Supplementary Material (ESI) for New Journal of Chemistry

The use of silsesquioxane cages and phage display technology to probe silicone-protein interactions

Alan R. Bassindale*, Antonio Codina-Barrios, Nunzianda Frascione and Peter G. Taylor*

Supplementary information

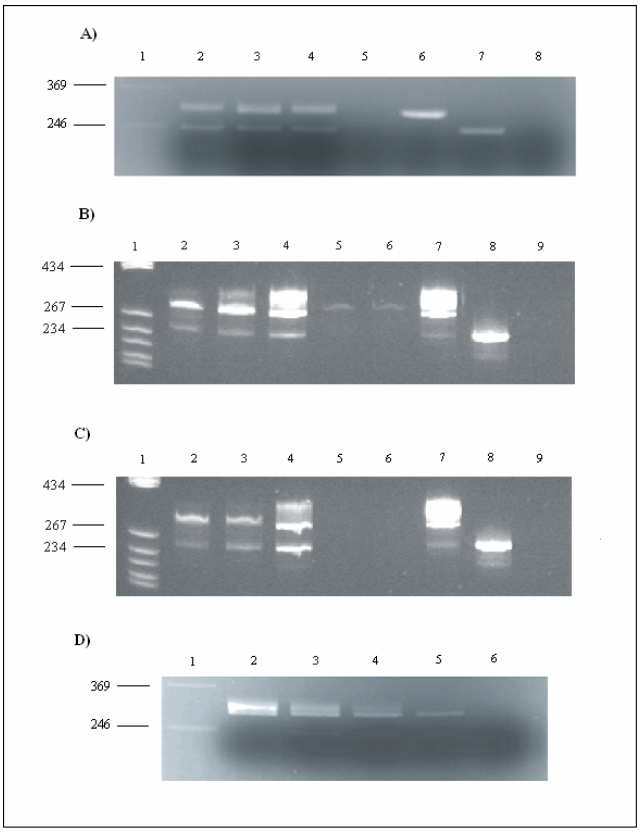

Fig. 1 Agarose gel of PCR products in the panning against **A**) methyl^{8}-T^{8}. 1 DNA ladder; 2 non-bound phages; 3 acid eluate; 4 acid washed solid; 5 acid washed solid after 50 ml surfactant wash; 6 PhD.-12 library positive control; 7 wild type positive control; 8 negative control; **B)** trifluoropropyl^{12}-T^{12}. 1 DNA ladder; 2 non-bound phages; 3 acid eluate; 4acid washed solid; 5 acid washed solid after 50 ml surfactant wash; 6 acid washed solid after 100 ml surfactant wash; 7 PhD.-12 library positive control; 8 wild type positive control; 9 negative control; **C)** phenyl^{8}-T^{8}. 1 DNA ladder; 2 non-bound phages; 3 acid eluate; 4 acid washed solid ; 5 acid washed solid after 50 ml surfactant wash; 6 acid washed solid after 100 ml surfactant wash;7 PhD.-12 library positive control; 8 wild type positive control; 9 negative control; **D)** H^{8}-T^{8}. 1 DNA ladder; 2 acid eluate; 3 acid washed solid; 4 acid washed solid after 100 ml surfactant; 5 acid washed solid after 300 ml of surfactant; 6 negative control

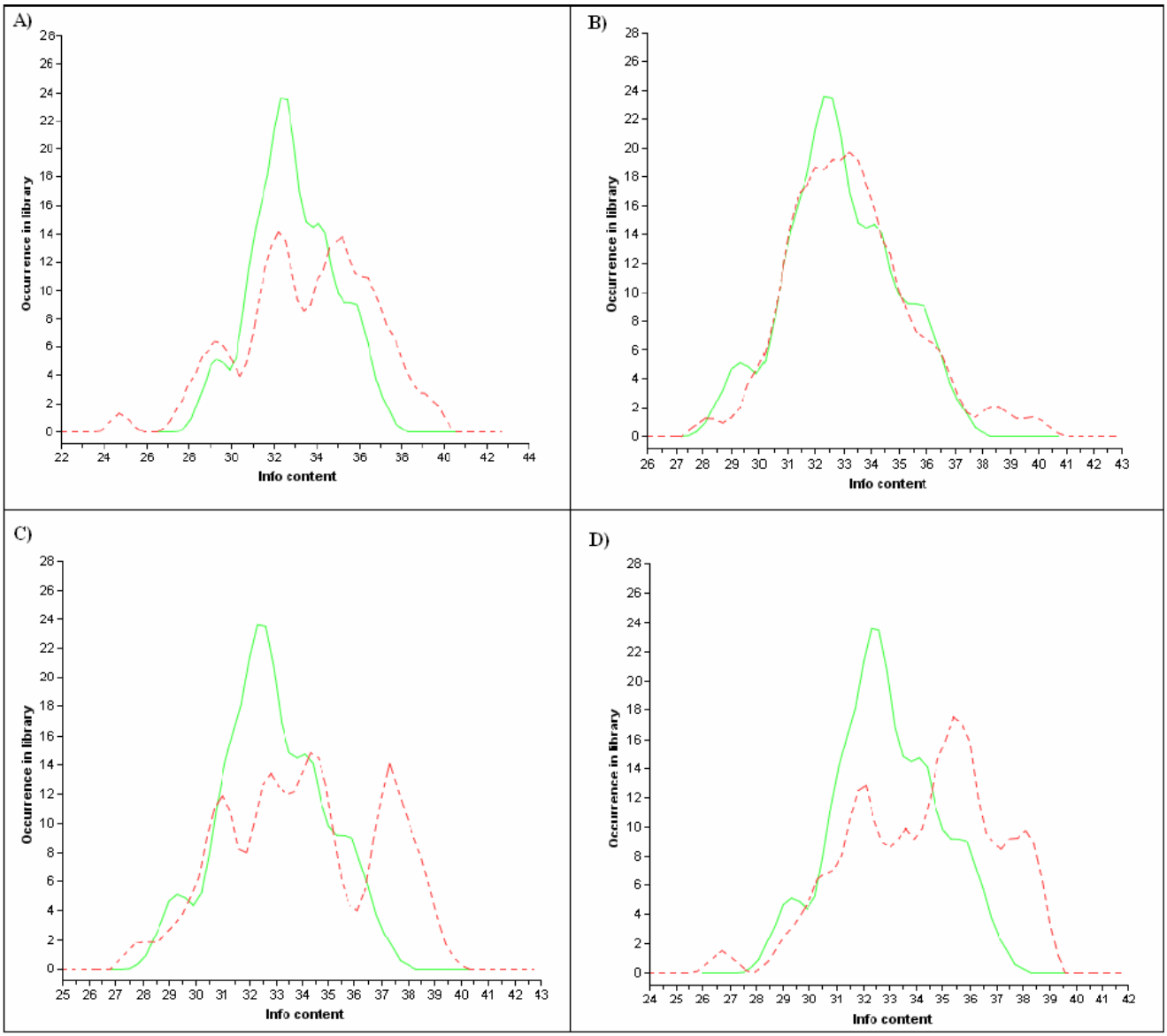

Table 1. Information content distribution of **A)** methyl^{8}-T^{8}, **B)** phenyl^{8}-T^{8}, **C)** H^{8}-T^{8} and **D)** trifluoropropyl^{12}-T^{12} tight binding peptides (dashed line) compared to the original library (solid line) as obtained from the RELIC program, INFO[30].

Methyl^{8}-T^{8}			
AWATLNLPPSPP	QKLSPRLSNLNI	HNLDIMSQSIPN	NHHLQMSFSQDS
FPPPNMLHPLPR	QLAPVPSPSSRV	ATGTSKQPYPHG	QNHNMKLTYHQF
SQLRMNSLPMFA (4)	SHPTPQWFRDKD (2)	WPLALNYPLYRK	HAVPVPVSSSTL
MTPHLSSYAPVW (3)	YGLAPSITVLRN	HRDNTTSSGNRL	QTISRQLTLPIV
AENKATHHPQPS (3)	HPWPLPIVYVPP	SSAGARQPPCHV	QERQPLPVTMHG
LPPLSRSSSLPL	VPPGLKDPAVGL	SFAKHIGTAGYR (2)	AKEMQVWHAFSR (2)
TTLWSHFTPLPT	VGTTLASSPRGQ	GL.NALRVPLLF	THHFEPHVSYHT
GIWTPGTPTLSD (3)	FMSATSYHRQSS	KLSRPTVQLPLT	DLIKKSLTIRQT
YNFNPSQSPSNT	WYPENTDNSIFY	DPLSNAVSMSQL	TAPQSKLYLGDE
LHTPPTLFRLVA	ANTFWLIRAVAD	KNPNVHRTYVVW	EPNDHESKDSYG
IISSPSATYSPN	FTQVHITGNNNK	YHPHANNKRDWS	FTRSLEKANMAA
TQMPPPTSLVPL	DSLHLPTIYLTP	DFSMTLYVPFHQ	RTDATSICISQH
LTSAETSARSFH	ICKKTTNPPVSH	TMSFHWEHTNNP (2)	FRYQEGAWWLPA
Phenyl^{8}-T^{8}			
SPTRTLQPLTWT	LHITTTTSRPDS	TNMYIQAICPLT	YPGASHELNSNL
TALPTLTQTMTS	NEPTGNRPSEPP	LSFDAQIFYRSR	HTSHSSPGQ.LP
QLSTTRDPNMVS	HNHLPTRLLAIS	GPTAFQLLSNSF	FVTLPSMRQMDL
SEPFPLIPLTQS	NPPLNFLSYANV	QPTPEPFVFFPP	RSW.TPNPLD..
VPDPTRMTPPST	APEHGRVSRQVS	TNHSLLSWPPRC	NIPALNKQTQWA
SQTHTMEPLSSP	QHPLTALTLVVV	VIHLPPVPTGFL	FPQQRLWWEHPP
TLSTLQTPRAFH	HPLFTRSSFGLI	QSLAMWFPPRAP	MGFQAPTLPMFS
SASHLLELKLFS	YLTLGPSAHLSE	FATPPRPLKMTA	.AYTPRAALQYM
TTSSWLYSRLSP	SLTNVENMLPSV	RPPSFILAGMHP	DSGQYSPPQQCH
DPKTNDPLALLS	VPSFVKC.APRS	ALPTSWRIAVAR	YAVSQLPHRYWL
LFLNTLLPIAST	VPAPQLQAAWGR	SMLLLPTHQTYR	HYVRGVPQLRPT
ELPPLRNRTPSN	HMPKILTIHPYR	YVYPLWMARAPL	QISRQYISSPGG
HPWTLSEFVRRS	LSLPMAKNASIR	TVKANIVLHQSF	QCHQVESTALSK
INPLNWDSARPS	HCDWKSCPFGMH	Q.PALMVMCPAC	NERLTG.TTVPV
YLNTNNRPSNTS	OTNSESARLHTY	TPGWIVPIEWTP	
TTANSSVYYPSS	YTPVLPIPISGM	TSAYNLHSSLKA	
H^{8}-T^{8}			
SYEPSLSLLKPE	TFRLHQPVPSTG	FRESPTPPYTLF	SLFSTTLVSRLL
IVSNLQTSMSAV	RHMNTILSTPSP	LQGPSMPCSHGA	GIWTIRYQSLHG
KLCAHSPPLLDL	ALLSNAQPLAWS	DPFPHSPPQPPM	MGQDHRDNHGAG
TTAHQVYAVSSP	HEPPSLARYAHT	KEMFYPSTHTRS	KVNMNLMWGAPR
QHLQLEGTWSGG	SINAHIFSFIQI	LTLSYPTPTHGL	HDSKRPVVYMFS
KAPETKPLVLLS	SSDVPQTAFVTH	SNAYGNLRIPWT	AMPHNLASMYRV
EAPSPKIGLRTL	NIQMFPKTWAPK	LSMQTLLEPQRA	GLRLPLHSWPHV
SFLTRQFLDILP	CWTNFWPFPLMI	GPTTRTTDLVVH	HHASSGQRGSTK
YIAHYPPTVDWP	WSTSLPPYSLSP	DAEDWMHKRTKL	APRTETLTLNLA
TGLEHYKPPSTR	RSADGGRSYPNK	YLSQTDTARATA	ESMNNIPMISMV
SPVQYKVQHLLS	HYVARANPHQTF	YTPILRLPTSPA	FTMHLGKSYTYD
DKQSNFAPTILW	TNPWLWSVHSNS	WPIQSNKPGPHL	STFRHEFTPQQL
NTSELRTYLRTL	SSMPTYWTNYPA	DYEVKNLSAIVV	KLPQTLRSHPYQ
Trifluoropropyl^{12}-T^{12}			
NSTKDQMRSFLY	TSKAMHGSSVTL	KAPPWPFQDSQS	QTIFTTTQFPTG
NSYVDWFPLGPT	EQTHPQSTWTAP	LAQTPHNPTSYS	QWLRQTTFYEPG
SSTSLSKHYHSP	VTQSDVMHTYPP	ERHGSVHLHHSL	GINAPQNQFMDT
TSPQNSSTDLPI	ITTPSHPWSSNI	ITRDYNDLHSLF	KTAMDMSWLHRA
SFQPRQFVNAWQ	FSKNAQTLSHNS	HTPGNITGLVLD	GASPVQKGHNTN
SPQPRYHDQFAS	FRVPSGGFPPSP	QVVSKPVLPRPD	DHHVHNLLPTAG
QTHAPHPAHPPP	GPQPYNRAQAAT	ADPAPDCFPLCT	FNTPVDYAIYLV
AYAPYSFHLQPS	NPHTRESAFPYT	VDPNLEWVAIAP	TYGTSKDHADQA
WPLNNRWLTSPP	AREPTPIAGTPT	IHNLKPTGLHAQ	VFNYKFAVTNIP
NLENWTLASPSP	FIGTNTMWQQLP	HTNTAADHMHSK	ETSRIDFAVRVA
LPNSYLASPSWT	THLVPSDRIPWP	NMGNPQFELRNL	YSNHLTMSMQVA
APLVLQTPSHRP	QFPFYKSLPTVT	TVISEPVAPLDL	LFAGMSASGANL
HATLQLPPSRWP	FQAPLPRAEHPW	LPTHSHPIYRML	MYGENHNPAGLW
SAPPRLPGLPSV	LEPTQAFTLLEL	TVEMQAKPPRHP	NQPDVYIPRQKG
SVHAMSLLDHPS	VKSTQLWPALSL	HNQDFKTMSLRP	GPHDDKKYQQLP
SAITQDHTHDQT (2)	STLATNTISMWF (2)	YTNETNSSSFSM	DSPLKPAVYQPP

Table 2. Tight binding peptide amino acid sequences found for exposure of PhD.12 library to silsesquioxanes using the PCR method. (The numbers represent the peptides found more than once).

A)

AA	1	2	3	4	5	6	7	8	9	10	11	12
A	2.02	1.71	2.02	2.89	1.53	1.71	2.89	1.71	1.71	1.53	1.47	1.53
C	1.11	1.59	1.11	1.59	1.34	1.11	1.59	1.59	1.59	2.01	1.59	1.59
D	2.98	1.34	1.11	1.11	2.01	1.59	1.11	1.59	1.11	1.11	1.34	1.11
E	1.34	2.01	1.59	2.01	1.34	1.34	1.11	1.59	2.01	1.34	1.59	1.59
F	1.34	1.11	2.01	2.01	1.11	1.59	1.11	1.11	1.11	1.11	1.34	1.11
G	2.02	3.23	1.53	2.02	1.47	1.53	3.23	1.53	3.23	2.02	1.47	1.53
H	1.34	2.01	2.98	1.11	4.19	**5.61**	2.01	**5.61**	1.59	1.59	1.11	1.34
I	2.98	2.01	1.59	1.59	1.34	1.11	1.34	1.59	1.11	1.59	2.01	2.01
K	2.98	1.11	1.34	1.59	1.34	2.01	1.34	1.34	1.59	1.34	1.11	1.34
L	2.38	1.84	1.84	2.35	1.65	1.84	1.84	1.84	1.70	2.38	3.70	1.95
M	1.34	1.34	2.01	1.59	1.11	1.34	1.34	1.11	1.59	1.34	1.59	1.34
N	2.01	4.19	1.11	2.01	1.34	**7.21**	1.34	**5.61**	2.01	1.11	1.34	2.01
P	3.23	**4.79**	**7.25**	2.89	1.71	1.71	2.20	3.76	**5.96**	2.20	1.71	**5.96**
Q	1.71	1.71	1.53	2.02	1.53	1.71	3.23	1.53	1.71	2.02	1.71	3.23
R	**4.92**	2.35	2.35	2.35	3.28	3.28	2.35	1.84	2.38	1.65	1.84	1.65
S	1.84	1.65	2.35	1.70	1.84	1.84	1.70	1.70	3.28	1.95	1.95	1.65
T	2.89	3.23	**5.96**	2.89	1.71	1.71	3.76	2.02	1.47	2.20	1.53	2.89
V	1.71	2.02	2.02	1.71	1.71	1.53	1.71	2.02	1.47	1.53	1.71	2.02
W	1.59	1.34	1.59	1.11	1.59	1.11	2.98	2.01	1.34	2.01	1.59	1.11
Y	1.59	1.34	2.01	2.01	2.01	1.59	1.59	2.01	2.01	1.34	1.11	1.11

B)

AA	1	2	3	4	5	6	7	8	9	10	11	12
A	1.53	1.47	2.20	1.47	3.23	1.53	1.47	1.47	2.02	2.20	1.53	2.20
C	1.11	1.59	1.11	1.59	1.59	1.59	1.59	1.11	1.59	1.59	1.59	1.59
D	2.98	1.11	1.11	2.01	1.59	1.11	1.11	1.11	1.11	1.11	1.11	1.11
E	1.11	1.34	2.98	2.01	1.59	1.34	1.59	1.11	1.59	1.59	1.59	1.11
F	1.34	1.34	2.01	1.11	1.34	1.11	2.01	1.11	1.34	1.59	1.11	1.34
G	1.47	1.53	2.02	3.23	1.53	1.47	2.02	2.02	1.47	2.02	1.47	1.71
H	2.98	2.01	1.59	2.98	**7.21**	1.59	1.34	1.59	**5.61**	1.34	2.98	1.34
I	1.11	2.98	1.11	1.11	1.11	1.34	1.11	1.59	1.11	2.98	1.59	1.34
K	4.19	1.11	1.59	1.11	1.11	2.01	2.98	1.11	1.59	1.11	1.11	2.01
L	1.84	1.95	1.70	2.35	2.38	2.38	1.95	1.84	1.95	1.95	1.95	2.97
M	1.11	1.11	4.19	1.34	1.59	1.34	1.11	1.59	1.34	1.11	1.11	1.11
N	1.34	1.34	1.34	2.01	2.98	2.01	1.11	1.11	1.11	1.59	1.34	1.59
P	3.23	1.71	3.76	2.20	1.71	2.89	**5.96**	**7.25**	2.20	3.76	2.89	2.20
Q	2.02	2.02	1.47	2.89	2.02	1.71	1.53	1.53	2.02	1.47	1.53	2.02
R	2.35	3.28	2.35	3.28	1.65	1.65	2.35	1.84	2.35	1.84	1.84	2.35
S	3.70	1.70	1.65	2.97	1.70	2.35	1.84	3.70	1.65	2.97	1.84	1.95
T	1.71	2.89	1.47	1.47	3.76	1.47	2.89	2.89	1.71	1.71	**4.79**	1.53
V	3.23	1.53	1.53	1.53	3.23	2.02	1.53	1.71	1.47	1.53	1.53	1.71
W	1.34	1.11	1.11	1.11	1.11	1.34	1.11	1.11	2.01	1.59	2.01	1.11
Y	2.01	2.01	1.59	1.11	2.98	1.34	1.34	1.34	4.19	1.34	1.34	1.59

C)

AA	1	2	3	4	5	6	7	8	9	10	11	12
A	1.71	2.20	1.53	2.20	1.53	1.47	1.53	3.76	1.53	1.53	2.20	3.23
C	1.59	3.23	3.23	3.23	3.23	3.23	2.02	3.23	3.23	3.23	2.02	3.23
D	1.11	1.53	3.23	1.47	1.71	1.53	1.47	2.02	1.47	2.02	2.02	1.53
E	1.34	2.02	1.47	2.02	2.02	1.53	3.23	2.02	2.02	3.23	2.02	3.23
F	2.98	1.53	3.23	2.02	2.02	3.23	2.89	1.53	2.02	1.47	3.23	1.53
G	1.34	3.23	1.53	1.53	3.23	2.02	1.53	1.47	2.02	2.02	3.23	3.23
H	2.98	1.53	2.20	1.53	3.23	2.20	1.47	1.71	1.71	4.79	2.02	3.23
I	2.01	2.02	1.53	3.23	3.23	2.02	2.02	1.53	2.02	2.02	3.23	1.53
K	1.11	2.02	1.53	2.02	1.47	1.47	1.47	3.23	3.23	3.23	3.23	2.02
L	2.98	2.02	1.71	1.47	1.71	1.71	1.53	3.76	3.76	2.89	2.20	3.76
M	1.59	2.02	3.23	2.02	1.53	3.23	1.47	2.02	2.02	2.02	2.02	2.02
N	4.19	2.02	1.71	2.20	1.71	1.71	2.02	3.23	2.02	3.23	1.47	3.23
P	1.59	**4.79**	**5.96**	**7.25**	2.89	3.76	2.20	2.89	3.76	2.89	**8.67**	**17.3**
Q	2.01	1.53	2.89	2.02	2.20	2.89	3.23	2.02	1.71	1.71	1.53	1.53
R	1.59	1.47	2.02	3.23	1.71	2.02	1.53	1.53	3.23	2.20	1.53	3.23
S	**7.21**	3.76	2.02	2.20	1.71	2.20	2.20	1.47	**7.25**	2.89	**4.79**	2.89
T	**4.19**	**4.79**	2.89	3.76	1.47	1.53	2.89	1.71	1.47	1.47	2.02	**4.79**
V	2.01	1.71	1.53	1.47	3.23	1.53	1.53	1.47	3.23	1.53	2.02	2.02
W	1.11	3.23	3.23	3.23	1.53	2.02	1.47	1.53	2.02	3.23	2.20	2.02
Y	1.11	2.02	2.02	3.23	2.20	2.02	3.23	2.02	1.47	2.02	1.53	2.02

Table 3. Information content[27] [-Ln (P)] of each amino-acid on each position for **A**) 50 random peptides from the PhD.-12 library; **B**) H^8-T^8 peptides; **C**) trifluoropropyl-T12 peptides. Values that indicate frequencies higher than at least one standard deviation above or below the expected value are

printed in bold. The number of observation of each amino acid residue at each position (Pr) *m* times in *n* peptides is calculated from the binomial distribution:

$$Pr = n!P^{m}Q^{(n-m)}/m!(n-m)!$$

where P is the *a priori* probability of occurrence for each amino acid (i.e. the probability of finding a particular amino acid at a given position in *n* peptides). We calculated it on the number of codons coding for a particular amino acid. Since the peptide libraries analyzed in this work used only 32 codons, the probability of a particular amino acid was set as the number of corresponding codons divided by 32. Q is the probability of any one failure (= (1-P)).