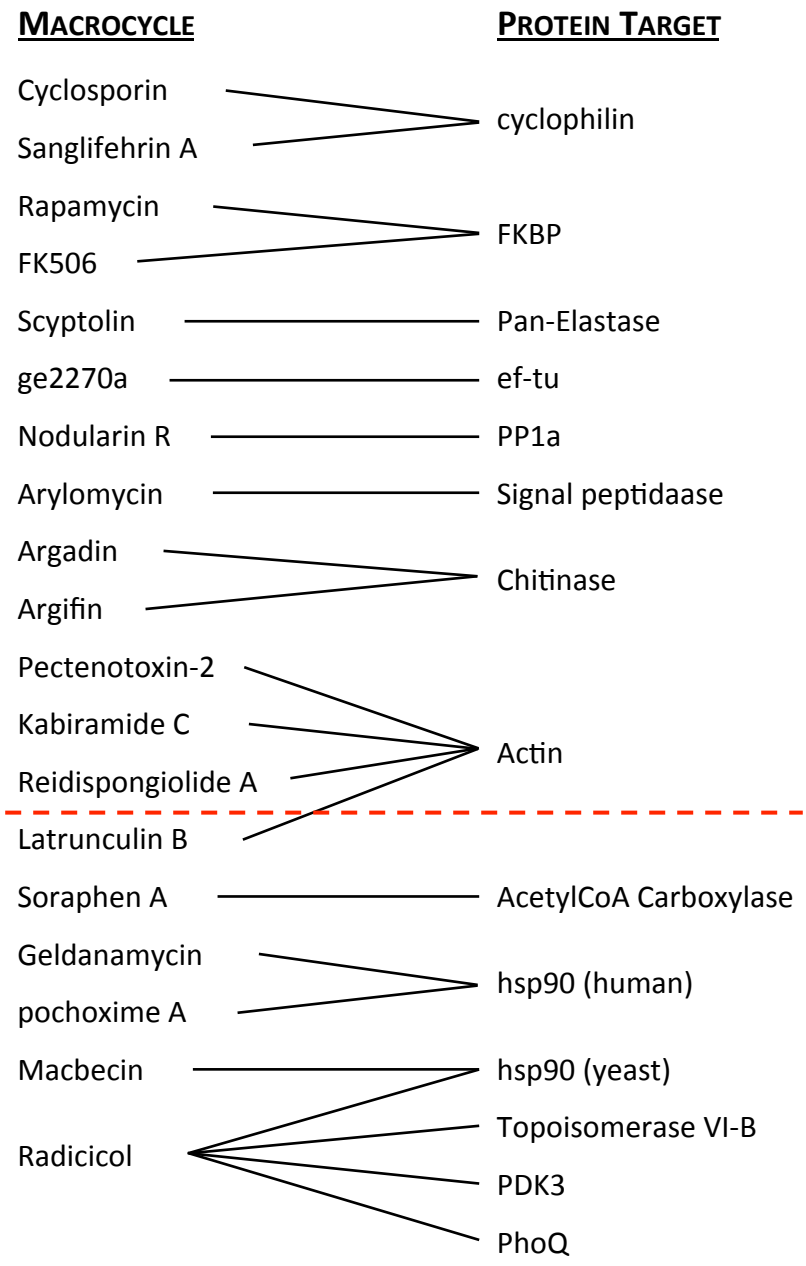
Supplementary Information

How Proteins Bind Macrocycles

Elizabeth A. Villar, Dmitri Beglov, Spandan Chennamadhavuni, John A. Porco Jr., Dima Kozakov*, Sandor Vajda* and Adrian Whitty*

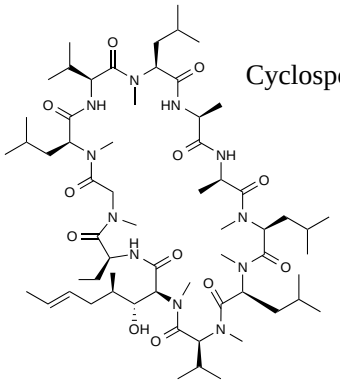
SUPPLEMENTARY RESULTS

Supplementary Figure 1A. The test set comprises 19 unique MCs and 13 unique proteins. MCs above the dashed line have molecular weights >600 Da., and thus are considered as "large MCs" for the purposes of this study. PDB codes for the complexes in the test set are given in Methods.

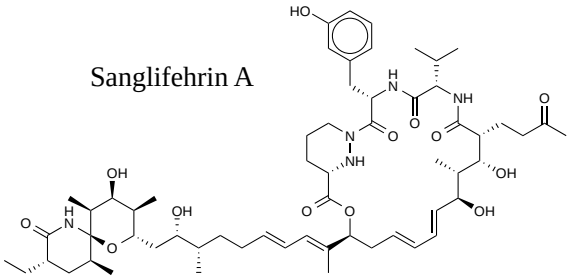


Supplementary Figure 1B. Chemical structures of the 19 macrocycles included in the test set.

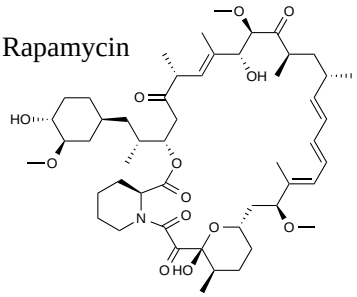
Large Macrocycles (MW > 600 Da)



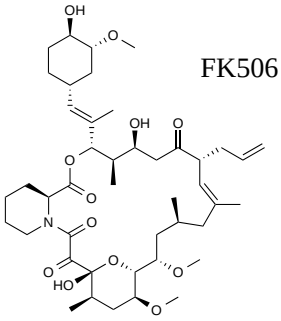
Cyclosporin A



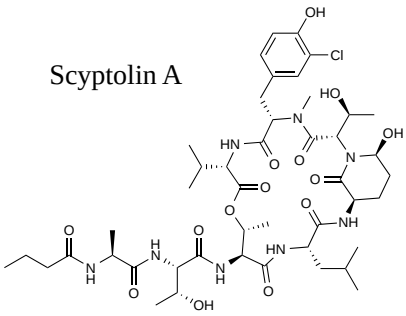
Sanglifehrin A



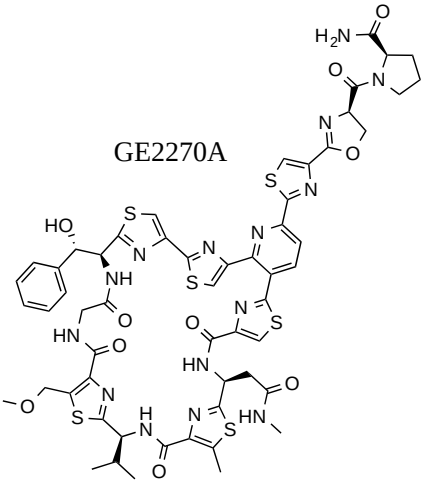
Rapamycin



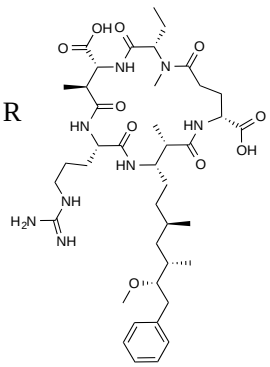
FK506



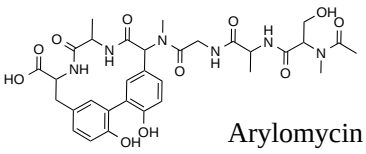
Scyptolin A



GE2270A



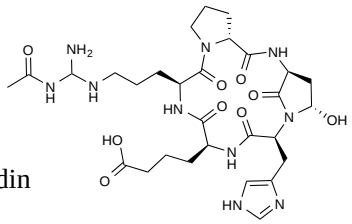
Nodularin R



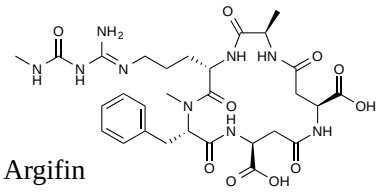
Arylomycin

Supplementary Figure 1B (Continued)

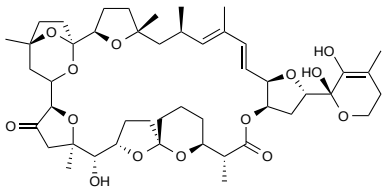
Argadin



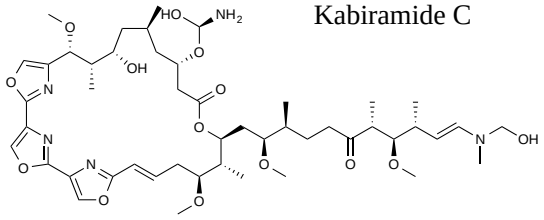
Argifin



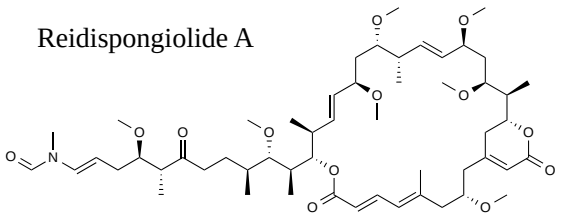
Pectenotoxin-2



Kabiramide C

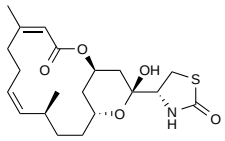


Reidispongiolide A

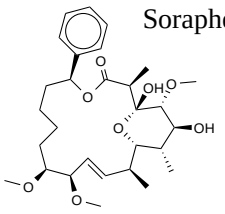


<u>Small Macrocycles (MW < 600 da)</u>

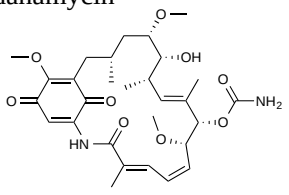
Latrunculin B

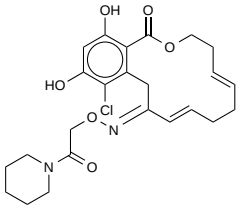


Soraphen A

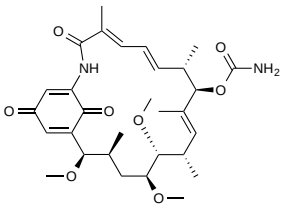


Geldanamycin

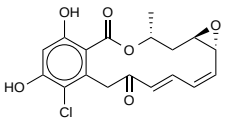




Pochoxime A

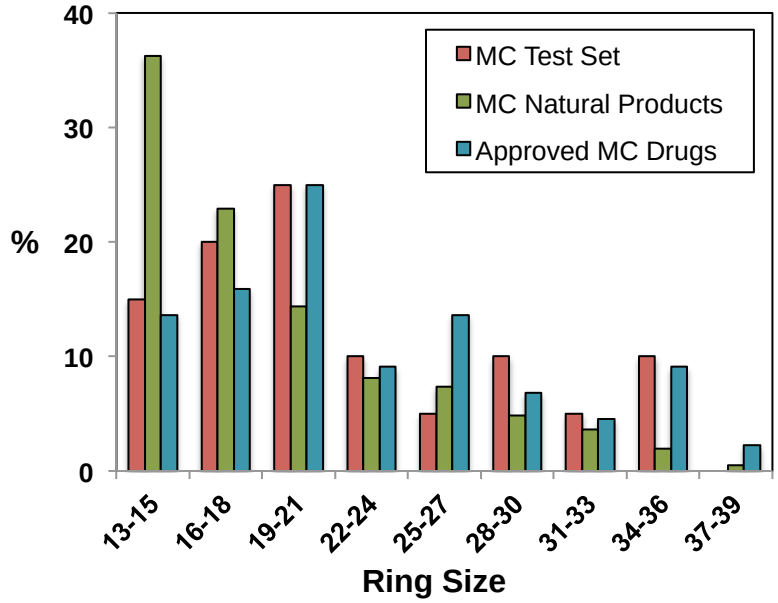


Macbecin



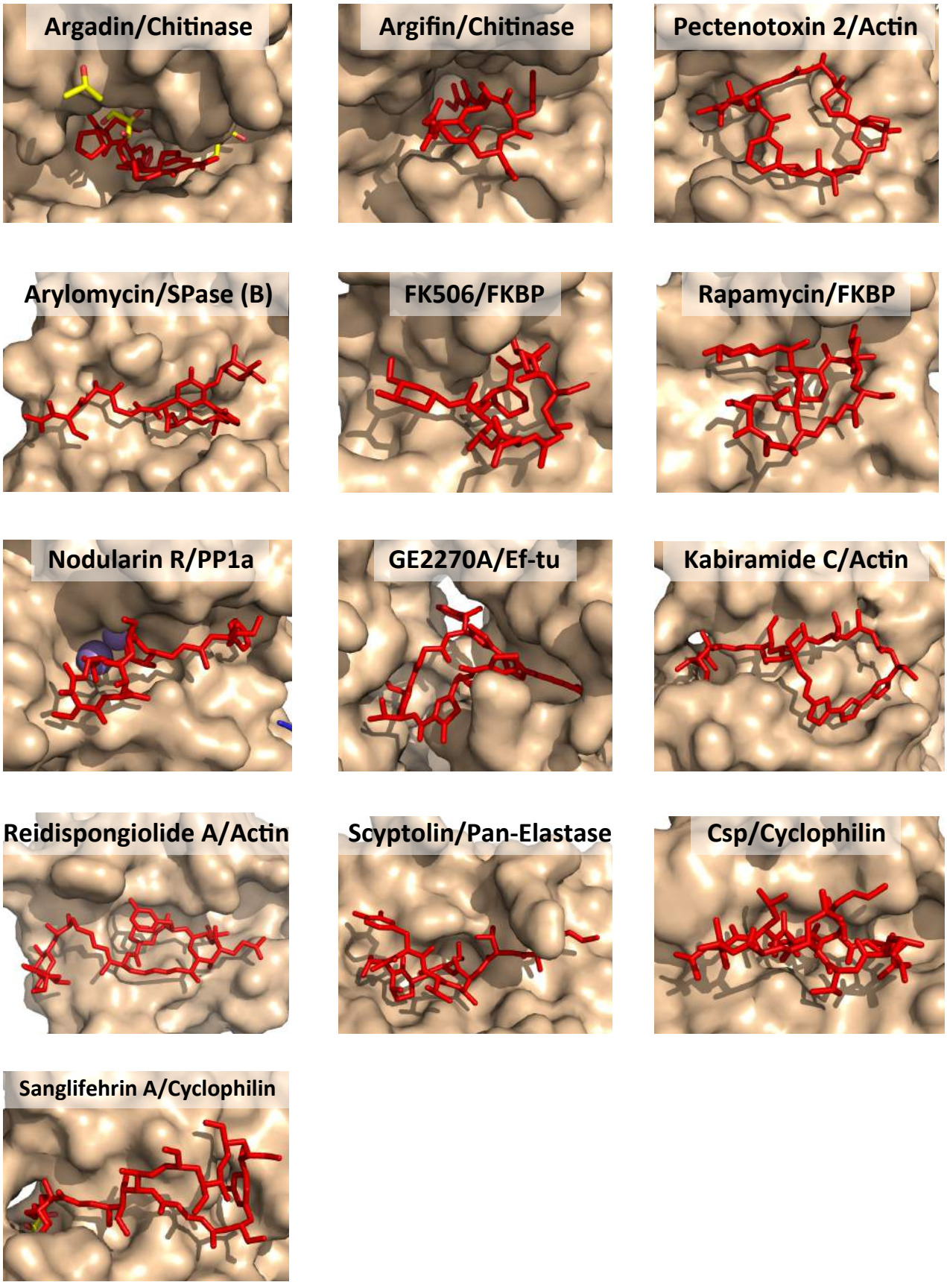
Radicicol

Supplementary Figure 2. Distribution of ring sizes for the MCs in the test set.

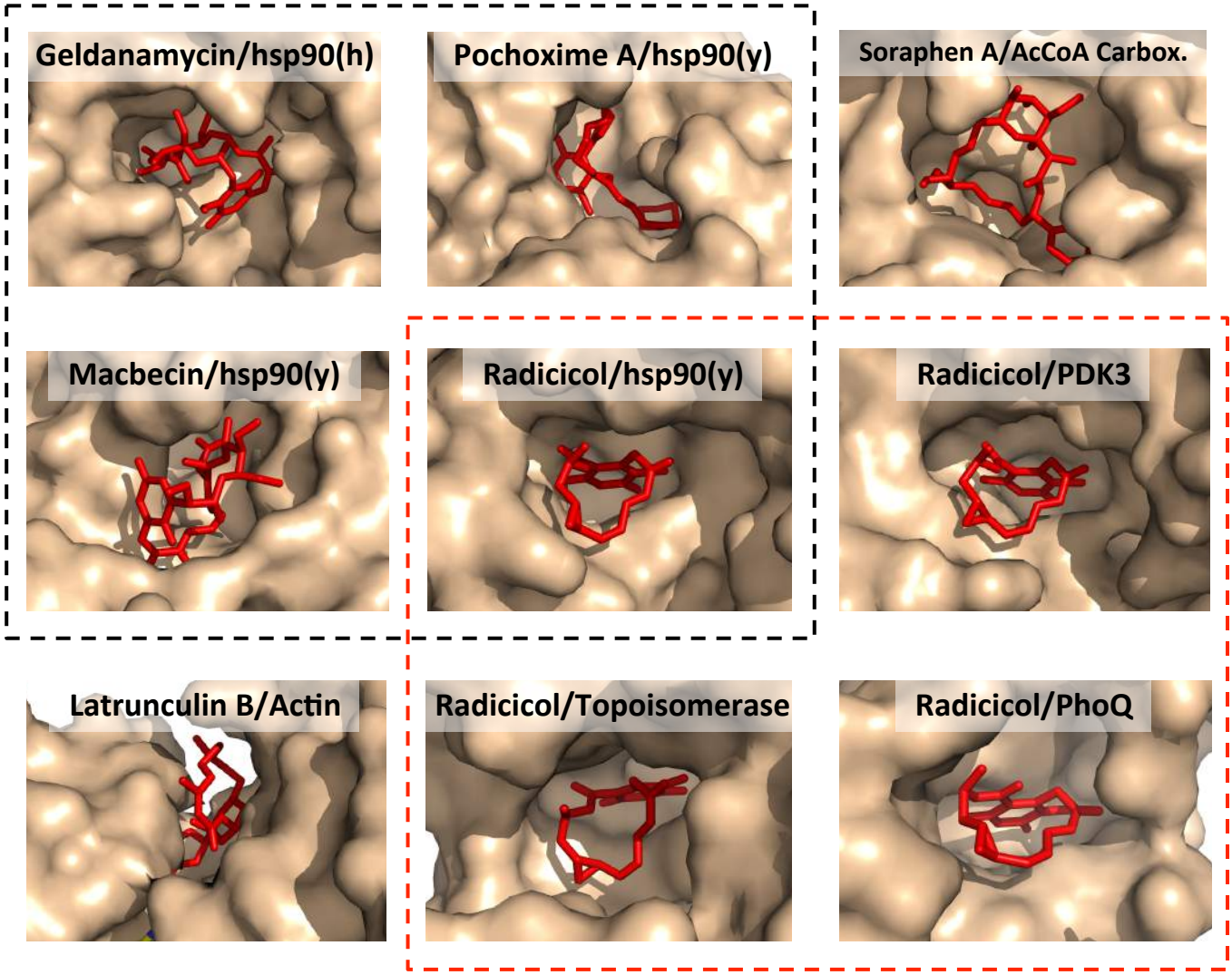


Distribution of ring sizes for the MCs in the test set compared to the 3747 natural product MCs described in reference *(15)* (main text), and the 44 approved MC drugs from Supplementary Table 1. The figure shows that, in terms of ring size, our test set is fairly representative of all natural product MCs, with two exceptions: (1) The smallest range of ring sizes, containing 13-15 atoms, is relatively under-represented in our study set. In the original publication (reference *15*, main text) the authors attributed the abundance of natural product MCs in this ring size range to the great variety of 14-membered macrocyclic terpenoids that exist in nature. The fact that our structurally characterized subset of MCs is not similarly biased towards 14-members rings is, in our opinion, desirable, given that our goal is to characterize protein-MC binding in general, and not to particularly focus on one size range or structural class. (2) Our test set contains a somewhat disproportionate number of very large-ring MCs, with ring size 34-36 atoms. This feature probably arises from the fact that a handful of large MCs such as cyclosporin A and FK506 have been of high interest as pharmaceuticals, leading to a strong drive to elucidate their X-ray crystal structures in complex with their protein targets. In terms of ring size our test set closely mirrors the size distribution found for approved MC drugs, even though only three compounds are common to both sets.

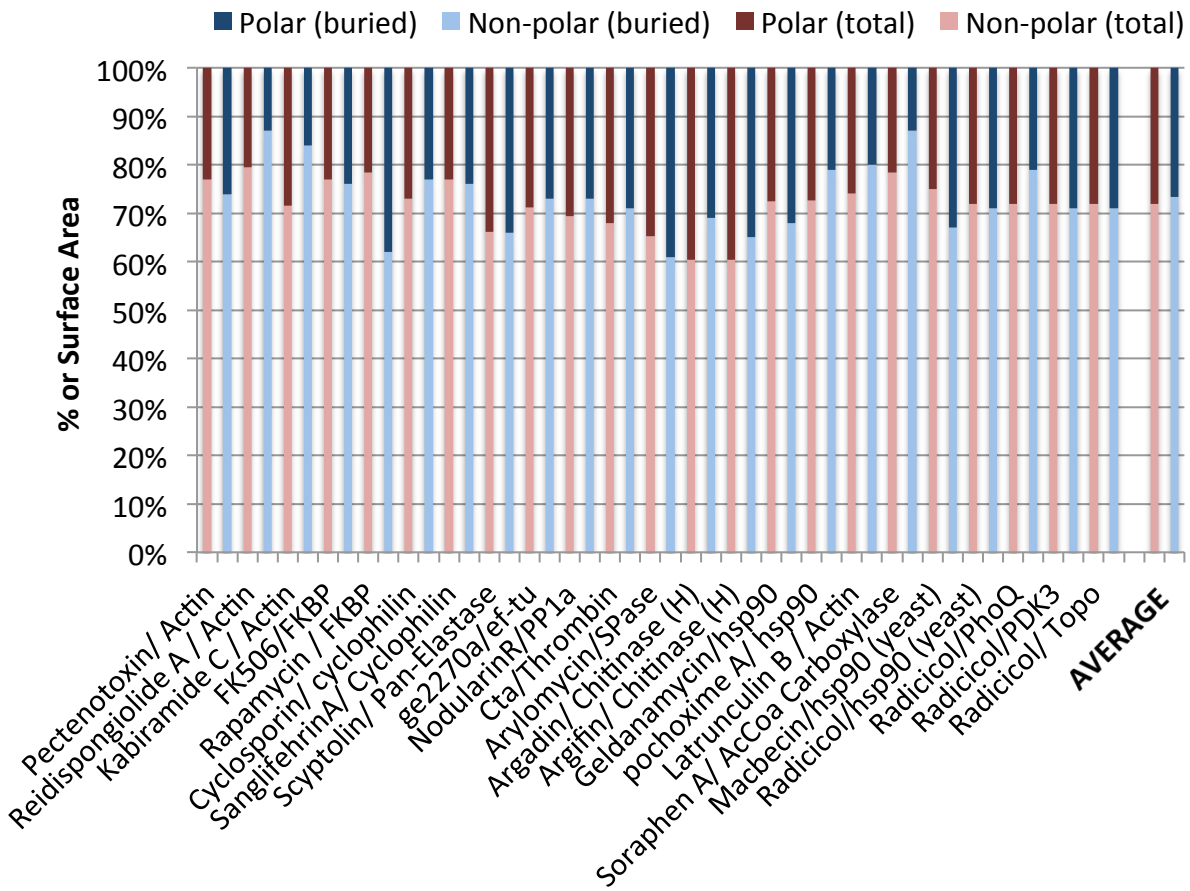
Supplementary Figure 3. Binding geometries for MCs in the test set. (A) Large MCs



Supplementary Figure 3, cont'd. Binding geometries for MCs in the test set. (B) Small MCs. The black dashed box indicates the four distinct MCs for which there are structures bound to hsp90; the red dashed box indicates the four complexes that contain Radicicol.

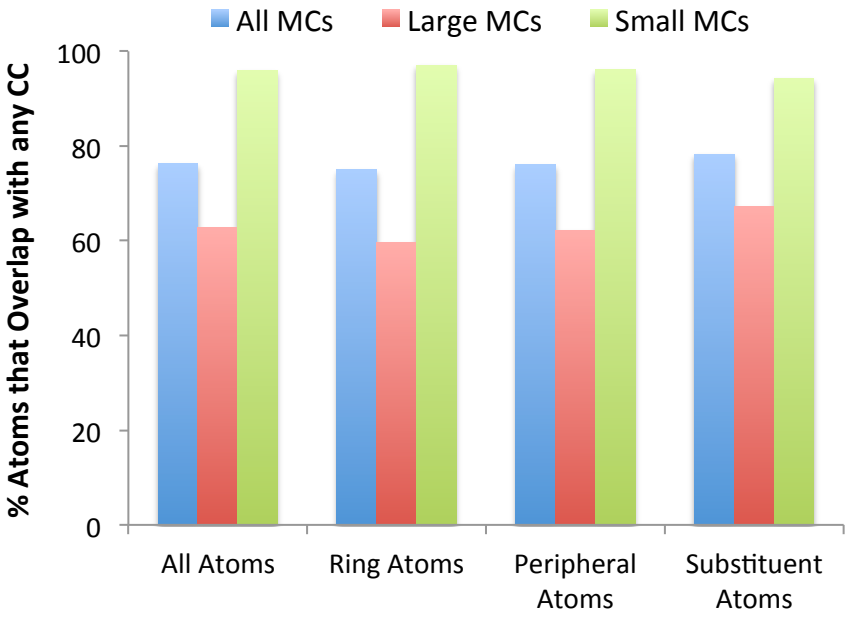


Supplementary Figure 4. Polar/non-polar breakdown for contact atoms versus all MC atoms, shown separately for each MC-protein complex.

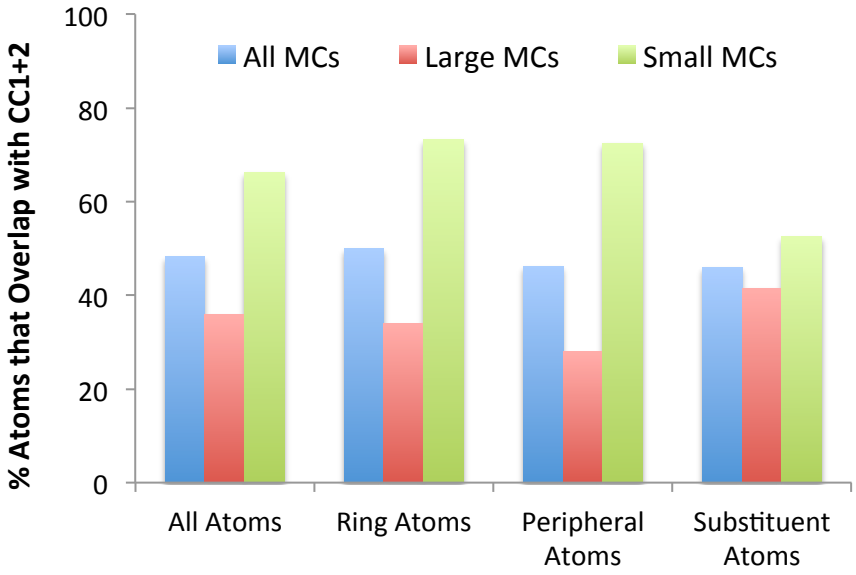


Supplementary Figure 5. Overlap of different regions of MC structure with FTMap Consensus Clusters (CCs)

A. Overlap with all CCs

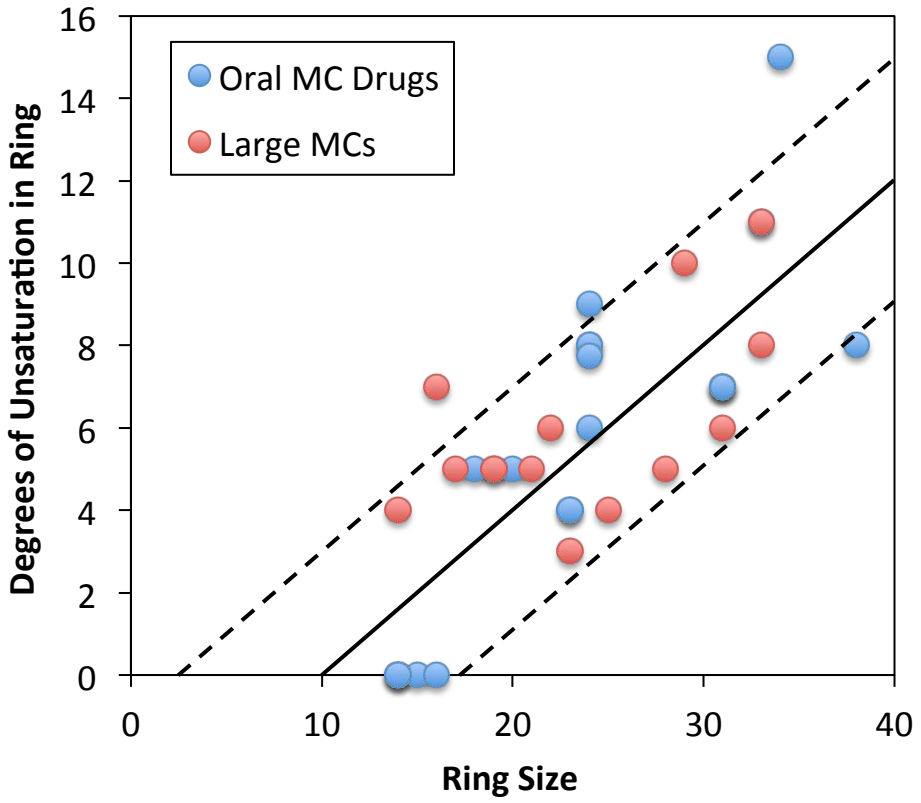


B. Overlap with with the two top-ranking CCs



An MC atom was considered to overlap with a hot spot if (i) the center of mass of the probes in a given CC fell within 2 Å of the MC atom, and (ii) at least 25% of the CC probe atoms were within 1.25 Å of the MC atom. Results were robust with respect to the distance constraint used to define hot spot occupancy, and also with respect to whether the analysis was restricted to the two or three highest-ranking hot spots, or whether all CCs were included.

Supplementary Figure 6. Plot showing how the number of degrees of unsaturation (DOU) in the MC ring varies with ring size (*R*), for the large MCs in the test set and also for oral MC drugs. Degrees of unsaturation include *pi* bonds or amide C-N bonds in the ring, or fused ring systems that inhibit dihedral vibration about a ring *sigma* bond. The data substantiate the intuitively expected trend that larger rings tend to have more DOU. The solid line represents the relationship DOU = 0.4*R* – 4, and the dashed lines show the boundaries for 3 DOU above and below the solid line, illustrating the best description of the relationship given the limited number of data points that are available.



Supplementary Table 1. Predicted properties of 44 macrocycles that are approved drugs.

DRUG	MOL. WT	PROPERTIES (Prediction from CHEMAXON[1])				Rotatable Bonds
		cLogP [2]	PSA (Å^2) [3]	HBA [4]	HBD [5]	
Orally Bioavailable						
Rifaximin	785.4	3.3	205.3	11	5	3
Amphotericin B	923.5	-2.3	324.1	17	12	3
Rifabutin	846.4	4.2	206.8	13	5	5
Rifampin	822.4	2.8	224.2	14	6	5
Sirolimus	913.6	7.5	195.4	12	3	6
Rifapentine	876.5	3.6	221.4	14	6	6
Tacrolimus	803.5	5.6	178.4	11	3	7
Azithromycin	748.5	2.4	182.5	13	5	7
Abamectin	874.5	5.8	170.1	13	3	8
Clarithromycin	747.5	3.2	184.1	13	4	8
Everolimus	957.6	7.4	204.7	13	3	9
Troleandomycin	813.5	4.3	185.4	12	0	12
roxithromycin	836.5	3.0	218.1	16	5	13
vancomycin	1447.4	-4.4	539.5	24	19	13
Erythromycin	861.5	3.4	227.5	14	4	14
Cyclosporin A	1201.8	3.6	278.8	12	5	15
Fidaxomicin	1056.4	8.6	269.5	15	7	15
Desmopressin	1068.4	-6.1	437.2	15	14	19
AVERAGE	**921.4**	**3.1**	**247.4**	**14.0**	**6.1**	**9.3**
Not Orally Bioavailable						
Romidepsin	540.2	1.1	142.7	5	4	2
Ixabepilone	506.3	3.4	112.1	6	3	2
Natamycin	665.3	-1.7	235.4	13	7	3
Nystatin	925.5	-1.9	324.1	17	12	3
Metocurine	652.4	-1.8	55.4	4	0	4
Eribulin	729.4	2.3	148.0	12	2	4
Pimecrolimus	809.4	6.8	158.1	10	2	6
Dalfopristin	690.3	1.6	177.6	9	2	7
Actinomycin D	1254.6	-0.1	355.5	16	5	8
Quinupristin	1021.5	2.2	232.4	12	4	10
Eptifibatide	831.3	-5.1	328.5	12	11	10
Capreomycin	668.3	-11.0	382.5	14	15	10
Viomycin	685.3	-11.0	395.3	15	16	10
Candicidin	1108.6	0.2	368.7	19	11	10
Temsirolimus	1029.6	7.1	242.0	14	4	11
Anidulafungin	1139.5	-1.5	377.4	17	14	14
Octreotide	1018.4	-1.4	335.5	12	13	17
Lanreotide	1095.5	-0.3	358.3	12	13	17
Oxytocin	1006.4	-5.0	401.2	13	12	17
Micafungin	1269.4	-6.3	513.0	22	16	18
Lypressin	1055.4	-6.8	428.8	14	13	19
Caspofungin	1092.6	-4.8	415.3	18	16	23
Polymyxin B	1202.7	-7.2	498.8	18	18	29
Telavancin	1753.6	-6.2	617.5	29	23	30
Bacitracin	1421.7	-7.2	541.0	20	17	31
Daptomycin	1619.7	-9.4	715.0	27	22	35
AVERAGE	**992.1**	**-2.5**	**340.8**	**14.6**	**10.6**	**13.5**

[1]Calculator plug-ins (Marvin 6.0, 2013) were used for structure property prediction, as described in Methods. [2]Log_{10} of the calculated 1-octanol/water partition coefficient. [3]Polar surface area, for the predominant ionization state at pH = 7.0. [4]Number of hydrogen bond acceptors. [5]Number of hydrogen bond donors.

Supplementary Table 2. Structural and physicochemical properties of the MC Test Set compared to macrocyclic drugs and all oral drugs.[a]

Class	Mol. Wt.	cLogP	PSA (Å^2)	HBD	HBA	Rotatable Bonds
MC Test Set (*n* = 19)	**778 ± 265** *(476-1090)* P<0.0001	**2.8 ±3 .9** *(-2.5-5.7)* P=0.37	**212 ± 84** *(104-296)* P<0.0001	**4.9 ± 3.2** *(2-9)* P<0.0001	**10.7 ± 3.4** *(6-13)* P<0.0001	**8.6 ± 5.7** *(1-15)* P<0.0001
Large (*n* = 13)	**915 ± 194** *(676-1202)* P<0.0001	**2.3 ± 4.7** *(-5.1-7.5)* P=1.0	**255 ± 60** *(178-322)* P<0.0001	**6.1 ± 3.2** *(3-10)* P<0.0001	**12.7 ± 1.7** *(11-15)* P<0.0001	**11.2 ± 5.0** *(6-18)* P<0.0001
Small (*n* = 6)	**479 ± 84** *(396-559)* P=0.0021	**3.9 ± 0.6** *(3.5-4.5)* P=0.103	**118 ± 33** *(96-143)* P=0.0278	**2.3 ± 0.8** *(2-2)* P=0.3083	**6.5 ± 2.1** *(5-7)* P=0.3653	**3.0 ± 2.1** *(1-5)* P=0.0937
MC Drugs[b] (*n* = 44)	**963 ± 268** *(668-1269)* P<0.0001	**-0.2 ± 5.3** *(-7.2-6.8)* P<0.0001	**303 ± 14** *(158-513)* P<0.0001	**8.7 ± 6.2** *(2-17)* P<0.0001	**14.4 ± 4.9** *(10-20)* P<0.0001	**11.8 ± 8.2** *(3-23)* P<0.0001
Orally Available[c] (*n* = 18)	**921 ± 177** *(785-1068)* P<0.0001	**3.1 ± 3.9** *(-2.3-7.4)* P= 0.166	**247 ± 98** *(183-324)* P<0.0001	**6.1 ± 4.6** *(3-12)* P<0.0001	**14.0 ± 3.0** *(12-16)* P<0.0001	**9.3 ± 4.7** *(5-15)* P<0.0001
Not Orally Available[d] (*n* = 26)	**992 ± 316** *(665-1269)* P<0.0001	**-2.5 ± 4.9** *(-7.2-2.3)* P<0.0001	**341 ± 160** *(148-513)* P<0.0001	**10.6 ± 6.6** *(2-17)* P<0.0001	**14.6 ± 5.9** *(9-20)* P<0.0001	**13.5 ± 9.6** *(3-29)* P<0.0001
All Oral Drugs[e] (*n* = 1193)	**344 ± 107** *(200-475)*	**2.3 ± 2.4** *(-0.8-5.2)*	**78 ± 44** *(22-134)*	**1.8 ± 1.2** *(0-3)*	**5.5 ± 2.7** *(2-9)*	**5.4 ± 3.5** *(1-10)*

[a]Numbers are mean ± standard deviation values, with the 10-90% range given in parentheses. Abbreviations are: clogP, $\log_{10}$ of calculated 1-octanol/water partition coefficient; PSA, polar surface area for the predominant ionization state at pH = 7.0; HBD, number of hydrogen bond donors; HBA, number of hydrogen bond acceptors. Properties calculated using ChemAxon Software Calculator plug-ins (Marvin 6.0, 2013), as described in Methods. P values represent the probability that the average property value for the MC set in question differs from the average value for all oral drugs, and were calculated using classical (non-paired) t-tests after establishing sample normality using the Anderson-Darling test (see Methods for details). [b]MC Drugs include all compounds in the CHEMBL Database that were listed as approved drugs and that have a macrocyclic structure, plus a

small number of additional compounds identified from web-based sources. [c]The subset of MC Drugs that can be administered orally with the exception of Nystatin which, though administered orally to treat GI tract fungal infections, is not systemically absorbed (*1*). [d]The subset of MC Drugs that cannot be administered orally, including Nystatin (see note e). [e]Values from Vieth et al. (*2*).

Supplementary Table 3. Extent and physicochemical character of the MC surface area that is buried in the complex.[a]

	MC Test Set[b] (*n* = 22)	***Large MCs***[c] (*n* = 13)	***Small MCs***[d] (*n* = 9)
ΔASA (Complex)[e]	984 ± 235 Å^2	1106 ± 199 Å^2	807 ± 163 Å^2
ΔASA (MC only)[f]	571 ± 127 Å^2	628 ± 116 Å^2	489 ± 97 Å^2
% of MC SASA buried[g]	67 ± 14 %	57 ± 8 %	82 ± 4 %
Buried MC PSA[h]	139 ± 42 Å^2	142 ± 43 Å^2	134 ± 43 Å^2
Buried MC NPSA[i]	432 ± 137 Å^2	486 ± 134 Å^2	355 ± 104 Å^2
% Buried MC SASA that is Polar/Nonpolar[j]	25/75 ± 9 %	23/77 ± 9 %	28/71 ± 9 %
Number of Hydrogen Bonds	3.8 ± 2	4.8 ± 4.8	2.3 ± 1

[a]SASA and ΔASA values calculated as described in Methods. Numbers given are mean values plus or minus the standard deviation. [b]Test Set of 22 MC-protein complexes, involving 19 distinct MCs and 13 distinct proteins (See Supplementary Figure 1). [c]Complexes containing MCs with MW > 700 Da., involving 13 distinct MCs bound to 8 distinct proteins (See Supplementary Figure 1). [d]Complexes containing MCs with MW < 600 Da., involving 6 distinct MCs and 5 distinct proteins (See Supplementary Figure 1). [e]Total buried solvent accessible surface area for protein plus MC ligand. [f]Amount of MC surface area buried in the complex. [g]Percentage of MC SASA buried in the complex. [h]Amount of MC polar SASA buried in the complex. [i]Amount of MC nonpolar SASA buried in the complex. [j]Percentage of MC ΔASA that is polar/nonpolar, expressed as a percentage of total MC ΔASA.

Supplementary Table 4. Contributions to MC contact area, and physicochemical nature of the contacts, made by ring atoms, "peripheral" atoms and substituent atoms (as defined in the main text).[a]

	Test Set[b] (*n* = 22)	***Large MCs***[c] (*n* = 13)	***Small MCs***[d] (*n* = 9)
Structural Composition of MCs			
Number of Ring atoms	22 ± 7	25 ± 7	17 ± 2
% of total HA	42 ± 10 %	38 ± 8 %	50 ± 6 %
Number of Peripheral atoms	8.4 ± 4.5	9.9 ± 4.5	5.2 ± 2.4
% of total HA	15 ± 6 %	15 ± 6 %	15 ± 6 %
Number of Substituent Atoms	24 ± 12	30 ± 10	12 ± 5
% of total HA	43 ± 14 %	47 ± 13 %	35 ± 11 %
Average Number of Substituents	4.1 ± 1.9	4.7 ± 1.7	2.8 ± 1.7
Ring atoms per Peripheral group	3.4 ± 2.6	2.8 ± 0.8	4.8 ± 4.5
Physicochemical Composition of MCs			
Total MC	28/72 ± 6 %	29/71 ± 7 %	26/74 ± 2 %
Ring	15/85 ± 11 %	19/81 ± 11 %	7/93 ± 3 %
Peripheral	60/40 ± 19 %	62/38 ± 16 %	57/43 ± 26 %
Substituent	30/70 ± 11 %	26/74 ± 10 %	37/63 ± 11 %
MC Regional Contributions to Contact with Protein			
% ΔASA due to Ring Atoms	21 ± 11 %	15 ± 7 % P<0.0001[k]	29 ± 9 % P=0.00016[k]
% ΔASA due to Peripheral Groups	24 ± 13 %	22 ± 14 % P<0.0001[k]	26 ± 10 % P<0.0001[k]
% ΔASA due to Substituents	55 ± 17 %	62 ± 16 %	46 ± 8 %
Percentage of Atoms in each MC Region that are buried			
Ring Atoms	35 ± 19 %	24 ± 15 %	51 ± 11 %
Peripheral Atoms	81 ± 18 %	72 ± 17 % P<0.0001[l]	94 ± 12 % P<0.0001[l]
Substituent Atoms	66 ± 21 %	54 ± 17 % P<0.0001[l]	83 ± 13 % P=0.00016[l]
Percentage of buried atoms by Region that are Polar/Nonpolar			
Ring Atoms	7/93 ± 13 %	10/90 ± 16 %	3/97 ± 7%
Peripheral Atoms	60/40 ± 22 %	59/41 ± 21 %	60/40 ± 24 %
Substituent Atoms	26/74 ± 14 %	21/79 ± 16 %	32/68 ± 10%

[a]Numbers given are mean values plus or minus the standard deviation. [b]Test Set of 22

MC-protein complexes, involving 19 distinct MCs and 13 distinct proteins (See Supplementary Figure 1A). [c]Complexes containing MCs with MW > 600 Da., involving 13 distinct MCs bound to 8 distinct proteins (See Figure 1A). [d]Complexes containing MCs with MW < 600 Da., involving 6 distinct MCs and 5 distinct proteins (See Supplementary Figure 1A). [e]Total buried solvent accessible surface area for protein plus MC ligand. [f]Amount of MC surface area buried in the complex. [g]Percentage of MC SASA buried in the complex. [h]Amount of MC polar SASA buried in the complex. [i]Amount of MC nonpolar SASA buried in the complex. [j]Percentage of MC ΔASA that is polar/nonpolar, expressed as a percentage of total MC ΔASA. [k]Probability that the mean value differs from that observed for the substituent atoms, calculated using the Mann-Whitney U (rank) test. [l]Probability that the mean value differs from that observed for ring atoms, calculated using the Mann-Whitney U (rank) test.

Supplementary Table 5. Analysis of intramolecular hydrogen bonds in the bound MCs from the test set.

Protein	MC Ligand	PDB[a]	MC (total)[b]		Intramolecular Hydrogen Bonds in Bound MC	
			HBD	HBA	Number	Description
Cyclophilin	Cyclosporin A	1cwa	5	12	**1**	OH to O (single OH) substituent to periphery
	Sanglifehrin A	1sfa	9	14	**2**	1: NH (ring) to O of OH (peripheral), 2: NH to O of OH (both within a substituent)
FKBP	FK50	2fke	3	12	**none**	
	Rapamycin	2dg3	3	13	**none**	
Pan. elastase	Scyptolin	1okx	10	14	**2**	1: NH (ring) to O (peripheral); 2: NH (ring) to O of OH (peripheral)
EF-TU	GE2270A	1d8t	8	17	**2**	1: NH (substituent) to O (peripheral); 2: NH (ring) to O (substituent)
PP1a	Nodularin R	3e7a	10	11	**2**	1: NH (ring) to O (substituent); 2: NH (ring) to O (peripheral)
Spase	Arylomycin	1t7d	8	11	**none**	
Chitinase	Argadin	1waw	10	11	**5**	2 within ring (ring NH to peripheral O), 1 from ring NH to substituent O, 2 between 2 pairs of substituent atoms
	Argifin	1wb0	10	12	**2**	1: ring NH to peripheral O, 1 within large substituent
Actin	Kabiramide C	1qz5	4	16	**none**	
	Reidispongiolide A	2asm	0	13	**none**	
	Latrunculin B	2q0u	2	6	**none**	
	Pectenotoxin 2	2q0r	3	14	**none**	
AcetylCoA Carboxylase	Soraphen A	3gid	2	8	**1**	Peripheral OH to O of substituent OH
HSP90 (human)	Geldanomycin	1yet	4	9	**none**	
	Pochoxime A	3inw	2	8	**none**	
HSP90 (yeast)	Macbecin	2vwc	2	7	**none**	
	Radicicol	1bgq	2	6	**none**	
Topoisomerase		2hkj			**1**	Substituent OH to peripheral O
PDK3		2q8i			**none**	
PhoQ		3cgy			**1**	Substituent OH to ring O

[a]Protein Data Bank code for the protein-MC complex structure used for the analysis. [b]Number of hydrogen bond donors (HBD) and acceptors (HBA) present in the MC compound. HBD are defined as NH or OH groups; HBA as N or O atoms with a lone pair that is not involved in an aromatic ring. Amide N atoms were also excluded from the HBA count.

Supplementary Table 6: Comparator Set of protein-ligand complexes containing traditional drug-like ligands. This set was primarily based on the "Astex Diverse Set" of protein-ligand complex structures that was originally developed as a test set for benchmarking docking software *(3)*, supplemented by additional complex structures from the EMBL-EBI index of approved drugs (http://www.ebi.ac.uk/thornton-srv/databases/drugport/). From this combined set we selected complexes encompassing a diverse set of protein folds bound to ligands with reported K_D values ≤100 nM or IC50 values ≤1 μM.

Ligand	Protein	PDB ID
Astex Diverse Set[a]		
Tomudex	Thymidylate Synthase	1hvy
JE-2147	HIV-1 Protease	1kzk
BCX-1812	Neuraminidase A	1l7f
Compound 4	Thrombin	1oyt
Indirubin-3'-monoxime	GSK-3	1q41
Resveratrol	Quinone Reductase 2	1sg0
Compound 2	Dihydrofolate Reductase	1s3v
IDD552	Aldose Reductase	1t40
Compound 1	Fatty acid-binding Protein	1tow
TPI	Thymidine Phosphorylase	1uou
Compound 59	NS5B Polymerase	1yvf
Compound 11	HSP90	2bsm
EMBL-EBI[b]		
Zanamivir	Neuraminidase	1a4g
Tranexamic Acid	Plasminogen	1ceb
Lovastatin	LFA-1 (CD11A)	1cqp
Trifluoperazine	Calmodulin	1ctr
Sitagliptin	Dipeptidyl Peptidase IV	1x70
Aliskiren	Renin	2v0z
Trimethoprim	Dihydrofolate Reductase	2w3a
Indomethacin	Prostaglandin Reductase 2	2zb8
Miglitol	Maltase-Glucoamylase	3l4w
Galantamine	Acetylcholinesterase	4ey6
Both Sets		
Fluvastatin	HMG-CoA reductase	1hwi
Tadalafil	Phosphodiesterase 5A	1xoz

[a]See supplementary reference *(3)*.
[b]http://www.ebi.ac.uk/thornton-srv/databases/drugport/

Supplementary Table 7. Evaluation of the druggability of the MC binding sites in the Test Set based on FTMap analysis of the unbound protein structures

Protein name	PDB[a]		Hot spots[b]		Druggable[c]	Comments[d]
	U	B	P	S		
Cyclophilin	2cpl	1cwa 1sfa	1(23)+5(6)	3(15)	Yes	MC binds edge-on in narrow, hydrophobic pocket. Site druggable, but large MC provides additional contact area.
FKBP	2ppn	2fke 2dg3	1(18)+3(14)+ 7(6)	6(7)	Yes	P in deep cavity; S on surface, binding a conserved ring; P to S distance 7 Å. Non-MC binding might be weak.
Pan. elastase	1esa	1okx	5(9)+7(7)	3(11), 4(10)	Yes	P in deep pocket, S close to surface, but P to S distance < 5.5 Å. MC with tail provides additional contact area.
EF-TU	1efc	1d8t	1(21)+6(5)	4(12)	No	P to S distance 11.3 Å.
PP1a	3egg	3e7a	2(16)	6(7)	No	P to S distance 12.3 Å.
SPase	1kn9	1t7d	2(18)	8(2)	No	S is very small; MC with tail provides large contact area.
Chitinase	1guv	1waw 1wb0	1(23)+5(6)	4(10)+6(6)	Yes	P to S distance < 7 Å. Very deep pocket, MC provides additional contact area with walls.
Actin Site 1	1nwk	1qz5 2asm	3(14)	None	No	Only end of MC tail reaches any hot spot; MC provides large contact area needed for binding.
Site 2		2q0u	2(16)	7(4)	No	S too weak; MC provides large contact area.
Site 3		2q0r	10(2)	None	No	No strong interaction with hot spot; MC provides large contact area.
AcetylCoA Carboxylase	3glk	3gid	3(15)+6(6)	4(9)+5(8)	No	P to S distance > 9 Å.
HSP90 (closed)	1yer	1yet 3inw	1(23)+3(10)	5(8), 6(8)+9(3)	Yes	MC binds edge-on; P to S distance < 6 Å, P and two S sites are in V-shaped arrangement; MC closes the loop and provides 3D character.
HSP90 (open)	1yes	1yet 3inw	1(22)+2(13) +5(9)	3(12)	Yes	MC binds edge-on; P to S distance < 7 Å.
Topoisomerase	1mu5	2hkj	1(21)+4(12)	6(4)	No	Compact site, pushing a tri-substituted ring into P; although P to S distance < 6 Å, S is weak, and MC is needed for additional contact area.
PDK3	2pnr	2q8i	1(28)+7(2)	None	No	Compact site, pushing a three-substituted ring into P; no additional hot spot, MC provides large contact area

[a]U and B are Protein Data Bank (PDB) access codes for the ligand-free and macrocycle-bound structures, respectively. [b]Data are shown as CC rank among all CCs identified (# probes in CC). P is the primary (i.e. strongest) consensus cluster in the region, or in some cases a pair of closely adjacent CCs. S represents nearby secondary hot spots. Note that global mapping identifies fewer CCs than does the site-specific mapping described in Fig. 4 and Table 1 of the main text. [c]Site meets druggability criteria of possessing one major CC (or a pair of closely adjacent CCs) containing at least 16 probe clusters, plus at least one other CC within 7 Å from the first that contains at least 5 probe clusters (see text). [c]Description of binding mode; explanation if site not conventionally druggable by a small molecule; advantage of an MC ligand if druggable.

SUPPLEMENTARY REFERENCES

1. Melkoumov, A., Goupil, M., Louhichi, F., Raymond, M., de Repentigny, L., and Leclair, G. Nystatin nanosizing enhances in vitro and in vivo antifungal activity against Candida albicans, *J. Antimicrobial Chemotherapy, 68*, 2099-2105 (2013).
2. Vieth, M., et al. Characteristic physical properties and structural fragments of marketed oral drugs, *J Med Chem 47*, 224-232 (2004).
3. Hartshorn et al., Diverse, high-quality test set for the validation of protein-ligand docking performance. *J. Med. Chem. 50*, 726-741 (2007).