Metal ions binding to recA inteins from *Mycobacterium tuberculosis*†

Liyun Zhang,[a] Yuchuan Zheng,[a] Zhaoyong Xi,[a] Zhaofeng Luo,[b] Xiaolong Xu,[a] Chunyu Wang[c] and Yangzhong Liu*[a]

Received 16th February 2009, Accepted 24th March 2009
First published as an Advance Article on the web 21st April 2009
DOI: 10.1039/b903144h

Zinc has been found in the crystal structures of inteins and the zinc ion can inhibit intein splicing both *in vitro* and *in vivo*. The interactions between metal ions and three minimized recA inteins have been studied in this work. Isothermal titration calorimetry (ITC) results show that the zinc binding affinity to three inteins is in the order of ΔI-SM > $\Delta\Delta I_{hh}$-SM ~ $\Delta\Delta I_{hh}$-CM, but is much weaker than to EDTA. These data explain the reversible inhibition and the presence of zinc only in the crystal structure of ΔI-SM of recA intein. A positive correlation between binding constants and inhibition efficiency was observed upon the titration of different metal ions. Single-site binding modes were detected in all interactions, except $\Delta\Delta I_{hh}$-CM which has two Zn sites. Zinc binding sites on $\Delta\Delta I_{hh}$-CM were analyzed by NMR spectroscopy and ITC titration on inteins with chemical modifications. Results indicate that the Cys1 and His73 are the second zinc binding sites in $\Delta\Delta I_{hh}$-CM. CD studies show the metal coordinations have negligible influence on protein structure. This work suggests that the mobility restriction of key residues from metal coordination is likely the key cause of metal inhibition of intein splicing.

Introduction

Self-catalytic protein splicing is a post translational process in which the intervening protein, intein, is cleaved from the precursor proteins with the concomitant ligation of flanking sequences.[1–4] Since its discovery in 1990,[5] intein has been widely applied in biotechnology and provides a novel tool in protein engineering.[3,6] Inteins are composed of 134 to 608 amino acids, and they are found in all three domains of life: eukaryotes, bacteria, and archaea.[7] Over 400 inteins have been identified to date;[8] most inteins possess two functional domains, the homing endonuclease domain and the splicing domain. The homing endonuclease domain, which recognizes and acts on host genes, normally inserts into the splicing domain between two sub-domains.[8] These two domains are functionally independent, and evidence from minimized inteins indicates that the splicing function is not prevented by the removal of the homing endonuclease domain.[9,10] Thus, such minimized inteins, composed of solely the splicing domain, are typically employed for studying the mechanism of protein splicing.

Zinc ions were found to interact with residues in the catalytic core in several crystal structures of intein.[11–13] Splicing activity studies showed that zinc reversibly inhibited protein splicing, while EDTA can relieve such inhibition.[14,15] In addition, several other divalent metal ions, such as Ni^{2+} and Co^{2+}, demonstrated weaker inhibitory activity than Zn^{2+} on protein splicing.[15] Nearly no inhibition was detected for Ca^{2+} and Mg^{2+}.[14,15] These observations indicate the inhibition efficiency is highly dependent on the properties of the metal ions. Thus it is of great interest to explore the mechanism of metal–intein interaction and their correlation with inhibition of protein splicing.

The three intein mutants studied in this paper, ΔI-SM, $\Delta\Delta I_{hh}$-SM and $\Delta\Delta I_{hh}$-CM, have been derived from the *Mycobacterium tuberculosis* recA intein. The wild-type recA intein consists of 440 amino acid residues. The deletion of the central endonuclease domain creates a minimized intein with 168 amino acid residues (ΔI).[16,17] This minimized intein has lower splicing activity than the full length protein. A V67L mutation can enhance the activity of the minimized intein and results in a splicing mutant ΔI-SM (see Scheme 1). Additionally, the D422G mutation could promote C-terminal cleavage activity with enhanced pH dependence, which results in the cleavage mutant (ΔI-CM).[16,17] It was observed in the crystal structure of ΔI-SM that residues in the linker region between the two sub-domains were disordered.[13] By replacing the 36 residues in the disordered region with a seven-amino-acid β-turn sequence (VRDVETG) from hedgehog proteins, a further minimized intein $\Delta\Delta I_{hh}$-SM was created. This 139-residue intein has comparable splicing activity to that of ΔI-SM.[17] $\Delta\Delta I_{hh}$-CM, generated in a similar way with 139 residues, also retains the cleavage properties of ΔI-CM (see Scheme 1).

The zinc coordination in the crystal structure of ΔI-SM is shown in Fig. 1.[13] Two zinc atoms are located at the interface of two protein molecules and each zinc binds to four residues from two different proteins, one His, one Glu and one Asn from one protein, and one His from the other protein.

[a] *Department of Chemistry, University of Science and Technology of China, Hefei, Anhui, China. E-mail: liuyz@ustc.edu.cn; Fax: +86-551-3600874; Tel: +86-551-3600904*
[b] *School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, China*
[c] *Department of Biology, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, New York, USA*

† Electronic supplementary information (ESI) available: Thermodynamic parameters, results from equilibrium dialysis and DLS measurements. Link to PDB visualization. See DOI: 10.1039/b903144h


```
                10        20        30        40        50        60        70
I       CLAEGTRIFD PVTGTTHRIE DVVDGRKPIH VVAAAKDGTL HARPVVSWFD QGTRDVIGLR IAGGAIVWAT
ΔI-SM   CLAEGTRIFD PVTGTTHRIE DVVDGRKPIH VVAAAKDGTL HARPVVSWFD QGTRDVIGLR IAGGAILWAT
ΔΔIhh-SM CLAEGTRIFD PVTGTTHRIE DVVGGRKPIH VVAAAKDGTL HARPVVSWFD QGTRDVIGLR IAGGAILWAT
ΔΔIhh-CM CLAEGTRIFD PVTGTTHRIE DVVDGRKPIH VVAAAKDGTL HARPVVSWFD QGTRDVIGLR IAGGAILWAT
          A-block                                                   B-block

                80        90       100       110          383          400
I       PDHKVLTEYG WRAAGELRKG DRVAQPRRFD GFGDSAPIPA Endonuclease RVQALADA LDDKFLHDML
ΔI-SM   PDHKVLTEYG WRAAGELRKG DRVAQPRRFD GFGDSAPIPA ..........  RVQALADA LDDKFLHDML
ΔΔIhh-SM PDHKVLTEYG WRAAGELRKG DRVA..... .......... ..VRDVETG... ........ ..........
ΔΔIhh-CM PDHKVLTEYG WRAAGELRKG DRVA..... .......... ..VRDVETG... ........ ..........
          B-block

                410       420       430       440
I       AEELRYSVIR EVLPTRRART FDLEVEELHT LVAEGVVVHN
I-SM    AEELRYSVIR EVLPTRRART FDLEVEELHT LVAEGVVVHN
ΔΔIhh-SM ..ELRYSVIR EVLPTRRART FDLEVEELHT LVAEGVVVHN
ΔΔIhh-CM ..ELRYSVIR EVLPTRRART FGLEVEELHT LVAEGVVVHN
                    F-block     G-block
```

Scheme 1 Protein sequences of the three minimized inteins used in this work, ΔI-SM, $\Delta\Delta I_{hh}$-SM and $\Delta\Delta I_{hh}$-CM. The bold letters denote the conserved residues in the active core, Cys1, His73, Asp422, His429, His439 and Asn440. The two crucial mutation sites, V67 and D422 are highlighted with a gray background.

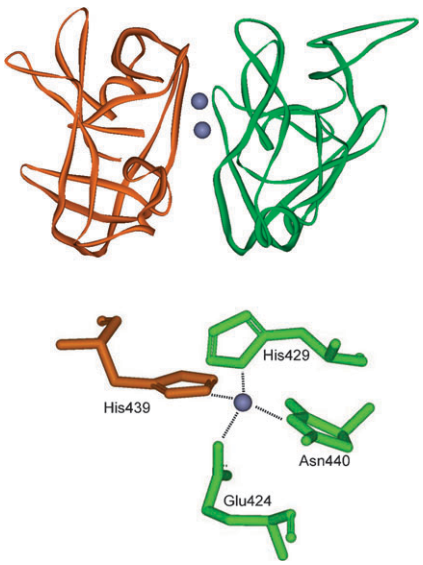

Fig. 1 Zinc coordination in the crystal structure of ΔI-SM (PDB: 2IMZ)†. (A) Ribbon representation of the protein. The two protein molecules are shown in different colors and the zinc atoms are shown as blue spheres. (B) Zn binding core in the crystal structure. Asn440 cyclized and formed succinimide in the crystal.

Such zinc coordination induces a pseudo-dimer with 1 : 1 stoichiometry for the ΔI-SM mutant in the asymmetric unit (Fig. 1). Three of the binding residues are highly conserved in the intein sequences, C-terminal residue (Asn440), penultimate histidine (His439) and F-block histidine (His429). Interestingly, zinc ions were present in the crystal structure of ΔI-SM, but not in the crystal structures of $\Delta\Delta I_{hh}$-SM and $\Delta\Delta I_{hh}$-CM.

Isothermal titration calorimetry (ITC) is an efficient technique for the determination of thermodynamic parameters of ligand binding to biomolecules.[18] The measurement detects the heat released or absorbed during the binding process, so that the enthalpy change (ΔH) and binding affinity (K) can be obtained directly, from which the free energy change (ΔG) and entropy change (ΔS) can be derived. This technique has also been used for studying the interactions of metal ions with proteins, which could quantify the stoichiometry and coordination constants of each binding step.[19]

Interactions of five metal ions, Zn^{2+}, Ni^{2+}, Co^{2+}, Mn^{2+}, and Cd^{2+}, with three intein mutants have been studied in this work using ITC. The binding sites were analyzed by NMR spectroscopy and ITC titration of the chemically modified proteins. The coordination constants were measured, and the correlation between binding affinities and the different inhibition efficiencies of metal ions was observed. Data suggested that the metal coordination to residues in the active core is the key cause of the inhibition of protein splicing, probably due to the immobilization of key residues.

Results

Thermodynamic studies of zinc binding to inteins

The interactions of the zinc ion with three intein mutants were measured by ITC at pH 7.4 in 20 mM buffer of Tris-HCl or MOPS. The representative ITC data in Fig. 2 demonstrate the isotherms of Zn binding to the three inteins. While similar profiles are seen for the interactions of ΔI-SM and $\Delta\Delta I_{hh}$-SM (Fig. 2A and B), $\Delta\Delta I_{hh}$-CM shows a rather different titration pattern (Fig. 2C). The least-square fitting of titration data resulted in a single-site binding model for ΔI-SM and $\Delta\Delta I_{hh}$-SM, however, the titration to $\Delta\Delta I_{hh}$-CM was fit with a two-site binding model (Table 1). This result indicates that an additional zinc binding site appears in the cleavage intein mutant (CM) relative to the splicing mutant (SM). The binding stoichiometry was confirmed by the measurements from equilibrium dialysis. The ratio of zinc : intein was detected as 1 : 1 for $\Delta\Delta I_{hh}$-SM and 2 : 1 for $\Delta\Delta I_{hh}$-CM in the equilibrated proteins (ESI Table S1†).

Transition metal ions typically coordinate to buffers and form metal–buffer complexes.[19] The dissociation of metal ions from buffer complexes occurs while the metal ions bind to the protein. Therefore, the buffer coordination to metal ions contributes to the overall enthalpy change (ΔH_{ITC}) and

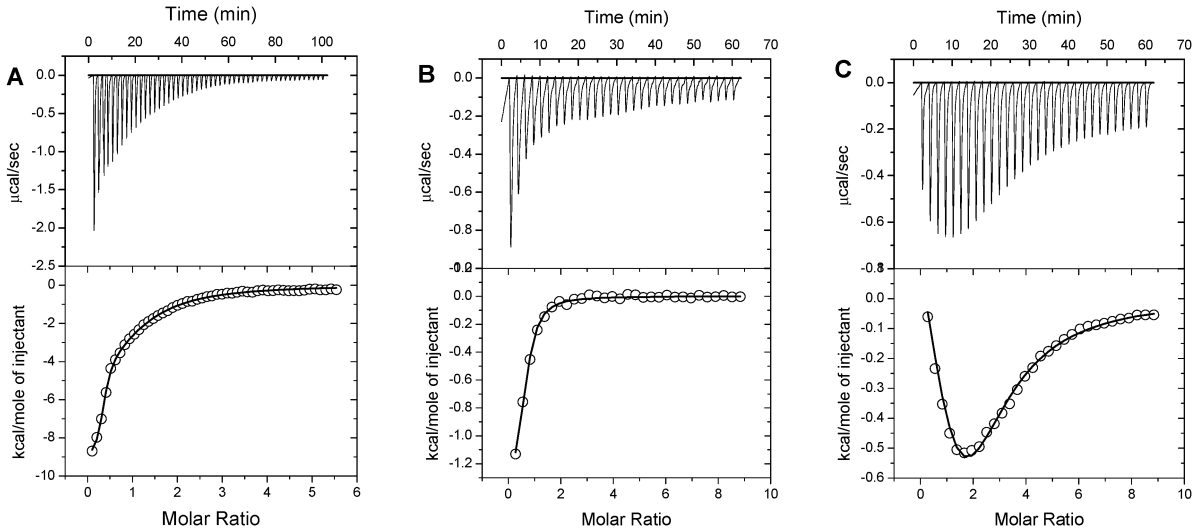

Fig. 2 ITC data for the titration of Zn^{2+} into intein mutants (A) ΔI-SM, (B) $\Delta\Delta I_{hh}$-SM, (C) $\Delta\Delta I_{hh}$-CM. In each panel, the upper portions contain the baseline corrected raw data, and the lower portions indicate the concentration normalized heat from titration at the molar ratio of intein.

Table 1 Binding constants of Zn^{2+} to three intein mutants measured by ITC titration at 25 °C

	Buffer	K_{ITC} (10^3)	K_d/nM[a]
ΔI-SM	Tris	$1.7 \pm 0.3 \times 10^3$	0.29 ± 0.05
	MOPS	$1.9 \pm 0.1 \times 10^3$	0.20 ± 0.01
$\Delta\Delta I_{hh}$-SM	Tris	$1.6 \pm 0.3 \times 10^2$	3.1 ± 0.6
	MOPS	$6.0 \pm 0.1 \times 10$	6.4 ± 0.1
$\Delta\Delta I_{hh}$-CM	Tris	$7.4 \pm 0.3 \times 10$	6.8 ± 0.2
		$2.0 \pm 0.4 \times 10^b$	25 ± 5^b
	MOPS	$1.1 \pm 0.2 \times 10^2$	3.5 ± 0.7
		$2.3 \pm 0.1 \times 10^b$	17 ± 1^b
ΔI-SM-m^{His}	MOPS	7.0 ± 0.9	55 ± 7
ΔI-SM-m^{Cys}	MOPS	$1.3 \pm 0.3 \times 10^2$	3.0 ± 0.7
$\Delta\Delta I_{hh}$-SM-m^{Cys}	MOPS	$5.1 \pm 0.1 \times 10$	7.5 ± 0.1
$\Delta\Delta I_{hh}$-CM-m^{Cys}	MOPS	7.2 ± 0.4	53 ± 3

[a] Dissociation constants K_d were calculated based on the formula $K_d = 1/K_{bind} = 1/(K_{ITC} \times K_{buffer})$. K_{ITC} is the constant obtained from ITC measurements and K_{buffer} is the Zn-buffer binding constant. $K_{buffer} = 2.0 \pm 0.1 \times 10^3$ and $2.6 \pm 0.2 \times 10^3$ for Zn-Tris and Zn-MOPS, respectively, and were measured by ITC titration in this work. Titrations were carried out in two different buffers to eliminate possible buffer-specific effects. m^{His} and m^{Cys} indicate the modified residues in the proteins, and their binding constants were measured only in the MOPS buffer because Tris could react with DEPC. [b] To the second binding site.

entropy change (ΔS_{ITC}) measured from the ITC titrations. The weak coordination of metal to buffer (K_{buffer}) has been measured by ITC titrations of metal ions into buffer. The metal binding constant to proteins is calculated as $K_{bind} = K_{ITC} \times K_{buffer}$, and the dissociation constant $K_d = 1/K_{bind}$.

Table 1 summarizes the best-fit binding constants from Zn^{2+} titration to inteins. Titrations were carried out in two different buffers to avoid possible buffer-specific effects. Results clearly show that the order of binding affinity for Zn^{2+} is ΔI-SM > $\Delta\Delta I_{hh}$-SM ~ $\Delta\Delta I_{hh}$-CM, with dissociation constants of 0.29 nM, 3.1 nM and 6.8 nM in Tris buffer, respectively. Comparable binding constants are observed in MOPS buffer. The binding affinity of Zn^{2+} to ΔI-SM is one order of magnitude higher than to $\Delta\Delta I_{hh}$-SM or $\Delta\Delta I_{hh}$-CM, which could explain the unique presence of zinc atoms in the crystal structure of ΔI-SM but not in $\Delta\Delta I_{hh}$-SM and $\Delta\Delta I_{hh}$-CM.[13]

The values of the enthalpy change (ΔH) and entropy change (ΔS) indicate different driving forces for zinc binding among the three intein mutants (ESI Fig. S1†). The zinc binding to ΔI-SM gives a large negative ΔH, indicating the enthalpy drive in the reaction. Whereas the entropy change is more important in the zinc binding to $\Delta\Delta I_{hh}$-SM and $\Delta\Delta I_{hh}$-CM. Data suggest that the unstructured loop region in ΔI-SM has a negative contribution to the overall entropy change during the zinc binding, compromising the entropy increase relative to the zinc binding to $\Delta\Delta I_{hh}$-SM and $\Delta\Delta I_{hh}$-CM. This contribution is compensated by the ΔH in the reaction of ΔI-SM, resulting in the largest value of ΔG (negative) in the zinc titration.

Mapping zinc binding sites in inteins

Although Zn coordination has been revealed in the crystal structure of ΔI-SM, crystal packing forces may alter the zinc coordination mode from solution. In addition, zinc binding was detected in solution to $\Delta\Delta I_{hh}$-SM and $\Delta\Delta I_{hh}$-CM, but not in the crystal structures. The zinc binding between two different molecules in ΔI-SM causes protein dimerization, while $\Delta\Delta I_{hh}$-SM and $\Delta\Delta I_{hh}$-CM are in monomeric conformations without zinc coordination in the crystal structures. To map the zinc binding sites in recA intein mutants in solution, ITC measurements were carried out on inteins with chemical modification on histidine and cysteine residues. NMR spectroscopy was also performed to provide site specific binding information.

Diethyl pyrocarbonate (DEPC) was used for histidine modification and iodoacetamide was used for cysteine modification.[20] Data in Table 1 shows that the histidine modification significantly decreased the zinc binding affinity to ΔI-SM (K_d from 0.20 nM to 55 nM), while the interference by the cysteine modification is much less (K_d = 3.0 nM) (see Table 1). This result indicates that histidines are critical

for Zn coordination while cysteine plays a less important role. This observation is consistent with the zinc coordination mode in the crystal structure of ΔI-SM (Fig. 1). Although cysteine is not directly involved in zinc binding in the crystal structure, Cys1 is close to the C-terminal residues (His439 and Asn440),[13] so that the cysteine modification may alter the local structure for zinc binding.

Two-dimensional (2D) ^{1}H–^{15}N HSQC NMR spectroscopy was employed to further identify the zinc binding sites. The complete assignment of the NMR spectra has been reported previously.[21] While the majority of the signals of intein were not affected in the presence of zinc ions, several peaks were considerably reduced by zinc binding (labeled in Fig. 3). Consistent with the coordination in the crystal structure, the zinc binding to Glu424, His429, His439 and Asn440 significantly affects the signals of Val425, Thr430 and Asn440 amides and the Asn440 sidechain. In addition to the decrease in these four signals, the His73 $N^{\varepsilon}H$ peak was reduced considerably by the second zinc binding to $\Delta\Delta I_{hh}$-CM (see the discussion below).

Structural analysis indicates that the zinc binding to the four residues in the intein (Glu424, His429, His439 and Asn440) must cause protein dimerization. Therefore, the protein size could provide an additional confirmation for the zinc binding to these four residues. Dynamic light scattering (DLS) detections were carried out to determine the protein size in solution. The measurements were performed on inteins in the absence and in the presence of 500 μM Zn^{2+}. The data were analyzed using Dynamics V6.2 software and the apparent molecular masses were calculated based on the hydrodynamic radii (R_h) detected (ESI Table S2†). Results show that all three intein mutants become dimers in the presence of zinc, suggesting the zinc coordination structure in the crystal structure of ΔI-SM remains in solution for all three intein mutants.

Table 2 Binding constants measured by ITC titration in Tris buffer at 25 °C

Metal	Intein	K_{ITC} (10^3)	K_d/nM[a]
Cd^{2+}	ΔI-SM	43 ± 8	137 ± 28
	$\Delta\Delta I_{hh}$-SM	46 ± 1	128 ± 3
	$\Delta\Delta I_{hh}$-CM	59 ± 4	100 ± 7
Ni^{2+}	ΔI-SM	32 ± 4	14 ± 2
	$\Delta\Delta I_{hh}$-SM	110 ± 10	4.1 ± 0.3
	$\Delta\Delta I_{hh}$-CM	45 ± 6	10 ± 1
Co^{2+}	ΔI-SM	17 ± 2	12 ± 2
	$\Delta\Delta I_{hh}$-SM	13 ± 2	16 ± 3
	$\Delta\Delta I_{hh}$-CM	4.5 ± 0.3	46 ± 3
Mn^{2+}	ΔI-SM	9.3 ± 0.1	57 ± 1
	$\Delta\Delta I_{hh}$-SM	2.7 ± 0.3	195 ± 20
	$\Delta\Delta I_{hh}$-CM	1.3 ± 0.4	404 ± 130

[a] K_d was calculated in the same way as in Table 1. K_{buffer} for Tris were measured by ITC titration in this work as Cd^{2+}: 1.7 ± 0.2 × 10^2; Ni^{2+}: 2.2 ± 0.9 × 10^3; Co^{2+}: 4.8 ± 0.5 × 10^3 and Mn^{2+}: 1.9 ± 0.1 × 10^3.

Other divalent metal binding to inteins

It has been reported that several divalent metals can inhibit protein splicing in addition to zinc, though with lower inhibitory efficiencies. Table 2 shows binding constants measured for four other metal ions to the three mutants (Mg^{2+} was not considered in this work due to its low capability for coordination to the relevant residues). Here the K_{ITC} is the relevant binding constant for the comparison because the inhibition assay has been performed in the buffer.[15] The K_{ITC} of the five metal ions to ΔI-SM roughly follows the order of Zn^{2+} >> Cd^{2+} > Ni^{2+} > Co^{2+} > Mn^{2+}, which correlates with the inhibitory efficiency.[15] The inhibition could result from either the metal coordination to the key residues (as observed with zinc binding), or the metal induced protein conformation change. CD spectra were measured for all reactions between the five metal ions and three inteins. Results exclude any

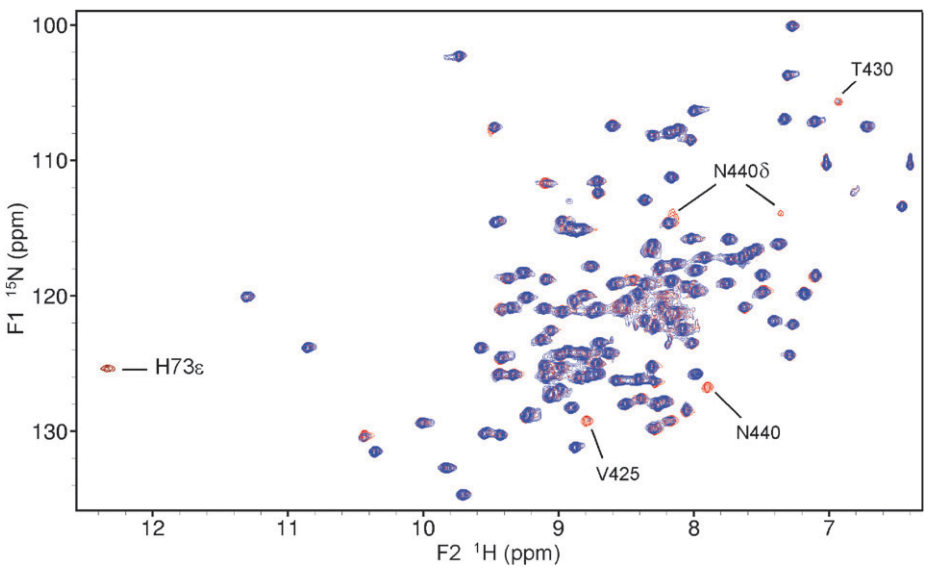

Fig. 3 Two-dimensional ^{1}H–^{15}N HSQC NMR spectra for determining the zinc binding sites in the intein $\Delta\Delta I_{hh}$-CM. Overlay of the spectrum of zinc bound intein (blue) on the apo-intein (red). Zinc binding reduces the NMR signal from residues near the binding site. Red peaks show through only if the blue peaks disappear upon zinc binding (labeled in the figure). The full spectra assignments can be found in ref. 21. For clarity, only signals interfered by zinc binding are labeled in the figure.

significant change in protein secondary structure upon metal binding (data not shown). Therefore, inhibition of protein splicing is likely caused by metal coordination to active site residues instead of structural changes in the protein.

Binding site analysis has shown that histidine residues are crucial for metal binding to inteins. Based on the HSAB principle (hard soft acid base), this binding site is obviously unfavorable for hard metal ions, such as Mg^{2+} and Ca^{2+}. Therefore, a low inhibition effect of these metal ions can be expected. Indeed, the activity study showed that Mg^{2+} had very low inhibition efficiency on the recA intein ($\sim$4% at 2 mM).[15] A crystal structure detected zinc binding to DnaE intein at a rather soft pocket (to Cys, His and Glu). Accordingly, no significant inhibition by Mg^{2+} and Ca^{2+} was observed on DnaE intein.[14] These results further confirm the correlation between binding affinity and intein inhibition.

Discussions

The effect of zinc on intein activity has been studied *in vivo* and *in vitro*.[14,15] Results showed that both *cis*- and *trans*-splicing protein splicing of recA intein were almost completely inhibited by 2 mM $ZnCl_2$ *in vitro*.[15] By using *E. coli* containing a mutated ALS gene, the protein splicing was assayed *in vivo*, and the inhibition of *Ssp* DnaE intein by $ZnCl_2$ was detected.[14] Both studies demonstrated that the inhibitory effect of the zinc ion could be reversed by the metal chelating agent EDTA.

In the crystal structure of minimized recA intein, zinc was detected in ΔI-SM. However, no zinc was added during the protein preparation and crystallization. The treatment of EDTA prior to crystallization did not result in a zinc-free crystal structure.[13] In the meantime, zinc was not detected in the other two crystal structures of $\Delta\Delta I_{hh}$-CM and $\Delta\Delta I_{hh}$-SM with the same preparation and crystallization methods.

The thermodynamics measurements from ITC titration in this work indicate that the zinc binding affinity to ΔI-SM is over one order of magnitude higher than to $\Delta\Delta I_{hh}$-SM or $\Delta\Delta I_{hh}$-CM (K_d of 0.29 nM relative to 3.1 nM and 6.8 nM), which accounts for the unique presence of zinc in the crystal structure of ΔI-SM. However, this binding constant is far smaller than that of Zn-EDTA (log $K \sim 16.5$). These data indicate that the binding affinity of zinc to intein is reasonably high for non-metallo-proteins, but is still much weaker than to EDTA, which explains the reversible inhibition by zinc.

Based on the coordination constants measured in this work, the treatment with EDTA during the intein purification should remove Zn^{2+} in the protein. Indeed, we measured the zinc concentration in proteins which have been treated by EDTA followed by dialysis. No zinc was detected by atomic absorption spectroscopy (AAS) (ESI Table S1†). The origin of the zinc in the crystal was discussed in the literature.[13] Based on our AAS results and the binding constants comparison, it can be clarified that the presence of zinc in the crystal of ΔI-SM was from contaminations.

Zinc has been detected in several intein crystal structures. However, structural comparison has shown large variation around the zinc binding sites among inteins,[13] which may arise from the substantial sequence difference among the inteins studied. For the recA intein ΔI-SM, zinc coordinates to Glu424, His429, the penultimate histidine and the C-terminal residue in the crystal structure. This is different from the zinc coordination in two inteins from PI-SceI VMA and DnaE, which have the thiol coordination from C-extein C+1.[11,12] The only cysteine residue in recA intein, Cys1, does not coordinate to zinc in the crystal structure of ΔI-SM. Consistent with this coordination structure, ITC measurements on the chemically modified intein indicated that histidines are critical for Zn binding to ΔI-SM, while cysteine modification affects the zinc binding to a much lesser extent (Table 1).

Although zinc was not present in the crystal structures of $\Delta\Delta I_{hh}$-SM and $\Delta\Delta I_{hh}$-CM due to the weak binding affinity, the interaction in solution indicates the zinc coordination also caused protein dimerization of these two inteins, based on the DLS measurements. This observation implies that the binding sites in ΔI-SM very likely remain in the two further minimized inteins (although an additional zinc binding site in $\Delta\Delta I_{hh}$-CM). This suggestion was confirmed by 2D ^{1}H–^{15}N HSQC NMR spectra, which provide detailed information of each residue in solution.

While most titration data can be fit with a single-site binding mode, a two-site binding mode was obtained for fitting the data from zinc titration to $\Delta\Delta I_{hh}$-CM (Fig. 2C). We propose the additional binding site is at the Cys1 thiol in $\Delta\Delta I_{hh}$-CM. In $\Delta\Delta I_{hh}$-SM and ΔI-SM, Cys1 coordination is prevented by a hydrogen bond between the Cys1 thiol and the carboxylate group of D422 (supported by unpublished NMR results). This hydrogen bond does not exist in $\Delta\Delta I_{hh}$-CM due to the D422G mutation, making the Cys1 thiol group more accessible for zinc coordination. To confirm this idea, we carried out ITC measurements in $\Delta\Delta I_{hh}$-CM with a cysteine modification. The Cys1 is the only cysteine in the recA intein. The restoration to one binding site with cysteine modification in $\Delta\Delta I_{hh}$-CM provides direct evidence for the second zinc binding site at Cys1, whereas no such effect was observed for $\Delta\Delta I_{hh}$-SM under the same conditions (only a minor change of K_d from 6.4 nM to 7.5 nM, see Table 1). The His73 imidazole ring is only a few Ångstrom away from the Cys1 thiol group in the structure, which could form a chelate for the zinc coordination. 2D ^{1}H–^{15}N HSQC NMR spectra of $\Delta\Delta I_{hh}$-CM indicate that the His73 $N^{\varepsilon}H$ signal becomes significantly weaker with zinc binding, which confirms His73 coordination to zinc (Fig. 3). Although the Cys1 amide is not detectable on the HSQC map, the Leu2 signal decreases $\sim$50% due to Cys1 binding.

The Zn binding affinity of $\Delta\Delta I_{hh}$-SM is about one order of magnitude lower than that of ΔI-SM. Comparing these two mutants, the $\Delta\Delta I_{hh}$-SM was generated by replacing the 36-residue peptide in an unstructured loop region from ΔI-SM with a 7-aa peptide β-turn. Although this mutation decreases the zinc binding affinity, the loop region is rather distant from the zinc binding sites. The overall structure and the position of the four coordination residues are nearly identical in the two inteins, therefore, the different binding affinity is unlikely from the local structural difference for the zinc coordination. CD spectroscopy confirmed that the protein conformation changes induced by zinc binding were negligible (data not shown). This conclusion is also supported

by NMR spectra, in which the zinc binding does not shift any peaks on the HSQC maps (Fig. 3). However, $T\Delta S$ changes from positive in $\Delta\Delta I_{hh}$-SM/Zn binding to negative in ΔI-SM/Zn binding, while the ΔH value is considerably larger for the zinc binding to ΔI-SM than to $\Delta\Delta I_{hh}$-SM (ESI Fig. S1†). These data suggest the loop region has a dominant contribution to the enthalpy change in the formation of the dimeric conformation with zinc binding. In the meantime, the negative entropy change (ΔS) in the ΔI-SM/Zn interaction may facilitate the protein crystallization. The ΔI-SM crystal was obtained slowly after zinc was recruited (most likely from glassware contamination),[13] showing the important role of zinc binding in the crystallization. In contrast, $\Delta\Delta I_{hh}$-SM could be crystallized without zinc binding.[13]

It is generally accepted that intein splicing consists of four steps: (1) N–S or N–O acyl shift and the formation of a linear ester intermediate; (2) transesterification and formation of a branched ester intermediate; (3) succinimide formation and peptide cleavage by the cyclization of the asparagine residue adjacent to the C-terminal splicing junction; (4) succinimide hydrolysis and rearrangement of the ester linking two extein segments to form a peptide bond.[2–4] Several bond rearrangements involved in the reaction process and mechanism show the significance of three conserved residues at the cleavage junctions, Cys1, C-terminal asparagine and Cys + 1. Even though not illustrated in this four-step mechanism, several other conserved residues play important roles in the protein splicing based on the activity studies, such as His73, Asp422 and His443 in the recA intein.[13,17] Although the detailed function of these three residues in the recA intein is not clear, NMR studies show that the protonation status of His73, Asp422 and His443 are crucial for the protein dynamic properties and the stabilization of protein structures (unpublished results).

While zinc binds to the active core residues in the intein, the coordination must restrict the mobility of these residues. Among the four zinc binding residues in ΔI-SM, three of them are conserved residues in intein sequences, C-terminal N440, penultimate H439 and F-block histidine H429. Therefore, the bond rearrangements for new peptide bond formation are restrained by zinc coordination and consequently the protein splicing is inhibited. In such conditions, the metal binding affinity must associate with the inhibition efficiency. ITC measurements in this work demonstrated that the observed binding constants K_{ITC} of different metal ions have reasonable correlation with the inhibition efficiency in the literature.[15]

Protein structures are not perturbed by metal ion coordination based on CD measurements. This is consistent with the result from crystallography study on the recA intein.[13] By comparison of the crystal structures of zinc-bound ΔI-SM and zinc-free $\Delta\Delta I_{hh}$-SM, the overall structures of the two mutants are nearly the same, even the local structures around the zinc binding site are very similar. This result indicates the zinc binding does not affect the protein structures. Therefore, we can conclude that the restraint of key residues by metal coordination is likely the key cause of the inhibition of the protein splicing instead of structural changes.

Conclusions

This work has quantified the binding affinity and thermodynamic properties of metal ions binding to *Mtu* recA intein mutants. The higher affinity of zinc binding to ΔI-SM relative to the other mutants is consistent with the unique existence of zinc in the crystal structure of this mutant. The binding constants indicated that the zinc binding to intein was strong but still much weaker than to EDTA, which explains the reversible inhibition. The restriction of key residues by metal coordination is probably the major cause of inhibition, based on the correlation of the inhibition efficiency and binding constants, while metal-induced protein structure changes were negligible. The zinc binding sites are the same in the crystal structure of ΔI-SM as in solution for ΔI-SM, $\Delta\Delta I_{hh}$-SM and $\Delta\Delta I_{hh}$-CM. However, an additional zinc binding site was detected in $\Delta\Delta I_{hh}$-CM with coordination to the Cys1 thiol and His73 imidazole.

Experimental

Protein expression and purification

The inteins were overexpressed in *E. coli* host strain JM101 in Luria-Bertani (LB) medium as described previously.[17] For preparing the ^{15}N isotope labeled intein for NMR measurements, the M9 minimal medium containing $^{15}NH_4Cl$ as the sole nitrogen source was used. The intein expression plasmids were generous gifts from Marlene Belfort. The expressed protein is a fusion of the chitin binding domain (CBD) at the N-terminus and intein at the C-terminus. The purification of the inteins was performed using affinity chromatography with chitin beads (New England Biolabs).[13] After washing out impurities, the inteins were released from the chitin beads by 50 mM (for CM) or 200 mM (for SM) dithiothreitol (DTT), which could induce the N-terminal cleavage between the CBD and inteins. 10 mM EDTA was present in the eluents for removing metal ion contaminations. Proteins were concentrated by ultrafiltration in a stirred cell (Millipore Amicon), and DTT and EDTA were removed by buffer exchange under N_2. To prevent the oxidation of cysteine residue in inteins, air must be avoided once DTT was removed from the proteins. No metal ions were detected in the purified inteins by atomic absorption spectrometry (AAS) measurements. The protein purity was verified by SDS gel, which gave single bands for all inteins. NMR spectra show no impurity peaks (Fig. 3, comparing with the ref. 21).

Protein modification

Chemical modification of inteins was carried out as described.[22] Diethyl pyrocarbonate (DEPC) and iodoacetamide were used for histidine and cysteine modification, respectively. Reactions were carried out with 50 μM protein in 20 mM MOPS buffer at pH 7.4 with 500 molar-equivalent of freshly prepared DEPC or iodoacetamide. The protein was treated with tris(2-carboxyethyl)phosphine (TCEP) prior to the reaction with iodoacetamide for the cysteine modification. The formation of *N*-carbethoxy imidazole with DEPC was monitored by UV spectroscopy at 25 °C by measuring the

absorbance at 242 nm. The excessive DEPC and iodoacetamide were removed from the protein by dialysis. Size-exclusion chromatography and gel electrophoresis were performed to verify the purity of the intein after modification.

ITC measurement and data processing

ITC measurements were carried out at 25 ± 0.2 °C on an isothermal titration calorimeter (VP–ITC MicroCal). All solutions were filtered through 0.22 μm filters and degassed prior to use. Samples were buffered at pH 7.4 with either 20 mM Tris or 20 mM MOPS, and the ionic strength was adjusted to 100 mM with NaCl. The metal ion concentrations of stock solutions were verified by titrations against standardized EDTA solutions. Typically, 30 or 50 portions of 8 or 5 μl metal solution were injected into the protein in the sample cell during each titration. A 120 s delay between injections was allowed for the equilibration. The heat generated or absorbed during interactions was measured by the instrument. The titration of metal ions to buffer solution was carried out along each experiment as a reference for baseline corrections in the data process. All experiments were repeated at least three times.

The ITC data were analyzed with the Origin 7.0 software package from MicroCal. A baseline correction was applied to each experiment by subtraction of data from the titration of the metal ion solution into a buffer blank correlating to the heat of dilution. A binding isotherm was fitted to the data using a nonlinear least squares method to minimize χ^2 values. The detected enthalpy change (ΔH_{ITC}) and association constant (K_{ITC}) were obtained from the best fit parameters.

The binding constants of metal ions to buffer (K_{buffer}) have been measured by ITC. The same experimental method and data process are used as for the titration to proteins. The titration was carried out by the injection of metal ion solution into buffer solution in the sample cell. Pure water was used as the reference correlating to the heat of dilution for the baseline corrections.

NMR measurements

NMR spectra were collected at 25 °C on a Bruker 600 MHz spectrometer equipped with a cryogenic probe. NMR samples were prepared in phosphate buffer at pH 7.0 containing 10% D_2O. 500 μL of 0.3 mM uniformly ^{15}N-labeled inteins were used. Two-dimensional ^{1}H–^{15}N HSQC spectra were recorded in the absence and presence of zinc ion. Data were obtained with a spectral width of 10 800 Hz in the ^{1}H dimension and 2400 Hz in the ^{15}N dimension and 32 scans of 2048 real time points for each of 128 t_1 increments were recorded. The ^{15}N dimension was zero-filled to 256 points. The data were processed with NMRpipe and analyzed with Sparky. Complete peak assignments have been reported previously.[21]

Dynamic light scattering (DLS)

DLS measurements were carried out at 25 °C on a DynaPro MSTC800 DLS instrument (Wyatt Technology Corporation, America) fitted with a 624.4 nm, 50 mW laser by measuring scattering at a 90° angle. 20 μM intein samples were prepared in 20 mM Tris-HCl buffer with 200 mM NaCl at pH 7.4. All samples were filtered through a 100 nm membrane before the measurements. The hydrodynamic radii (R_h) of inteins were measured in the absence and in the presence of 500 μM Zn^{2+}. The data were analyzed using Dynamics V6.2 software.

Equilibrium dialysis

The stoichiometry of metal binding to inteins was determined by equilibrium dialysis. Metal ions were measured by either an atomic absorption spectrometer (AAS on AAnalyst 800, Perkin Elmer, USA), or inductively coupled plasma atomic emission-mass spectroscopy (ICP-MS on PlasmaQuad 3, VG Elemental, USA). Every sample was prepared by adequate equilibrium dialysis and the reference samples were collected from the solution outside of the dialysis bag.

Acknowledgements

This work was supported by the Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China (No. 707036), NSFC (No. 20873135) and NIH (GM44844 & GM081408). Y. L. would like to thank the Chinese Academy of Sciences for the One Hundred Talent Project. We are grateful to Dr Marlene Belfort for the continuous support and helpful discussions.

References

1 H. Paulus, *Annu. Rev. Biochem.*, 2000, **69**, 447–496.
2 M. Q. Xu and T. C. Evans, *Curr. Opin. Biotechnol.*, 2005, **16**, 440–446.
3 V. Muralidharan and T. W. Muir, *Nat. Methods*, 2006, **3**, 429–438.
4 H. Paulus, *Chem. Soc. Rev.*, 1998, **27**, 375–386.
5 P. M. Kane, C. T. Yamashiro, D. F. Wolczyk, N. Neff, M. Goebl and T. H. Stevens, *Science*, 1990, **250**, 651–657.
6 M. E. Hahn and T. W. Muir, *Trends Biochem. Sci.*, 2005, **30**, 26–34.
7 J. P. Gogarten, A. G. Senejani, O. Zhaxybayeva, L. Olendzenski and E. Hilario, *Annu. Rev. Microbiol.*, 2002, **56**, 263–287.
8 F. B. Perler, *Nucleic Acids Res.*, 2002, **30**, 383–384.
9 V. Derbyshire, D. W. Wood, W. Wu, J. T. Dansereau, J. Z. Dalgaard and M. Belfort, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 11466–11471.
10 K. Shingledecker, S. Q. Jiang and H. Paulus, *Gene*, 1998, **207**, 187–195.
11 B. W. Poland, M. Q. Xu and F. A. Quiocho, *J. Biol. Chem.*, 2000, **275**, 16408–16413.
12 P. Sun, S. Ye, S. Ferrandon, T. C. Evans, M. Q. Xu and Z. H. Rao, *J. Mol. Biol.*, 2005, **353**, 1093–1105.
13 P. Van Roey, B. Pereira, Z. Li, K. Hiraga, M. Belfort and V. Derbyshire, *J. Mol. Biol.*, 2007, **367**, 162–173.
14 I. Ghosh, L. Sun and M. Q. Xu, *J. Biol. Chem.*, 2001, **276**, 24051–24058.
15 K. V. Mills and H. Paulus, *J. Biol. Chem.*, 2001, **276**, 10832–10838.
16 D. W. Wood, W. Wu, G. Belfort, V. Derbyshire and M. Belfort, *Nat. Biotechnol.*, 1999, **17**, 889–892.
17 K. Hiraga, V. Derbyshire, J. T. Dansereau, P. Van Roey and M. Belfort, *J. Mol. Biol.*, 2005, **354**, 916–926.
18 S. Leavitt and E. Freire, *Curr. Opin. Struct. Biol.*, 2001, **11**, 560–566.
19 D. E. Wilcox, *Inorg. Chim. Acta*, 2008, **361**, 857–867.
20 Y. Rahimi, S. Shrestha, T. Banerjee and S. K. Deo, *Anal. Biochem.*, 2007, **370**, 60–67.
21 Z. M. Du, Y. Z. Liu, Y. C. Zheng, S. McCallum, J. Dansereau, V. Derbyshire, M. Belfort, G. Belfort, P. Van Roey and C. Y. Wang, *Biomol. NMR Assignments*, 2008, **2**, 111–113.
22 C. Follmer and C. R. Carlini, *Arch. Biochem. Biophys.*, 2005, **435**, 15–20.

