



OPEN

## Rational design of a helical peptide inhibitor targeting c-Myb–KIX interaction

Shunji Suetaka<sup>1</sup>, Yoshiki Oka<sup>1</sup>, Tomoko Kuniyama<sup>1</sup>, Yuuki Hayashi<sup>1</sup> & Munehito Arai<sup>1,2</sup>✉

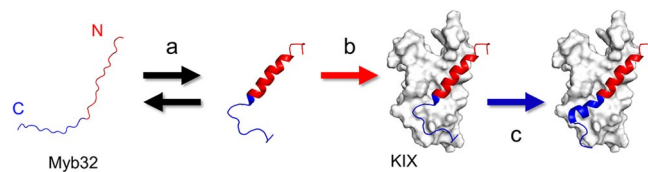
The transcription factor c-Myb promotes the proliferation of hematopoietic cells by interacting with the KIX domain of CREB-binding protein; however, its aberrant expression causes leukemia. Therefore, inhibitors of the c-Myb–KIX interaction are potentially useful as antitumor drugs. Since the intrinsically disordered transactivation domain (TAD) of c-Myb binds KIX via a conformational selection mechanism where helix formation precedes binding, stabilizing the helical structure of c-Myb TAD is expected to increase the KIX-binding affinity. Here, to develop an inhibitor of the c-Myb–KIX interaction, we designed mutants of the c-Myb TAD peptide fragment where the helical structure is stabilized, based on theoretical predictions using AGADIR. Three of the four initially designed peptides each had a different Lys-to-Arg substitution on the helix surface opposite the KIX-binding interface. Furthermore, the triple mutant with three Lys-to-Arg substitutions, named RRR, showed a high helical propensity and achieved three-fold higher affinity to KIX than the wild-type TAD with a dissociation constant of 80 nM. Moreover, the RRR inhibitor efficiently competed out the c-Myb–KIX interaction. These results suggest that stabilizing the helical structure based on theoretical predictions, especially by conservative Lys-to-Arg substitutions, is a simple and useful strategy for designing helical peptide inhibitors of protein–protein interactions.

Intrinsically disordered proteins (IDPs) play important roles in living cells, including transcription, translation, and cell cycle regulation<sup>1–4</sup>. Moreover, IDPs are often associated with serious diseases, such as cancer, through interactions with other proteins<sup>5</sup>. Therefore, protein–protein interactions (PPIs) involving IDPs are recognized as novel targets for drug discovery. One of the IDPs related to cancer is c-Myb, a transcription factor that binds the KIX domain of the transcriptional coactivator CREB-binding protein (CBP)<sup>3</sup>. The c-Myb–KIX interaction is essential for promoting the proliferation of hematopoietic cells and regulating hematopoiesis<sup>6–8</sup>. However, its aberrant expression causes leukemia<sup>9</sup>. In addition, overexpression of c-Myb also promotes the proliferation of tumor cells in breast and colon cancers<sup>10–12</sup>. Importantly, the growth and survival of tumor cells are strongly dependent on the activity of c-Myb<sup>13</sup>. Hence, suppressing the c-Myb activity by inhibiting the c-Myb–KIX interaction is a key strategy for cancer therapeutics. An inhibitor of the interaction is potentially useful as an antitumor drug<sup>14–17</sup>.

Several studies have been conducted to develop small-molecule inhibitors targeting the c-Myb–KIX interaction<sup>18–20</sup>. However, small-molecule compounds have both low affinity and specificity toward target proteins<sup>21</sup>. Therefore, peptide-based inhibitors have attracted considerable attention to overcome these problems. Furthermore, despite their relatively low molecular weights, peptides have higher affinity and specificity toward target proteins than small molecules<sup>22,23</sup>. Thus, efficient methods for designing peptides that inhibit target PPIs are strongly desired. One of the simple and promising strategies to design such peptides is to use a peptide fragment involved in the target PPI and modify it to further improve its binding affinity to the partner protein.

c-Myb, a 640-residue protein, consists of three domains: the DNA-binding domain, transactivation domain (TAD), and negative regulatory domain<sup>9,24–28</sup>. The TAD of c-Myb (residues 275–327) is intrinsically disordered but folds into a helical structure upon binding to the KIX domain of CBP. KIX consists of 87 residues and has the structure of a three-helix bundle. Despite its small size, KIX has two different binding sites, “the c-Myb/pKID site” and “the MLL site.” Both sites interact with the intrinsically disordered regions of various transcription factors, including c-Myb, CREB, and p53 on the c-Myb/pKID site<sup>29–31</sup> and a mixed lineage leukemia protein (MLL), c-Jun, p53, HIV-1 Tat, and HBZ on the MLL site<sup>32–36</sup>. These IDPs are also involved in the development of serious

<sup>1</sup>Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo 153-8902, Japan. <sup>2</sup>Department of Physics, Graduate School of Science, The University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo 153-8902, Japan. ✉email: arai@bio.c.u-tokyo.ac.jp



**Figure 1.** Binding mechanism of Myb32 to KIX<sup>38,40</sup>. (a) Myb32 is in equilibrium between a disordered and partially helical structure in the free form. (b) The N-terminal helical region of Myb32 (red) binds KIX by a conformational selection mechanism. (c) After binding, the C-terminal region (blue) forms a helical structure by an induced-fit mechanism.

diseases. Among the IDPs, c-Myb most tightly binds KIX<sup>31,37,38</sup>. Thus, inhibitors of the c-Myb–KIX interaction, which bind the c-Myb/pKID site of KIX, may also inhibit the diseases caused by PPIs at this site.

Although the c-Myb TAD is intrinsically disordered, the region of residues 284–315, called Myb32, has a high helical propensity<sup>38</sup>. Previous studies showed that Myb32 binds KIX with a dissociation constant ( $K_d$ ) of 0.21  $\mu$ M, tighter than the binding of the c-Myb TAD full length to KIX ( $K_d = 0.5 \mu$ M)<sup>30,38,39</sup>. Moreover, Myb32 is known to bind KIX by a conformational selection mechanism. This information may provide a clue to improve the binding affinity to KIX<sup>38,40</sup>. Therefore, Myb32 is a suitable template for designing peptide inhibitors.

Here, we rationally designed peptide inhibitors targeting the c-Myb–KIX interaction using a Myb32 fragment of the c-Myb TAD. Based on theoretical predictions, we introduced mutations into Myb32 and successfully developed a peptide inhibitor of the c-Myb–KIX interaction that has a higher affinity to KIX than wild-type Myb32 ( $K_d = 80$  nM). Our results suggest that stabilizing the helical structure based on theoretical predictions, especially Lys-to-Arg substitutions, is a useful strategy for designing helical peptide inhibitors of PPIs.

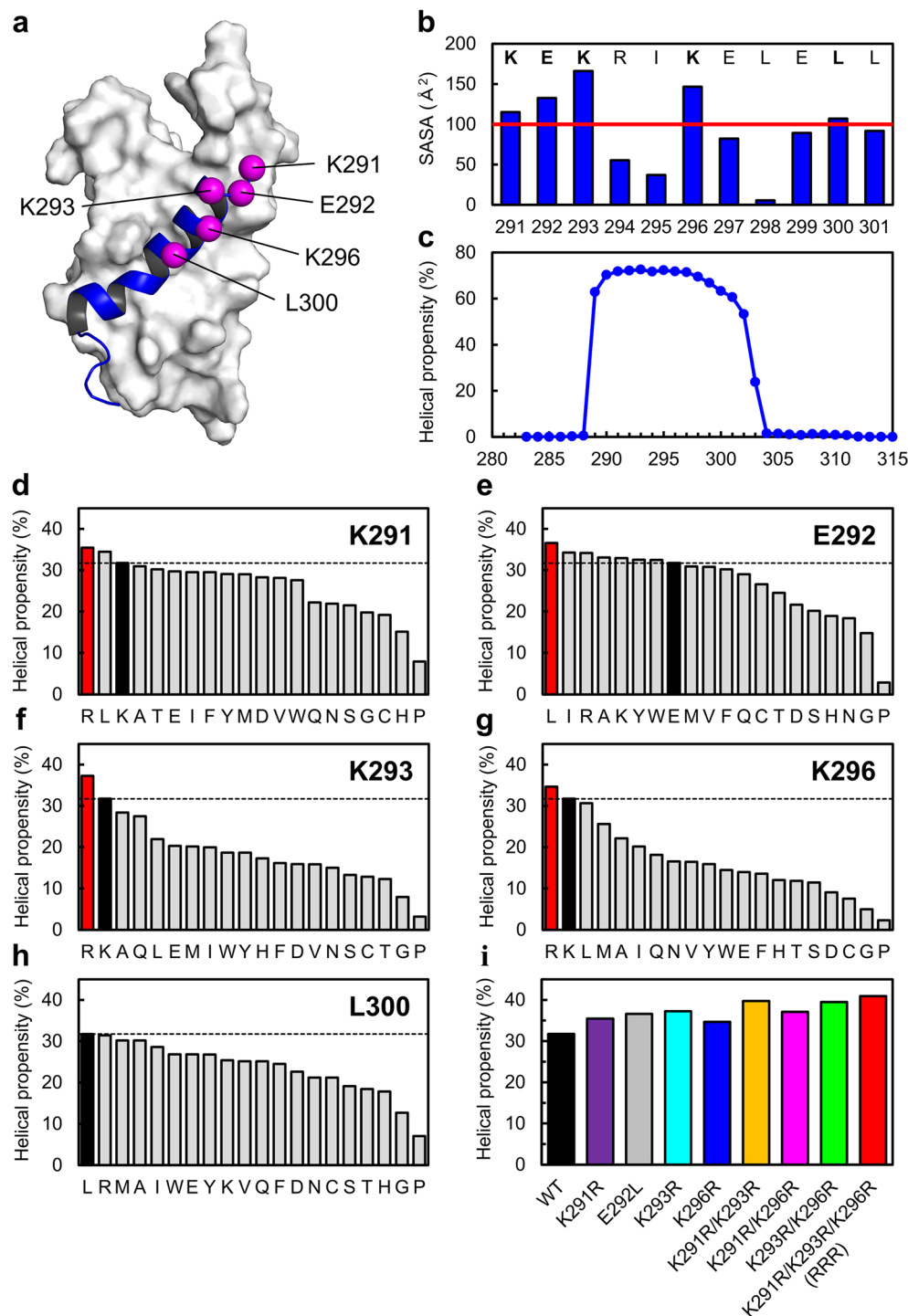
## Results

**Rational design of Myb32 mutants.** Although Myb32 is intrinsically disordered, previous studies have shown that it has a high helical propensity at the N-terminal region (residues 290–301) and is in equilibrium between a disordered and partially helical structure<sup>38</sup> (Fig. 1). Moreover, the N-terminal helical region of Myb32 selectively binds KIX via a conformational selection mechanism<sup>38</sup>. Since helix formation is a prerequisite for binding via the conformational selection mechanism, stabilizing the Myb32 helical structure should accelerate the apparent binding rate and increase the binding affinity to KIX. Based on this hypothesis, we rationally designed Myb32 mutants to stabilize their helical structure.

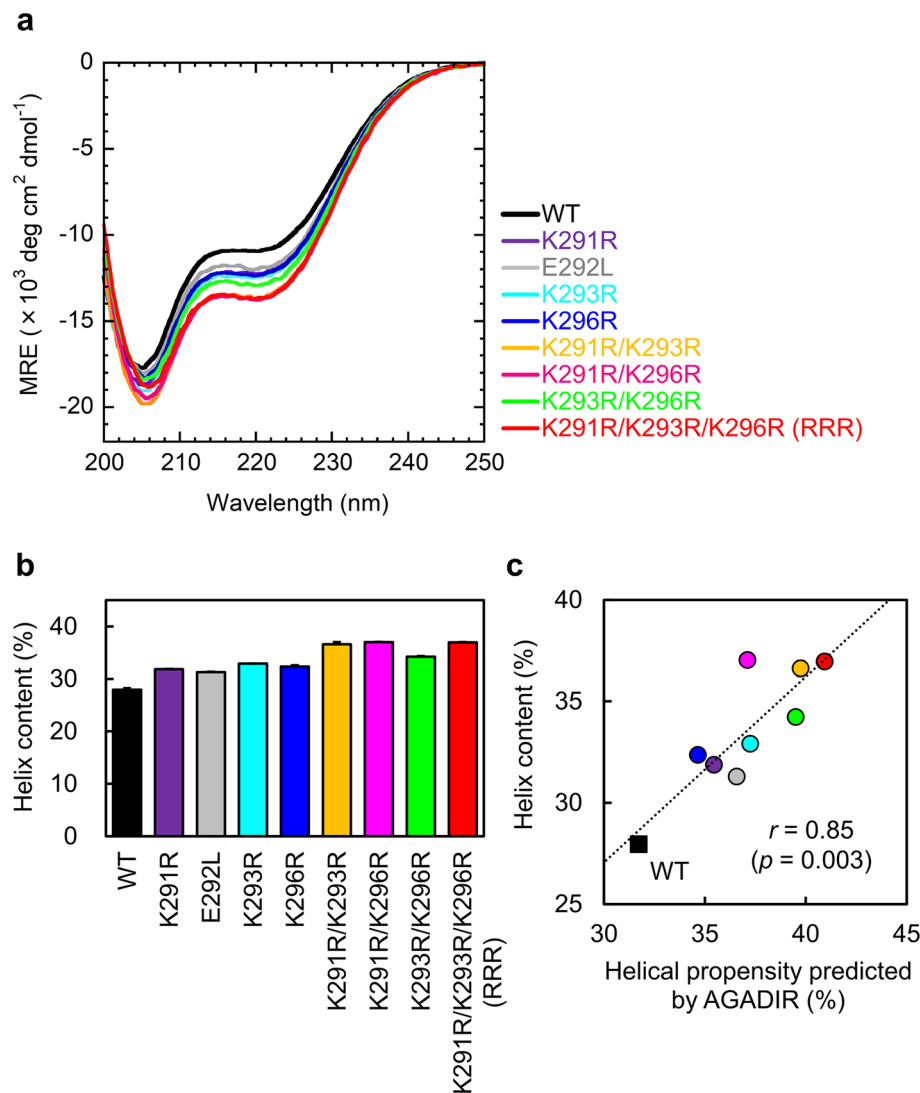
First, we selected mutation sites located on the helix surface opposite the KIX-binding interface to avoid interfering with the interactions between Myb32 and KIX. Using the c-Myb–KIX complex structure (PDB ID: 2AGH, model 1), we calculated the solvent-accessible surface area (SASA) of the residues involved in the N-terminal helix (residues 291–301) (Fig. 2a,b). Five residues (K291, E292, K293, K296, and L300) were highly exposed to solvent when c-Myb TAD was bound to KIX and considered candidates for the mutation sites.

Next, we predicted the helical propensity of the wild type and mutants of Myb32 using AGADIR<sup>41–43</sup>. AGADIR predicts helical propensities from amino acid sequences based on the helix/coil transition theory and has been used to design helical peptides<sup>44–47</sup>. AGADIR prediction showed that the helical propensity of wild-type Myb32 was 31.7% and that the N-terminal half of Myb32 had a high helical propensity (Fig. 2c). These results are consistent with the previous experimental results of the isolated Myb32 fragment<sup>38</sup>, indicating that AGADIR can correctly predict the helical propensity of c-Myb. Then, we calculated the helical propensities of the mutants in which 20 types of amino acid substitutions were introduced at the residues of the five solvent-exposed sites one at a time (Fig. 2d–h). The prediction showed that the wild-type residue had the highest helical propensity at residue L300. Similarly, for residues K291, K293, and K296, only one or two mutations were predicted to increase the helical propensity of the Myb32 peptide. Interestingly, the substitution of Lys into Arg was predicted to increase the helical propensity at these residues. Finally, at residue E292, substitution of Leu was predicted to best stabilize the helical structure. Therefore, we selected four mutations (K291R, E292L, K293R, and K296R) for experimental characterization.

**Helix contents of Myb32 mutants.** The wild type and four single mutants of Myb32 were overexpressed in *Escherichia coli* as a fusion with a His-tag and the B1 domain of protein G (GB1), both removed by thrombin cleavage during purification (see “Methods”). The peptides were purified to high homogeneity (Supplementary Figs. 1–2), and their molecular weights were in excellent agreement with the expected values (Supplementary Table 1). To estimate the helix contents of the designed peptides, we measured the far-ultraviolet (UV) circular dichroism (CD) spectra of the wild type and Myb32 mutants (Fig. 3a). The CD spectra showed a strong negative band at  $\sim 205$  nm and a small minimum at  $\sim 222$  nm, suggesting a partially helical structure. The helix contents were evaluated using the mean residue ellipticity (MRE) at 222 nm according to Eq. (2) (see “Methods” and Fig. 3b). The helix content of the wild-type Myb32 was 27.9%, consistent with the AGADIR prediction (Fig. 2). Compared with the wild type, all mutants increased the helix content to 31% or more. These results are well correlated with the AGADIR predictions (Fig. 3c); the correlation coefficient for the correlation between the helical propensity predicted by AGADIR and the helix content estimated from CD spectra was 0.85 ( $p = 0.003$ ).



**Figure 2.** Rational design of Myb32 mutants. **(a)** Mutation sites (magenta spheres) of Myb32 (blue) mapped onto the c-Myb–KIX complex structure. The figure was drawn using the PyMOL Molecular Graphics System, Version 2.4.0 Schrödinger, LLC. **(b)** Solvent accessible surface area (SASA) of the c-Myb residues in the c-Myb–KIX complex. **(c)** Residue-specific helical propensity of Myb32 predicted by AGADIR. The helical propensity of a whole peptide was obtained as an average of these values. **(d–h)** Helical propensities of the wild type (WT, black) and mutants (red or gray) of Myb32 predicted by AGADIR. For each site, mutations are sorted in the decreasing order of the helical propensity. **(i)** Helical propensities predicted by AGADIR for the single, double, and triple mutants of Myb32 used in the present study.



**Figure 3.** Helix contents of Myb32 mutants. **(a)** Far-ultraviolet (UV) circular dichroism (CD) spectra. **(b)** Helix contents estimated from MRE at 222 nm. The mean and standard error are shown. **(c)** Correlation between the helical propensity predicted by AGADIR and the helix content estimated from CD spectra. The correlation coefficient was 0.85 ( $p = 0.003$ ).

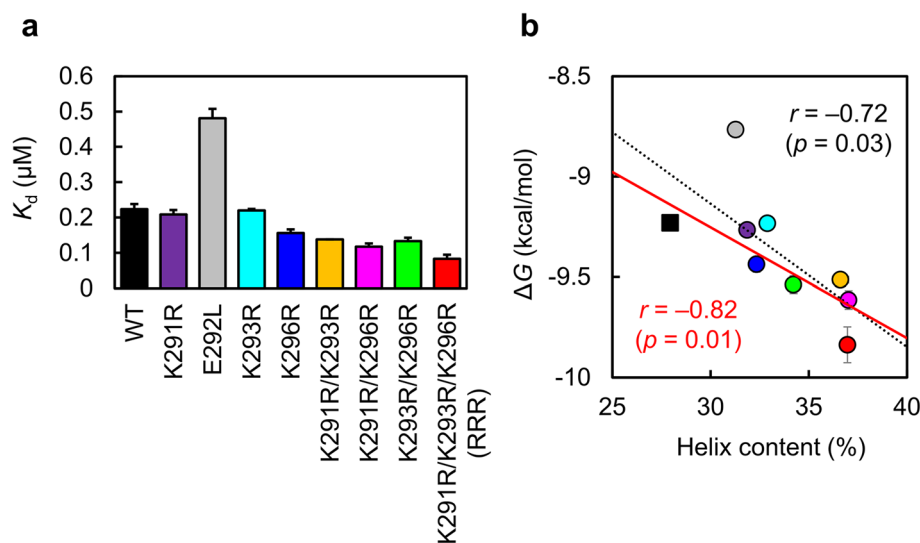
**KIX-binding affinity of Myb32 mutants.** To confirm that the designed peptides bind KIX, we measured the CD spectra of the mixture of KIX and one of the peptides (Supplementary Fig. 3). In all cases, the CD intensities of the mixture were larger than the summation of the CD intensities of KIX and the peptide alone, indicating that binding occurs between KIX and the designed peptides and induces helix formation.

Next, we performed isothermal titration calorimetry (ITC) measurements to investigate the KIX-binding affinity of the peptides. Although Myb32 primarily binds the c-Myb/pKID site on KIX, previous studies have shown that it also binds the MLL site, located opposite the c-Myb/pKID site<sup>38</sup>. Therefore, to prevent the secondary binding of Myb32, we conducted all ITC experiments in the presence of the MLL TAD peptide (residues 2842–2869, called MLL28) with a two-fold excess of MLL over KIX<sup>38</sup>.

ITC experiments showed that wild-type Myb32 binds KIX with a  $K_d$  of 0.22  $\mu\text{M}$ , consistent with a previous report<sup>38</sup> (Table 1; Fig. 4a and Supplementary Fig. 4). The KIX-binding affinity of the Lys-to-Arg mutants, which had higher helix contents than the wild-type Myb32, was similar to or higher than that of the wild type. The  $K_d$  values for K291R, K293R, and K296R were 0.21, 0.22, and 0.16  $\mu\text{M}$ , respectively. Unexpectedly, the E292L mutant showed a ~twofold decrease in affinity than the wild type ( $K_d = 0.48 \mu\text{M}$ ), although the mutation increased the helix content (Fig. 3). Because this mutation altered the polar surface into a hydrophobic one, it may interfere with the hydrophobic interactions present in the binding between KIX and Myb32, resulting in a decrease in the binding affinity. In accordance with this, E292L had a smaller enthalpy change ( $\Delta H$ ) upon KIX binding than the wild type, suggesting that the E292L mutation interfered with the intermolecular interactions between KIX and E292L.

Myb32	$K_d$ ( $\mu\text{M}$ )	$N$	$\Delta H$ ( $\text{kcal mol}^{-1}$ )	$-T\Delta S$ ( $\text{kcal mol}^{-1}$ )	$\Delta G$ ( $\text{kcal mol}^{-1}$ )
WT	$0.22 \pm 0.01$	$0.97 \pm 0.01$	$-6.49 \pm 0.05$	$-2.74 \pm 0.04$	$-9.23 \pm 0.04$
K291R	$0.21 \pm 0.01$	$0.97 \pm 0.01$	$-6.46 \pm 0.04$	$-2.80 \pm 0.01$	$-9.27 \pm 0.04$
E292L	$0.48 \pm 0.03$	$0.92 \pm 0.02$	$-5.45 \pm 0.03$	$-3.32 \pm 0.01$	$-8.77 \pm 0.04$
K293R	$0.22 \pm 0.01$	$0.94 \pm 0.01$	$-7.03 \pm 0.01$	$-2.21 \pm 0.02$	$-9.23 \pm 0.01$
K296R	$0.16 \pm 0.01$	$0.96 \pm 0.02$	$-6.3 \pm 0.1$	$-3.2 \pm 0.2$	$-9.44 \pm 0.04$
K291R/K293R	$0.14 \pm 0.01$	$0.84 \pm 0.02$	$-7.31 \pm 0.04$	$-2.21 \pm 0.03$	$-9.51 \pm 0.01$
K291R/K296R	$0.12 \pm 0.01$	$0.90 \pm 0.01$	$-7.56 \pm 0.03$	$-2.05 \pm 0.03$	$-9.62 \pm 0.05$
K293R/K296R	$0.13 \pm 0.01$	$0.96 \pm 0.01$	$-7.31 \pm 0.04$	$-2.23 \pm 0.08$	$-9.54 \pm 0.04$
RRR	$0.08 \pm 0.01$	$0.99 \pm 0.01$	$-7.57 \pm 0.03$	$-2.27 \pm 0.07$	$-9.84 \pm 0.09$
WT (SPR) <sup>a</sup>	$0.51 \pm 0.02$	$0.99 \pm 0.02$	$-13.79 \pm 0.03$	$5.07 \pm 0.05$	$-8.72 \pm 0.02$
RRR (SPR)	$0.26 \pm 0.01$	$0.99 \pm 0.01$	$-14.80 \pm 0.03$	$5.67 \pm 0.02$	$-9.14 \pm 0.01$

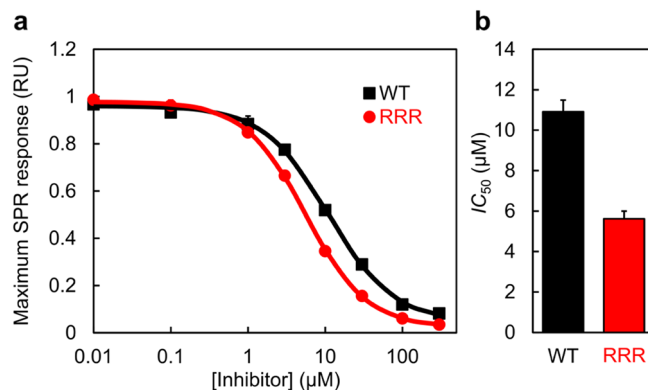
**Table 1.** Thermodynamic parameters for the KIX binding of Myb32 mutants. The dissociation constant ( $K_d$ ), stoichiometry of binding ( $N$ ), and changes in enthalpy ( $\Delta H$ ), entropy ( $\Delta S$ ), and Gibbs free energy ( $\Delta G$ ) upon binding were obtained by ITC measurements (20 mM Tris-acetate [pH 7.0] and 50 mM NaCl).  $T$  is the temperature. All measurements were performed more than twice, and the means and standard errors are shown. <sup>a</sup>The ITC measurement was performed under the conditions for SPR measurement (20 mM sodium phosphate [pH 7.0] and 300 mM NaCl).



**Figure 4.** KIX-binding affinity of Myb32 mutants. **(a)** Dissociation constants ( $K_d$ ) for the KIX binding of the wild type and mutants of Myb32 measured by isothermal titration calorimetry (ITC). **(b)** Correlation between the helix content estimated from CD spectra and the KIX-binding affinity represented as the Gibbs free energy change ( $\Delta G$ ) for the Myb32–KIX interaction. The correlation coefficient was  $-0.72$  ( $p = 0.03$ ) (black dotted line). When only the Lys-to-Arg mutants and the wild type were used (when the E292L mutant [gray] was excluded), the correlation coefficient was  $-0.82$  ( $p = 0.01$ ) (continuous red line).

**Combination of multiple mutations.** We combined two or three Lys-to-Arg substitutions (i.e., K291R, K293R, and K296R) to further increase the helical propensity and binding affinity. As a result, AGADIR predicted that double mutations (K291R/K293R, K291R/K296R, and K293R/K296R) increase the helical propensity of Myb32 compared to single mutations (Fig. 2i). Moreover, the triple mutant (K291R/K293R/K296R, denoted as RRR) was predicted to have the highest helical propensity (Fig. 2i). As expected, CD measurements showed that the double and triple mutants had helix contents higher than single mutants (Fig. 3). The helix contents were 36.6% for K291R/K293R, 37.0% for K291R/K296R, 34.2% for K293R/K296R, and 37.0% for RRR, respectively. Moreover, the CD spectra of the mixture of KIX and one of the double or triple mutants showed additional helix formation upon binding (Supplementary Fig. 5).

Interestingly, the KIX-binding affinity was also improved for the double and triple mutants compared to the single mutants (Table 1, Fig. 4a). Thus, the  $K_d$  values were  $0.14 \mu\text{M}$  for K291R/K293R,  $0.12 \mu\text{M}$  for K291R/K296R, and  $0.13 \mu\text{M}$  for K293R/K296R. In particular, the triple mutant, RRR, had a  $K_d$  of  $80 \text{ nM}$ , which is  $\sim$  threefold higher affinity to KIX than the wild-type Myb32.



**Figure 5.** Surface plasmon resonance (SPR) competitive binding assay. The wild-type Myb32 was immobilized on a sensor chip, and a mixture of KIX and various concentrations of an inhibitor was injected. **(a)** The maximum level of SPR response was plotted against the concentration of the wild-type Myb32 (black) and the designed RRR inhibitor (red). The inhibitor concentration is plotted in a logarithmic scale. The continuous lines were obtained by fitting to Eq. (5). The experiments were performed in triplicate, and the mean and standard error are shown. **(b)** The  $IC_{50}$  values. The errors are standard errors.

**Correlation between helical propensity and binding affinity.** In this study, we assumed that stabilizing the helical structure of the Myb32 peptide enhances its binding affinity to KIX. To check the validity of this assumption, we performed a correlation analysis between the helix content estimated from the CD spectra and the KIX-binding affinity measured by ITC (Fig. 4b). The KIX-binding affinity was represented as the Gibbs free energy change for the Myb32–KIX interaction calculated from a  $K_d$  value (see “Methods”). The correlation coefficient was  $-0.72$  ( $p=0.03$ ), indicating a good correlation between the helical propensity and binding affinity. Furthermore, when only the Lys-to-Arg mutants were selected, the correlation coefficient became  $-0.82$  ( $p=0.01$ ), showing an excellent correlation.

**Inhibition of c-Myb–KIX interaction.** Finally, to confirm that the designed inhibitor, that is, the RRR Myb32 mutant, can prevent the c-Myb–KIX interaction, we conducted a competitive binding assay using surface plasmon resonance (SPR). In this experiment, we immobilized wild-type Myb32 on a sensor chip and injected a mixture of KIX and various concentrations of the designed inhibitor. To reduce the non-specific binding of KIX to the sensor chip, we used the KIX protein without a His-tag and added high concentrations (300 mM) of sodium chloride (NaCl) during the measurements. Under this condition, the  $K_d$  for the KIX binding measured by ITC was  $0.51 \pm 0.02 \mu\text{M}$  for the wild-type Myb32 and  $0.26 \pm 0.01 \mu\text{M}$  for RRR, indicating that RRR has ~two-fold higher affinity with KIX than the wild type (Table 1; Supplementary Fig. 6).

At low concentrations of the RRR inhibitor, KIX was bound to the immobilized wild-type Myb32, resulting in a high SPR response (Fig. 5a). By increasing the concentration of the inhibitor, the SPR response levels gradually decreased. The maximum response level plotted against the inhibitor concentration gave a half-maximal inhibitory concentration ( $IC_{50}$ ) of  $5.6 \pm 0.4 \mu\text{M}$ , ~two-fold smaller than the value obtained when the wild-type Myb32 was injected in the same way ( $10.9 \pm 0.6 \mu\text{M}$ ) (Fig. 5b). Thus, these data show that our designed peptide, RRR, can effectively inhibit the c-Myb–KIX interaction. The ~two-fold higher inhibitory effect of RRR than the wild-type Myb32 is consistent with the difference in  $K_d$  for KIX binding under the same conditions (see above).

## Discussion

This study rationally designed a helical peptide inhibitor that binds KIX with high affinity using the Myb32 peptide as a template. According to the conformational selection mechanism, the intrinsically disordered Myb32 requires helix formation before binding to KIX. Thus, we assumed that stabilizing the helical structure enhanced the binding affinity to KIX. Our strategy is simple: introducing mutations opposite a binding interface to increase the helical propensity. The helical propensities of the Myb32 peptides were accurately predicted using AGADIR. As expected, there was a good correlation between the helical propensity and binding affinity, indicating that our strategy is effective and has a high success rate. Furthermore, since we selected mutation sites opposite the binding interface and introduced conservative Lys-to-Arg substitutions, the increase in the binding affinity is probably due to stabilizing the helical structure rather than forming new interactions with KIX. Although K291 of Myb32 has electrostatic interactions with KIX, the K291R mutation enhanced both the helix content and the binding affinity to KIX, suggesting that the conservative Lys-to-Arg substitution forms additional intramolecular interactions that stabilize the helical structure without affecting the intermolecular interactions present in the wild-type Myb32–KIX complex.

Our present results are consistent with previous reports on the mutational effects of the c-Myb–KIX interaction. For example, Poosapati et al. showed that the KIX-binding affinity was decreased by destabilizing the N-terminal helix of c-Myb TAD due to amino acid substitutions<sup>39</sup>. Moreover, Parker et al. showed that the L301P mutant, which destabilizes the helical structure of c-Myb TAD, significantly impaired KIX binding and target

gene activity, while the L301A mutation had little effects<sup>48</sup>, indicating that the loss of c-Myb activity is mainly due to the reduction in helical propensity.

In addition to our results, several studies have succeeded in increasing the binding affinity of IDPs to their partners by stabilizing their helical structures. Langlois et al. designed short peptides that mimic p53 TAD by introducing a capping motif and a Leu-Leu bridge to enhance the helical propensity and increase the binding affinity to its target<sup>49</sup>. Borchers et al. designed a p53 mutant (P12A/P13A/P27A) with a higher helical propensity than the wild type and achieved an increase in the binding affinity to MDM2<sup>50</sup>. Similarly, Iešmantavičius et al. introduced mutations into the activation domain of the activator for thyroid hormone and retinoid receptors and showed an increase in the affinity to the nuclear coactivator-binding domain of CBP<sup>51</sup>. In contrast, some examples do not show a correlation between the helical propensity and binding affinity. For example, Crabtree et al. showed that the P2846A/P2858A mutant of MLL TAD slightly increased the helical propensity but significantly decreased the binding affinity to KIX<sup>52</sup>, inconsistent with the conformational selection mechanism. Dahal et al. observed no correlation between the helical propensity and binding affinity for eight BH3 peptides<sup>53</sup>. Moreover, while many IDPs fold upon binding, some IDPs bind their targets without a well-defined structure, forming a “fuzzy complex”<sup>54</sup>. Therefore, it is not always possible to increase the binding affinity to a partner by stabilizing the helical structure of an IDP. However, our results suggest that stabilizing the helical structure effectively enhances binding affinity when binding via a conformational selection mechanism. Furthermore, it has been suggested that stabilizing helical structures in IDPs can shift the binding mechanism from induced-fit to conformational selection<sup>40</sup>. Thus, the binding affinity of IDPs may increase if the helical structure is sufficiently stabilized.

An interesting finding in this study is that conservative Lys-to-Arg substitutions and their combination increased the helical propensity of the Myb32 peptide, as predicted by AGADIR. Although both Lys and Arg have similar properties, Arg has a higher intrinsic helical propensity than Lys and is more likely to form a salt bridge with acidic residues<sup>55</sup>. Therefore, a Lys-to-Arg substitution in the helical region may be an effective way to stabilize helices, thereby enhancing PPIs involving the binding of helical structures.

Since many PPIs are mediated by  $\alpha$ -helices<sup>56–58</sup>, the use of helical fragments taken from target PPIs combined with Lys-to-Arg substitutions for helix stabilization can be a promising strategy for designing PPI inhibitors. Moreover, although there are many chemical methods for stabilizing helical structures, such as side-chain crosslinks and hydrogen bond surrogates<sup>59,60</sup>, conservative amino acid substitutions are the simplest ways to stabilize the helical structure. Stabilization of helical structures may also help resolve in vivo short half-lives of peptide-based drugs in addition to the use of D-amino acids for improving proteolytic resistance<sup>61</sup>. Future challenges will include the translocation of the RRR peptide into cells because the c-Myb-KIX interaction occurs inside cells. This may be accomplished by conjugating cell-penetrating peptides (CPPs)<sup>62</sup> and by making shorter peptides for efficient cellular uptake<sup>63,64</sup>. Recently, Ramaswamy et al. developed a peptidomimetic inhibitor of the c-Myb-KIX interaction, named MYBMIM, using a D-amino acid peptide corresponding to residues 293–310 of c-Myb fused with a CPP<sup>61</sup>. MYBMIM accumulated in the nucleus of acute myeloid leukemia (AML) cells and successfully impeded leukemia growth<sup>61</sup>. Since our RRR inhibitor has a higher affinity to KIX than MYBMIM, modified versions of the RRR peptide that are fused with a CPP and composed of D-amino acids might effectively work as an AML inhibitor. Although further studies will be required to solve these issues, our strategy will be useful for designing peptide inhibitors of PPIs involved in various diseases.

## Methods

**Theoretical predictions.** SASA of the c-Myb TAD residues in the complex with KIX was calculated using the GetArea server (<http://curie.utmb.edu/getarea.html>)<sup>65</sup>. Prediction of helical propensity was performed using the AGADIR server (<http://agadir.crg.es/>)<sup>41–43</sup>. In the AGADIR prediction, the Myb32 sequence with Gly at N-terminus was used as an input sequence, and pH, ionic strength, and temperature were set to 7.0, 0.05 M, and 298 K, respectively, to match them with the experimental conditions.

**Protein expression and purification.** The DNA fragments encoding Myb32 (residues 284–315 of mouse c-Myb; YNDEDPEKEKRIKELELLLMSTENELKGQQAL), MLL28 (residues 2842–2869 of human MLL), and KIX (residues 586–672 of mouse CBP) were constructed by overlap extension PCR. The amino-acid sequence of mouse KIX is the same as that of human KIX. In addition, the amino acid sequence of mouse Myb32 is the same as that of human Myb32 except that A314 in mouse Myb32 is V314 in human; however, residue 314 is not involved in the binding interface between KIX and Myb32, and the difference in this residue may not affect the binding affinity. The codons were optimized for high-level expression in *E. coli*<sup>66</sup>. The DNA fragment of Myb32 or MLL28 with a 6 × His-tag, GB1, and the thrombin cleavage site at the N-terminus was inserted into the NcoI and BamHI restriction sites of the pET-15b expression vector (Millipore Sigma, Burlington, MA, USA). A single Tyr residue was attached at the C-terminus of MLL28 for protein concentration determination by UV absorption. The DNA fragment of KIX with a 6 × His-tag and the thrombin cleavage site at the N-terminus (MGHHH-HHHSSGLVPR + G586 of KIX) was also inserted into the pET-15b vector. All mutations of Myb32 were introduced using the protocol of the QuikChange mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA).

For the expression of Myb32 and MLL28, *E. coli* BL21(DE3) competent cells were transformed with the corresponding plasmid and cultivated at 37 °C in 2 × YT liquid medium. When the cell culture reached an optical density at 600 nm of ~0.7, 1 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside was added to induce protein expression. After an additional 4 h of cultivation, the cells were harvested by centrifugation.

The *E. coli* cell pellet was resuspended in 35 mL of buffer HP (50 mM sodium phosphate [pH 7.4], 300 mM NaCl, 10 mM imidazole, and 6 M guanidine hydrochloride) and sonicated gently on ice using a Branson Sonifier 250D Advanced (Branson, Danbury, CT, USA). The lysate was centrifuged at 35,140 ×g for 1 h at 4 °C. The supernatant was filtered through a 0.45  $\mu$ m syringe filter (Millipore Sigma) and loaded onto a glass Econo-Column®

gravity column (50 mL; Bio-Rad Laboratories, Inc., Hercules, CA, USA) containing 5 mL of nickel-nitrilotriacetic acid (Ni-NTA) resin (Qiagen, Hilden, Germany). After the column was shaken at 35 rpm for 15 min at room temperature, it was washed with buffer HP three times and was sufficiently washed with buffer AP (50 mM sodium phosphate [pH 7.4], 300 mM NaCl, and 10 mM imidazole). The column was filled with 40 mL of buffer EP (50 mM sodium phosphate [pH 7.4], 300 mM NaCl, and 250 mM imidazole) and shaken at 35 rpm for 15 min at room temperature before elution to obtain the His-tagged Myb32 protein (His-tag-GB1-Myb32) for SPR measurement. On the other hand, to obtain the Myb32 or MLL28 protein without a His-tag and GB1, the Ni-NTA column, in which a protein was bound, was washed with 40 mL buffer TP (50 mM sodium phosphate [pH 7.4] and 300 mM NaCl) for three times to remove imidazole. The column was filled with 20 mL of buffer TP, and 100–150 units of thrombin were added. The column was shaken at 35 rpm for 16 h at room temperature to remove the His-tag and GB1 by thrombin cleavage. After cleavage, Myb32 contained an additional Gly residue at the N-terminus, resulting in 33-residue peptides. Myb32 or MLL28 was eluted from the column. The proteins were further purified by anion exchange chromatography using a HiPrep DEAE FF 16/10 column (20 mL; Cytiva, Marlborough, MA, USA) equilibrated with buffer A (10 mM sodium phosphate [pH 7.0]) for Myb32 and with buffer B (10 mM sodium phosphate [pH 6.0]) for MLL28. The proteins were eluted by gradually increasing the NaCl concentration from 0 to 0.2 M for Myb32 and from 0 to 0.3 M for MLL28 using the ÄKTAprime plus chromatography system (Cytiva).

Expression and purification of the KIX protein with or without a His-tag were performed as described above for Myb32 and MLL28, except that *E. coli* BL21(DE3)pLysS competent cells were used. In addition, instead of anion exchange chromatography, size exclusion chromatography was performed using a Superdex 200 pg column (Cytiva) and the ITC buffer (20 mM Tris-acetate [pH 7.0] and 50 mM NaCl).

After purification, the purity of all proteins was determined by sodium dodecyl sulfate–polyacrylamide gel electrophoresis and analytical reverse-phase high-performance liquid chromatography (HPLC). The reverse-phase HPLC experiments were performed on a LC-20AP HPLC system (Shimadzu, Kyoto, Japan) using a TSKgel Protein C<sub>4</sub>-300 column (C4 column, 4.6 × 150 mm; Tosoh, Tokyo, Japan) equilibrated with 90% of solvent A containing 0.05% (v/v) trifluoroacetic acid (TFA) in water and 10% of solvent B containing 0.05% (v/v) TFA in acetonitrile. The bound peptides were eluted by a linear gradient of solvent B from 10 to 80% at a flow rate of 1 mL/min for 45 min. The peak fractions were collected to determine the molecular weights by electrospray ionization mass spectrometry using micrOTOF-Q II (Bruker, Yokohama, Japan). All proteins were concentrated using an Amicon® Ultra-4 centrifugal filter (MWCO 3000; Millipore Sigma) and dialyzed with ITC buffer before measurements.

**CD measurements.** Far-UV CD spectra were measured using a J-805 spectropolarimeter (JASCO, Tokyo, Japan) from 200 to 250 nm in a quartz cuvette with a 1-mm path length at 25 °C. The scan speed was 50 nm/min, and a response time was 1 s. Samples contained 50 μM Myb32 protein with or without 50 μM KIX dissolved in ITC buffer. Raw CD data (mdeg) were converted into MRE using the following equation:

$$MRE \text{ (deg cm}^2 \text{ dmol}^{-1}\text{)} = \frac{CD \text{ (mdeg)}}{N \cdot c \text{ (M)} \cdot l \text{ (mm)}} \quad (1)$$

where  $N$  is the number of residues,  $c$  is the protein concentration, and  $l$  is the path length. The helix contents ( $f_H$ ) (%) of Myb32 peptides were evaluated using MRE at 222 nm according to the following equation<sup>67</sup>:

$$f_H = -(MRE_{222} + 2,340)/30,300 \times 100 \quad (2)$$

**ITC measurements.** ITC measurements were conducted using a MicroCal iTC200 (Malvern Panalytical Ltd., Malvern, United Kingdom) in ITC buffer at 30 °C. All samples were dialyzed sufficiently before the experiments. 600 μM Myb32 was titrated into the cell containing 45 μM KIX and 90 μM MLL28. The data were analyzed using Origin software (OriginLab Co., Northampton, MA, USA) to obtain the  $K_d$ , stoichiometry of binding ( $N$ ), and enthalpy change upon binding ( $\Delta H$ ). The Gibbs free energy change ( $\Delta G$ ) and entropy change ( $\Delta S$ ) upon binding were obtained as follows:

$$\Delta G = -RT \ln (1/K_d) \quad (3)$$

$$\Delta S = \frac{\Delta H - \Delta G}{T} \quad (4)$$

where  $R$  is the gas constant, and  $T$  is the temperature.

**SPR competitive binding assay.** The SPR competitive binding assay was conducted with OpenSPR (Nicoya, Ontario, Canada), which uses localized surface plasmon resonance. Ligand immobilization was performed using a buffer containing 20 mM sodium phosphate (pH 7.0) and 100 mM NaCl. First, 10 mM HCl was injected at 150 μL/min to clean the surface of the sensor chip for amine coupling (a high sensitivity carboxyl sensor, Nicoya). Next, a mixture of *N*-hydroxysuccinimide and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (included in the amine coupling kit, Nicoya) was injected at 20 μL/min for surface activation. Then, 0.05 mg/mL of the Myb32 protein with a His-tag and GB1, which was dissolved in activation buffer (Nicoya), was injected at 20 μL/min. Finally, blocking solution (Nicoya) was injected to deactivate the remaining active carboxyl groups on the sensor chip.



The competitive binding assay was conducted in 20 mM sodium phosphate (pH 7.0) and 300 mM NaCl at 30 °C. The buffer conditions were optimized to reduce the non-specific binding. As an analyte, we injected 1 μM KIX without a His-tag or a mixture of 1 μM KIX and various concentrations of the RRR inhibitor (from 1 nM to 300 μM) at a flow rate of 50 μL/min. After every injection, 10 mM HCl was injected for regeneration. The maximum value of the SPR response, which was obtained after subtracting the value in the absence of the ligand, was plotted against the concentration of the inhibitor and fitted with the four-parameter logistic curve:

$$y_{\text{obs}} = y_{\text{min}} + \frac{y_{\text{max}} - y_{\text{min}}}{1 + \left(\frac{[\text{Inhibitor}]}{IC_{50}}\right)^H} \quad (5)$$

where  $y_{\text{max}}$  and  $y_{\text{min}}$  are the maximum and minimum values of the titration curve, [Inhibitor] is the concentration of the inhibitor,  $IC_{50}$  is the half-maximal inhibitory concentration, and  $H$  is the Hill coefficient.

**Statistics and reproducibility.** CD spectra were measured three times with 10 scans each, and the averaged spectra were shown. Number of replicates in ITC, SPR, and mass spectrometry measurements are reported in the figure and table captions. Error bars denote the standard errors. Statistical analyses were performed using Microsoft Excel. Data fitting was conducted using KaleidaGraph 4.1 software (HULINKS).

## Data availability

The authors declare that all data supporting the findings of this study are available within the paper and its supplementary information file.

Received: 6 August 2021; Accepted: 22 December 2021

Published online: 20 January 2022

## References

- Wright, P. E. & Dyson, H. J. Intrinsically unstructured proteins: Re-assessing the protein structure–function paradigm. *J. Mol. Biol.* **293**, 321–331 (1999).
- Dunker, A. K. *et al.* Intrinsically disordered protein. *J. Mol. Graph. Model.* **19**, 26–59 (2001).
- Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell. Biol.* **6**, 197–208 (2005).
- Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell. Biol.* **16**, 18–29 (2015).
- Uversky, V. N., Oldfield, C. J. & Dunker, A. K. Intrinsically disordered proteins in human diseases: Introducing the D2 concept. *Annu. Rev. Biophys.* **37**, 215–246 (2008).
- Greig, K. T., Carotta, S. & Nutt, S. L. Critical roles for c-Myb in hematopoietic progenitor cells. *Semin. Immunol.* **20**, 247–256 (2008).
- Ramsay, R. G. & Gonda, T. J. MYB function in normal and cancer cells. *Nat. Rev. Cancer* **8**, 523–534 (2008).
- Wang, X., Angelis, N. & Thein, S. L. MYB: A regulatory factor in hematopoiesis. *Gene* **665**, 6–17 (2018).
- Oh, I. H. & Reddy, E. P. The myb gene family in cell growth, differentiation and apoptosis. *Oncogene* **18**, 3017–3033 (1999).
- Biroccio, A. *et al.* c-Myb and Bcl-x overexpression predicts poor prognosis in colorectal cancer: Clinical and experimental findings. *Am. J. Pathol.* **158**, 1289–1299 (2001).
- Drabsch, Y. *et al.* Mechanism of and requirement for estrogen-regulated MYB expression in estrogen-receptor-positive breast cancer cells. *Proc. Natl. Acad. Sci. USA* **104**, 13762–13767 (2007).
- Miao, R. Y. *et al.* MYB is essential for mammary tumorigenesis. *Cancer Res.* **71**, 7029–7037 (2011).
- Zuber, J. *et al.* An integrated approach to dissecting oncogene addiction implicates a Myb-coordinated self-renewal program as essential for leukemia maintenance. *Genes Dev.* **25**, 1628–1640 (2011).
- Pattabiraman, D. R. & Gonda, T. J. Role and potential for therapeutic targeting of MYB in leukemia. *Leukemia* **27**, 269–277 (2013).
- Liu, X., Gold, K. A. & Dmitrovsky, E. The Myb-p300 interaction is a novel molecular pharmacologic target. *Mol. Cancer Ther.* **14**, 1273–1275 (2015).
- Mitra, P. Transcription regulation of MYB: A potential and novel therapeutic target in cancer. *Ann. Transl. Med.* **6**, 443 (2018).
- Ciciro, Y. & Sala, A. MYB oncoproteins: Emerging players and potential therapeutic targets in human cancer. *Oncogenesis* **10**, 19 (2021).
- Best, J. L. *et al.* Identification of small-molecule antagonists that inhibit an activator: coactivator interaction. *Proc. Natl. Acad. Sci. USA* **101**, 17622–17627 (2004).
- Li, B. X. & Xiao, X. Discovery of a small-molecule inhibitor of the KIX-KID interaction. *ChemBioChem* **10**, 2721–2724 (2009).
- Uttarkar, S. *et al.* Targeting acute myeloid leukemia with a small molecule inhibitor of the Myb/p300 interaction. *Blood* **127**, 1173–1182 (2016).
- Lu, H. *et al.* Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials. *Signal Transduct. Target. Ther.* **5**, 213. <https://doi.org/10.1038/s41392-020-00315-3> (2020).
- Tsomaia, N. Peptide therapeutics: Targeting the undruggable space. *Eur. J. Med. Chem.* **94**, 459–470 (2015).
- Lee, A. C., Harris, J. L., Khanna, K. K. & Hong, J. H. A comprehensive review on current advances in peptide drug development and design. *Int. J. Mol. Sci.* **20**, 2383. <https://doi.org/10.3390/ijms20102383> (2019).
- Sakura, H. *et al.* Delineation of three functional domains of the transcriptional activator encoded by the c-myb protooncogene. *Proc. Natl. Acad. Sci. USA* **86**, 5758–5762 (1989).
- Ogata, K. *et al.* Solution structure of a DNA-binding unit of Myb: A helix–turn–helix-related motif with conserved tryptophans forming a hydrophobic core. *Proc. Natl. Acad. Sci. USA* **89**, 6428–6432 (1992).
- Aziz, N. *et al.* Modulation of c-Myb-induced transcription activation by a phosphorylation site near the negative regulatory domain. *Proc. Natl. Acad. Sci. USA* **92**, 6429–6433 (1995).
- Bies, J. & Wolff, L. Oncogenic activation of c-Myb by carboxyl-terminal truncation leads to decreased proteolysis by the ubiquitin–26S proteasome pathway. *Oncogene* **14**, 203–212 (1997).
- Bies, J., Markus, J. & Wolff, L. Covalent attachment of the SUMO-1 protein to the negative regulatory domain of the c-Myb transcription factor modifies its stability and transactivation capacity. *J. Biol. Chem.* **277**, 8999–9009 (2002).
- Radhakrishnan, I. *et al.* Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: A model for activator:coactivator interactions. *Cell* **91**, 741–752 (1997).

30. Zor, T., De Guzman, R. N., Dyson, H. J. & Wright, P. E. Solution structure of the KIX domain of CBP bound to the transactivation domain of c-Myb. *J. Mol. Biol.* **337**, 521–534 (2004).
31. Lee, C. W., Arai, M., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP. *Biochemistry* **48**, 2115–2124 (2009).
32. Van Orden, K., Giebler, H. A., Lemasson, I., Gonzales, M. & Nyborg, J. K. Binding of p53 to the KIX domain of CREB binding protein: A potential link to human T-cell leukemia virus, type I-associated leukemogenesis. *J. Biol. Chem.* **274**, 26321–26328 (1999).
33. Campbell, K. M. & Lumb, K. J. Structurally distinct modes of recognition of the KIX domain of CBP by Jun and CREB. *Biochemistry* **41**, 13956–13964 (2002).
34. Vendel, A. C. & Lumb, K. J. Molecular recognition of the human coactivator CBP by the HIV-1 transcriptional activator Tat. *Biochemistry* **42**, 910–916 (2003).
35. De Guzman, R. N., Goto, N. K., Dyson, H. J. & Wright, P. E. Structural basis for cooperative transcription factor binding to the CBP coactivator. *J. Mol. Biol.* **355**, 1005–1013 (2006).
36. Yang, K. *et al.* Structural basis for cooperative regulation of KIX-mediated transcription pathways by the HTLV-1 HBZ activation domain. *Proc. Natl. Acad. Sci. USA* **115**, 10040–10045 (2018).
37. Zor, T., Mayr, B. M., Dyson, H. J., Montminy, M. R. & Wright, P. E. Roles of phosphorylation and helix propensity in the binding of the KIX domain of CREB-binding protein by constitutive (c-Myb) and inducible (CREB) activators. *J. Biol. Chem.* **277**, 42241–42248 (2002).
38. Arai, M., Sugase, K., Dyson, H. J. & Wright, P. E. Conformational propensities of intrinsically disordered proteins influence the mechanism of binding and folding. *Proc. Natl. Acad. Sci. USA* **112**, 9614–9619 (2015).
39. Poosapati, A., Gregory, E., Borchers, W. M., Chemes, L. B. & Daughdrill, G. W. Uncoupling the folding and binding of an intrinsically disordered protein. *J. Mol. Biol.* **430**, 2389–2402 (2018).
40. Arai, M. Unified understanding of folding and binding mechanisms of globular and intrinsically disordered proteins. *Biophys. Rev.* **10**, 163–181 (2018).
41. Muñoz, V. & Serrano, L. Elucidating the folding problem of helical peptides using empirical parameters. *Nat. Struct. Biol.* **1**, 399–409 (1994).
42. Muñoz, V. & Serrano, L. Elucidating the folding problem of helical peptides using empirical parameters II. Helix macrodipole effects and rational modification of the helical content of natural peptides. *J. Mol. Biol.* **245**, 275–296 (1995).
43. Lacroix, E., Viguera, A. R. & Serrano, L. Elucidating the folding problem of  $\alpha$ -helices: Local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters. *J. Mol. Biol.* **284**, 173–191 (1998).
44. Petukhov, M. *et al.* Design of stable  $\alpha$ -helices using global sequence optimization. *J. Pept. Sci.* **15**, 359–365 (2009).
45. Garcia-Aranda, M. I. *et al.* Helical peptides from VEGF and Vammin hotspots for modulating the VEGF-VEGFR interaction. *Org. Biomol. Chem.* **11**, 1896–1905 (2013).
46. Bonache, M. A. *et al.* De novo designed library of linear helical peptides: An exploratory tool in the discovery of protein-protein interaction modulators. *ACS Comb. Sci.* **16**, 250–258 (2014).
47. Karoyan, P. *et al.* Human ACE2 peptide-mimics block SARS-CoV-2 pulmonary cells infection. *Commun. Biol.* **4**, 197. <https://doi.org/10.1038/s42003-021-01736-8> (2021).
48. Parker, D. *et al.* Role of secondary structure in discrimination between constitutive and inducible activators. *Mol. Cell. Biol.* **19**, 5601–5607 (1999).
49. Langlois, C. *et al.* Structure-based design of a potent artificial transactivation domain based on p53. *J. Am. Chem. Soc.* **134**, 1715–1723 (2012).
50. Borchers, W. *et al.* Disorder and residual helicity alter p53-Mdm2 binding affinity and signaling in cells. *Nat. Chem. Biol.* **10**, 1000–1002 (2014).
51. Iešmantavičius, V., Dogan, J., Jemth, P., Teilum, K. & Kjaergaard, M. Helical propensity in an intrinsically disordered protein accelerates ligand binding. *Angew. Chem. Int. Ed. Engl.* **53**, 1548–1551 (2014).
52. Crabtree, M. D. *et al.* Conserved helix-flanking prolines modulate intrinsically disordered protein: Target affinity by altering the lifetime of the bound complex. *Biochemistry* **56**, 2379–2384 (2017).
53. Dahal, L., Kwan, T. O. C., Hollins, J. J. & Clarke, J. Promiscuous and selective: How intrinsically disordered BH3 proteins interact with their pro-survival partner MCL-1. *J. Mol. Biol.* **430**, 2468–2477 (2018).
54. Tompa, P. & Fuxreiter, M. Fuzzy complexes: Polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* **33**, 2–8 (2008).
55. Wolny, M. *et al.* Characterization of long and stable de novo single alpha-helix domains provides novel insight into their stability. *Sci. Rep.* **7**, 44341. <https://doi.org/10.1038/srep44341> (2017).
56. Mohan, A. *et al.* Analysis of molecular recognition features (MoRFs). *J. Mol. Biol.* **362**, 1043–1059 (2006).
57. Guharoy, M. & Chakrabarti, P. Secondary structure based analysis and classification of biological interfaces: Identification of binding motifs in protein-protein interactions. *Bioinformatics* **23**, 1909–1918 (2007).
58. Jochim, A. L. & Arora, P. S. Assessment of helical interfaces in protein-protein interactions. *Mol. Biosyst.* **5**, 924–926 (2009).
59. Klein, M. A. Stabilized helical peptides: A strategy to target protein-protein interactions. *ACS Med. Chem. Lett.* **5**, 838–839 (2014).
60. Klein, M. Stabilized helical peptides: Overview of the technologies and its impact on drug discovery. *Expert Opin. Drug Discov.* **12**, 1117–1125 (2017).
61. Ramaswamy, K. *et al.* Peptidomimetic blockade of MYB in acute myeloid leukemia. *Nat. Commun.* **9**, 110. <https://doi.org/10.1038/s41467-017-02618-6> (2018).
62. Wang, F. *et al.* Recent progress of cell-penetrating peptides as new carriers for intracellular cargo delivery. *J. Control. Release* **174**, 126–136 (2014).
63. Yang, N. J. & Hinner, M. J. Getting across the cell membrane: An overview for small molecules, peptides, and proteins. *Methods Mol. Biol.* **1266**, 29–53 (2015).
64. Mattsson, P. & Kihlberg, J. How big is too big for cell permeability?. *J. Med. Chem.* **60**, 1662–1664 (2017).
65. Fraczkiewicz, R. & Braun, W. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comput. Chem.* **19**, 319–333 (1998).
66. Hoover, D. M. & Lubkowsky, J. DNAWorks: An automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res.* **30**, e43 (2002).
67. Chen, Y. H., Yang, J. T. & Martinez, H. M. Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion. *Biochemistry* **11**, 4120–4131 (1972).

## Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers JP16H02217, JP19H02521, and JP21K18841 Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (Y.H. and M.A.).

### Author contributions

S.S. and M.A. conceived the ideas and designed the study. S.S., Y.O., and M.A. conducted theoretical predictions. S.S., Y.O., T.K., and Y.H. conducted the experiments. S.S. and M.A. performed data analysis. S.S. and M.A. wrote the manuscript. All authors approved the final manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-04497-w>.

**Correspondence** and requests for materials should be addressed to M.A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022