NMR Structure of the Single QALGGH Zinc Finger Domain from the *Arabidopsis thaliana* SUPERMAN Protein

Carla Isernia,[a] Enrico Bucci,[b] Marilisa Leone,[b] Laura Zaccaro,[b] Paola Di Lello,[a] Giuseppe Digilio,[c] Sabrina Esposito,[a] Michele Saviano,[b] Benedetto Di Blasio,[a] Carlo Pedone,[b] Paolo V. Pedone,[a] and Roberto Fattorusso*[a]

Zinc finger domains of the classical type represent the most abundant DNA binding domains in eukaryotic transcription factors. Plant proteins contain from one to four zinc finger domains, which are characterized by high conservation of the sequence QALGGH, shown to be critical for DNA-binding activity. The Arabidopsis thaliana *SUPERMAN protein, which contains a single QALGGH zinc finger, is necessary for proper spatial development of reproductive floral tissues and has been shown to specifically bind to DNA. Here, we report the synthesis and UV and NMR spectroscopic structural characterization of a 37 amino acid SUPERMAN region complexed to a Zn^{2+} ion (Zn – SUP37) and present the first high-resolution structure of a classical zinc finger domain from a plant protein. The NMR structure of the SUPERMAN zinc finger domain consists of a very well-defined $\beta\beta\alpha$ motif, typical of all other Cys_2-His_2 zinc fingers structurally characterized. As a consequence, the highly conserved QALGGH sequence is located at the N terminus of the α helix. This region of the domain of animal zinc finger proteins consists of hypervariable residues that are responsible for recognizing the DNA bases. Therefore, we propose a peculiar DNA recognition code for the QALGGH zinc finger domain that includes all or some of the amino acid residues at positions -1, 2, and 3 (numbered relative to the N terminus of the helix) and possibly others at the C-terminal end of the recognition helix. This study further confirms that the zinc finger domain, though very simple, is an extremely versatile DNA binding motif.*

KEYWORDS:

DNA recognition · NMR spectroscopy · structure elucidation · SUPERMAN · zinc finger domain

Introduction

The zinc finger is one of the major structural motifs involved in eukaryotic protein – nucleic-acid interaction.[1, 2] Several zinc finger families differing in secondary structure, metal coordination, and modularity have been identified.[3] Zinc finger domains of the classical type (also named Cys_2-His_2), which were first discovered in the *Xenopus* transcription factor III A (TFIIIA),[4, 5] represent the most abundant DNA binding domains in eukaryotic transcription factors and are characterized by the consensus sequence (Phe, Tyr)-X-Cys-X_{2-5}-Cys-X_3-(Phe, Tyr)-X_5-ψ-X_2-His-X_{3-5}-His, where X represents any amino acid and ψ is a hydrophobic residue.[6] The three-dimensional structures of the Cys_2-His_2 zinc finger domains described so far consist of two antiparallel β strands faced by an α helix ($\beta\beta\alpha$ motif); the cysteine and histidine side chains coordinate a zinc ion with a tetrahedral geometry and the other three conserved residues are packed to form a hydrophobic core.[2] Solution and solid-state structural studies have revealed that the α helix makes contact with the major groove of DNA.[6, 7] Typically, more than one zinc finger binds contiguous sets of target DNA triplets and the fingers are separated by short spacers of seven amino acids, known as HC links. However, it was recently shown that the single zinc finger domain of the *Drosophila* GAGA factor, if preceded by two highly basic regions, is sufficient for high-affinity specific DNA binding.[8, 9]

The first TFIIIA-type zinc finger found in plants was that in the DNA-binding protein EPF1 of *Petunia*.[10] More than 80 plant genes encoding proteins containing Cys_2-His_2 zinc finger domains have been reported and the DNA-binding activities of some multiple-zinc-finger proteins that belong to the EPF family

[a] *Prof. R. Fattorusso, Prof. C. Isernia, Dr. P. Di Lello, Dr. S. Esposito, Prof. B. Di Blasio, Prof. P. V. Pedone*
Dipartimento di Scienze Ambientali
Seconda Università di Napoli
81100 Caserta (Italy)
Fax: (+39) 0823-274605
E-mail: fattorusso@chemistry.unina.it

[b] *Dr. E. Bucci, Dr. M. Leone, Dr. L. Zaccaro, Dr. M. Saviano, Prof. C. Pedone*
Dipartimento di Chimica Biologica
Sezione Biostrutture & Istituto
di Biostrutture e Bioimmagini – CNR
Università di Napoli "Federico II"
80134 Napoli (Italy)

[c] *Dr. G. Digilio*
Bioindustry Park del Canavese
10010 Colleretto Giacosa, Torino (Italy)

  1439-4227/03/04/2-3 $ 20.00+.50/0

from *Petunia* have been characterized.[11] These plant proteins contain from one to four zinc fingers and they differ from their counterparts in eukaryotic organisms by two remarkable structural features. The first feature is the spacer between consecutive fingers, which in plants is of different lengths and usually much longer than in animals. The other structural difference is the high degree of conservation of the sequence QALGGH, which has been shown to be critical for DNA-binding activity.[12] This conserved sequence is presumably located in the putative DNA binding helix, which is, in contrast, constituted of hypervariable residues in animal zinc finger proteins.[11] Remarkably, no high-resolution structure of the QALGGH plant zinc finger motif is so far available in the literature.

In 1995, Sakai et al.[13] demonstrated that the *Arabidopsis thaliana* SUPERMAN protein is necessary for the proper spatial development of reproductive floral tissues. In fact, recessive mutations of the SUPERMAN gene cause extra stamens to form within the normal third whorl stamens at the expense of fourth whorl carpel development. The open reading frame of the SUPERMAN gene encodes a protein of 204 amino acids that contains a single zinc finger domain with the highly conserved QALGGH sequence. Carboxy-terminal to the SUPERMAN zinc finger is a region rich in serine and proline residues, similar to the proline-rich sequences present in the transcriptional activation or repression regions of many *Drosophila* transcription factors.[14, 15] Furthermore, a cluster of basic residues, which may act as a nuclear localization signal, and a domain that contains two overlapping leucine-zipper-like motifs are present at the C-terminal end of the protein. All these protein motifs suggest that SUPERMAN may act as a transcription factor. Indeed, we recently demonstrated[16] that the *Arabidopsis* SUPERMAN protein is able to specifically bind the DNA through its single Cys_2-His_2 zinc finger domain and two flanking basic regions.

In this paper we report the synthesis and UV and NMR spectroscopic structural characterization of a 37 amino acid SUPERMAN region, which represents the first high-resolution structure of a classical zinc finger domain from a plant protein. Analysis of the three-dimensional structure reveals that the DNA recognition mechanism of classical plant zinc finger domains significantly differs from that observed in animal counterparts.

```
           1                                   37
 SUPERMAN  WPPRSYTCSF CKREFRSAQA LGGHMNVHRR DRARLRL
   Q9SLB8  WPPKNYTCSF CRREFRSAQA LGGHMNVHRR DRAKLRQ
   O80942  WPPRSYTCSF CRREFKSAQA LGGHMNVHRR DRARLKQ
   Q9LHS9  WPPRSYSCSF CGREFKSAQA LGGHMNVHRR DRARLKQ
   O23621  WPPRSYTCNF CRREFRSAQA LGGHMNVHRR DRASSRA
   Q9SR34  NQSRSYVCSF CIRGFSNAQA LGGHMNIHRR DRAKLRQ
   Q9M344  SQARPYICEF CERGFSNAQA LGGHMNIHRK DRAKLRQ
      LIF  HTTKSYECNF CKRGFSNAQA LGGHMNIHRK DKAKLKK
   Q9FIY6  VTRRMYECTF CKRGFTNAQA LGGHMNIHRR DRLNKAK
   Q9LTD6  ENRRTYDCDI CKRGFTNPQA LGGHNNIHRR ERERYPS
     ZFP7  ANPRVFSCNY CRRKFYSSQA LGGHQNAHKR ERTMAKR
     ZFP1  ADPRVFSCNY CQRKFYSSQA LGGHQNAHKR ERTLAKR
     ZFP4  VSKRVFSCNY CQRKFYSSQA LGGHQNAHKR ERTLAKR
     ZFP2  EQPRVFSCNY CQRKFYSSQA LGGHQNAHKL ERTLAKK
     ZFP3  TEQKLFSCNY CQRTFYSSQA LGGHQNAHKR ERTLAKR
   Q9LEU3  TEQKLFSCNY CQRTFYSSQA LGGHQNAHKR ERTLAKR
   Q9FKA9  SSLRIFPCLF CSRKFHSSQA LGGHQNAHKK ERTAARR
   Q9FFX4  ASHRLFPCQY CPRKFYTSQA LGGHQNAHKR ERAAARR
   Q9LZW0  AKARVFACTF CKKEFSTSQA LGGHQNAHKQ ERSLAKR
   Q9C9H1  GGERKYECQY CCREFGNSQA LGGHQNAHKK ERQQLKR
   Q943S6  GDGRRYECQY CCREFANSQA LGGHQNAHKK ERQQLKR
     ZFP6  SESRKYECQY CCREFANSQA LGGHQNAHKK ERQLLKR
     ZFP5  TVMKRHECQY CGKEFANSQA LGGHQNAHKK ERLKKKR
   P93751  NNNRRFECHY CFRNFPTSQA LGGHQNAHKR ERQHAKR
     ZFP8  SNNRRFECHY CFRNFPTSQA LGGHQNAHKR ERQHAKR
   Q9M2P1  SHNRRFECHY CFRNFPTSQA LGGHQNAHKR ERQLAKR
   Q9FG05  NKEKRFKCHY CFRNFPTSQA LGGHQNAHKR ERQQTKR
   Q9FF62  IENPIYKCKY CPRKFDKTQA LGGHQNAHRK EREVEKQ
   Q9LG97  EAGKVYECRF CSLKFCKSQA LGGHMNRHRQ ERETETL
   Q9LIA9  QSKSSHKCKI CGKSFECYQA LGGHQRIHRP IKEKLKR
Consensus  ...r.y.C.. C.r.F..sQA LGGHqNaHr. #r...kr
```

Figure 1. *Sequence alignment of SUP37 with the proteins present in the databases that contain a single QALGGH zinc finger domain. Where the protein name is not properly given in the database, the entry name is reported. All the proteins are from* Arabidopsis thaliana, *with the exception of LIF, which is from* Petunia, *and of Q9LG97 and Q943S6, which are from* Oryza sativa. *In the consensus sequence, amino acids conserved in more than 90% of the sequences are indicated with capital letters, amino acids conserved in more than 50% are indicated with small letters, and # indicates the conservation of any of the following amino acids: Asn, Asp, Gln, Glu.*

Results

Sequence alignment and peptide synthesis

MultAlin software[17] was used to align the SUPERMAN sequence with all the sequences of plant zinc finger proteins present in the sequence databases (SWISS-PROT, TrEMBL, TrEMBL-new) that show high conservation of the QALGGH sequence located between the second cysteine and the first histidine residue of the zinc finger domain. The alignment with plant proteins containing a single QALGGH zinc finger is reported in Figure 1 and revealed the presence of a 37 amino acid sequence comprised of residues 42 – 78 of the SUPERMAN protein that is almost identical in five proteins and highly conserved in all the proteins. This fragment, which contains the QALGGH zinc finger domain and a C-terminal region rich in basic residues, was therefore selected for the structural determination.

The peptide (termed SUP37 hereafter) was synthesized by the solid-phase method with standard 9-fluorenylmethoxycarbonyl (Fmoc) chemistry, with the N and C termini acetylated and amidated, respectively. SUP37 was cleaved from the resin by treatment with trifluoroacetic acid (TFA) and purified by reversed-phase high-performance liquid cromatography (RP-HPLC) in the presence of dithiothreitol (DTT) to obtain the product in the reduced form. The purity and identity of the product were confirmed by means of analytical RP-HPLC and MALDI-TOF mass spectrometry.

The zinc ion was then added to the peptide as previously reported,[9] to form the complex (Zn – SUP37). Successful folding of the zinc finger motif was demonstrated by NMR spectroscopy.

UV/Vis experiments

The interaction between a Co^{II} species and the peptide was studied by a UV/Vis quantitative titration to probe the metal-binding properties of SUP37.[18] A series of UV spectra were acquired for samples containing Co^{II}/peptide at a ratio of 0.1 – 2.0 at pH 6.4.

The cobalt binding affinity of SUP37 was investigated by following the increase in intensity of the thiolate-to-metal charge transfer absorbance band at 290 nm. To estimate the dissociation constant of the cobalt – SUP37 complex, the absorbance at 290 nm (A_X in the following equation) was collected and the fractional saturation values (f_{pCo}) were calculated. The f_{pCo} values obtained were fitted by the following binding isotherm[19] (Figure 2):

$$f_{pCo} = \frac{A_X - A_0}{A_{max} - A_0} = \frac{K_d^{Co} + [P]_{tot} + [Co]_{tot} - \sqrt{([P]_{tot} + [Co]_{tot} + K_d^{Co})^2 - 4[P]_{tot}[Co]_{tot}}}{2[P]_{tot}}$$

where f_{pCo} is the fractional saturation; A_0 and A_{max} are the absorbance values in the absence and in the presence of 2.0 equivalents cobalt, respectively; $[Co]_{tot}$ is the total concentration of cobalt added; $[P]_{tot}$ is the total concentration of the peptide to be titrated and K_d^{Co} is the dissociation constant of the cobalt – peptide complex. A dissociation constant, K_d^{Co}, of 5.59×10^{-7} M was obtained.

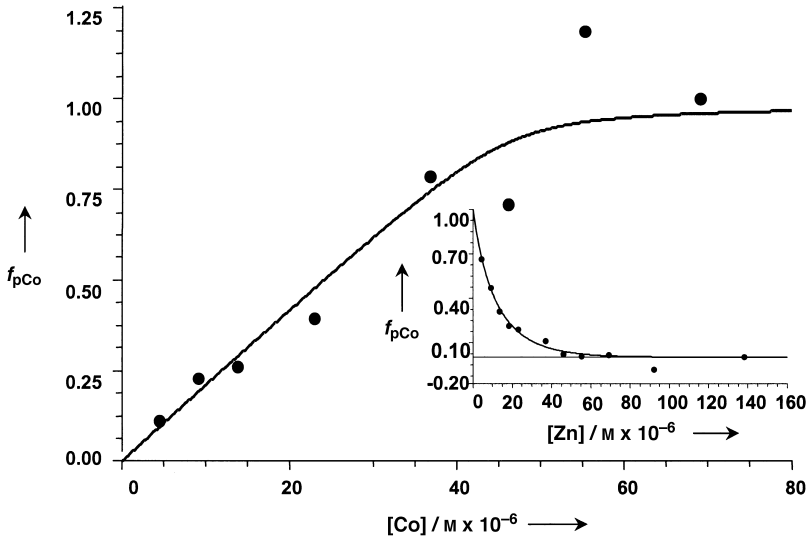

Figure 2. *Titration of SUP37 with $CoCl_2$ monitored by recording the absorbance at 290 nm. The cobalt fractional saturation f_{pCo} is plotted against the total cobalt concentration. The fit of the data to the equation given in the text is shown. The inset shows the titration of the SUP37 – Co(II) complex with a Zn(II) species, monitored by the decrease of the absorbance at 290 nm. The solid line shows a fit of the data assuming simple competition between Co(II) and Zn(II) species for binding to the peptide. The dissociation constant for the SUP37 – Co(II) complex was recorded as 5.59×10^{-7} M.*

The affinity of SUP37 for Zn^{II} ions could be determined by competition with cobalt because of the ability of this ion to displace cobalt from the complex. Titration of the SUP37 – cobalt complex with zinc in the presence of 2.0 equivalents cobalt induces a decrease in the absorption band at 290 nm (A_X in the following equation). The K_d^{Zn} was obtained by fitting the collected data with the following equation (Figure 2, inset):

$$f_{pCo} = \frac{A_X - A_{max}}{A_0 - A_{max}} = \frac{(K_d^{Zn}[P]_{tot} + K_d^{Zn}[Co]_{tot} + K_d^{Co}[Zn]_{tot} - K_d^{Co}[P]_{tot})}{2[P]_{tot}(K_d^{Zn} - K_d^{Co})}$$
$$- \frac{\sqrt{(K_d^{Zn}[P]_{tot} + K_d^{Zn}[Co]_{tot} + K_d^{Co}[Zn]_{tot} - K_d^{Co}[P]_{tot})^2 - 4[P]_{tot}(K_d^{Zn} - K_d^{Co})K_d^{Zn}[Co]_{tot}}}{2[P]_{tot}(K_d^{Zn} - K_d^{Co})}$$

where f_{pCo} is the fractional saturation; A_0 and A_{max} are the absorbance values in the absence and in the presence of 3 equivalents zinc, respectively; $[Co]_{tot}$ is the total cobalt concentration; $[Zn]_{tot}$ is the total concentration of zinc added; K_d^{Co} and K_d^{Zn} are the dissociation constants of the cobalt – peptide complex and zinc – peptide complex, respectively. The K_d^{Zn} value obtained was 2.17×10^{-8} M. These K_d^{Co} and K_d^{Zn} values are comparable to those obtained for other naturally occurring zinc binding domains.[20, 21]

Resonance assignment, collection of conformational constraints, and structure determination of Zn – SUP37

Sequence-specific backbone assignment of Zn – SUP37 was obtained on the basis of sequential NOE connectivities[22] established by an accurate analysis of 2D [^{1}H,^{1}H]-NOESY,[23] 2D [^{1}H,^{1}H]-DQFCOSY,[24] and 2D [^{1}H,^{1}H]-CleanTOCSY[25] spectra acquired by using a sample dissolved in 90% H_2O/10% ^2H_2O. The chemical shifts of the α and β protons provided the starting point for the complete ^{1}H assignment of all CH_n moieties in aliphatic and aromatic side chains, which was obtained by using 2D-[^{1}H,^{1}H]-CleanTOCSY, 2D [^{1}H,^{1}H]-NOESY, and 2D [^{1}H,^{1}H]-DQFCOSY spectra recorded with a sample dissolved in pure ^2H_2O. Stereospecific assignments for $2 \times \alpha CH_2$, $8 \times \beta CH_2$, and the methyl groups of Leu21 and Val27 were obtained with the DYANA software.

Comparison with the so-called fingerprint chemical shifts for zinc fingers, reported by Lee et al. in 1992,[26] provides a valuable tool for preliminary structure identification. These authors examined the proton chemical shifts of several zinc fingers and identified fixed positions in the sequence that show primary-structure-independent variations in the chemical shifts, mostly due to ring current effects. This highly conserved pattern observed for Cys_2-His_2 classical zinc fingers has been proposed as a clear marker of the characteristic $\beta\beta\alpha$ fold of this structural motif.[26, 27] In Table 1, ^{1}H NMR chemical shifts of Zn – SUP37 key residues are listed together with the calculated differences, $\delta_{experim} - \delta_{random\ coil}$ with respect to random coil chemical shifts[22] and the variability ranges gathered by Lee. It is worth noting that the Zn – SUP37 $\delta_{experim} - \delta_{random\ coil}$ values lie within the typical ranges of the classical zinc finger domains, and in the case of Phe15 H_ζ, Leu21 H_α, and His28 H_N the upfield shifts are even more pronounced. This analysis indicates, therefore, that the SUP37 peptide should assume the typical $\beta\beta\alpha$ fold of a classical zinc finger domain.

Conformational constraints were obtained from NOE-derived upper-limit distances and from 15 scalar spin – spin coupling constants (Figure 3). Of the total of 516 NOE cross-peaks assigned, 456 resulted from the

Table 1. *Classical zinc finger key conserved ^1H chemical shifts observed for Zn – SUP37 and their differences from the values for a random coil conformation[22].*

Residue	Proton	Chemical shift [ppm]	$\delta_{experim} - \delta_{random\ coil}$ [ppm]	Variability range[26] [ppm]
Phe10	H_N	8.95	0.72	+0.49/+1.65
Glu14	H_α	4.80	0.51	−0.35/+0.63
Phe15	H_ζ	5.95	−1.39	−1.26/−0.68
Ala18	H_α	3.35	−1.00	−1.43/+0.10
Leu21	H_N	7.35	−1.07	−1.57/−0.58
	H_α	3.12	−1.26	−1.21/+0.01
His28	H_N	6.95	−1.46	−1.37/−0.78
	H_α	4.57	−0.06	−0.25/+0.63
	H_δ	6.77	−0.37	−0.83/−0.55

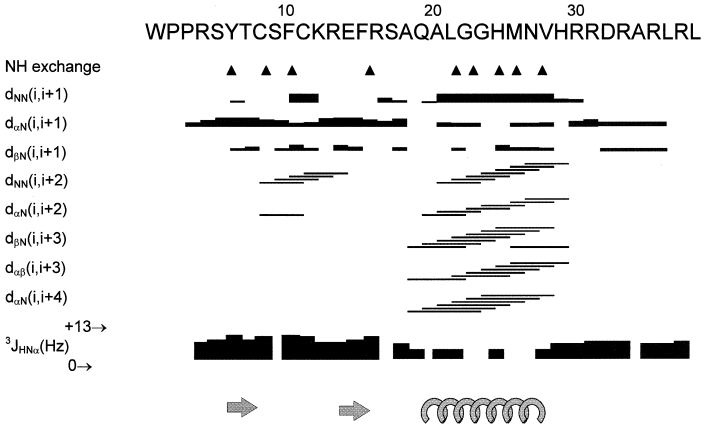

Figure 3. *Summary of short- and medium-range NOEs involving the H_N, H_α, and H_β protons of Zn – SUP37. Evaluations of $^3J_{HNH\alpha}$ coupling constants and the slowly exchanging amide protons are also reported. The secondary structure elements are indicated.*

2D [^{1}H,^{1}H]-NOESY spectrum (acquired with a mixing time of 100 ms) obtained from the sample dissolved in 90% H_2O/10% ^2H_2O, and 60 from the 2D [^{1}H,^{1}H]-NOESY spectrum (mixing time of 100 ms) recorded with the sample in pure ^2H_2O. The final input for the torsion angle dynamics program DYANA[28] for calculation of the structure of Zn – SUP37 consisted of 338 upper distance constraints (90 intraresidual, 83 sequential, 83 medium range, and 82 long range, according to DYANA definitions) and 123 dihedral angle constraints, which were derived from intraresidue and sequential NOEs and *J* couplings by using the program DYANA. No hydrogen bonding constraints were used in the calculations.

Among one hundred calculated structures, the twenty with the lowest target function were selected and subjected to further energy minimization. The small size and small number of residual violations indicate that the input data set is self-consistent and show that the constraints are well satisfied by the structures obtained (Table 2). Analysis of the calculated structures, for residues 1 – 29 with the PROCHECK-NMR program[29] revealed that 98.1% of backbone ϕ/ψ pairs lie within the most favoured or additional favoured allowed regions of the Ramachandran plot. Data concerning hydrogen bonds are reported in Table 3.

Table 2. *Input for the structure calculation and characterization of the energy-minimized NMR structures of the Zn – SUP37 peptide.*

Quantity	Value
NOE upper distance limits	338
Dihedral angle constraints	123
Residual target function [Å^2]	1.02 ± 0.01
Residual NOE violations	
Number > 0.1 Å	1 ± 1
Maximum [Å]	0.22 ± 0.08
Residual angle violations	
Number > 2.0°	2 ± 1
Maximum [°]	7.0 ± 0.5
AMBER energies [KJ/mol]	
Total	−2404 ± 41
van der Waals	−835 ± 21
Electrostatic	−2171 ± 36
RMSD of the mean coordinates [Å]	
N, Cα, C′ (5 – 28)	0.097 ± 0.020
N, Cα, C′ (5 – 28) plus zinc coordinating side chains (8, 11, 24, 28)	0.084 ± 0.018
N, Cα, C′ (5 – 28) plus 8, 10, 11, 15, 21, 24, 28 residue side chains	0.60 ± 0.13
All heavy atoms (5 – 28)	0.74 ± 0.16

Table 3. *Hydrogen bond distances and angles derived from the 20 energy-minimized structures of Zn – SUP37.*

Location	Donor – Acceptor	H – O distance [Å]	N-H···O Angle [°]
N-terminal region	Arg4 H_N – Pro2 O	2.0 ± 0.2	154 ± 8
β hairpin	Tyr6 H_N – Phe15 O	2.3 ± 0.4	150 ± 8
	Thr7 Hγ – Arg13 O	2.6 ± 0.4	147 ± 8
	Cys8 H_N – Arg13 O	2.2 ± 0.2	153 ± 4
	Phe10 H_N – Cys8 Sγ	2.4 ± 0.3	153 ± 15
	Phe10 H_N – Cys8 O	2.6 ± 0.2	139 ± 9
	Cys11 H_N – Cys8 Sγ	2.8 ± 0.1	161 ± 5
	Lys12 H_N – Cys8 O	2.5 ± 0.1	132 ± 2
	Arg13 H_N – Cys11 Sγ	2.8 ± 0.2	150 ± 9
	Phe15 H_N – Tyr6 O	1.8 ± 0.1	177 ± 2
α helix	Ala20 H_N – Ser17 O	2.2 ± 0.2	149 ± 8
	Leu21 H_N – Ser17 O	2.4 ± 0.4	152 ± 22
	Leu21 H_N – Ala18 O	2.3 ± 0.3	147 ± 16
	Gly22 H_N – Ala18 O	2.0 ± 0.2	171 ± 8
	Gly23 H_N – Gln19 O	2.6 ± 0.3	165 ± 8
	His24 H_N – Ala20 O	2.4 ± 0.3	142 ± 4
	His24 H_N – Leu21 O	2.2 ± 0.1	145 ± 2
	Met25 H_N – Leu21 O	2.2 ± 0.1	167 ± 2
	Asn26 H_N – Gly22 O	2.5 ± 0.2	161 ± 6
	Asn26 H_N – Gly23 O	2.6 ± 0.1	135 ± 4
	Val27 H_N – Gly23 O	2.6 ± 0.1	157 ± 5
	Val27 H_N – His24 O	2.7 ± 0.2	136 ± 6
	His28 H_N – His24 O	2.8 ± 0.1	155 ± 1
	His28 H_N – Met25 O	2.3 ± 0.4	147 ± 11
	Arg29 H_N – Met25 O	2.0 ± 0.2	162 ± 13
	Arg29 H_N – Val27 O	2.2 ± 0.2	143 ± 13

The NMR structure of Zn – SUP37

The NMR structure of Zn – SUP37 (Figure 4) consists of a very well-defined core region (residues 5 – 28), a less-ordered N-terminal end (residues 1 – 4), and a highly disordered basic tail (residues 29 – 37). Figure 5 shows a ribbon diagram of a representative structure. The overall fold of the QALGGH zinc finger domain from the SUPERMAN protein is very similar to the

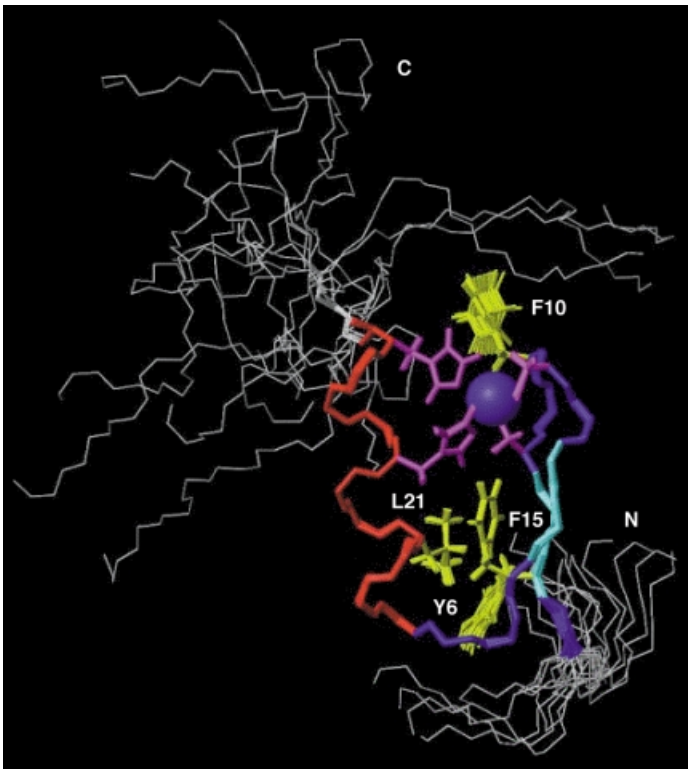

Figure 4. *Superposition of the 20 energy-minimized conformers of Zn – SUP37 used to obtain the minimal root mean square deviation (RMSD) of the backbone heavy atoms of residues 5 – 28. The polypeptide backbone and the side chains of Tyr6, Cys8, Phe10, Cys11, Phe15, Leu21, His24, and His28 are shown. The β-sheet region is shown in cyan, the α-helical region in red, and the segments 1 – 4 and 29 – 37 in grey; the coordinating side chains are depicted in magenta, the hydrophobic side chains in yellow, and the zinc ion in blue.*

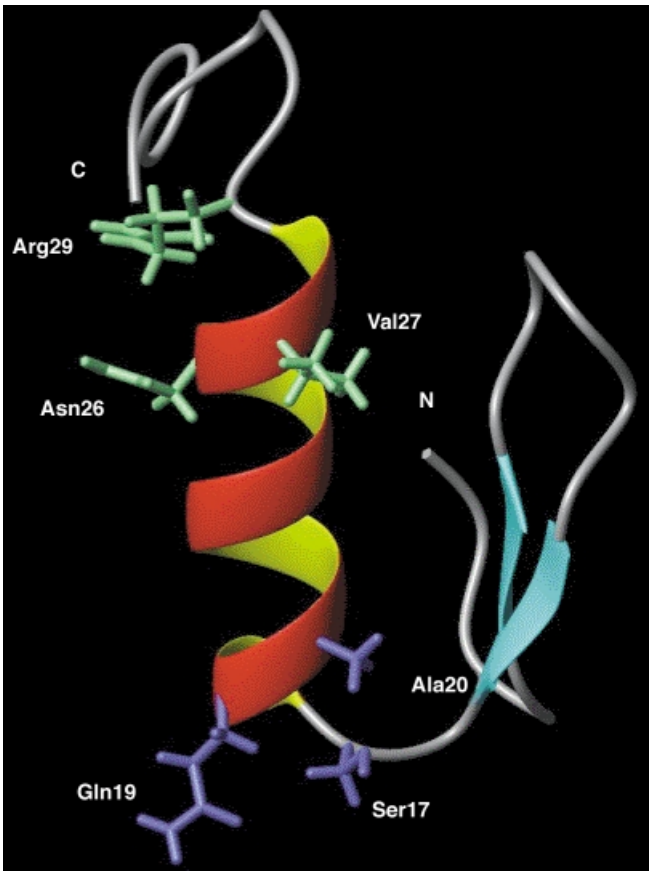

Figure 5. *Ribbon drawing of one of the 20 energy-minimized conformers of Zn – SUP37 used to represent the NMR structure. Side-chain conformations of Ser17, Gln19, Ala20, Asn26, Val27, and Arg29 of Zn – SUP37 are also shown.*

classical zinc finger domains that have been structurally characterized (Table 4); the structure includes two successive short β strands composed of residues 6 – 7 and 13 – 14 that form a β sheet, which is packed against the subsequent α helix composed of residues 18 – 28. The β sheet and the α helix are held together by the zinc ion and by a small hydrophobic core constituted of the side chains of the conserved residues Tyr6, Phe15, and Leu21, which are very well-defined in the NMR structure (Figure 4). Interestingly, the well-ordered side chain of Phe10 stacks against the aromatic ring of His28 and protects the zinc coordination site from solvent exposure (Figure 4).

Several hydrogen bonds connect the first β strand with the second (see Table 3). Additional hydrogen bonds are found from Phe10 H_N to Cys8 S_γ and from Arg13 H_N to Cys11 S_γ; these interactions form the well-known "knuckle" structure common to many classical zinc finger domains.[3] The slow proton/deuterium exchange of Tyr6, Cys8, Phe10, and Phe15 amide protons (Figure 3) confirms their involvement in hydrogen bonds with the Phe15 and Arg13 carbonyl oxygen, Cys8 sulphur, and Tyr6 carbonyl oxygen atoms, respectively.

The helix is essentially in α conformation and includes entirely the highly conserved sequence QALGGH. Remarkably, two glycine residues occupy the helix positions +5 and +6, a feature not yet found in any zinc finger domain high-resolution structure. The C terminus of the α helix comprises the Arg29 H_N

Table 4. *Comparison of one representative conformer of Zn – SUP37 with the natural zinc finger domains structurally characterized that have only two amino acids between the two cysteine residues (C-X-X-C in the sequence).*[a]

Protein	PDB Code[b]	Technique	RMSD [Å]
GAGA complexed to DNA[9]	1YUJ	NMR	1.44
Tramtrack first finger complexed to DNA[46]	2DRP	X-ray	1.15
Tramtrack second finger complexed to DNA[46]	2DRP	X-ray	1.34
ADR1 first finger[47]	2ADR	NMR	0.96
ADR1 second finger[47]	2ADR	NMR	1.13
EPB[48]	4ZNF	NMR	1.27
Zif268 second finger complexed to DNA[49]	1AAY	X-ray	1.27
Zif268 third finger complexed to DNA[49]	1AAY	X-ray	1.43
Xfin31[50]	1ZNF	NMR	1.23
YY1 second finger complexed to DNA[51]	1UBD	X-Ray	1.32
ZFY[52]	5ZNF	X-Ray	1.34

[a] The RMSD values refer to the superposition of the backbone atoms of Zn – SUP37 residues 5 – 28 with those of the corresponding residues of each zinc finger domain. [b] PDB = Protein Data Bank.

atom, which is, in half of the structures, hydrogen bonded to the Met25 carbonyl oxygen atom. A detailed analysis of the hydrogen-bond interactions (see Table 3 and Figure 3) underlines some 3_{10} character in the turn comprising the two zinc-coordinating histidine residues. This feature is often observed when the fingers have an H-X_3-H sequence pattern.[6]

The N terminus (residues 1 – 4) of Zn – SUP37, although relatively disordered, tends to fold into a turned conformation (Figure 4) and we could argue that this folding results in a more ordered structural element within the whole SUPERMAN protein.

In Table 4, the structure of Zn – SUP37 is compared with those of the classical zinc finger domains, which contain the $-C-X_2-C-$ moiety (see the consensus sequence) in the β hairpin. The tertiary fold of Zn – SUP37 is generally very similar to the selected domains, especially when compared to the structures of zinc finger domains, which are not complexed to DNA.

Discussion

More than 80 plant genes have been reported containing one, two, three, or four Cys_2-His_2 zinc finger domains. Zinc fingers in plant proteins are characterized by spacers of very different lengths between adjacent fingers (with a range of 19 – 65 amino acids) and a highly conserved QALGGH sequence.[11] These two important features strongly distinguish plant zinc fingers from their counterparts in other eukaryotes. In animal proteins, the spacer between successive zinc finger domains is mostly short and of invariant length, and includes the sequence TGEKP in about half of the known Cys_2-His_2 zinc finger proteins.[30] This common structural feature allows the recognition of contiguous sets of three DNA base pairs, with each triplet recognized by one finger. Conversely, it has been demonstrated that the two-fingered ZPT-2 protein from *Petunia*, which has a spacer of 44 amino acids, binds to two separate DNA triplet core sites and that each finger makes a specific contact solely with one site; the protein binds with the highest affinity when the target sites are spaced by 10 base pairs.[31] Furthermore, the base determinant positions of TFIIIA-type zinc finger proteins in animals lie in the α-helical region and are hypervariable in their amino acid composition in order to allow recognition of a large number of possible DNA triplets.[6] In striking contrast, almost all the classical zinc finger domains from plant proteins so far reported include the highly conserved QALGGH sequence.[32] These important features make the structural characterization of the QALGGH Cys_2-His_2 zinc finger domain a first essential step to fully understanding the DNA recognition mechanism used by these proteins.

We have focused our attention upon the SUPERMAN protein of *Arabidopsis thaliana*, which comprises a single QALGGH zinc finger domain and was found to play a role in maintaining the *Arabidopsis* floral whorl boundaries. Recently, we reported[16] that a 64-residue fragment of SUPERMAN comprised of the residues in the region 15 – 78 and containing the single zinc finger domain and two extra N- and C- terminal basic regions, is able to specifically bind a 20 base pair DNA sequence. However, smaller fragments of the protein that lack any of the basic regions do not bind DNA with high affinity. In this paper we report the synthesis, UV/Vis characterization, and NMR spectroscopy solution structure of Zn – SUP37, a smaller fragment of SUPERMAN that includes the QALGGH zinc finger domain and only the C-terminal basic region, and thus provide a structural description of this particular DNA-binding zinc finger domain. Though Zn – SUP37 lacks the N-terminal basic region and fails to display specific high-affinity DNA binding (data not shown), nonetheless this SUPERMAN region is structurally of a great interest since it includes the zinc finger motif and the C-terminal basic extension, which are involved in the recognition of DNA by the SUPERMAN protein.[16] Moreover, the high degree of conservation of the zinc finger domain sequence in plant zinc finger proteins makes it possible to extend the structural features presented in this paper to the other QALGGH zinc-finger-containing proteins.

UV/Vis data and structure calculation based on the NMR data confirmed that two cysteine sulphur and two histidine N_ε nitrogen atoms tetrahedrally coordinate the zinc ion in Zn – SUP37. The measured values of the Co – SUP37 and Zn – SUP37 dissociation constants (5.59×10^{-7} M and 2.17×10^{-8} M, respectively) are within the range typical for zinc finger domains.

The global fold of the domain is essentially identical to all the other Cys_2-His_2 zinc fingers structurally characterized so far (Figure 4 and 5, Table 4). In particular, the $\beta\beta\alpha$ motif is preserved, along with the knuckle structure in the β hairpin; the small hydrophobic core between the α helix and β sheet consists of the side chains of the conserved residues Tyr6, Phe15, Leu21. These structural features are additionally validated by the analysis of Zn – SUP37 key conserved ^{1}H chemical shifts (Table 1), which lie within the ranges typical for zinc finger domains. The structural features cause the QALGGH sequence, highly conserved in plant zinc finger proteins, to be located at the N terminus of the α helix (Figure 6). Moreover, the basic C-terminal tail, demonstrated to be critical for high-affinity binding to DNA,[16] appears to be largely disordered in solution and presumably adopts a more rigid conformation when bound to the DNA.

Hypotheses for QALGGH zinc finger DNA recognition

In the animal zinc-finger – DNA complexes solved so far, specific recognition of the target site by the protein is achieved by contacts between the helix of each finger and the major groove of the DNA.[33] All the structural and functional studies conducted on animal zinc finger proteins interacting with DNA led to the proposal of a generalized consensus zinc finger recognition code in which three of the four residues at positions -1, $+2$, $+3$, and $+6$ (numbered relative to the N terminus of the helix) recognize the third, the second, and the first base of a DNA triplet (reading the DNA sequence in the 5′ to 3′direction). In particular, in the case of the GAGA DNA binding domain, whose structure complexed with DNA is the only high-resolution characterization of a natural single Cys_2-His_2 zinc finger bound to its target sequence,[9] the three amino acid residues of the finger that specifically make contacts with the DNA bases (Arg47, Asn48, and Arg51) occupy the helix positions $+2$, $+3$, and $+6$, respectively.

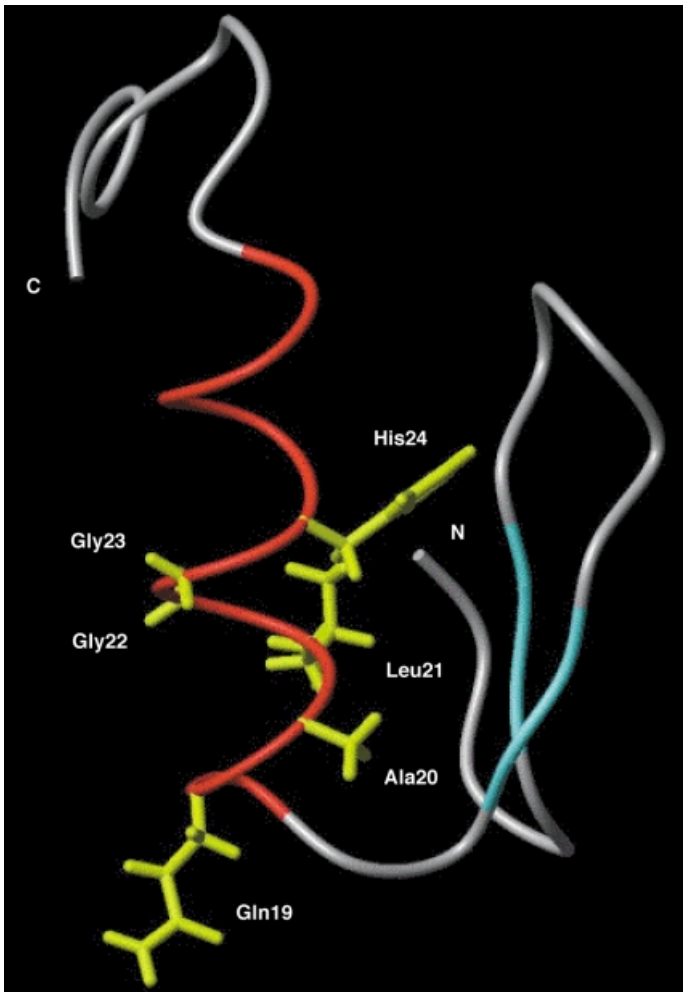

Figure 6. *Location of the QALGGH sequence with respect to the tertiary fold of Zn – SUP37. The side chains of Gln19, Ala20, Leu21, His24, and H_α protons of Gly22 and Gly23 are shown in yellow. The β-sheet region is depicted in cyan, the α-helical region in red.*

Indeed, in Zn – SUP37 the QALGGH sequence occupies positions 2 – 7 of the helix, that is, it includes all the three residues ($+2$, $+3$, and $+6$) corresponding to the base determinant positions in the GAGA DNA complex. In SUPERMAN, position 2 is occupied by a glutamine residue, position 3 by an alanine, and position 6 by a glycine residue. Even though glutamine and more infrequently alanine residues are reported to act as base determinants,[34] a glycine residue could hardly be involved in base-specific DNA contacts. This observation suggests that SUPERMAN uses residues present in different positions with respect to the GAGA factor to recognize DNA. It must be noted that in GAGA no apolar amino acid residues are involved in the base recognition, and two of the base-contacting residues are charged and basic.

Position -1 in Zn – SUP37 is occupied by serine 17, whose side chain has already been shown to specifically contact DNA.[6] Thus, residues Ser17, Gln19, and Ala20, which correspond to positions -1, 2, and 3 with respect to the α helix, could act as base determinants in the interaction of SUPERMAN with the DNA in a noncanonical docking arrangement,[6] and recognize the third, second, and first base of the DNA triplet, respectively. This recognition hypothesis is supported by the recent paper by Takatsuj and co-workers,[31] in which the authors found that the two fingers in a double QALGGH zinc finger protein from *Petunia* (ZPT2 – 2) bind to two different DNA sequences. The most favorable binding site found for the N-terminal zinc finger includes the AGC sequence, while that for the C-terminal finger comprises the AGT triplet. Interestingly, the two zinc finger domains differ at position -1 with respect to the α helix (they have a serine and a threonine residue, respectively), while positions 2 and 3 of both the fingers are occupied by the glutamine and alanine residues of the QALGGH sequence, respectively. Consequently, the above proposed recognition mechanism could explain the diverse optimal binding DNA sequences for the two zinc finger domains as a result of the difference in one of the three putative base determinant residues. In Figure 5, the conformation of Ser17, Gln19, and Ala20 (positions -1, $+2$, and $+3$, respectively) is shown with respect to the zinc finger folding. As is clearly evident, the Ser17 and Gln19 side chains are largely solvent exposed and in principle available for a specific interaction with DNA bases. Conversely, the Ala20 side chain is partly folded towards the small hydrophobic core present between the α helix and the β hairpin and would need a big conformational change to form proper contacts with a DNA base. For Gly23 (Figure 6), which lacks a side chain, we can imagine a sort of passive role in DNA recognition in which its presence allows the nucleic acid to closely approach the α helix.

Alternatively, we could assume that SUPERMAN binds DNA in a more peculiar manner, significantly different from binding observed in the zinc finger DNA complexes characterized so far. In this case, some of the base determinant residues might lie in the C terminus of the α helix. Indeed, previous experiments indicate that helix position 10 is base determinant in EPF2 – 5 and EPF2 – 7, two double QALGGH zinc finger proteins from *Petunia*.[35] In Zn – SUP37, this position is taken by Val27, which is partially folded towards the β sheet and, as in the case of Ala20, capable of freely contacting a DNA base only if DNA binding causes a side chain conformational change. On the contrary, the side chain of the preceding residue, Asn26, as well as that of the residue just at the end of the helix, Arg29, is fully solvent exposed (Figure 5) and could readily form specific strong interactions such as hydrogen bonds, with DNA bases. This alternative recognition hypothesis would imply a very atypical DNA – zinc-finger docking arrangement. Intriguingly, in such a interaction mechanism, the presence of a couple of highly conserved glycine residues within the α helix could provide the conditions required for considerable α helix bending during DNA binding.

In this paper we have presented the synthesis, UV/Vis characterization, and NMR solution structure of a 37 amino acid region (Zn – SUP37) of the SUPERMAN DNA binding protein, which contains a single QALGGH zinc finger motif. The analysis of the solution structure of Zn – SUP37, which is the first high-resolution characterization of a classical zinc finger domain from a plant protein, clearly establishes that the zinc finger motif in plants recognizes the target DNA base pairs in a manner peculiar

to plants. This consensus recognition code presumably includes all or some of the amino acid residues at the helix positions −1, 2, and 3, and possibly other positions at the C-terminal end of the recognition helix. The structural details of SUP37 confirm that the zinc finger domain, though very simple, is an extremely versatile DNA binding motif and indicate that a complete description of all the possible recognition codes used in DNA–zinc-finger-domain interaction is still to be obtained.

Experimental Section

Peptide synthesis: SUP37 was synthesized on a MilliGen 9050 peptide synthesizer with 5-(4-*N*-Fmoc-aminomethyl-3,5-dimethoxyphenoxy)valeryl-polyethylene glycol-polystyrene (Fmoc-PAL-PEG-PS) as the starting resin (substitution 0.195 mmol g^{-1}) in order to obtain a C-amidated sequence. Glass beads were added to the resin at a ratio of 1:1 (w/w) with the initial dry solid support to improve the swelling properties.

The side chain protecting groups used were: Trp(tert-butyloxycarbonyl (Boc)), Arg(2,2,5,7,8-pentamethylchroman-6-sulfonyl), Ser(triphenylmethyl (Trt), Tyr(*t*Bu), Glu(O*t*Bu), Cys(Trt), Gln(Trt), His(Trt), Lys(Boc), Asp(O*t*Bu), Gln(Trt), Thr(*t*Bu), and Asp(O*t*Bu). The synthesis was carried out in the presence of a fourfold excess of amino acid at every cycle and the residue was twice recirculated through the resin for 1 h.

The *N*-α-amino acids were activated in situ according to the standard 1*H*-benzotriazol-1-yl-oxy-tris-pyrrolidino-phosphonium (PyBop)/*N*-hydroxybenzotriazole (HOBt) procedure[36] except in the case of the difficult couplings of Pro36, Pro35, Arg34, Cys30, Arg25, Glu24, Ser21, Ala20, Gln19, Ala18, Gly16, Gly15, Met13, His10, Arg9, Arg8, Arg4, Arg2, for which the more effective *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium (HATU)/HOBt mixture was used.[37] A final acetylation step on the resin gave an *N*-acylated product. Cleavage of the peptide from the resin was accomplished by treating the peptidyl resin with a mixture of TFA and scavengers (TFA, thioanisole, phenol, water, 1,2-ethanedithiol, and triisopropylsilane 81.5:5:5:5:2.5:1 v/v) for 4 h. The resin was filtered off and the crude peptide was precipitated with diethyl ether, dissolved in water, and lyophilized.

The peptide was analyzed by RP-HPLC performed on a Shimadzu LC instrument equipped with a diode array SPDM AV-10 and a SIL-10A autosampler. A Phenomenex C_{18} column (4.6 × 250 mm, 5 μm, 300 A) was used and eluted with a H_2O/0.1% TFA (A) and CH_3CN/0.1% TFA (B) linear gradient, from 5 to 70% B over 30 min, at 1 mLmin^{-1} flow rate. Analytical analysis by RP-HPLC showed a main peak at R_t = 20.67, which corresponds to the desired product.

The crude peptide was reduced prior to purification by incubation at 60 °C for 3 h in the presence of 10 equivalents DTT/cysteine residue in phosphate buffer at pH 7. The peptide was purified on a Waters Delta Prep 4000 HPLC equipped with a UV Waters Mod 441 detector (Vydac C_{18} column 22 × 150 mm, 15 μm, 300 A) by using the linear gradient described above.

The collected fractions were lyophilized and analyzed by a MALDI-TOF Perseptive spectrometer. The MALDI-TOF analysis of the samples containing the peptide gave a molecular ion peak $[M-H]^+$ of 4502.8, as was expected.

All Amino acids used were in L configuration. The amino acids and HOBt were purchased from NovaBiochem; *N*,*N*-diisopropylethylamine and TFA were from Applied BioSystems; piperidine, pyridine, DTT and scavengers were from Fluka; HATU and Fmoc-PAL-PEG-PS resin from Primm. The solvents used in the synthesis, purification, and characterization of the peptide were anhydrous and HPLC grade and were supplied by LabScan Analytica.

UV/Vis studies: Optical absorption spectra were recorded with a UV/Vis Jasco model 550 spectrophotometer from 800 to 200 nm at room temperature and by using a cell with a 1-cm path length. The free thiol concentration of SUP37 was measured before cobalt titration with 5-5′-dithiobis (2-nitrobenzoic acid) to confirm that the peptide was fully reduced throughout the experiment.[38]

The spectra were acquired in water containing 0.05% TFA, with the pH value adjusted to 6.4 with NaOH. $CoCl_2$, $ZnCl_2$, and SUP37 peptide stock solutions were prepared by dissolving a weighed amount of the product in water containing 0.05% TFA. All TFA solutions were degassed by bubbling with nitrogen for 30 minutes prior to use. The peptide concentration was determined by using averaged extinction coefficients at 280 nm of 1490 $M^{-1}cm^{-1}$ and 5500 $M^{-1}cm^{-1}$ for Tyr and Trp, respectively.[39]

For the cobalt titration experiment the peptide stock solution (1.82×10^{-5} M) was divided into nine samples (1 ml) to which different sized aliquots of $CoCl_2$ (4.8×10^{-3} M) solution were added to give a Co^{+2}/Sup37 ratio of up to 2.0. The pH value was slowly raised to 6.4 with NaOH solution (0.1 M). The samples were lyophilized and then resuspended in water (400 μL) to give the final solutions (4.61×10^{-5} M). The absorbance at 290 nm was recorded and used to calculate the cobalt binding affinity constant.

To estimate the zinc binding constant, the competition procedure was used as a convenient indirect spectroscopic method for monitoring the binding of zinc to the peptide. Titration of the Co(II)–SUP37 complex with a Zn(II) species was carried out by the addition of aliquots of $ZnCl_2$ solution (2.4×10^{-3} M) to the sample containing 2.0 Co^{+2}/ 1.0 peptide to give a Zn^{+2}/peptide ratio of up to 3.0. The decrease of the 290 nm absorbance maximum was monitored.

Preparation of NMR samples: The NMR samples of SUP37 and the Zn–SUP37 complex were prepared according to the following procedure: SUP37 (about 6 mg) was dissolved in H_2O containing 0.05% TFA to give a final solution concentration of 1.5×10^{-4} M; for the preparation of the complex, $ZnCl_2$ was slowly added until a 1.1:1 ratio of Zn^{2+} to SUP37 was obtained. For both samples, the pH value was corrected with NaOH until a value of 6.4 was reached, then the solutions were lyophilized. Finally, the samples were prepared by dissolving SUP37 in H_2O/^2H_2O (9:1, 650 μL; Isotec Inc., USA) and Zn–SUP37 in H_2O/^2H_2O (9:1, 650 μL) or neat ^2H_2O (650 μL).

NMR spectroscopy: NMR experiments were carried out at the Bioindustry Park del Canavese (Colleretto Giacosa Torino, Italy) on a Bruker Avance 600 MHz spectrometer equipped with a triple axis-PFG probe optimized for ^1H detection and operating at 14 T, which corresponds to a proton Larmor frequency of 600 MHz. The 2D spectra were recorded by using the time-proportional phase-incrementation method to obtain complex data points in the t_1 dimension. Water suppression was achieved by means of presaturation of the solvent line during the recycle delay (2.5 s) in the case of ^2H_2O solutions or by means of the WATERGATE PFG[40] technique in the case of H_2O/^2H_2O solutions.

The 2D-TOCSY[25] experiments were recorded at 28 °C with a MLEV17 mixing scheme of 70 ms with 9 kHz spin-lock field strength (spectral width 8000 Hz both along f_1 and f_2, 2048 × 256 data points in t_2 and t_1, respectively, recycle delay 3 s, 16 scans per t_1 increment). The 2D-NOESY[23] spectra, used to derive the geometric constraints, were carried out at 28 °C by the standard pulse sequence with a mixing time of 100 ms (spectral width 8000 Hz along both f_1 and f_2,

4096 × 512 data points in t_2 and t_1, respectively, recycle delay 3 s, and 72 scans per t_1 increment). The 2D-DQFCOSY[24] spectra were obtained at 28 °C (spectral width 8000 Hz along both f_1 and f_2, 2048 × 650 data points in t_2 and t_1, respectively, recycle delay 2.5 s, 128 scans per t_1 increment). The data were typically apodized with a square cosine window function and zero-filled to a matrix of size 4096 × 1024 prior to Fourier transformation and baseline correction. Chemical shifts were referenced to external tetramethylsylane ($\delta =$ 0 ppm). The programs PROSA[41] and XEASY[42] were used for data processing and spectral analysis.

Determination of the amide protons exchange rates was performed by dissolving a perfectly dried sample of Zn – SUP37 in pure ^2H_2O and recording a series of monodimensional ^1H spectra at 28 °C on a Varian Inova spectrometer operating at 14 T, 600 MHz for the ^1H nucleus, located at the Centre for Design and Structure in Biology (CDSB) of Jena (Germany).

Determination of the three-dimensional structure: Experimental distance restraints for structure calculations were derived from the cross-peak intensities in NOESY spectra recorded in H_2O and ^2H_2O. NOESY cross-peaks were manually integrated by using the XEASY software and converted to upper distance constraints according to an inverse sixth power peak volume-to-distance relationship for the backbone and an inverse fourth power function for side chains by using the CALIBA module of the DYANA program.[28] Distance constraints together with 15 scalar coupling constants were then used by the GRIDSEARCH module, implemented in DYANA, to generate a set of allowable dihedral angles. Structure calculations, which used the torsion angle dynamics protocol of DYANA, were then started from 100 randomized conformers. The inspection of these structures clearly identified the cysteine S_γ and the histidine N_ε atoms as Zn^{2+} ion binding ligands. Additional upper and lower distance limits were successively inserted between the two cysteine S_γ and the two histidine N_ε atoms to better restrain the tetrahedral geometry. These limits were obtained from measurement of the corresponding distances in the crystal structure of different zinc fingers (Zif268 (PDB code 1ZAA), TFIIIA (PDB code 1TF6)).

The 20 structures with the lowest target functions were further refined by unrestrained energy minimization with the GROMOS96 program.[43] Several cycles of steepest descent[44] were repeated until the energy difference between two successive steps was less then 0.001 KJ mol^{-1}. No zinc ion was included, and the zinc coordinating residues were kept fixed during energy minimization. The obtained structures were examined with the programs MOLMOL[45] and PROCHECK-NMR.[29] The color figures were obtained by using the program MOLMOL.

The complete assignment of the protein has been deposited at the BioMagResBank under accession number 5342. The coordinates of Zn – SUP37 (entry 1NJQ) have been deposited in the Protein Data Bank.

This work was supported by the Training and Mobility of Researchers Program of the European Commission with NMR experiments performed at the European Large Scale Facility, Centre for Design and Structure in Biology at the Institut für Molekulare Biotechnologie in Jena, Germany (Contract no. ERB FMGE CT98 0121). This work was also supported by Ministero dell'Istruzione, Università e Ricerca (Grant no. PRIN), awarded in 1998 to B. Di Blasio. The authors thank Dr. Oliver Ohlenschläger and Dr. Ramadurai Ramachandran of the Centre for Design and Structure in Biology of Jena for the acquisition and processing of the NMR spectra for the determination of the amide proton exchange rates, and Mr. Leopoldo Zona, Mr. Marco Mammucari, and Mr Maurizio Muselli for excellent technical assistance.

[1] J. M. Berg, *Acc. Chem. Res.* **1995**, *28*, 14 – 19.
[2] J. M. Berg, H. A. Godwin, *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 357 – 371.
[3] J. W. R Schwabe, A. Klug, *Nat. Struct. Biol.* **1994**, *1*, 345 – 349.
[4] J. S. Hanas, D. J. Hazuda, D. F. Bogenhagen, F. Y.-H. Wu, C.-W. Wu, *J. Biol. Chem.* **1983**, *258*, 14 120 – 14 125.
[5] J. Miller, A. D. McLachlan, A. Klug, *EMBO J.* **1985**, *4*, 1609 – 1614.
[6] S. A. Wolfe, L. Nekludova, C. O. Pabo, *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 183 – 212.
[7] M. P. Foster, D. S. Wuttke, I. Radhakrishnan, D. A. Case, J. M. Gottesfeld, P. E. Wright, *Nat. Struct. Biol.* **1997**, *4*, 605 – 608.
[8] P. V. Pedone, R. Ghirlando, G. M. Clore, A. Gronenborn, G. Felsenfeld, J. Omichinski, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 2822 – 2826.
[9] J. G. Omichinski, P. V. Pedone, G. Felsenfeld, A. Gronenborn, G. M. Clore, *Nat. Struct. Biol.* **1997**, *4*, 122 – 132.
[10] H. Takatsuji, M. Mori, P. N. Benfey, L. Ren, N.-H. Chua, *EMBO J.* **1992**, *11*, 241 – 249.
[11] H. Takatsuji, *Plant Mol. Biol.* **1999**, *39*, 1073 – 1078.
[12] K. Kubo, A. Sakamoto, A. Kobuyashi, Z. Rybka, Y. Kanno, H. Nakagawa, T. Nishino, H. Takatsuji, *Nucleic Acids Res.* **1998**, *26*, 608 – 616.
[13] H. Sakai, L. J. Medrano, E. M. Meyerowitz, *Nature* **1995**, *378*, 199 – 203.
[14] H. P. Gerber, K. Seipel, O. Georgiev, M. Hofferer, M. Hug, S. Rusconi, W. Schaffner, *Science* **1994**, *263*, 808 – 811.
[15] W. C. Soeller, C. E. Oh, T. B. Kornberg, *Mol. Cell. Biol.* **1993**, *13*, 7961 – 7970.
[16] N. Dathan, L. Zaccaro, S. Esposito, C. Isernia, J. G. Omichinski, A. Riccio, C. Pedone, B. Di Blasio, R. Fattorusso, P. V. Pedone, *Nucleic Acids. Res.* **2002**, *30*, 4945 – 4951.
[17] F. Corpet, *Nucleic Acids Res*. **1988**, *16*, 10 881 – 10 890.
[18] I. Bertini, C. Luchinat, *Adv. Inorg. Chem*. **1984**, *6*, 71 – 111.
[19] M. S. Wisz, C. Z. Garrett, H. W. Hellinga, *Biochemistry* **1998**, *37*, 8269 – 8277.
[20] G. Roesijadi, R. Bogumil, M. Vašák, H. R. Kägi, *J. Biol. Chem*. **1998**, *273*, 17 425 – 17 432.
[21] J. M. Berg, D. L Merkle, *J. Am. Chem. Soc.* **1989**, *111*, 3759 – 3761.
[22] K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, **1986**.
[23] Anil Kumar, R. R. Ernst, K. Wüthrich, *Biochem. Biophys. Res. Commun*. **1980**, *95*, 1 – 6.
[24] M. Rance, O. W. Sørensen, G. Bodenhausen, G. Wagner, R. R. Ernst, K. Wüthrich, *Biochem. Biophys. Res. Commun.* **1983**, *117*, 479 – 485.
[25] C. Griesinger, G. Otting, K. Wüthrich, R. R. Ernst, *J. Am. Chem. Soc.* **1988**, *110*, 7870 – 7872.
[26] M. S. Lee, R. J. Mortishire-Smith, P. Wright, *FEBS Lett.* **1992**, *309*, 29 – 32.
[27] G. Barbato, D. O. Cicero, E. Bianchi, A. Pessi, R. Bazzo, *J. Biomol. NMR* **1996**, *8*, 36 – 48.
[28] P. Güntert, C. Mumenthaler, K. Wüthrich, *J. Mol. Biol.* **1997**, *273*, 283 – 298.
[29] R. A. Laskowski, J. A. Rullmannn, M. W. MacArthur, R. Kaptein, J. M. Thornton, *J. Biomol. NMR* **1996**, *8*, 477 – 486.
[30] J. H. Laity, B. M. Lee, P. Wright, *Curr. Opin. Struct. Biol.* **2001**, *11*, 39 – 46.
[31] K. Yoshioka, S. Fukushima, T. Yamazaki, M. Yoshida, H. Takatsuji, *J. Biol. Chem*. **2001**, *276*, 35 802 – 35 807.
[32] H. Takatsuji, *Cell. Mol. Life Sci.* **1998**, *54*, 582 – 596.
[33] O. C. Pabo, E. Peisach, R. A. Grant, *Annu. Rev. Biochem.* **2001**, *70*, 313 – 340.
[34] Y. Choo, A. Klug, *Curr. Opin. Struct. Biol.* **1997**, *7*, 117 – 125.
[35] H. Takatsuji, *Biochem. Biophys. Res. Commun.* **1996**, *224*, 219 – 213.
[36] J. Coste, D. Le-Nguyen, B. Castro, *Tethaedron Lett.* **1990**, *31*, 201 – 208.
[37] L. A. Carpino, *J. Am. Chem. Soc.* **1993**, *115*, 4397 – 4398.
[38] P. W. Riddles, R. L Blakeley, B. Zerner, *Methods Enzymol.* **1983**, *91*, 49 – 60.
[39] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, *Protein Sci.* **1995**, *4*, 2411 – 2423.
[40] M. Piotto, V. Saudek, V. Sklenár, *J. Biomol. NMR* **1992**, *2*, 661 – 665.
[41] P. Güntert, V. Dötsch, G. Wider, *J. Biomol. NMR* **1992**, *2*, 619 – 629.
[42] C. Bartels, T. Xia, M. Billeter, K. Wüthrich, *J. Biomol. NMR* **1995**, *5*, 1 – 10.
[43] W. F. Van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hunenberger, P. Kruger, A. E. Mark, W. R. P. Scott, I. G. Tironi, *Biomolecular simulation: the GROMOS96 manual and user guide*, BIOMOs b. v. Zürich, Groningen, The Netherlands, **1996**.

[44] G. Arfken, *Mathematical Methods for Physicists*, Academic Press, Orlando, USA, **1985**, pp. 428 – 436.
[45] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graphics* **1996**, *14*, 29 – 32.
[46] L. Fairall, J. W. R. Schwabe, L. Chapman, J. T. Finch, D. Rhodes, *Nature* **1993**, *366*, 483 – 487.
[47] M. P. Bowres, L. E. Schaufler, R. E. Klevit, *Nat. Struct. Biol.* **1999**, *6*, 478 – 485.
[48] J. G. Omichinski, G. M. Clore, E. Appella, K. Sakaguchi, A. M. Gronenborn, *Biochemistry* **1990**, *29*, 9324 – 9334.
[49] M. Elrod-Erickson, M. A. Rould, L. Nekludova, C. O. Pabo, *Structure* **1996**, *4*, 1171 – 1180.
[50] M. S. Lee, G. P. Gippert, K. V. Soman, D. A. Case, P. E. Wright, *Science* **1989**, *245*, 635 – 637.
[51] H. B. Houbaviy, A. Husheva, T. Shenk, S. K. Burley, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 13 577 – 13 582.
[52] M. Kochoyan, T. F. Havel, D. T. Nguyen, C. E. Dahl, H. T. Keutmann, M. A. Weiss, *Biochemistry* **1991**, *30*, 3371 – 3386.

Received: September 16, 2002 [F 487]