SUPPLEMENTARY DATA

Table 2D Constitution of metal coordination groups.

(i) Composition of metal coordination groups in 30% cull set with duplicate chains excluded*.

Numbers with total metal coordination number, N, in coordination groups with two or more protein donors.

N	2	3	4	5	6	7	≥8	all N
Ca	2	6	13	36	110	22	1	190
Mg	3	8	7	16	46	1	-	81
Mn	-	-	3	13	18	3	-	37
Fe	-	-	12	10	15	-	-	37
Cu	-	9	17	11	-	-	-	37
Zn	7	19	89	31	3	-	-	149
Na	2	5	6	14	13	2	-	42
K	-	2	5	5	6	8	2	28

Numbers with N protein donor groups in metal coordination group.

N	0**	1	2	3	4	5	6	7	all N
Ca	11	27	29	26	45	61	27	2	228
Mg	54	71	41	31	6	3	-	-	206
Mn	5	6	11	14	11	1	-	-	48
Fe	1	24	7	8	15	5	2	-	62
Cu	-	1	1	21	14	1	-	-	38
Zn	2	33	21	51	76	-	-	-	184
Na	4	28	10	15	9	6	-	1	74
K	4	-	8	3	7	7	2	1	32

Numbers with N non-protein donors in coordination groups with two or more protein donors.

N	0	1	2	3	4	5	6	all N
Ca	51	60	39	26	8	5	1	190
Mg	10	12	9	25	24	1	-	81
Mn	1	3	15	13	4	1	-	37
Fe	12	10	6	5	4	-	-	37
Cu	17	17	3	-	-	-	-	37
Zn	71	59	18	1	-	-	-	149
Na	12	10	9	8	3	-	-	42
K	5	5	11	5	-	2	-	28

* Allowance has not been made here for the cases where a metal is coordinated by donor groups or solvent molecules which are in neighbouring asymmetric units in the crystal, related by symmetry to those in the PDB file. The numbers are small.

** These could be related to duplicate chains, there is no criterion on which to exclude them; the large numbers for Mg relate to protein/DNA complexes, and/or phosphate complexes.

(ii) Distribution of chelate loop sizes.

seqdif	0	1	2	3	4	5	6-10	11-19	20-29	30-49	50-99	100-199	200-499	all
Ca	31	56	237	68	14	38	16	29	28	48	22	16	3	606
Zn	9	9	37	69	29	13	26	38	30	31	40	18	5	354
Mg, Fe, Cu	2	30	56	35	25	24	12	24	34	47	54	34	16	
Na, K	16	19	37	30	0	5	2	4	15	12	11	7	3	

(iii) Most commonly occurring chelate loops for each metal. Lists of *all* chelate loops for each metal are in Table 2W (at http://tanna.bch.ed.ac.uk/arch/).

metal	total number coordination groups	number of donor pairs	commonest and	chelate loops (number)			
Ca	190	606	DD 2 (35) DN 2 (16)	DO 1 (19) DO 2 (38) NO 2 (15)	OE 5 (27) OO 2 (38) OO 3 (20)	OD 0 (12) OD 2 (32) OD 3 (12)	 [ON 2 (6)] [ON 3 (6)]
Mg	81	133	DD 2 (7)	OO 3 (14)			
Mn	37	76	DD 2 (4)	DO 1 (3)			
Fe	37	96	EH 3 (7)	HH 5 (4)	CC 3 (5)	(non-haem,	non Fe/S)
Cu	37	88	HH 2 (6)	HM 5 (5)	CH 4 (4)	CH 5 (6)	CC 4 (5)
Zn	149	354	HH 2 (11)	HH 4 (16)	CC 2 (9)	CC 3 (53)	CC 5 (9)
Na	42	93	OO 1 (4)	OO 2 (12)	OO 3 (16)		
K	28	78	OO 1 (8)	OO 2 (5)	OO 3 (5)	OD 2 (3)	OT 0 (4)

Table 4 D (i) Conformations of most commonly occurring Ca chelate loops. The conformations are defined by letters as in Fig 2(b), The range of conformations within a group is indicated by the (sample) standard deviation, s.d., of the torsion angles ϕ,ψ within the group. The local conformation is that from relseq = -10, to 10 amino-acids beyond the end of the chelate loop. (Small discrepancies between the numbers of chelate loops found here and in Table 2 f) are due to the different ways of dealing with duplicate chains in the asymmetric unit.)

Ca	No.	conformation and number in each	s.d. of ϕ,ψ (^{o})	fold families	local confns. same ?	example (and resolution/Å)	Comments
DO 2	38	25 kgb	10-17	various	various	2pvb at 94A (0.9)	
13-ring		13 bkb or bab	19-38	various	various	1nls at 10 (0.94)	
NO 2	13	8 kgb	10-14	more than one	various	1bfd at 455 (1.6)	like Ca DO 2
13-ring		5 xkb	8-14	more than one	various	1gci at 79 (0.78)	like Ca DO 2
DD 2	28	15 dak	10-21	mostly same	mostly same	2pvb at 51A (0.91)	1st dpair** in 2 2 2 5
14-ring		8 kgk	9-11	mostly same	most same	2pvb at 92A (0.91)	clear subsets
		3 bkb or bdb	2-5	all same	all same		
		2 misc					
DN 2	15	9 dak or bak	7-17	various	most same	1gca at 134 (1.7)	like Ca DD 2
14-ring		5 kgk	8-12	all same	all same	1g4y at 58R (1.6)	like Ca DD 2
		1 bbb					
OO 2	40	16 kb	9-23	various	various	1i76 at 169A (1.20)	clear subsets, see ****
10-ring		11 bj	8-22	various	various	1i76 at 155A (1.20)	clear subsets, see ****
		9 jb	7-27	some same	some same	1gci at 79 (0.78)	clear subsets
		3 misc					

OD 0	12	5 b	(ψ) 9	mostly same	various	2msb at 206A (1.7)	
7-ring		7 a	(ψ) 9	various	various	1pa2 at 43A (1.45)	
OS/T 0	13	4 k	(ψ) 5	various	various	1pa2 at 170A (1.45)	
6-ring		9 b or d	(ψ) 30	various	various	1d2v at 168C (1.75)	range of conformations
OO 3	21	6 kkb	4-12	various	some same	2btc at 72E (1.50)	} three are fairly close - i.e.
13-ring		4 akb	5-11	various	various	1slu at 72B (1.8)	} subsets of one confn
		4 aab	9-26	various	some same	1g5c at 118F (2.10)	} with s.d. 11-30^{o}
		7 misc					
OD 2	32	27 bb, ba, bd	16-31	various	various	see Table 4 W	clear subsets, s.d. 4-10^{o}
11-ring		5 misc					
OE 2	8	7 ba or bd	4-17		various	2pvb at 57A (0.9 Å)	like Ca OD 2 counterparts
12-ring		1 misc					
DO 1	19	17 bb	19-35		various	2sus at 40 (1.50 Å)	clear subsets s.d. 4-13^{o}
10-ring		2 misc					

** all but one are first chelate loop in a 2 2 2 5 coordination group (EF hand)
*** all but one are second chelate loop in a 2 2 2 5 coordination group (EF hand)
****see also 1kap at 253P for kb conformation, and at 370P for bj conformation(1.64A resolution)

(Table 4 D continued)

(ii) Distribution of conformations of most commonly occurring Zn chelate loops - details as in Table 6(a)

Zn	No.	conformation and number in each	s.d. of ϕ,ψ (^{o})	fold families	local confns. same ?	example (and resolution/Å)	comments
CC 2 11-ring	9	6 bkd 2 + 1 misc	15-22	?	no	1h7n at 133A (1.6)	
CC 3 14-ring	50**	47 baak or baaa	9-26	various	no	1vfy at 176A (1.15) 1vfy at 222A (1.15)	) two clear subsets,) s.d. < 11^o, rest between
		3 agab	4-21			1het at 97A (1.15)	
CC 5 17-ring	10	5 bkbakb 2 kgbaad 2+ 1 misc	4-18	?	no	1ali at 107A (1.6)	
HH 2 11-13 ring	11	5 bbb 2 bdb 2 bbb 2 misc	6-20	3 same	no	1hzy at 55B (1.30)	N^{ε} in all his his are N^{ε} N^{δ} N^{ε} in both his his are N^{δ} N^{ε} and N^{ε} N^{δ}
HH 4 19-ring	18	18 aaaaa	5-14	11 same, rest various	no	1c7k at 83A (1.00)	HH4 is α-helix N^{ε} in all his

** 3 groups removed because of doubts about structure determination (in 1ile)

(Table 4 D continued)

(iii) Examples of conformations in other less common chelate loops for Zn and Ca

	X or XX	number	conformation	
Zn XX 3	CH, HC, HD, HE	7	baaa	like *Zn CC 3*
	EE, EH	3	aaaa	helix
	CH, DD, HC, RH	6		at least 3 other conformations
Zn XX 4	DH, ED, HD , HE	7	aaaaa	like *Zn HH 4*
	CH, EE	2		similar to *Zn HH 4* above
	CC, CH, EE, HC	4		all different
Ca XX 3	DD, EE, NE, TQ	5	aaaa	helix
	DD, DE, EE, EN, HD, NE, QE	11		at least 5 other conformations
Ca XX 4	DD, ED	3	aaaaa	like *Zn HH 4*
	DE, DT, EE, QE	6		three other conformations
Ca XO 1	E, N, S	7	bb	like *Ca DO 1*
Ca OX 2	N, Q, S, T	7	bd	close to *Ca OD 2* above (s.d. 7-11^{o})
	E, N, Q, T	11		quite close to *Ca OD 2* above (s.d. 14-29^{o})

Fig. 5D One protein chain may provide the donors for more than one metal coordination group; sometimes the two metal atoms are close, as a result of the sharing of donor groups. These schemes a)-f) illustrate the interactions found when Zn…Zn approaches between 3 and 6 Å were investigated, and Ca…Ca between 3 and 7.5 Å.

a) is found in 1zme, 1hwt, Zn…Zn distance ~3.0 Å.

b) is found in 1lam, 1ush with distances 3.0, 3.3 Å

c) is found in 1cg2, 1bf6, 1aol, 1ah7, 1ak0 with distances 3.3 - 3.7 Å.

d) is found in 1qtw, 1qh5, 1amp with distances 3.4 - 3.5 Å.

e) is found in 1rmd, 4mt2 with distance 3.9 Å.

f) is found in1hzy, 1j79, 1qq9, 1ew2 with distances 3.4 - 4.0 Å.

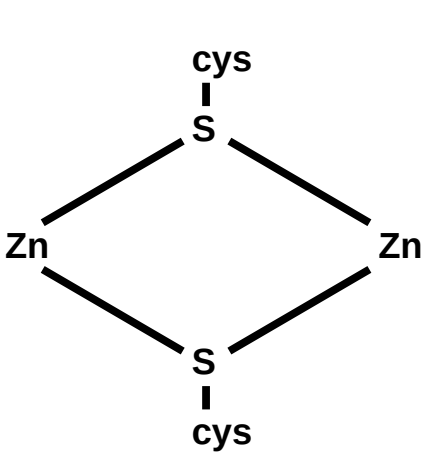
In 1sml there are two water molecules shared , giving a Zn…Zn distance 3.5 Å.

In five cases with Zn…Zn in the range 4 - 6 Å there was no obvious shared group, and in 1b0n an imidazole appears to be shared, giving Zn…Zn 5.8 Å.

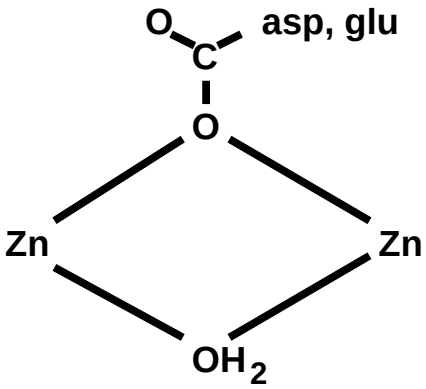
There are fewer examples of close Ca…Ca approaches, and they all involve the sharing of carboxylate groups as in b), c) or d).

When two or three carboxylate groups are shared, Ca…Ca distances are 3.9 - 4.5 Å: 1acc, 1e43 (2) , 1q0h, 1nls, 1sac, 2msb.

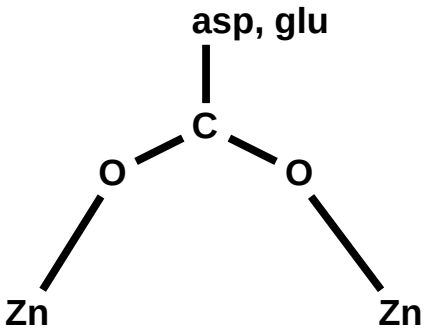
When only one carboxylate group is shared the distances are longer: 4.7-5.2 Å in 1kap (4 examples), 6.6 Å in 1ava.



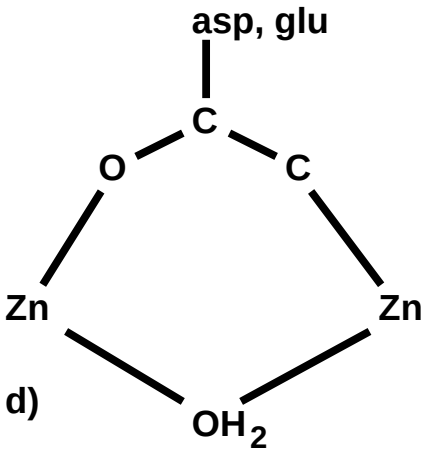
a)



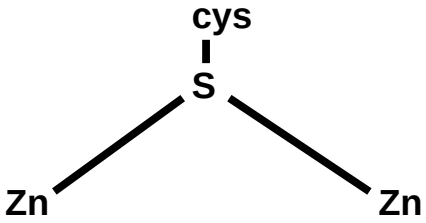
b)



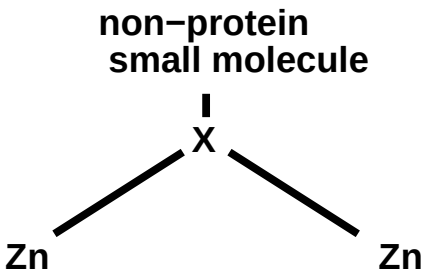
c)



d)



e)



f)

Fig. 5D