor.
The binding solvation free energy was calculated by the molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) (Srinivasan et al., 1998; Kollman et al., 2000) method using the v.32b2 of the CHARMM package. The grid size used was 32 Å in each of the two protein systems for hydration and water molecules. The system was simulated in NPT ensemble mode with periodic boundary conditions and full electrostatics computed using the particle mesh Ewald method (Baker et al., 2001) and viewed by the v.0.99 of PyMOL software (DeLano, San Carlos, CA). The root mean square deviation (RMSD), atom distance, and interaction energy were analyzed using the v.32b2 of the CHARMM software package.

The binding solvation free energy was calculated by the molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) (Srinivasan et al., 1998; Kollman et al., 2000) method using the v.32b2 of the CHARMM package. The grid size used was 32 Å. Solute and solvent dielectric constants were taken as 72 Å and 80.0, respectively. Ionic strength was 150 mM. Temperature was 310 K. The linearized Poisson–Boltzmann equation was solved with the optimized successive over-relaxation method. A commonly used surface tension coefficient of 0.02 kcal/mol Å−2 was used for the calculation of the solvent-accessible surface area free energy (Fogolari et al., 2003). The entropy change was

**MATERIALS AND METHODS**

**Preparation of the structures**

The structures were prepared using the v.32b2 of the program CHARMM (Brooks et al., 1983), with v.27 of the CHARMM force field (Mackerell et al., 1998; Mackerell et al., 2004). The protein system of human SHC1 SH2 domain complexed with the tyrosine-phosphorylated portion of the third ITAM of CD247 was built using the structure 1TCE deposited in PDB database (Zhou et al., 1995b). Another protein system was built straight by removing the phosphate group directly from the CD247 peptide and CD247 protein. The residue number in the original structure was used in this study for consistency. To determine the protonation states of histidine residues, the potential energy of the complex was calculated with all of them in HSD (ND1-protonated) form. Then the energy of the complex was calculated with all histidines in HSD form except for one in HSE (NE2-protonated) and HSP (ND1 and NE2-protonated) form, respectively. After comparing the three energy values, the one corresponding to the lowest energy was chosen as the form of this histidine residue. The procedure was repeated for each histidine residue. At last, the C-terminal His107 in SHC1 SH2 domain was determined to be in HSP form, which may due to its exposure to solvent and the negative electrostatic potential at the C-terminus, and all other histidines in HSE form. Then the protein systems were energy minimized to remove unfavorable contacts and reduce the strain. A pre-equilibrated cube of water molecules (72 × 72 × 72 Å3 dimensions) was superimposed on each of the two protein systems for hydration and water molecules. The structure was determined in a solution with much lower pH value 6.5 and ionic strength 0.05 mol/kg than physiological conditions. So, what is the conformation of the nonphosphorylated CD247? How does the phosphate group at Tyr(P)206 lead to the binding conformation of the phosphorylated CD247? What is the truth under the physiological conditions? Structural studies alone cannot answer these questions. To understand the physical basis for the binding process and thus the role of the phosphate group, we performed molecular dynamics (MD) simulations on the CD247/SHC1 complexes with the CD247 Tyr206 phosphorylated and nonphosphorylated, respectively. The simulations were performed in explicit water solvent with similar ionic strength to the physiological conditions. Confirmed by the affinity pull-down experiment, the simulations revealed the binding of nonphosphorylated CD247 to SHC1 for the first time. The behavior and role of the CD247 Tyr(P)206 in the binding process were also revealed by the simulations.

**MD simulations and analysis**

Both whole systems were heated to and equilibrated at 310 K for 100 ps. Hereafter, they were simulated at 310 K under 1 atmosphere for 26.7 ns. The MD simulations were carried out using the v.32b2 of the program CHARMM (Brooks et al., 1983), with v.27 of the CHARMM force field (Mackerell et al., 1998; Mackerell et al., 2004). All bonds involving hydrogen atoms were constrained during the simulations using SHAKE and an integration time-step of 2 fs was used. The system was simulated in NPT ensemble mode with periodic boundary conditions and full electrostatics computed using the particle mesh Ewald method (Luty et al., 1994). Constant pressure was maintained using the Langevin piston method (Feller et al., 1995) with a 1 kDa pressure piston, a piston collision frequency of 10 ps−1, and a piston bath temperature of 310 K. Constant temperature was maintained using the Nose–Hoover algorithm (Nosé, 1984; Hoover, 1985) with a Hoover reference temperature of 310 K. Langevin dynamics was used throughout both simulation runs.

The trajectories were viewed by v.1.8.5 of VMD program (Humphrey et al., 1996). The electrostatic potential was calculated by the v.0.5.0 of adaptive Poisson–Boltzmann solver (APBS) tool (Baker et al., 2001) and viewed by the v.0.99 of PyMOL software (DeLano, San Carlos, CA). The root mean square deviation (RMSD), atom distance, and interaction energy were analyzed using the v.32b2 of the CHARMM software package.
THE BINDING OF CD247 TO SHC1

approximated by a quasi-harmonic analysis, in which the average coordinate set was used as a reference and the coordinate trajectory has been reoriented with respect to the reference structure to remove the translation-rotation motion by the MERGE command using all atoms and mass weighting.

Bioinformatics

The protein sequences of SHC1 SH2 domains and the CD247 peptides of various species available were retrieved from Genbank. Multiple sequence alignments were performed using ClustalX program (Thompson et al., 1997) with default parameters.

Expression of glutathione S-transferase (GST) and GST-SHC1 fusion protein

In view of the 98% homology between the human and rat SHC1 SH2 domains, we used the rat SHC1 SH2 domain, corresponding to residues 330–469 of the full rat SHC1 protein, in the pull-down experimental verification. RNA was harvested from fresh rat lung tissue using the SV Total RNA Isolation System and protocol (Promega, Madison, NJ, USA), and the coding sequence of the SH2 domain was amplified using a forward primer including the EcoRI site (underlined), 5'-AAAGAATTCGGACCCCCAAATCCC-TTCTGT-3', and a reverse primer including the XhoI site, 5'-TTTCTGAGTCACACCTTTGTCCACGGG-3'. The amplified fragment was digested with EcoRI and XhoI and subcloned into pGEX-6P-1 (Amersham Biosciences, Piscataway, NJ, USA), a N-terminal GST fusion vector. The encoded GST-SHC1 fusion protein contained a GST sequence at the N-terminal to aid in western blot detection. The intact pGEX-6P-1 vector was used for pull-down experiment.

Pull-down experiment

The peptide (Gly-His-Asp-Gly-Leu-Tyr-Gln-Gly-Leu-Ser-Thr-Ala-Thr-Lys) and its Tyr(P)-containing counterpart (Gly-His-Asp-Gly-Leu-Tyr(P)-Gln-Gly-Leu-Ser-Thr-Ala-Thr-Lys) corresponding to the portion of the third ITAM of CD247 were carried out: phosphorylated CD247 peptide with GST and GST-SHC1, respectively, non-phosphorylated CD247 peptide with GST and GST-SHC1, respectively, His-tag peptide with GST-SHC1, and none peptide with GST-SHC1.

Western blot

The eluted protein samples were separated by a 12% polyacrylamide gel and electrotransferred to a nitrocellulose membrane in the Trans-Blot® Semi-dry Electrophoretic Transfer Cell (Bio-Rad, USA). Nonspecific bands were blocked in TBS-T (25 mM Tris, 150 mM NaCl, 0.05% Tween20, pH 7.5) containing 5% skimmed milk for 1 hour at room temperature. Membranes were subsequently incubated with the primary mouse antibody of GST (1:2000) at room temperature for 1.5 h and followed by anti-mouse IgG horseradish peroxidase conjugate (1:1000). The immunoreactive proteins were detected by the enhanced chemiluminescence system (ECL, Amersham Pharmacia) and serial exposures were made to X-ray film (X-Omat BT Film, Kodak).

In-gel digestion, peptides purification and concentration, matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS)

The eluted protein samples were separated by a 12% polyacrylamide gel and the protein bands were stained with Coomassie Brilliant Blue R250. The protein band corresponding to the immunologically detected band around the 43 KDa position was excised, destained in 50 mM NH₄HCO₃/acetonitrile (50:50) and dried by vacuum centrifugation. Hereafter, modified bovine trypsin (10 ng/μl) dissolved in 50 mM NH₄HCO₃ digestion buffer was added to the dry gel piece and incubated on ice for 1 h for re-swelling and then changed to 37°C for 18 h. The peptide mixture was then concentrated and purified to remove detergents like SDS. After being eluted in matrix solution (0.5% trifluoroacetic acid, 50% acetonitrile, and 20 mg/ml α-cyano-4-hydroxy cinnamic acid), the peptide mixture were deposited on the stainless steel MALDI probe to dry slowly at ambient temperature. Mass spectrometry was performed using a Bruker Daltonics autoflex MALDI-TOF-MS (Bruker Daltonics, Bremen, Germany). Mass spectra were detected in the reflectron mode and recorded by the flexControl 2.4 software (Bruker Daltonics). Monoisotopic peptide masses were labeled by the Xmass 5.1.1 software (Bruker Daltonics). The average mass accuracy was less than 0.2 Da in the mass scanning range and the resolution was about 15,000. All spectra were externally calibrated by Peptide Calibration Standard (Bruker Daltonics) which was composed of Angiotensin II ([M+H]^+ 1046.5418), Angiotensin I ([M+H]^+ 1296.6848), Substance P ([M+H]^+ 1347.7354), Bombesin ([M+H]^+ 1619.8223), ACTH clip 1-17 ([M+H]^+ 2093.0862), ACTH clip 18-39 ([M+H]^+ 2465.1983), and Somatostatin 28 ([M+H]^+ 3147.4710).

Analysis of peptide mass values

As the GST-SHC1 fusion protein does not exist in any protein databases, we used the FindPept tool (http://www.expasy.org/tools/findpept.html) to analyze the correspondence between the MS data and GST-SHC1 fusion protein sequence. Peptide modifications were set to include partial oxidation of methionine and partial carbamidomethylation of cysteine. The experimental peptide mass values were set to monoisotopic including the mass...
of the charge carrier, MH+. The error window on experimental peptide mass values was 0.3 Dalton. The enzyme used for protein digestion was set to bovine cationic trypsin. Under the aforementioned criteria, the GST-SHC1 protein sequence (Figure 2) and the experimental peptide mass values were submitted for analysis.

RESULTS

MD simulations

One striking difference between the two simulations is the behavior of the terminal ends of the CD247 peptide. For the system with CD247 phosphorylated at Tyr206, the CD247 peptide extended to achieve a thorough binding with the SH2 domain (Figure 3F). The Tyr(P)-binding pocket in SH2 domain, containing Arg16, Arg32, His53 and Leu55, and Arg17 are positively charged (Figure 3A). The Tyr(P)206 and Asp203 in CD247 are negatively charged (Figure 3B). Due to the electrostatic attraction (Figure 4A, arrow 1), the negative Tyr(P)206 moved closer to the positive Arg32 and Arg16 at about 2.8 ns, and simultaneously the negative Asp203 was attracted to the positive Arg16 and Arg17 soon (Figure 4C, arrow 6). After an equilibration of about 10.8 ns, the interaction of Asp203 with the Arg16 and Arg17 tended to a stable state (Figure 4A, arrow 2, C, arrow 7). After another equilibration of about 10.4 ns, there began to form stable interaction of the negative Tyr(P)206 with the positive Arg16 and Arg32 (Figure 4A, arrow 3, C, arrow 8). As shown by the RMSD time series of the CD247 peptide backbone atoms with the structure of the complex reoriented on the backbone of the SH2 domain alone (Figure 3H), the peptide ultimately reached a stable binding conformation with the SHC1 protein.

However, with the phosphate group removed from the Tyr(P) in CD247, the N-terminal end (residues 201–206) and C-terminal end (residues 211–214) of CD247 became dissociated from the SH2 domain and aggregated with each other (Figure 3G). As shown in Figure 3C, remove of the phosphate group resulted in the loss of the strong negative electrostatic potential and even made the side chain of Tyr206 positive charged slightly. The transition from the strong electrostatic attraction to slight repulsion between the Tyr206 and the Tyr(P)-binding pocket (Figure 4B, arrow 4) pushed the Tyr206 away from Arg32 and Arg16 immediately at about 420 ps (Figure 4D, arrow 9). This dissociation was accompanied by the immediate aggregation of the Tyr206 and Ala212 (Figure 4D, arrow 9). After a long equilibration of about 17 ns, the Tyr206 tended to dissociate from Arg32 and Arg16 completely and aggregate with Ala212 stably (Figure 4B, arrow 5 and D, arrow 10). Furthermore, the amino of the N-terminal Gly201 approached to the carboxyl of the C-terminal Lys214 at about 5 ns and kept closing to it from about 20 ns (Figure 4D), which should be due to their electrostatic attraction (Figure 4B). As shown by the RMSD time series of the CD247 peptide backbone atoms with the structure of the complex reoriented on the backbone of the SH2 domain alone (Figure 3I), the nonphosphorylated peptide moved much with respect to the SHC1 protein and ultimately reached a stable binding conformation with the SHC1 protein.

No matter if the Tyr206 was phosphorylated, Leu209 of CD247 remained inserted into the hydrophobic binding pocket of SH2 domain during both simulations, as revealed both through visualization of the trajectories and analysis of the interaction energies between the Leu209 residue and the SH2 domain (Figure 3G, Figure 4A, B). The analysis of interaction energy also showed that there was a stronger interaction between Glu207 and the SH2 domain, compared with that of the phosphorylated peptide (Figure 4A, B). The interaction energy of Glu207 with the SH2 domain was even lower than that of Leu209 in the nonphosphorylated peptide (Figure 4B).

In order to understand the detailed interaction between SHC1 SH2 domain and the CD247 peptide, the structure possessing the smallest average RMSD of CA atoms in the last 1.6 ns of the simulation during which the RMSD was relatively steady (Figure 3D, E), was taken as the most representative conformation of the ensemble. The structures chosen were those at 26.293 and 26.228 ns for tyrosine and phosphotyrosine containing simulations, respectively. For the structure containing phosphotyrosine, the interaction energy of SH2 domain with the Tyr(P)206 (−162 kcal/mol) and Asp203 (−168 kcal/mol) residues made up 71% of the total interaction energy (−465 kcal/mol) with CD247 peptide. As shown in Figure 3F, the phosphate group of Tyr(P)206 could form hydrogen bonds with the side chains of Arg32 and Arg16, whereas the side chain of Asp203 could form hydrogen bonds with the side chains of Arg16 and Arg17. At the C-terminal end, the backbone nitrogen atoms of Ser210 and Thr211 can form hydrogen bonds with the backbone oxygen atom of the Arg64, and the carboxyl oxygen atom of Lys214 can form hydrogen bond with the zeta nitrogen atom of Lys66. In addition to the hydrophobic interactions between Leu209 and the
hydrophobic binding pocket (examined by Leu76), hydrophobic interactions were also found for Tyr(P)206 with the side chain carbon atoms of Gln40 and Leu55. For the structure with Tyr(P)206 replaced by Tyr, all these interactions dismissed except for the hydrophobic contact between Leu209 and the hydrophobic binding pocket (examined by Leu76) (Figure 3G). However, the backbone oxygen and nitrogen atoms of Gln207 can form hydrogen bonds with the backbone nitrogen atom of Leu55 and the backbone oxygen atom of Ala90, respectively. In the peptide, the amido nitrogen atom of Gly201 can form hydrogen bond with the carboxyl oxygen atom of Lys214. And the Tyr206 and Ala212 aggregated together forming hydrophobic interactions between their side chain carbons. As shown by the dotted arrows in Figure 3G, the residues 201–206 and 211–214 were displaced from the SH2 domain and turned to each other. The interaction energy of the CD247 peptide with the SH2 domain increased to −115 kcal/mol which was 75% higher than that of the Tyr(P)-containing CD247 peptide.

The binding affinities of the two peptides with the SH2 domain were further evaluated by calculating the binding solvation free energy using the MM/PBSA method and the trajectory with ultimately stable conformation of the complex (from 25 200 to 26 800 ps). The binding solvation free energy of the nonphosphorylated peptide and the protein was calculated to be −33 ± 7 kcal/mol, which should favor their binding. However, it is a rather weak binding, compared with the much lower binding solvation free energy, −252 ± 20 kcal/mol, of the phosphorylated peptide and the protein.
Experimental verification of the binding of the nonphosphorylated CD247 peptide to the SHC1 SH2 domain

Pull-down experiment of His-tagged CD247 peptides was used to validate the different binding of SHC1 SH2 domain to the CD247 peptide and its tyrosine-phosphorylated version. The GST fusion protein of SHC1 SH2 domain was successfully expressed in soluble form, as shown in lane 3 of Figure 5A. The bacteria proteins, containing the GST-SHC1 fusion protein, were incubated with the Ni-agarose chelated with the two versions of the CD247 peptides respectively. After washing off the non-bound proteins, the remained proteins were separated by SDS-PAGE. And the GST-SHC1 fusion protein was detected by highly sensitive immunological blot by using antibody against GST. As shown in Figure 5B and C, a great deal of GST-SHC1 fusion protein was specifically pulled down by the Tyr(P)-containing CD247 peptide, whereas a small quantity, detected only by western blot, was pulled down by the nonphosphorylated CD247 peptide. For a control of the non-specific binding to the Ni-agarose, another pull-down experiment was also performed with the same procedure except for the adding of either CD247 peptide. Few proteins were pulled down by the Ni-agarose and were detected as three light bands by SDS-PAGE. Among them, one band was detected to contain a trace of GST-SHC1 fusion protein by Western blot. However, the three bands were not detected by SDS-PAGE when the Ni-agarose was enveloped with the nonphosphorylated CD247 peptide. So the GST-SHC1 fusion proteins obtained from the Ni-agarose enveloped with nonphosphorylated CD247 peptides should be mostly pulled down by the CD247 peptides. To eliminate the influence of the GST and (His)$_6$ tags on the pull-down results, another three pull-down experiments were accomplished. The bacterial proteins containing GST were subjected to both CD247 peptide pull-down systems, and the bacterial proteins containing GST-SHC1 were subjected to His-tag pull-down system. No GST or GST-SHC1 protein was detected by SDS-PAGE and western blot methods. So the detected GST-SHC1 proteins should be pulled down by their specific interactions with the CD247 peptides. To confirm the result, the pull-down experiments were repeated three times and got the same result. To confirm the identity of the detected protein band, the protein band pulled down by the Tyr(P)-containing CD247 peptide was cut and subjected to MALDI-TOF-MS analysis. The correspondence between the MS data and GST-SHC1 fusion protein sequence was evaluated by the FindPept tool. As long as the number of allowed missed cleavage sites was set to one, 18 peptide mass values matched total 13 peptides from specific cleavage of GST-SHC1 fusion protein by bovine trypsin (Figure 2, Figure 5D), 24 matched peptides from unspecific cleavage of GST-SHC1, 1 matched peptide resulting from trypsin autolysis, 5 matched peptides resulting from cleavage of contaminants, essentially keratin, and the last 6 unmatched. Among the matched specifically cleaved peptides, eight were from the GST part and the remained five were from the SHC1 part. The matched specifically cleaved peptides covered 46.2% of the Ni-agarose was enveloped with the nonphosphorylated CD247 peptide. So the GST-SHC1 fusion proteins obtained from the Ni-agarose enveloped with nonphosphorylated CD247 peptides should be mostly pulled down by the CD247 peptides. To eliminate the influence of the GST and (His)$_6$ tags on the pull-down results, another three pull-down experiments were accomplished. The bacterial proteins containing GST were subjected to both CD247 peptide pull-down systems, and the bacterial proteins containing GST-SHC1 were subjected to His-tag pull-down system. No GST or GST-SHC1 protein was detected by SDS-PAGE and western blot methods. So the detected GST-SHC1 proteins should be pulled down by their specific interactions with the CD247 peptides. To confirm the result, the pull-down experiments were repeated three times and got the same result. To confirm the identity of the detected protein band, the protein band pulled down by the Tyr(P)-containing CD247 peptide was cut and subjected to MALDI-TOF-MS analysis. The correspondence between the MS data and GST-SHC1 fusion protein sequence was evaluated by the FindPept tool. As long as the number of allowed missed cleavage sites was set to one, 18 peptide mass values matched total 13 peptides from specific cleavage of GST-SHC1 fusion protein by bovine trypsin (Figure 2, Figure 5D), 24 matched peptides from unspecific cleavage of GST-SHC1, 1 matched peptide resulting from trypsin autolysis, 5 matched peptides resulting from cleavage of contaminants, essentially keratin, and the last 6 unmatched. Among the matched specifically cleaved peptides, eight were from the GST part and the remained five were from the SHC1 part. The matched specifically cleaved peptides covered 46.2% of the

Figure 4. The interaction energy and atom distances during the simulations. (A), (B) the interaction energy between the SHC1 and the phosphorylated and nonphosphorylated CD247 peptides, respectively, during the two MD simulations. (C), (D) the distance of some certain atom pairs during the simulations of the SHC1 complexed with the phosphorylated and nonphosphorylated CD247 peptides, respectively. In the legends, the S letters mean SHC1, the C letters after the hyphens mean the CD247 peptide, the numbers after S and C mean the residue numbers, the E letters mean the electrostatic interaction energy, and the letters after the residue numbers mean atom names.
whole GST-SHC1 fusion protein sequence. The protein pulled down by the CD247 peptides should be the GST-SHC1 fusion protein.

DISCUSSION AND CONCLUSIONS

In the present study, we have used MD simulations to obtain atomic-level insight into the effect of phosphorylation of CD247 peptide on its binding to the SHC1 SH2 domain. No matter if the Tyr206 was phosphorylated, the CD247 peptide was found to be able to bind to the SHC1 SH2 domain, however, in dramatically different manners. When unphosphorylated at Tyr206, the ITAM peptide of CD247 aggregated to a curled conformation which withdrew the sidechain of Tyr206 from the Tyr(P)-binding pocket of the SH2 domain and embedded the hydrophobic atoms, especially the CA and CB atoms of Ala212 and the CB and CG atoms of Tyr206, inside the curled peptide structure. However, due to the interaction between Leu209 and the hydrophobic binding pocket of the SH2 domain, and the interactions between Gln207 and two residues, Ala90 and Leu55, of the SH2 domain, the CD247 peptide could still bind to the SH2 domain (Figure 3G). As shown by the curves after 14 ns in Figure 4B and D, the electrostatic attraction between the negative charged Asp203 of CD247 and the positive charged Tyr(P)-binding pocket of SH2 domain could not lead to a stable binding between them. Aided by the electrostatic attraction between Gly201 and Lys214, the Asp203 was displaced from the SH2 domain at last with other N-terminal residues to aggregate with the C-terminal residues of CD247. The MD analysis therefore offers a clear representation of a less stable CD247/SHC1 complex, which was further proved by the binding solvation free energy comparison and pull down experiment.

In a previous study (Laczko et al., 1998), the conformation of a CD247 peptide 6 residues longer at the C-terminus and two residues shorter at the N-terminus than the CD247 peptide in this work was studied. In water, 51% of the Tyr206-phosphorylated peptides were found to form β-sheet conformation, and the rest

Figure 5. Experimental verification of the molecular dynamics simulation results. (A) SDS-PAGE of the proteins expressed in the E. coli transformant of GST-SHC1 fusion expression vector, lane 1 and 2, the whole bacteria before and after inducement, lane 3, the supernate after inducement and sonication, the 42.5 kDa bands marked by arrows corresponded to the expressed GST-SHC1 fusion protein. (B) SDS-PAGE of the proteins pulled down. GST-SHC1 proteins pulled down by the peptide corresponding to the third ITAM motif of CD247 (lane 5), its tyrosine-phosphorylated counterpart (lane 4), and the Ni-agarose (lane 6); GST-SHC1 proteins pulled down by the (His)$_6$ peptide (lane 7); GST proteins pulled down by the CD247 peptide (lane 8) and its tyrosine-phosphorylated version (lane 9). (C) Western blot detection of the GST-SHC1 fusion protein or the GST protein pulled down lanes 10–12 and 13–15 correspond to lanes 4–6 and 7–9, respectively. The arrows in B and C panels indicated the 42.5 kDa positions. (D) The MALDI-TOF peptide mass fingerprints of the band in lane 4 of B panel, the matched fingerprints were indicated by solid and dotted arrows for SHC1 and GST, respectively.
were in unordered conformation. For the nonphosphorylated peptides, the β-sheet conformation reduced to 29%, the unordered conformation populated to 61%, and the rest 9% were in α-helix conformation. In this study, upon dephosphorylation, the peptide transformed from an extended conformation to a curled conformation, which mimicked the transformation from β-sheet to unordered conformation found in the previous study.

Based on the two simulations, the binding process of the phosphorylated CD247 to the SH2 domain can be deduced as follows. Upon phosphorylation at Tyr206, the negative charged Tyr(P)206, and soon the negative charged Asp203, was attracted strongly toward the positive charged Tyr(P)-binding pocket of the SH2 domain, which destroyed the curled conformation of the CD247 peptide and then released the C-terminal residues to bind to the SH2 domain around its hydrophobic binding pocket. Additionally, the Asp203 was found to play a synergetic role in the binding process by facilitating the binding of the N-terminal residues with the SH2 domain. So we proposed two key factors controlling the transition between the two binding states. One is the electrostatic interaction between Tyr(P)-binding pocket and the N-terminal residues of CD247 got strong enough to stretch CD247 peptide to the extended conformation to accommodate its binding to SHC1.

In an evolutionary view, the residues with important functional roles should conserve during the evolution process. The protein sequences of the CD247 peptide (Figure 6A), as well as those of SHC1 SH2 domain (Figure 6B), were compared among various species. And the residues referred to in the interaction analysis were all conserved highly, which provided evolutionary evidences for the simulated results.

In contrast to the finding in the MD simulation, it had once been reported that the nonphosphorylated CD247 peptide did not exhibit any detectable binding affinity to the SHC1 SH2 domain by NMR titration (Zhou et al., 1995a). In the previous study, the 15N/1H chemical shift changes of the residues located in phosphotyrosine binding pocket were monitored as a function of the concentration of the bound peptide. As simulated in this work, the nonphosphorylated peptide coiled up and bound only to the hydrophobic binding pocket. So the chemical shift of the residues in the phosphotyrosine binding pocket should not be influenced by the bound nonphosphorylated peptide. We think this may explain why they had not detected any binding affinity of the unphosphorylated peptide to the SHC1 protein in the previous work. In this study, the experiment result agreed with the simulated result well, which proposed the possibility of the weak binding of CD247 to SHC1 in vivo and thus implied that nonphosphorylated CD247 can recruit SHC1 in advance to meet...
the instant needs for SHC1 upon TCR stimulation. However, this hypothesis needs to be validated by further experiment in vivo.

In summary, the present study revealed the weak binding of CD247 to SHC1 SH2 domain and the possible transition mechanism to the strong binding induced by CD247-phosphorylation. Reversible phosphorylation of proteins is a major pathway in which key cellular activities are regulated (Cohen, 2002). In recent years, phosphorylated sites in many proteins have been identified (Beausoleil et al., 2004; Di Pancrazio et al., 2006; Khan et al., 2006; Kim et al., 2006; Villar et al., 2006; Chen et al., 2007; Palmisano et al., 2007). However, to establish the functional behaviors of these phosphate groups is still a challenge for interactome research. MD simulation is an important and widely used theoretical tool for studying detailed micro- and macroscopic behavior of protein–protein systems. We think this study also exemplified the use of MD simulation combined with experimental measurements as a means for understanding the functional role of phosphorylation in future interactome researches.

Acknowledgements

This work was supported by The Chinese National Natural Science Foundation (30430310, 30670791) and The Chinese National Key Project of Basic Research (001CB510206).

REFERENCES


REFERENCES


