

Table S1. Crystallographic data of the test structure with a number of non-hydrogen atom in the asymmetric unit (N_{asym}) lower than 80 (small-size molecules). ID is the identification number in Crystallography Open Database or the reference number.

Name	ID	Space group	N_{asym}	Cell content
alley1	2104682	P -1	66	$C_{112} H_{104} O_{16} Fe_4$
alley	2225745	R 3c	54	$C_{828} H_{720} N_{72} O_{36} Fe_{36}$
allof	2230280	P 6 ₁	22	$C_{78} H_{126} N_6 O_{48}$
amide	(1)	P bc2 ₁	24	$C_{56} H_{72} N_{24} O_{16}$
apapa	(2)	P 4 ₁ 2 ₁ 2	69	$C_{240} H_{296} N_{168} O_{128} P_{16}$
av2pr	2103455	P 6 ₁	56	$C_{205} H_{374} N_{48} O_{83}$
ax118	(3)	P ccn	25	$C_{152} H_{168} N_{16} O_{24} Cl_8$
azet	(2)	P ca2 ₁	48	$C_{168} H_{128} N_8 O_8 Cl_8$
bats	2226585	P -42 ₁ c	16	$C_{120} H_{104} I_8$
becker	2017537	P 6 ₃ /m	40	$C_{354} H_{432} N_{18} O_{60} P_{12} Cl_{30} C_6$
bed	(2)	I 4	34	$C_{208} H_{208} N_{32} O_{32}$
blake	2016921	C 2/c	72	$C_{300} H_{220} N_{60} O_{160} Cl_{40} Pb_{20}$
bobby	(2)	P 2 ₁ 3	5	$C_{24} H_{24} N_4 O_{24} Na_4 Ca_4$
butyl	2222707	R -3	22	$C_{342} H_{342} N_{18} O_{36}$
carbamoto	2226134	I 2/a	59	$C_{344} H_{320} N_{32} O_{48} P_{16} S_{16} Au_{16}$
carbamide	2219630	P n	15	$C_{24} H_{20} N_4 O_2$
cephal	(3)	C 2	46	$C_{144} H_{168} N_8 O_{32}$
choror	2104300	P 3 ₂ 2 ₁	54	$C_{264} H_{180} N_{12} O_{48}$
cime	(4)	C c	18	$C_{40} H_{72} N_{24} O_4 S_4$
cubane	2222097	I -42d	27.5	$C_{316} H_{288} N_{32} O_{76} Zn_{16}$
cuimid	(2)	P 3 ₂ 2 ₁	12	$C_{36} H_{48} N_{24} Cl_6 Cu_6$
diam	(2)	P 4 ₂ /n	15	$C_{112} H_{160} O_8$
diben	2228788	F dd2	32	$C_{448} H_{448} O_{32} S_{32}$
dirole	(2)	I -42d	12.25	$C_{160} H_{288} O_{36}$
eg1scrat	(5)	C 12/c1	12	$O_{60} Cs_8 Si_{24} Ti_4$
ergo	(6)	P 2 ₁ 2 ₁ 2 ₁	58	$C_{224} H_{352} O_8$
erica	(7)	P 2 ₁	43	$C_{74} H_{86} O_8 P_2 Fe_2$
euclor	(8)	C 2/c	21	$O_{104} Na_8 S_{24} K_8 Cu_{24}$
ewa	2017580	I mmm	6	$Mg_{14} Zn_{50} Ce_{32}$
fegas	(9)	P 6 ₃ /mmc	0.75	$S_{10} Fe_4 Ga_4$
fisk	2017323	P 6 ₃ /mmc	0.42	$P_4 Ca_2 In_4$
flubo	2017493	P -31c	2.17	$B_4 O_{12} F_2 K_2 Zn_4 Cd_2$
fnltk	(10)	I 2/c	44	$C_{304} H_{192} N_{40} As_8$
freies	(11)	P 2 ₁ /a	6	$S_{12} Ag_4 Sb_4 Pb_4$
furan	2222612	P nma	4	$C_{16} H_{12} O_{12} Cl_4$
giac	(12)	P 2 ₁ /c	21	$C_{68} H_{32} N_4 O_8 S_4$
gold	(2)	C c	56	$C_{224} H_{128}$
golen	2230759	I 4 ₁ /a	44	$C_{640} H_{704} N_{32} Cl_{32}$
gra4	(3)	P -1	36	$C_{60} H_{44} N_4 O_8$
hexakis	2226269	C mca	28.5	$C_{96} H_{400} N_{48} O_{232} V_{80}$
hfac	2103750	I 4 ₁ cd	33	$C_{224} H_{96} N_{32} O_{64} F_{192} Cu_{16}$

hkw0m	2016712	P n2 ₁ m	7	C ₄ H ₂₄ N ₈ O ₈ Na ₄ Cl ₄
hov1	(2)	C 2/m	15.5	Si ₃₆ Pr ₅₆ Ni ₃₂
hwk3	2016711	I 2	7	C ₄ H ₂₄ N ₈ O ₈ Na ₄ Cl ₄
indide	2225931	I 4/mcm	4.38	Ru ₄₁ In ₁₅ La ₈₄
inos	(13)	P 2 ₁ /n	26	C ₄₈ H ₁₁₂ O ₅₆
jing	2230031	I 4	33.5	C ₁₉₂ H ₁₇₆ N ₂₄ O ₄₄ Cu ₈
kirillov	2218097	F d-3m	2	C ₃₈₄ H ₇₆₈ N ₁₉₂ O ₁₉₂ P ₁₉₂ Cl ₁₉₂ Cu ₁₉₂
kumar	2227526	P -42 ₁ c	38.38	C ₂₆₈ H ₂₄₆ N ₁₆ O ₂₃
lei	2225988	P bca	52	C ₃₀₄ H ₄₃₂ N ₄₈ O ₄₈ S ₁₆
litho	(14)	P 2 ₁ 2 ₁ 2 ₁	27	C ₉₆ H ₁₂₀ O ₁₂
loganin	(2)	P 2 ₁ 2 ₁ 2 ₁	27	C ₆₈ H ₁₀₄ O ₄₀
mag255	(15)	P 2 ₁ 2 ₁ 2 ₁	68	C ₂₂₄ H ₂₄₈ O ₂₄ N ₂₄
mbh2	(2)	P 1	54	C ₄₅ H ₇₂ O ₉
mema	2223241	R 3c	14.11	C ₂₁₆ H ₄₀₀ N ₁₈ O ₂ Cl ₁₈
mlx12b	2017265	I 4/m	2.38	O ₂₈ H ₁₆ P ₄ V ₄ Co ₂
munic	(2)	C 2	40	C ₁₆₀ H ₁₂₈
nabei_1	2016619	C c	67	B ₈ C ₁₈₀ H ₁₂₄ N ₃₆ O ₈ F ₃₂ Co ₄
nabei_2	2016620	C 2/c	74	B ₂₄ C ₃₆₄ H ₂₆₀ N ₇₆ O ₂₄ F ₉₆ Co ₈
newqb	(2)	P -1	62	C ₉₆ H ₈₀ N ₈ O ₂₀
no55	(2)	F dd2	24	C ₃₂₀ H ₃₈₄ N ₆₄
pge2	(2)	P 1	25	C ₂₀ H ₃₂ O ₅
photo	(16)	P 2 ₁ 2 ₁ 2 ₁	17	C ₄₈ H ₅₂ N ₄ O ₁₆
picol	2217894	P 6 ₅ 22	41.5	C ₂₈₈ H ₂₇₆ N ₄₈ O ₁₃₈ K ₁₂ Tb ₁₂
pocro	(17)	B 112/m	3.5	K ₂ Se ₁₆ Cr ₁₀
pyrid	2219167	I 4 ₁ /acd	10.5	C ₂₀₈ H ₁₄₄ N ₈₀ O ₃₂ Mn ₁₆
quinol	(2)	R -3	24	C ₃₂₄ H ₃₂₄ O ₁₀₈
quinone	2103571	P ccn	8	C ₄₀ H ₃₂ O ₈ S ₁₆
ramesh	2219800	R -3	40	C ₅₅₈ H ₆₁₂ N ₃₆ O ₁₂₆
rauf4n	2229326	P 4 ₃ 2 ₁ 2	33	C ₂₄₀ H ₃₆₈ O ₂₄
rc62	(3)	P 2 ₁	80	C ₁₃₆ O ₂₄
rifolo	(18)	P 2 ₁	53	C ₇₈ H ₉₈ N ₂ O ₂₆
ringe	2016743	P 4/nmm	0.44	Cu ₁ Sb ₄ U ₂
sahlberg	2016995	P 4/mbm	0.62	Mg ₂ Sc ₄ Ga ₄
salex	(19)	P -3	15.83	H ₆₀ O ₆₉ Na ₄ S ₁₂ K ₄ Fe ₆
schwarz	(20)	P 1	73	C ₄₆ H ₇₀ O ₂₇
seidel	2228989	P 4 ₂ /n	49	C ₃₂₀ H ₁₉₂ N ₆₄ Zn ₈
selenid	(2)	P 2 ₁	25	C ₄₄ H ₅₆ O ₄ Se ₂
skn1	(21)	P 3 ₁	13	C ₂₁ H ₄₈ N ₃ O ₁₂ Cl ₃
sliva	2225554	P -43n	24.5	C ₃₃₆ H ₅₀₄ N ₇₂ O ₁₂₀ Na ₆ P ₂₄ S ₂₄ Nd ₆
solokha	2016696	F -43m	5	Mg ₇₆ Zn ₃₂₄ Ce ₈₀
suoa	(2)	P 2 ₁ 2 ₁ 2 ₁	47	C ₁₁₂ H ₁₅₂ O ₇₆
sutovi	2016784	P 6 ₃ /m	7	H ₃ O ₄₁ Cl ₇ As ₁₂ Cd ₂₄
teoh	2226240	I -4	50	C ₃₃₆ H ₃₂₀ O ₄₈ Sn ₁₆
tetrakis	2219750	I -43d	19.58	C ₆₇₂ H ₄₈₀ N ₁₉₂ O ₄₈ P ₈ S ₄ Ni ₁₆
thiourea	2230193	I bca	18	C ₂₂₄ H ₃₀₄ N ₄₈ S ₁₆
tiny3	(22)	P 2 ₁ 2 ₁ 2 ₁	33	C ₈₈ H ₁₀₀ Cl ₄ N ₈ O ₃₂
togr44	2017713	C 2/c	59	C ₃₉₂ H ₄₉₆ O ₃₂ Cl ₄₈
totc	(2)	P 6 ₁	42.5	C ₂₁₉ H ₂₆₀ O ₃₆
tpala	(2)	P 2 ₁	39	C ₅₆ H ₈₄ N ₈ O ₁₄
tph	(2)	B 22 ₁ 2	39	C ₂₈₈ H ₂₄₀ N ₂₄

triosic	2101410	C 222 ₁	43.5	C ₂₂₄ H ₃₀₈ N ₅₆ O ₅₆ S ₈ Cl ₄
tungsto	2218160	R 3	61.33	C ₃₂₄ H ₂₁₆ N ₅₄ O ₁₂₆ Fe ₁₂ W ₃₆
tur10	(2)	P 6 ₃ 22	17.33	C ₁₈₀ H ₂₈₈ O ₂₈
wall26	2218446	P -43n	8	C ₁₂₈ H ₂₈₈ O ₃₂ Tl ₃₂
xanthene	2220578	P na2 ₁	67	C ₂₂₀ H ₃₂₀ N ₄ O ₄₄
y75	(23)	P 21/n	34	C ₁₀₀ H ₁₃₆ N ₁₂ O ₂₄
yah140	2230578	P 4/n	19.5	C ₉₆ H ₈₀ N ₃₂ O ₂₄ Mo ₂ Tb ₂
zhang	2222739	C 2/c	64.5	C ₃₀₄ H ₄₀₈ N ₄₈ O ₁₄₈ S ₈ Mn ₈

- (1) D. Viterbo, personal communication;
- (2) see reference quoted by Sheldrick, G.M. (1982). *Testing Structures for Direct Methods*. Univ. of Cambridge, UK.
- (3) crystallographic group at York University, personal communication;
- (4) Kojić-Prodić, B., Ružić-Toroš, Ž., Bresciani-Pahor, N. & Randaccio, L. (1980). *Acta Cryst.* **B36**, 1223–1225;
- (5) D. Watkin, personal communication;
- (6) Hull, S.E., Leban, I., Main, P., White, P.S. & Woolfson, M.M. (1976). *Acta Cryst.*, **B32**, 2374-2381;
- (7) D. Watkin, personal communication;
- (8) Scordari, F. & Stasi, F. (1990a). *Neues Jahrb Miner. Abh.*, **161**, 241-253;
- (9) Cascarano, G., Dogguy-Smiri, L. & Dung, N.-H. (1987). *Acta Cryst.* **C43**, 2050–2053;
- (10) Fares & Flamini, personal communication;
- (11) Ito, T. & Nowacki, W. (1974). *Z. Kristallogr.*, **139**, 85-102.
- (12) F. Stasi, personal communication;
- (13) Langs, personal communication;
- (14) Arora, S.K., Germain, G. & Declercq, J.P. (1976). *Acta Cryst.* **B32**, 415–419;
- (15) Graham, M.A., Wadsworth, A.H., Thornton-Pett, M., Carrozzini, B., Cascarano, G.L. & Rayner, C.M. (2001) - *Tetrahedron Letters*, **42**, 2865-2868;
- (16) Karle, I.L., Karle, J. & Estlin, J.A. (1967). *Acta Cryst.* **23**, 494–500;
- (17) Dung Nguyen-Huy, Tien Vo-Van, Behm, H.J. & Beurskens, P.T. (1987). *Acta Cryst.* **C43**, 2258-2260;
- (18) Cerrini, S., Lamba, D., Burla, M. C., Polidori, P. & Nunzi, A. (1988). *Acta Cryst.* **C44**, 489±495.
- (19) Scordari, F. & Stasi, F. (1990b). *Z. Kristallogr.*, **190**, 47-62;
- (20) B. Schweizer, personal communication;
- (21) D. Watkin, personal communication;
- (22) L. Cranswick, personal communication;
- (23) L. Di Costanzo, personal communication.

Table S2. Crystallographic data of the test structure with a number of non-hydrogen atom in the asymmetric unit (N_{asym}) ranging from 81 to 400 (medium-size molecules). ID is the identification number in Crystallography Open Database, the Protein Data Bank code or the reference number. The 24 medium size structures used in the paper IV [Burla, M.C., Carrozzini, B., Cascarano, G.L., Giacovazzo, C., Polidori, G. (2011). *J. Appl. Cryst.*, **44**, 1143-1151] are highlighted in bold.

Name	ID	Space group	N_{asym}	Cell content
alley2	2104681	P bca	158	C ₉₆₀ H ₆₈₈ Na ₈₀ O ₁₈₄ Fe ₄₀
amphisin	2012193	P 2 ₁ 2 ₁ 2 ₁	102	C ₂₆₄ H ₄₆₄ N ₄₈ O ₉₆
at119	2228715	P 2 ₁ 2 ₁ 2 ₁	259	C ₇₆₄ H ₈₉₂ N ₄₈ O ₂₀₈ S ₁₆
babu	2222265	C c	204	C ₅₂₈ H ₅₇₆ N ₉₆ O ₁₉₂
bcdimp	(1)	P 2 ₁	96	C ₁₁₀ H ₁₅₂ N ₈ O ₇₄
bk53m	2228724	P -1	156	C ₂₁₆ H ₁₆₉ N ₅₆ O ₃₃ S ₈
brock	2104241	B 2 ₁	208	C ₃₂₀ H ₇₆₈ N ₆₄ O ₄₁₆ Fe ₃₂
budka	2229223	P 2 ₁ /c	153	C ₅₆₄ H ₈₀₈ O ₄₀ Cl ₈
bulged	1P79	P 4 ₂ 2	138	C ₄₅₆ N ₁₆₈ O ₄₈₈ P ₄₈ K ₂₄
c8new	(2)	P 2 ₁	130	C ₁₉₆ H ₂₇₂ N ₃₂ O ₃₂
cah	(3)	P 2 ₁	155	C ₂₁₀ H ₂₆₀ N ₂₀ O ₃₀ F ₅₀
carbru_1	2016443	P 2 ₁ 2 ₁ 2 ₁	121	C ₄₂₈ H ₄₃₂ N ₁₆ O ₂₀ Cl ₁₆ Ru ₄
carbru_2	2016444	P 2 ₁ 2 ₁ 2 ₁	121	C ₄₂₈ H ₄₃₂ N ₁₆ O ₂₀ Cl ₁₆ Ru ₄
cb3_pap	(4)	P 2 ₁ /n	92	C ₃₀₄ H ₂₈₈ O ₆₄
ceholz	(5)	P 2 ₁	116	C ₁₂₈ H ₁₆₀ N ₁₆ O ₈₈
ceho2z	(5)	P 2 ₁	232	C ₂₆₄ H ₃₀₄ N ₃₂ O ₁₆₈
cemc1z	(6)	P 2 ₁	228	C ₂₄₈ H ₂₉₆ N ₃₂ O ₁₆₈ Cl ₈
cemc2z	(6)	P 2 ₁	207	C ₂₁₀ H ₄₄₄ N ₄ O ₂₀₀
cephaibol_a	1OB4	P 2 ₁ 2 ₁ 2	113	C ₃₀₈ N ₇₂ O ₇₂
cephaibol_c	1OB7	P 2 ₁ 2 ₁ 2 ₁	112	C ₃₀₀ N ₇₂ O ₇₂
cofi5a	2218002	P 2 ₁ /c	107	C ₃₈₈ H ₅₂₀ O ₃₂ Cl ₈
cyclo_bnz	2103881	P 2 ₁	206	C ₂₁₆ H ₂₂₄ N ₈ O ₁₈₉
cyclo_dba	2103880	P 2 ₁ 2 ₁ 2 ₁	219	C ₅₀₈ H ₄₄₈ N ₁₆ O ₃₅₂
dasco6	2200556	P 2 ₁	124	C ₁₆₂ H ₂₇₀ N ₃₈ O ₄₀ S ₈
dext	(3)	P 2 ₁	94	C ₁₁₀ H ₁₅₂ O ₇₄ N ₄
diene	2226177	P 2 ₁ 2 ₁ 2 ₁	90	C ₂₆₄ H ₄₈₀ B ₁₂ F ₄₈ P ₂₄ Rh ₁₂
dioxolan	2011222	P 2 ₁	141	C ₂₂₀ H ₂₅₂ N ₁₂ O ₄₈ Ru ₂
dodeca	(7)	P 2 ₁ 2 ₁ 2 ₁	200	C ₅₆₈ H ₉₁₂ N ₉₆ O ₁₃₆
echino	2101407	p 65	89	C ₃₆₃ H ₄₈₆ N ₆₀ O ₁₀₂ S ₁₂
ergop21	(8)	P 2 ₁	116	C ₂₂₄ H ₃₅₂ O ₈
gago	2227102	P 2 ₁ /n	129	C ₂₄₀ H ₂₈₈ N ₄₈ O ₁₅₆ S ₂₄ Cl ₂₄ Mo ₂₄
giorgi	(9)	P 2 ₁	88	C ₁₄₈ H ₂₂₀ N ₄ O ₂₄
glici2	(10)	C 2	139	C ₄₈₀ H ₇₆₀ O ₇₆
gna	2WNA	C 222 ₁	146	C ₄₅₆ O ₃₈₄ N ₁₉₂ P ₉₆ Co ₁₆ Mg ₁₆ Br ₈
guest	2202059	P 2 ₁ 2 ₁ 2 ₁	105	C ₂₂₄ H ₄₄₀ O ₁₈₈ S ₈
h104	(11)	P 2 ₁ /c	84	C ₂₅₂ H ₃₈₄ N ₁₂ O ₆₀ Cl ₁₂
hao	2226460	A ba2	84	C ₄₁₆ H ₄₉₆ N ₈₈ O ₁₄₄ S ₁₆ Gd ₈
helix	1VRZ	C 2	164	C ₄₇₂ H ₇₃₆ N ₉₂ O ₉₂
hepta	2103676	P 2 ₁ 2 ₁ 2 ₁	109	C ₂₈₈ H ₅₅₀ O ₁₅₀
hexa	2103675	P 2 ₁ 2 ₁ 2 ₁	100	C ₂₅₂ H ₅₉₉ O ₁₄₉

hsy115	2225069	P na ₂ ₁	120	C ₃₈₄ H ₄₈₀ N ₃₂ O ₆₄
iak48bm	2225123	P 2 ₁ /c	140	C ₄₁₆ H ₂₉₆ O ₉₆ P ₈ As ₁₆ Ru ₂₄
iled	(12)	P 2 ₁ 2 ₁ 2 ₁	84	C ₂₄₀ H ₄₀₈ N ₂₄ O ₇₂
inclu_rt	(13)	P 1	194	C ₉₇ H ₁₅₇ O ₉₇
jamilas	(14)	P 1	100	C ₆₄ H ₆₈ O ₂₀ N ₈ S ₄ K ₄
jebas	2220423	P 2 ₁ /c	121.5	C ₃₄₂ H ₄₅₆ N ₃₂ O ₈₀ S ₃₂
kemer	2015969	C 2/c	126	C ₆₄₈ H ₁₀₅₆ N ₁₆₈ O ₉₆ S ₇₂ Cr ₁₂ Er ₁₂
kfb44	2103879	P -1	368	C ₆₅₆ H ₅₇₆ N ₁₆ O ₃₂ Si ₃₂
lasso	2NJW	P 2 ₁ 2 ₁ 2 ₁	146	C ₆₀₈ N ₉₂ O ₁₀₄ S ₈
macro	(15)	P 2 ₁ 2 ₁ 2 ₁	123	C ₄₂₀ H ₄₃₂ O ₇₂
mghex	(16)	P 3 ₁	95	C ₁₆₈ H ₂₄₀ N ₄₈ O ₆₀ Cl ₆ Mg ₃
mor59	(17)	P 2 ₁	126	C ₁₉₂ H ₃₂₈ N ₂₄ O ₃₆
oe410	2204692	P 2 ₁ /n	135	C ₄₃₂ H ₃₂₀ N ₄₈ O ₁₂ P ₁₆ Te ₁₆ Re ₁₆
ohba	2202173	P 2 ₁	92	C ₁₆₂ H ₁₇₄ N ₂ O ₂₀
ohba_p1	2203226	P 1	188	C ₁₆₆ H ₁₅₈ N ₂ O ₂₀
pnib	1AKG	P 2 ₁ 2 ₁ 2 ₁	114	C ₂₈₀ N ₇₂ O ₈₈ S ₁₆
quail_a	2012678	P 2 ₁ 2 ₁ 2 ₁	81	C ₂₂₀ H ₃₉₆ N ₃₆ O ₆₈
quail_b	2012679	P 2 ₁ 2 ₁ 2 ₁	83	C ₂₂₈ H ₄₁₂ N ₃₆ O ₆₈
rac	2211608	P -1	93	C ₁₄₀ H ₁₃₈ N ₂₂ O ₁₆ Cl ₄ Cu ₄
rauf20	2220764	P 2 ₁ /n	104	C ₂₃₂ H ₁₀₄ N ₃₂ O ₁₆ S ₁₆ Cl ₁₂₀
shen	2230103	P c	90	C ₁₂₈ H ₁₈₄ N ₈ O ₄₀ Co ₄
so99	2016923	P -1	144	C ₂₂₄ H ₂₃₂ N ₃₂ O ₃₂
sr166	(18)	P 2 ₁	112	C ₁₉₂ H ₁₉₂ O ₁₆ N ₁₆
s6	(14)	P -1	89	C ₁₃₂ H ₁₄₁ N ₁₀ O ₂₄ S ₁₂
ta	(19)	P 2 ₁	142	C ₁₇₀ H ₂₂₄ O ₁₁₄
tato	(20)	P 2 ₁ 2 ₁ 2 ₁	109	C ₂₂₄ H ₃₆₈ N ₈ O ₂₀₄
tb	(19)	P 2 ₁ 2 ₁ 2 ₁	186	C ₅₂₀ H ₇₈₄ O ₂₂₄
tb02rlm	(21)	P 1	197	C ₉₆ H ₂₂₄ O ₁₀₁
tensin	2010528	P 2 ₁ 2 ₁ 2 ₁	107	C ₂₈₄ H ₅₀₈ N ₄₈ O ₉₆
theo	2227537	P 2 ₁ 2 ₁ 2 ₁	110.8	C ₂₈₀ H ₅₀₀ N ₅₂ O ₁₁₁
thebo	2008512	P 2 ₁ /c	112	C ₃₆₀ H ₃₆₈ N ₄₈ O ₄₀
thio2	2104680	P 2 ₁ /c	144	C ₄₄₄ H ₂₈₈ N ₁₂ F ₇₂ P ₂₄ S ₁₂ Ag ₁₂
tp	(19)	P 2 ₁	161	C ₂₁₀ H ₃₀₈ O ₁₁₂
triazole	2221962	P -1	243	C ₃₇₈ H ₄₁₄ N ₇₂ S ₁₈ Cl ₁₈
trip04e	2100208	P 2 ₁	223.5	C ₂₈₃ H ₂₉₅ O ₁₆₀ Cl ₄
triostin_a	2101409	P 2 ₁ 2 ₁ 2 ₁	94	C ₂₅₆ H ₃₆₀ N ₄₈ O ₆₄ S ₈
triostin_c	2101408	P 2 ₁ 2 ₁ 2 ₁	83.5	C ₂₁₉ H ₂₈₃ N ₄₈ O ₅₁ S ₈ Cl ₈
tval	(16)	P 1	156	C ₁₀₈ H ₁₈₀ N ₁₂ O ₃₆
winter	(16)	P 2 ₁	88	C ₁₁₀ H ₁₇₈ N ₂₂ O ₃₂ Cl ₁₂
x124	(22)	P 1	120	C ₉₂ H ₇₆ N ₄ O ₁₂ S ₄ Cl ₈
y54	(23)	P 2 ₁ 2 ₁ 2 ₁	100	C ₂₆₄ H ₃₄₄ N ₅₆ O ₈₀
zaib	(24)	P -1	141	C ₁₉₃ H ₃₁₆ N ₃₆ O ₅₃
zippercl	2WQ0	I 2 ₁ 3	275	C ₄₀₀₈ N ₁₂₇₂ O ₁₂₂₄ S ₂₄ Cl ₇₂
zipperbr	2WQ1	I 2 ₁ 3	275	C ₄₀₀₈ N ₁₂₇₂ O ₁₂₂₄ S ₂₄ Br ₇₂

- (1) Impellizzeri, G., Pappalardo, G., D'Alessandro, F., Rizzarelli, E., Saviano, M., Iacovino, R., Benedetti, E. & Pedone, C. (2000). *Eur. J. Org. Chem.*, **6**, 1065–1076;
- (2) Saviano, M., Isernia, C., Rossi, F., Di Blasio, B., Iacovino, R., Mazzeo, M., Pedone, C. & Benedetti, E. (2000). *Biopolymers*, **53**, 189–199.
- (3) L. Toupet, personal communication;
- (4) R. Spagna, personal communication;
- (5) Onagi, H., Carrozzini, B., Cascarano, G.L., Easton, C.J., Edwards, A.J., Lincoln, S.F. & Rae, A.D. (2003). *Chem. Eur. J.*, **9**, 5971 – 5977;
- (6) A.J. Edwards, personal communication;
- (7) Saviano, M., Improta, R., Benedetti, E., Carrozzini, B., Cascarano, G. L., Didierjean, C., Toniolo, C. & Crisma, M. (2004). *ChemBioChem.*, **5**, 541–544.
- (8) G. Leban, personal communication;
- (9) M. Camalli, personal communication;
- (10) J.F. Piniella, personal communication;
- (11) J.C. Bryan, personal communication;
- (12) Pletnev, V.Z., Galitskii, N.M., Smith, G.D., Weeks, C.M. & Duax, W.L. (1980). *Biopolymers*, **19**, 1517-1534.
- (13) Makedonopoulou, S. & Mavridis, I.M. (2000), *Acta Cryst. B***56**, 322-331;
- (14) D. Watkin, personal communication;
- (15) Dondoni, A., Marra, A., Scherrmann, M. C. & Bertolasi, V. (2001). *Chem. Eur. J.*, **7**, 1371–1382;
- (16) see reference quoted by Sheldrick, G.M. (1982). *Testing Structures for Direct Methods*. Univ. of Cambridge, UK.
- (17) Polese, A., Formaggio, F., Crisma, M., Valle, G., Toniolo, C., Bonora, G.M., Broxterman Q.B. & Kamphuis, J. (1996). *Eur. J. Chem.*, **2**, 1104-1111.
- (18) M. Thorton-Pett, personal communication;
- (19) K. Gessler, personal communication;
- (20) Tellini, V.H.S., Jover, A., Galantini, L., Meijide, F. & Tato, J.V. (2004), *Acta Cryst. B***60**, 204-210;
- (21) H. Gronitzka, personal communication;
- (22) R.C. Hatiwang, personal communication;
- (23) L. Di Costanzo, personal communication;
- (24) Geßmann, R., Brückner, H. & Kokkinidis, M. (1998). *Acta Cryst. B***54**, 300-307;

Table S3. Crystallographic data of the protein test structure with a number of non-hydrogen atom in the asymmetric unit (N_{asym}) ranging from 200 to about 2000. HA indicates the number and species of the heavy atoms ($Z \geq 16$) in the asymmetric unit. Data at atomic resolution (up to 1.2 Å).

PDB code	Space group	N_{asym}	HA
1AA5	P 4 ₃ 2 ₁ 2	200	8Cl
1SH0	P 4 ₃ 2 ₁ 2	207	8Cl
1HHY	P 6 ₃ 22	208	4Cl
1P9I	C 222 ₁	213	None
1ICK	P 2 ₁ 2 ₁ 2 ₁	250	None
1A0M	I 4	255	10S
2BF9	C 2	305	1Zn
2ERL	C 2	305	7S
1A7Z	P 2 ₁ 2 ₁ 2 ₁	307	2Cl
1ALZ	P 2 ₁ 2 ₁ 2 ₁	313	None
1CBN	P 2 ₁	330	6S
1HHZ	P 3 ₂ 21	354	6Cl
1BX7	P 4 ₃ 2 ₁ 2	365	11S
2FDN	P 4 ₃ 2 ₁ 2	373	16S 8Fe
8RXN	P 2 ₁	392	6S 1Fe
1IRO	R 3	411	5S 1Fe
1IRN	R 3	411	5S 1Zn
1NKD	C 2	439	4S
1IGD	P 2 ₁ 2 ₁ 2 ₁	468	1 S
1C75	P 2 ₁ 2 ₁ 2 ₁	528	3S 1Fe
1B0Y	P 222	593	9S 4Fe
1CTJ	R 3	672	3S 1Fe
2PVB	P 2 ₁ 2 ₁ 2 ₁	814	3S 2Ca
1A6N	P 2 ₁	892	3S 1Fe

1D4T	P 2 ₁	911	3S
3PYP	P 6 ₃	983	6S
1MFM	P 2 ₁ 2 ₁ 2 ₁	1106	2S 2Cl 1Cu 1Zn 4Cd
1A6G	P 2 ₁	1224	3S 1Fe
1CKU	P 2 ₁ 2 ₁ 2 ₁	1229	18S 8Fe
1SWZ	P 3 ₂ 21	1254	2Cl 3Rb
1I76	P 2 ₁ 2 ₁ 2 ₁	1315	3S 2Ca 2Zn
1JM1	P 6 ₁	1496	10 S 2Fe
1FY2	C 2	1685	12S 1Cl 1Cd
1NLS	I 222	1810	2S 1Ca 1 Mn
1GY0	P 3 ₁	2006	26S 4 Fe