

ORIGINAL ARTICLE

A phosphorylation pattern-recognizing antibody specifically reacts to RNA polymerase II bound to exons

Jungwon Han^{1,2}, Jong-Hyuk Lee^{1,2}, Sunyoung Park^{1,5}, Soomin Yoon^{1,6}, Aerin Yoon^{1,7}, Do B Hwang³, Hwa K Lee¹, Min S Kim^{1,8}, Yujean Lee^{1,6}, Won J Yang¹, Hong-Duk Youn^{1,2}, Hyori Kim⁴ and Junho Chung^{1,2,3}

The C-terminal domain of RNA polymerase II is an unusual series of repeated residues appended to the C-terminus of the largest subunit and serves as a flexible binding scaffold for numerous nuclear factors. The binding of these factors is determined by the phosphorylation patterns on the repeats in the domain. In this study, we generated a synthetic antibody library by replacing the third heavy chain complementarity-determining region of an anti-HER2 (human epidermal growth factor receptor 2) antibody (trastuzumab) with artificial sequences of 7–18 amino-acid residues. From this library, antibodies were selected that were specific to serine phosphopeptides that represent typical phosphorylation patterns on the functional unit (YSPTSPS)₂ of the RNA polymerase II C-terminal domain (CTD). Antibody clones pCTD-1stS2 and pCTD-2ndS2 showed specificity for peptides with phosphoserine at the second residues of the first or second heptamer repeat, respectively. Additional clones specifically reacted to peptides with phosphoserine at the fifth serine of the first repeat (pCTD-1stS5), the seventh residue of the first repeat and fifth residue of the second repeat (pCTD-S7S5) or the seventh residue of either the first or second repeat (pCTD-S7). All of these antibody clones successfully reacted to RNA polymerase II in immunoblot analysis. Interestingly, pCTD-2ndS2 precipitated predominately RNA polymerase II from the exonic regions of genes in genome-wide chromatin immunoprecipitation sequencing analysis, which suggests that the phosphoserine at the second residue of the second repeat of the functional unit (YSPTSPS)₂ is a mediator of exon definition.

Experimental & Molecular Medicine (2016) **48**, e271; doi:10.1038/emm.2016.101; published online 18 November 2016

INTRODUCTION

The synthetic antibody library approach to antibody generation provides advantages over conventional methods that utilize either naive or immunized animals to generate hybridomas and phage display libraries (for a review of those methods, see Miersch and Sidhu¹). The use of synthetic antibody libraries is especially valuable for developing antibodies for phosphorylated protein motifs, because it is difficult to produce these antibodies from naive or immunized sources.² Previously, one antiphosphopeptide motif antibody was successfully

engineered into multiple antibodies that were reactive to diverse phosphopeptides by inserting an anion-binding nest in the heavy chain complementarity-determining region 2 (HCDR2) and randomizing HCDR3 residues.³

We constructed a synthetic scFv library with a trastuzumab backbone^{4,5} and artificial HCDR3 with 7–18 amino-acid residues (Figure 1). We adopted the anti-HER2 antibody trastuzumab as the scaffold, because it can harbor reactivity to multiple antigens by modifying CDRs. Previously, trastuzumab has been engineered to display substantial affinity

¹Department of Biochemistry and Molecular Biology, Seoul National University College of Medicine, Seoul, Republic of Korea; ²Department of Biomedical science, Seoul National University College of Medicine, Seoul, Republic of Korea; ³Cancer Research Institute, Seoul National University College of Medicine, Seoul, Republic of Korea and ⁴Asan Institute for Life Sciences, University of Ulsan College of Medicine, Asan Medical Center, Seoul, Republic of Korea

⁵Current address: SK Telecom Co., Ltd., Seoul 04539, Republic of Korea.

⁶Current address: Abclone, Seoul 08381, Republic of Korea.

⁷Current address: Mogam Institute, Yongin-si, Gyeonggi-do 16924, Republic of Korea.

⁸Current address: Sorrento Therapeutics, San Diego, CA 92121, USA.

Correspondence: Dr H Kim, Asan Institute for Life Sciences, University of Ulsan College of Medicine, Asan Medical Center, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Republic of Korea.

E-mail: hyorikim@amc.seoul.kr

or Professor J Chung, Department of Biochemistry and Molecular Biology, Seoul National University College of Medicine, Samsung Cancer Research Building, Suite 510, 101 Daehak-ro, Jongno-gu, Seoul 00380, Republic of Korea.

E-mail: jjhchung@snu.ac.kr

Received 3 May 2016; revised 20 May 2016; accepted 25 May 2016

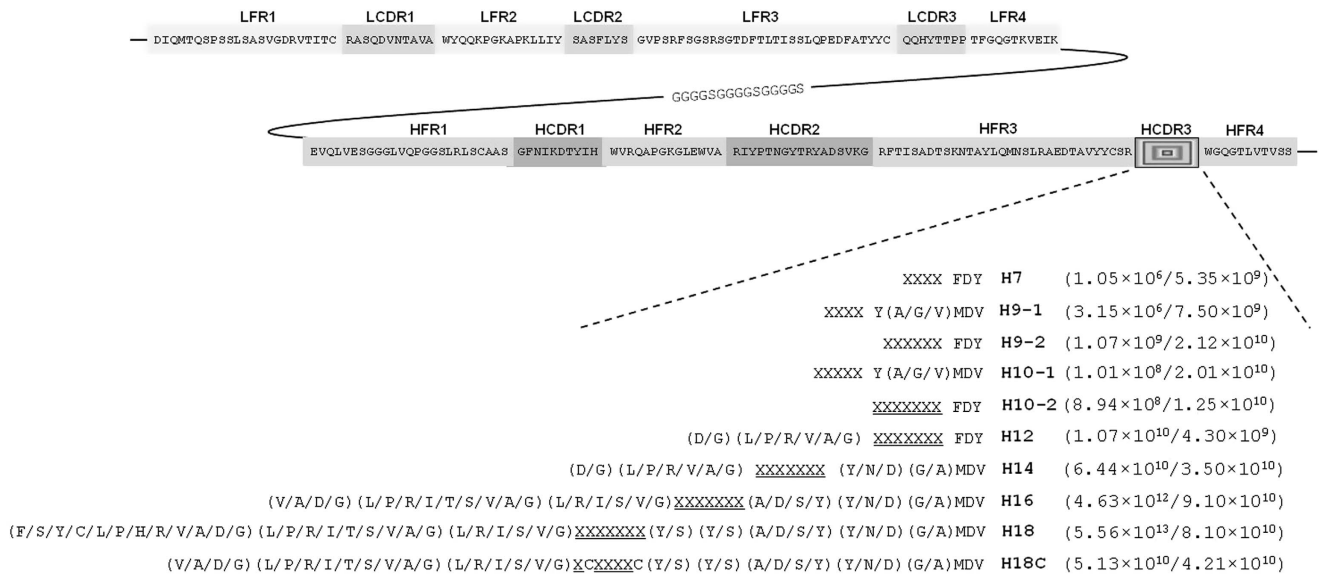


Figure 1 Sequences of constructed synthetic antibody libraries. The amino-acid sequence of scFv, except HCDR3, was adopted from trastuzumab. The length of HCDR3 varies from 7 to 18 amino acids. The H18C library contains two cysteines to form an intra-HCDR3 disulfide bond. 'X' represents residues randomized by the NNK nucleotide sequence, and 'X' (underlined) represents residues randomized by a mixture of trinucleotides encoding 19 amino acids except for cysteine. When a position is diversified with <19 amino acids by partially degenerate codons, the encoded amino acids are shown inside parentheses. The complexity required to cover all possible sequences and the actual complexities achieved in this study are also shown in parentheses (expected complexity/achieved complexity).

to vascular endothelial growth factor while maintaining its reactivity to HER-2 by introducing mutations to light chain CDRs.⁶ Therefore, its antigen reactivity likely can be preferentially localized to either the heavy or the light chain. We confined the artificial amino-acid sequences to HCDR3, because we have successfully generated integrin-specific antibodies from an antibody library with synthetic genes that were introduced only to HCDR3.⁷

The RNA polymerase II C-terminal domain (CTD) is an unusual extension appended to the C-terminus of the largest subunit of RNA polymerase II, which serves as a flexible binding scaffold for numerous nuclear factors.⁸ CTD contains multiple repeats of the YSPTSPS motif, which can be simultaneously phosphorylated at multiple residues (potentially at any residues except proline) and yield diverse phosphorylation patterns that can coordinate the binding of various nuclear factors.⁹ A twice-repeated sequence of the motif, YSPTSP-SYSTSPS, is a functional unit and has an optimal amino-acid length for epitope functionality¹⁰ that could facilitate the use of phosphopeptide-specific antibodies for functional assays. Because CTD has a flexible three-dimensional structure,¹¹ the domain can accommodate the structural changes necessary for binding with a synthetic antibody. Several antibodies react to the serine-phosphorylation pattern of CTD, but their specificity has been only minimally characterized, and their amino-acid sequences are not yet publically available.¹²

In this study, we constructed a synthetic antibody library with HCDR3-confined artificial sequences and selected antibodies specific to serine-phosphorylation patterns by biopanning nine phosphopeptides that represent serine-phosphorylated CTD (Figure 2). The selective reactivity of

antibodies to these phosphopeptides and a non-phosphorylated peptide was tested in enzyme immunoassays, and their reactivity to RNA polymerase II CTD was evaluated through immunoblot analysis and genome-wide chromatin immunoprecipitation (ChIP) sequencing analysis using mammalian cell lysates.

MATERIALS AND METHODS

Cell culture and peptide synthesis

HEK 293F cells (Invitrogen, Carlsbad, CA, USA) were grown in FreeStyle 293 Expression medium (GIBCO, Grand Island, NY, USA) containing 100 U ml^{-1} penicillin and $100 \mu\text{g ml}^{-1}$ streptomycin at 37°C in 7% CO_2 and 95% relative humidity in an orbital shaking incubator (Minitron, INFORS HT, Bottmingen, Switzerland) set at 135 r.p.m. HEK 293T and HeLa cells were obtained from the American Type Culture Collection (Manassas, VA, USA) and grown in Dulbecco's modified Eagle's medium (Welgene, Seoul, South Korea) supplemented with 10% fetal bovine serum (GIBCO), 100 U ml^{-1} penicillin and $100 \mu\text{g ml}^{-1}$ streptomycin at 37°C in a humidified atmosphere of 5% CO_2 . Phospho-YSPTSPSYSPS-BSA conjugates were synthesized by Pepton, Inc. (Daejeon, South Korea). The phosphorylation patterns of the peptides are described in Figure 2a.

Construction of the synthetic scFv library

Amino-acid sequences of the trastuzumab variable regions (V_H and V_L) were obtained from the PDB database (1N8Z). The sequence encoding the V_L domain, a linker (Glycine-Glycine-Glycine-Glycine-Serine)₃ and the part of the V_H domain from FR1 to FR3 was chemically synthesized (Integrated Device Technology, Inc., San Jose, CA, USA). The region encoding partial FR3, synthetic HCDR3 and FR4 of V_H was generated by PCR using degenerate

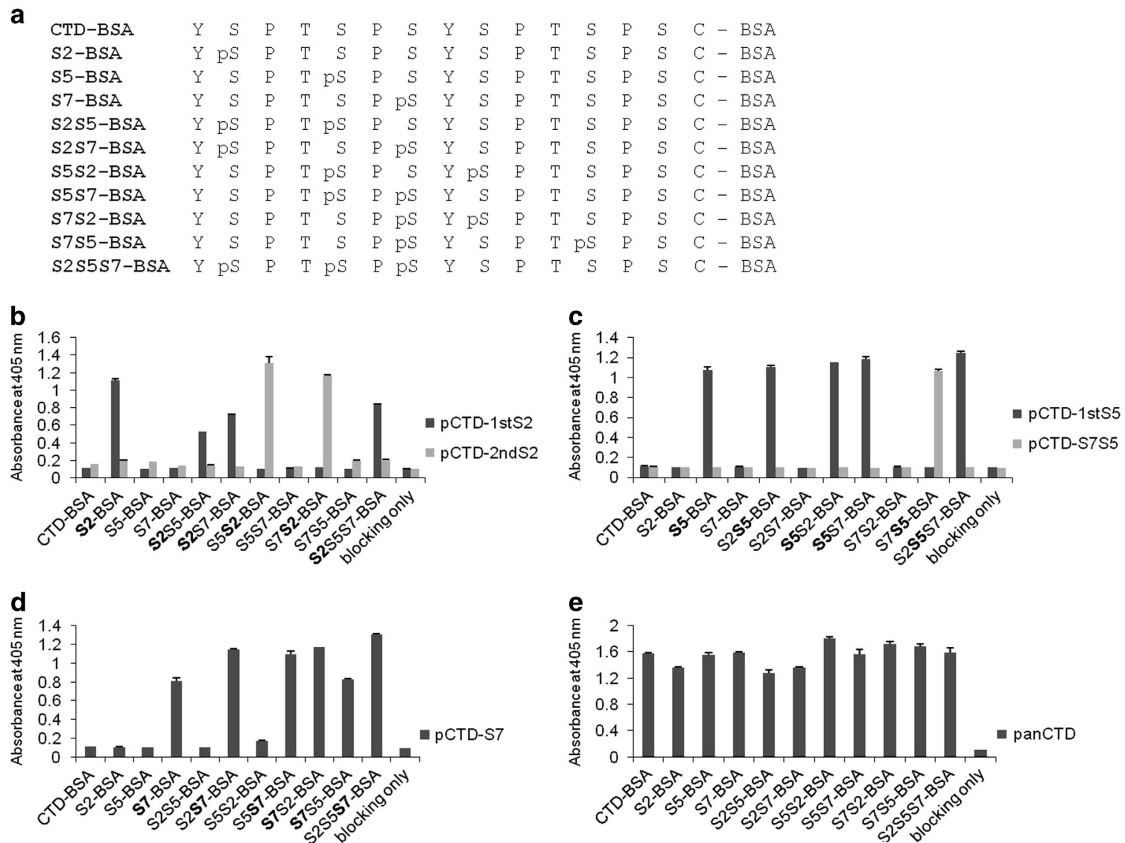


Figure 2 Reactivity of selected antibodies to BSA-conjugated CTD phosphopeptides. Phosphopeptides were conjugated to BSA (a). Each well of a microtiter plate was coated with a phospho-YSPSPSPSPSPS-BSA conjugate, and scFv-Fc fusion proteins were added to each well: pCTD-1stS2 and pCTD-2ndS2 (b), pCTD-1stS5 and pCTD-S7S5 (c), pCTD-S7 (d), and PanCTD (e). The plate was washed and incubated with anti-rabbit IgG Fc antibody conjugated with horseradish peroxidase. After washing, the amount of bound antibody was determined using a colorimetric substrate.

primers and a Her-VH-R primer (5'-CCGGCCGGCCTGGCCGGAG GACACGGTCACCAGG-3'). Ten degenerate primers were used to introduce artificial sequences into HCDR3 (Supplementary Table S1). Two DNA fragments were combined by overlap extension PCR using HAX-VL-F (5'-GGCCCAGGCGCCGACATCC-3') and Her-VH-R primers to generate the scFv gene with *Sfi*I restriction sites at both termini. The genes encoding scFv and the pComb3XSS phagemid vector were subjected to *Sfi*I restriction enzyme digestion and subsequent ligation. The recombinant vector was then transformed into the *Escherichia coli* strain ER2738 (New England BioLabs, Ipswich, MA, USA), as previously described.¹³ After overnight culture, a phage-displayed combinatorial antibody library was prepared.¹³

Biopanning and phage enzyme immunoassays

To enrich specific binders from the library, five rounds of biopanning were performed as previously described.¹³ One-hundred and fifty micrograms of each phospho-YSPSPSPSPSPS-BSA conjugate was mixed with 3.0×10^8 paramagnetic beads (M-270 Epoxy DynaBeads, Invitrogen) following the manufacturer's instructions. After overnight incubation at room temperature on a rotator, the beads were washed four times with 0.5% bovine serum albumin (BSA) in phosphate-buffered saline (PBS) and blocked with 3% BSA in PBS (w/v). The phage-displayed combinatorial antibody library was added to the beads and incubated for 2 h at room temperature on a rotator.

The beads were washed once with 0.05% Tween-20 in PBS (PBST, v/v) for the first round, which increased to three washes in subsequent rounds. After washing the beads, 0.1 M glycine-HCl (pH 2.2) was added, and the mixture was incubated for 10 min at room temperature to elute bound phages from the beads. The eluate was neutralized by adding 2 M Tris-Cl (pH 9.1), and the eluted phages were transduced into *E. coli* strain ER2738. Phages displaying scFv were rescued by adding VCSM13 helper phages (Stratagene, La Jolla, CA, USA). Individual phage clones were selected from the output titration plate from the last round of biopanning, and scFv-displaying phages were prepared for the phage enzyme immunoassay as previously described.¹³

To select binders, the reactivity of antibody-displaying phages was tested in phage enzyme immunoassays using phage-containing culture supernatant and horseradish peroxidase (HRP)-conjugated anti-M13 antibody (GE Healthcare, Piscataway, NJ, USA). The phage clones with positive signals were selected ($OD_{405} \geq 0.5$), and their nucleotide sequences were determined as previously described.¹³

Expression and purification of scFv-Fc fusion proteins

An expression vector modified from pCEP4 (Invitrogen) was used for cloning and has been previously described.^{14,15} The vector carries an expression cassette composed of the leader sequence of the human Ig κ -chain, two *Sfi*I sites for insertion of the antibody gene of interest, the hinge region of human IgG₁ and the C_{H2}-C_{H3} domains of rabbit IgG.

Table 1 HCDR3 amino-acid sequences of antibody clones

Phosphopeptides used in biopanning		HCDR3 sequence (amino-acid length)
	Clone	
S2-BSA	pCTD-1stS2	GAFWFSRSHFDY (12)
S2S5-BSA	pCTD-1stS5	GPWPGSHHKYAMDV (14)
S2S5S7-BSA	pCTD-S7	GISYPWQRGYYYAMDV (16)
S5S2-BSA	pCTD-2ndS2	DRAQWWEYKNAMDV (14)
S7S5-BSA	pCTD-S7S5	DPWSRWAKVFDY (12)
S2-BSA	panCTD	YGRRGVFDY (9)

The phagemid DNA of selected clones and the expression vectors were digested with *Sfi*I, and the scFv genes were cloned into the expression vector. Recombinant pCEP4 was transfected into HEK 293F cells using 25-kDa linear polyethylenimine (Polysciences, Warrington, PA, USA) as previously described.¹⁶ Overexpressed scFv-Fc fusion proteins were purified by affinity chromatography using protein A Sepharose columns (Repligen, Waltham, MA, USA) according to the manufacturer's instructions.

Enzyme immunoassay

Microtiter 96-well plates (Corning Costar, Cambridge, MA, USA) were coated with each phospho-YSPSPSYSPSPS-BSA conjugate overnight at 4 °C and blocked with 3% BSA in PBS. ScFv-Fc fusion proteins were diluted in 3% BSA in PBS at a concentration of 0.05 µg ml⁻¹ for pCTD-1stS2, pCTD-1stS5, pCTD-S7 and pCTD-S7S5 or 1 µg ml⁻¹ for pCTD-2ndS2 and panCTD (Table 1) and added to each well. After incubation for 2 h at 37 °C, the plate was washed with 0.05% PBST three times and incubated with HRP-conjugated anti-rabbit IgG Fc-specific antibody (Jackson ImmunoResearch, Inc., West Grove, PA, USA) for 1 h at 37 °C. The plate was washed three times with 0.05% PBST, and the amount of bound antibody was determined by adding 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (Amresco, Solon, OH, USA) in 0.05 M citric acid buffer (pH 4.0) with 1.0% H₂O₂. The resulting OD₄₀₅ was measured using a plate reader (Labsystems S.L., Barcelona, Spain).

Immunoblot analysis

HEK 293T cell lysates were subjected to electrophoresis on a NuPage 4–12% Bis-Tris gel (Invitrogen), and the resolved proteins were transferred to a nitrocellulose membrane. The membrane was blocked with 5% non-fat dry milk (NFD; BD Biosciences, Sparks, MD, USA) in Tris-buffered saline with 0.1% Tween-20 (TBST) at room temperature for 1 h. ScFv-Fc fusion proteins diluted in 5% NFD-0.1% TBST at 0.01 µg ml⁻¹ for pCTD-1stS5 and pCTD-S7, 0.05 µg ml⁻¹ for pCTD-S7S5, 1 µg ml⁻¹ for pCTD-1stS2 and PanCTD and 4 µg ml⁻¹ for pCTD-2ndS2 were added to the membrane and incubated overnight at 4 °C. After washing the membrane four times with 0.1% TBST, the membrane was incubated with HRP-conjugated anti-rabbit IgG Fc antibody (Jackson ImmunoResearch, Inc.) diluted in 5% NFD-0.1% TBST for 1 h at room temperature and then washed four times with 0.1% TBST. The target-bound scFv-Fc antibody was visualized with SuperSignal Pico West chemiluminescent substrate (Thermo Fisher Scientific, Rockford, IL, USA). A commercially available phosphorylated S2 CTD antibody (Abcam, Cambridge Science Park, Cambridge, UK) was used as a positive control.

Chromatin Immunoprecipitation assay

The ChIP assay was performed using the method of Lee *et al.*¹⁷ with slight modifications. Briefly, HeLa cells were cross-linked in a solution of 1% formaldehyde in PBS for 5 min at room temperature. The cross-linking reaction was stopped by adding glycine to a final concentration of 0.125 M. The cells were harvested and washed twice with cold PBS, and cytosolic fractions were eliminated with buffer A (5 mM PIPES (pH 8.0), 85 mM KCl, 0.5% NP-40, protease inhibitor cocktail (GenDEPOT, Katy, TX, USA)). Nuclear pellets were resuspended in buffer B (100 mM Tris-Cl (pH 8.1), 1% sodium dodecyl sulfate (SDS), 10 mM EDTA, protease inhibitor cocktail), and the chromatin was sheared with an S220 focused ultrasonicator (Covaris, Woburn, MA, USA). The prepared chromatin fraction (500 µg total sheared DNA per sample) was diluted 1/10 in IP buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-Cl (pH 8.1), 167 mM NaCl and a protease inhibitor cocktail) and incubated with 200 µg of scFv-Fc fusion proteins (pan or phosphorylation-specific RNA polymerase II CTD antibodies) overnight at 4 °C. Samples were incubated for 2–4 h at 4 °C with protein A or G beads precoated with salmon sperm DNA. Then the beads were washed with TSE150 (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-Cl (pH 8.1), 150 mM NaCl), TSE500 (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-Cl (pH 8.1), 500 mM NaCl), Buffer III (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-Cl (pH 8.1)) and two times with TE (pH 8.0) for 10 min in each solution. Bead-bound chromatin was eluted with elution buffer (1% SDS, 0.1 M NaHCO₃ (pH 8.0)) and incubated overnight at 65 °C with 200 mM NaCl to reverse cross-linking. Five hundred microliters of the sample were incubated at 50 °C after adding 10 µl of 0.5 M EDTA, 20 µl of 1 M Tris (pH 6.5) and 4 µl of Proteinase K (20 mg ml⁻¹) and then purifying the sample with phenol/chloroform/isoamyl alcohol, as previously described.¹⁸ Nucleic acids were precipitated by centrifugation for 30 min at 4 °C after mixing the sample with 1 µl of glycogen solution (20 mg ml⁻¹), 20 µl 5 M NaCl and 500 µl isopropanol. Purified nucleic acid pellets were washed with 70% ethanol, dried and dissolved in nuclease-free water.

ChIP sequencing

Sequencing and analysis of ChIPed DNA fragments were performed by Macrogen (Seoul, South Korea). Briefly, DNA fragments were ligated to a pair of adaptors for sequencing with the HiSeq 2500 sequencing system (Illumina, San Diego, CA, USA). The ligation products were size-fractionated separated on a 2% agarose gel to obtain 200–300 bp fragments and were PCR-amplified for 18 cycles. Each library was diluted to 8 pM for 76 cycles of single-read sequencing on the HiSeq 2500 following the manufacturer's recommended protocol. After read trimming, Bowtie V1.1.1 (read mapping), Picard V1.133 (remove duplicates), MACS 2 V2.1.0.20150420 (peak calling) and ChIPseeker (peak annotation) were used for downstream analysis.

RESULTS

Our synthetic scFv library was constructed using degenerate primers that yielded artificial sequences of HCDR3 on the backbone of trastuzumab. The primers were designed to encode 7, 9, 10, 12, 14, 16 and 18 amino-acid residues of HCDR3 (Figure 1), and all the amino-acid sequences of V_L and V_H (except HCDR3) were equivalent to the sequences of trastuzumab.⁴ The sequence of the V_L–V_H linker (Glycine–Glycine–Glycine–Serine)₃ has been commonly used to generate scFv.¹⁹ We constructed two different versions of

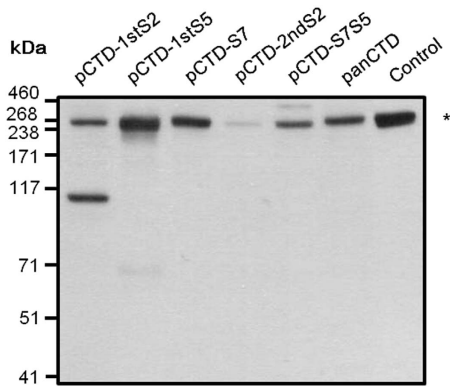


Figure 3 Reactivity of selected antibodies to RNA polymerase II CTD. HEK 293T cell lysates were subjected to SDS-polyacrylamide gel electrophoresis and transferred to a nitrocellulose membrane. After blocking, the membrane was probed with scFv-Fc fusion proteins. The location of bound scFv-Fc fusion protein was visualized using horseradish peroxidase-conjugated anti-rabbit IgG Fc antibody and a chemiluminescent reagent. A commercially available phosphorylated S2 CTD antibody was used as a positive control. The asterisk indicates the expected location of RNA polymerase II CTD.

HCDR3 with amino-acid lengths of 9 and 10 to enhance combinatorial diversity. We also included sequences containing two cysteine residues to potentially form a disulfide-constraint-loop inside HCDR3, because this structural motif was frequently used in our previous experiments⁷ and it exists in nature.²⁰ The expected and achieved complexities of the individual scFv libraries are described in Figure 1. The achieved complexity of the antibody library was 3.20×10^{11} .

The functional unit of RNA pol II CTD is two repeats of YSPTSPS,¹⁰ of which Tyr1, Ser2, Thr4, Ser5 and Ser7 can all be reversibly modified by phosphorylation at the side chain hydroxyl group.^{21,22} However, phosphorylation of Ser2, Ser5 and Ser7 may have a more critical role in molecular interactions between CTD and CTD-binding proteins.^{8,23} The phosphorylation pattern of CTD at serine residues has an affinity for certain binding proteins, which are classified into three different groups (writers/effectors, readers and erasers).²⁴ In this context, we synthesized 10 YSPTSPSYSPTSPS peptides with different serine-phosphorylation patterns and used them as antigens for antibody generation.

From five rounds of biopanning on 10 phospho-YSPTSPSYSPTSPS-BSA conjugates (Figure 2a), we enriched scFv clones and evaluated each clone's reactivity to the individual phosphopeptides and non-phosphorylated peptide control (data not shown) through a phage enzyme immunoassay. After nucleotide sequencing the clones (Table 1), we selected a subset of clones showing specificity for serine phosphopeptides for further analysis. None of the clones carried mutations in the regions adopted from trastuzumab. The selected clones were expressed as scFv-rabbit Fc fusion proteins, and their phosphopeptide specificities were confirmed by enzyme immunoassay (Figure 2).

Clone pCTD-1stS2 specifically reacted to peptides with phosphoserine at the second amino-acid position of the first

YSPTSPS repeat (Figure 2b), whereas clone pCTD-2ndS2 reacted to peptides with phosphoserine at the second amino-acid residue of the second repeat (Figure 2b). Clone pCTD-1stS5 reacted to peptides with phosphoserine at the fifth amino-acid residue of the first repeat (Figure 2c). Clone pCTD-S7S5 reacted to peptides with phosphoserine at the seventh position of the first repeat and the fifth position of the second repeat (Figure 2c). Clone pCTD-S7 reacted to peptides with phosphoserine at the seventh position of either the first or second repeat (Figure 2d). One clone, PanCTD, was reactive to all phospho- and non-phosphopeptides tested (Figure 2e).

To confirm the reactivity of the clones to native RNA polymerase II CTD, we performed immunoblots and found that all clones reacted to RNA polymerase II CTD from HEK 293T cell lysates (Figure 3). Nearly every antibody showed high specificity to RNA polymerase II CTD except for pCTD-1stS2, which reacted nonspecifically with a 100 kDa protein. Although the reactivity of pCTD-S7S5 ($0.05 \mu\text{g ml}^{-1}$) to S7S5-BSA (Figure 2c) was higher than the reactivity of pCTD-S7 ($0.05 \mu\text{g ml}^{-1}$) to S7S5-BSA (Figure 2d), the intensity of the band in the immunoblot lane probed with pCTD-S7S5 ($0.05 \mu\text{g ml}^{-1}$) was lower than the lane probed with pCTD-S7 ($0.01 \mu\text{g ml}^{-1}$), which was expected because the epitope of pCTD-S7S5 is less prevalent than the epitope of pCTD-S7.

We also examined whether these antibodies could be used for ChIP analysis. Four clones failed to co-immunoprecipitate enough genomic DNA fragments to be analyzed by sequencing (data not shown). Clone pCTD-2ndS2 successfully immunoprecipitated enough genomic DNA to be analyzed by sequencing. Interestingly, most of the DNA segments that were co-immunoprecipitated were distributed in the exonic regions of housekeeping genes, such as ACTB and GAPDH (glyceraldehyde 3-phosphate dehydrogenase; Figure 4a). This particular pattern was very different from patterns achieved using predefined RNA polymerase II panCTD or phosphorylated RNA polymerase II CTD-Ser2 antibodies, which are available in public databases (panCTD: GSM935395, CTD-Ser2: GSM935383). Additionally, DNA fragments were confined to the exonic regions of specific isoforms in some genes, such as HSPA8 and MYL9 (Figure 4b).

DISCUSSION

Considering the diverse patterns of RNA pol II CTD phosphorylation and its functional role in determining interactions with binding proteins, antibodies specific to a certain pattern of CTD serine phosphorylation are valuable tools for studying the physiological role of CTD phosphorylation. We previously showed that an intracellular anti-phospho-STAT3 antibody binds successfully to the phosphorylated motif of STAT3 (signal transducer and activator of transcription factor 3) and blocks its physiological function not only in *in vitro* but also in an *in vivo* setting.²⁵ When screening antibody clones, we attempted to isolate antibodies that could distinguish specific CTD phosphorylation patterns but realized that these antibodies are scarce (data not shown). We identified only one clone (pCTD-S7S5) that was specific to the S7S5 peptide but not reactive to the other nine phosphopeptides and non-

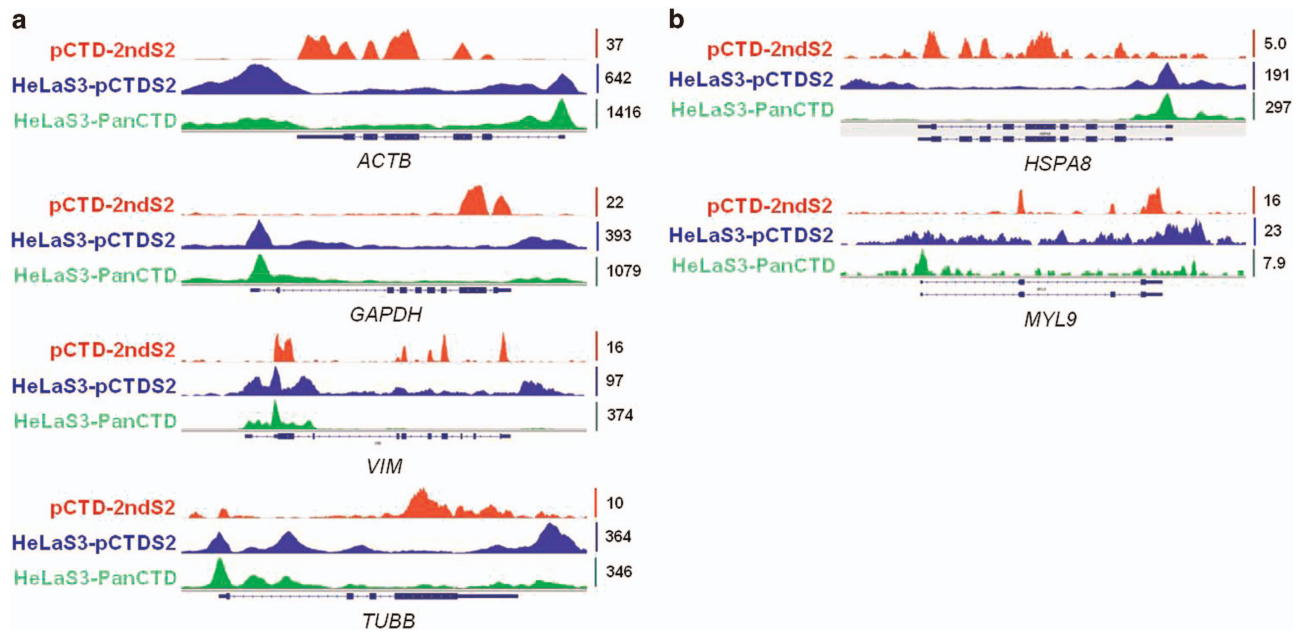


Figure 4 ChIP-binding profiles. pCTD-2ndS2 ChIP results showing an exonic pattern of distribution were compared with publicly available data (phosphorylated S2, HeLaS3-pCTDS2; panCTD, HeLaS3-PanCTD) (a, b). pCTD-2ndS2 profiles showing isoform-specific patterns (b). Scale data ranges are indicated on the right side of each individual track. RefSeq gene structures are depicted below the profiles.

phosphopeptide control (Figure 2). All other clones reacted to more than one phosphopeptide. Biopanning on peptides with multiple phosphoserine residues resulted in clones with broader specificity to phosphopeptides other than individual peptides used in biopanning (Table 1). These data also strongly suggest that high-throughput screening of antibody sequences using next-generation sequencing would be required to select rare antibody clones specific to peptides with multiple phosphoserines in future studies.^{26,27} The library we constructed in this study is optimal for this purpose because only nucleotide sequencing of HCDR3 is required, and the scFv gene can be easily synthesized using nucleotide sequence information obtained from next-generation sequencing. Alternatively, developing antibodies with a more stringent specificity may be possible by generating daughter libraries with randomized key amino acids in other CDRs.⁷

We selected five antibodies showing different specificities based on the phosphorylated patterns of CTD in the enzyme immunoassay (Figure 2) for further analysis. All of these antibodies successfully reacted to RNA polymerase II CTD in an immunoblot analysis, which proved that their reactivity was not limited to phosphopeptides. We also tested whether these antibodies could be used for ChIP analysis, which was quite challenging because there was only a limited number of monoclonal antibodies available for ChIP analysis. It is well known that monoclonal antibodies have a disadvantage in ChIP experiments because they only recognize a single epitope, and therefore there was a higher likelihood that the epitope will be masked by transcription factors or genomic DNA.²⁸ Consistent with these challenges, four antibody clones failed to co-immunoprecipitate enough DNA material to be analyzed by sequencing. However, the clone pCTD-2ndS2 successfully immunoprecipitated RNA polymerase

II that interacted with genomic segments confined to exonic regions (Figure 4). To the best of our knowledge, no RNA polymerase II CTD phosphorylation pattern has been reported to mark exonic regions. It is well known that DNA segments co-immunoprecipitated with the RNA polymerase II panCTD antibody show a widely distributed pattern throughout genomic loci with a greater abundance at the transcription start site. RNA polymerase II CTD phosphorylated at Ser2 shows another distribution pattern by predominantly interacting with 3' of transcription termination sites of genes, and it can recruit transcription termination factors.⁸

The mRNA reactions of capping, splicing and polyadenylation are believed to occur co-transcriptionally. The transcribed pre-mRNA itself contains several consensus elements in *cis* that are essential for the splicing reaction.²⁹ Therefore, it is a reasonable assumption that RNA polymerase II CTD with a certain phosphorylation code may recruit mRNA processing factors to facilitate this reaction. Because exonic regions are marked with cross-exon recognition complexes that contain serine/arginine-rich (SR) proteins,³⁰ and SR proteins are known to interact with phosphorylated RNA polymerase II CTD, it would be interesting to determine whether pCTD-2ndS2 phosphorylation in the double heptamer CTD unit is a code for SR protein recruitment. In addition, mammalian splicing factor Spt6 was reported to selectively interact with phosphorylated S2 through its SH2 domain and mediate hlws1-dependent splicing.³¹ However, Spt6 did not bind to a single CTD repeat,³² which suggested that a pattern of multiple phosphorylated sites over multiple heptamer repeats create a docking platform for a protein. It is possible that our pCTD-2ndS2 antibody is reacting to a similar modification on the chromatin, as observed in our genome-wide ChIP-seq data (Figure 4).

In summary, we constructed a synthetic antibody library with HCDR3-confined diversity and selected antibodies recognizing serine-phosphorylation patterns of RNA polymerase II CTD. The antibodies successfully reacted to CTD in immunoblots of cell lysate. One antibody discriminated one serine-phosphopeptide among the nine other peptides that were tested. Our data clearly showed that an antibody library constructed with synthetic sequences confined to HCDR3 provided enough structural diversity to encompass various phosphoserine-motif-specific antibodies and may enable screening of phosphorylation-pattern-recognizing antibodies using next-generation sequencing in a high-throughput manner. Most interestingly, clone pCTD-2ndS2 precipitated RNA polymerase II on the exonic regions of genes in a genome-wide ChIP sequencing analysis, which suggests the possibility that a phosphoserine at the second residue of the second repeat on the functional unit (YSPTSPS)₂ is a mediator of exon definition.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by a National Research Foundation of Korea (NRF) grant (No. 2012R1A5A2A44671346), the Bio and Medical Technology Development Program of the NRF (2014M3A9D3034034) and a National Research Foundation of Korea grant (No. 2012R1A3A2048767 to H-DY).

- 1 Miersch S, Sidhu SS. Synthetic antibodies: concepts, potential and practical considerations. *Methods* 2012; **57**: 486–498.
- 2 Adams JJ, Sidhu SS. Synthetic antibody technologies. *Curr Opin Struct Biol* 2014; **24**: 1–9.
- 3 Koerber JT, Thomsen ND, Hannigan BT, Degrado WF, Wells JA. Nature-inspired design of motif-specific antibody scaffolds. *Nat Biotechnol* 2013; **31**: 916–921.
- 4 Carter P, Presta L, Gorman CM, Ridgway JB, Henner D, Wong WL *et al*. Humanization of an anti-p185HER2 antibody for human cancer therapy. *Proc Natl Acad Sci USA* 1992; **89**: 4285–4289.
- 5 Magdelaine-Beuzelin C, Kaas Q, Wehbi V, Ohresser M, Jefferis R, Lefranc MP *et al*. Structure-function relationships of the variable domains of monoclonal antibodies approved for cancer treatment. *Crit Rev Oncol Hematol* 2007; **64**: 210–225.
- 6 Bostrom J, Yu SF, Kan D, Appleton BA, Lee CV, Billeci K *et al*. Variants of the antibody herceptin that interact with HER2 and VEGF at the antigen binding site. *Science* 2009; **323**: 1610–1614.
- 7 Chung J, Rader C, Popkov M, Hur YM, Kim HK, Lee YJ *et al*. Integrin alphaIIb beta3-specific synthetic human monoclonal antibodies and HCDR3 peptides that potently inhibit platelet aggregation. *FASEB J* 2004; **18**: 361–363.
- 8 Phatnani HP, Greenleaf AL. Phosphorylation and functions of the RNA polymerase II CTD. *Gene Dev* 2006; **20**: 2922–2936.
- 9 Heidemann M, Hintermair C, Voss K, Eick D. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim Biophys Acta* 2013; **1829**: 55–62.
- 10 Stiller JW, Cook MS. Functional unit of the RNA polymerase II C-terminal domain lies within heptapeptide pairs. *Eukaryot Cell* 2004; **3**: 735–740.
- 11 Meinhart A, Cramer P. Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature* 2004; **430**: 223–226.

- 12 Chapman RD, Heidemann M, Albert TK, Mailhammer R, Flatley A, Meisterernst M *et al*. Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. *Science* 2007; **318**: 1780–1782.
- 13 Barbas CF. Phage display: a laboratory manual. *Cold Spring Harbor Laboratory Press*. Cold Spring Harbor, NY, USA, 2001.
- 14 Park S, Lee DH, Park JG, Lee YT, Chung J. A sensitive enzyme immunoassay for measuring cotinine in passive smokers. *Clin Chim Acta* 2010; **411**: 1238–1242.
- 15 Kim H, Park S, Lee HK, Chung J. Application of bispecific antibody against antigen and hapten for immunodetection and immunopurification. *Exp Mol Med* 2013; **45**: e43.
- 16 Lee Y, Kim H, Chung J. An antibody reactive to the Gly63-Lys68 epitope of NT-proBNP exhibits O-glycosylation-independent binding. *Exp Mol Med* 2014; **46**: e114.
- 17 Lee JH, Kang BH, Jang H, Kim TW, Choi J, Kwak S *et al*. AKT phosphorylates H3-threonine 45 to facilitate termination of gene transcription in response to DNA damage. *Nucleic Acids Res* 2015; **43**: 4505–4516.
- 18 Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987; **162**: 156–159.
- 19 Trinh R, Gurbaxani B, Morrison SL, Seyfzadeh M. Optimization of codon pair use within the (GGGS)₃ linker sequence results in enhanced protein expression. *Mol Immunol* 2004; **40**: 717–722.
- 20 Johnson G, Wu TT. Kabat Database and its applications: 30 years after the first variability plot. *Nucleic Acids Res* 2000; **28**: 214–218.
- 21 Zhang J, Corden JL. Identification of phosphorylation sites in the repetitive carboxyl-terminal domain of the mouse RNA polymerase-II largest subunit. *J Biol Chem* 1991; **266**: 2290–2296.
- 22 Baskaran R, Dahmus ME, Wang JYJ. Tyrosine phosphorylation of mammalian RNA polymerase-II carboxyl-terminal domain. *Proc Natl Acad Sci USA* 1993; **90**: 11167–11171.
- 23 Eglhoff S, Murphy S. Cracking the RNA polymerase II CTD code. *Trends Genet* 2008; **24**: 280–288.
- 24 Eick D, Geyer M. The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem Rev* 2013; **113**: 8456–8490.
- 25 Koo MY, Park J, Lim JM, Joo SY, Shin SP, Shim HB *et al*. Selective inhibition of the function of tyrosine-phosphorylated STAT3 with a phosphorylation site-specific intrabody. *Proc Natl Acad Sci USA* 2014; **111**: 6269–6274.
- 26 Georgiou G, Ippolito GC, Beausang J, Busse CE, Wardemann H, Quake SR. The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat Biotechnol* 2014; **32**: 158–168.
- 27 Ravn U, Gueneau F, Baerlocher L, Osteras M, Desmurs M, Malinge P *et al*. By-passing in vitro screening-next generation sequencing technologies applied to antibody display and in silico candidate selection. *Nucleic Acids Res* 2010; **38**: e193.
- 28 Wardle FC, Tan H. A ChIP on the shoulder? Chromatin immunoprecipitation and validation strategies for ChIP antibodies. *F1000Res* 2015; **4**: 235.
- 29 Proudfoot NJ, Furger A, Dye MJ. Integrating mRNA processing with transcription. *Cell* 2002; **108**: 501–512.
- 30 Ast G. How did alternative splicing evolve? *Nat Rev Genet* 2004; **5**: 773–782.
- 31 Cho EJ. RNA polymerase II carboxy-terminal domain with multiple connections. *Exp Mol Med* 2007; **39**: 247–254.
- 32 Yoh SM, Cho H, Pickle L, Evans RM, Jones KA. The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. *Genes Dev* 2007; **21**: 160–174.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Supplementary Information accompanies the paper on Experimental & Molecular Medicine website (<http://www.nature.com/emm>)