DBLMSP DBL variants	DNA	Protein	Boundaries
3D7	XM_001347596	XP_001347632	S135-D435
7G8	FJ406853	ACR09793	S103-D405
Dd2	FJ406861	ACR09800	S103-D403
FCR3	FJ406860	ACR09799	S103-D403
P. reichenowi	FJ406845	ACR09785	S103-D403
Isolate 002	FJ556030	ACR49847	S103-D403
Isolate 017	FJ556042	ACR49858	S103-D403
Isolate 026	FJ556049	ACR49864	S103-D405
Isolate 067	FJ556070	ACR49882	S103-D403
Isolate 028	HM000310	ADJ58681	S103-D405
Isolate 384	HM000317	ADJ58688	S103-D405
DBLMSP2 DBL variants	DNA	Protein	Boundaries
3D7	XM_001347603	XP_001347639	T154-V454
Isolate 020	HM000330	ADI99310	H134-V436
Isolate 023	HM000332	ADI99312	S131-V430
Isolate 028	HM000334	ADI99314	T131-V431
Isolate 082	HM000336	ADI99316	S131-V431
Isolate 093	HM000339	ADI99319	T131-V430
Isolate 097	HM000340	ADI99320	T131-V430
Isolate 384	HM000390	ADI99370	S131-V431
Dd2 105 ALTAI KYGFSDW FCR3 105 ALTAI KYGFSDW Isolate067 105 ALTAI KYGFSDW Isolate017 105 ALTAI KYGFSDW Isolate017 105 ALTAI KYGFSDW 3D7 105 ALTAI KYGFSDW	GDI I KGT DLI DYQI TKNI NR ALDKI LGNE / GDI I KGT DLI DYQI TKNI NR ALDKI LGNE / GDI I KGT DLI DYQI TKNI NR ALDKI LGNE / GDI I KGT DLI DYQI TKNI NR ALDKI LGNE / GDI I KGT DLI DYQI TKNI NR ALDKI LGNE / GDI I KGT DLI DYQI TKNI NR ALDKI LGNE /	S N. DKI K KR VDIWE ANKAAF WDAF MCGYK VHIGN SS N. DKI K KR VDIWE ANKAAF WDAF MCGYK VHIGN SS N. DKI K KR VDIWE ANKAAF WDAF MCGYK VHIGN SS N. DKI K KR VDIWE ANKAAF WDAF MCGYK VHIGN SS N. AEI K NR VDIWE ANKAF WDAF MCGYK VHIGN KN VDAP KDAKT WUTGNKHR V WDA WCCY O'S AKK WTEN NHH VWER MCCY O'S AK KD UKN NDK PK DAKK WTEN KHH VWER MCCY O'S AK KD	KPCPEHDNWDRIPQYLRWFREWGT KPCPEHDNWDRIPQYLRWFREWGT KPCPEHDNWDRIPQYLRWFREWGT KPCPEHDNWDRIPQYLRWFREWGT KPCPEHDNWDRIPQYLRWFREWGT KRCPEHDNWDRIPQYLRWFREWGT KRCPEHDNWDRIPQYLRWFREWGT KRCFUGYGNIYDIPQYLRWFREWGT KRCFUGYGNIYDIPQYLRWFREWGT
Dd2 209 YKNKFENVIELC FCR3 209 YKNKFENVIELC Isolate067 209 YKNKFENVIELC Isolate017 209 YKNKFENVIELC 3D7 209 YKNKFENVIELC 3D7 209 YKNKFENVIELC 7G8 211 YKNKFENVIELC 1solate384 211 YKNKFENVIKEV 1solate384 211 YKNKFEVVIKEV	NV HQI T NQDDS QL LEISK BKCK GALKHYI NV RQI T NQDDS QL LEISK BKCK GALKHYI NV RQI T NQDDS QL LEISK BKCK GALKHYI NV RQI T NQDDS QL LEISK BCKCK GALKHYI NV QQI T NQDDS QL LEISK BCKCK GALKHYI NI QQI T NQDDS QL LEISK BKCK GALKHYI	E WWNRRR PE WKGQC DKF E KEKSKYE DT KSRT AE I E WWNRRPE WKGQC DKF E KEKSKYE DT KSRT AE I E WWNRRPE WKGQC DKF E KE KSKYE DT KSRT AE I E WWNRRPE WKGQC DKF E KE KSKYE DT KSRT AE I E WWNRRPE WKGQC DKF E KE KSKYE DT KSRT AE I E WWNRRPE WKGQC DKF E KE KSKYE DT KSIT AE K E WWNRRPE WKGQC DKF E KE KSKYE DT KSIT AE K E WWNRRPE WKGQC DKF E KE KSKYE DT KSIT AE K E WWNRRPE WKGQC DKF E KE KSKYE DT KSIT AE E WWNRRPE WKGQC DKF E KE KSKYE DT KSIT AE E WWNRRPE WKGQC DKF E KE KSKYE DT KSIT AE E WWNRRPE WKGQC DKF E KE KSKYE DT KSRT AE E WWNRRPE WKGQC DKF E KE KSKYE DT KSRT AE E WWNRRPE WKGQC DKF E KE KSKYE DT KSRT AE	YLKEI CSECDCKYKDLD YLKEI CSECDCKYKDLD
C 3D7 1 TINLGENKCPTEEI Isolate384 1 SINNLGPNKCPVEKI Isolate093 1 TINLGENKCPTEEI	CKDFSNLPQCRKNVHE RNNWLGSSVKI CKDFGPLPQCRKNVHE RNNWLSSNVKI CKDFSNLPQCRKNVHE - RNNWLGSSVK CKDFSNLPQCRKNVHE - RNNWLGSSVK	NFASDNKGVLVPPRROSLCLRITLODFRTKKKKEG FSSDNKGVLVPPRROSLCLRITLODFRTKKKKEG FASDNKGVLVPPRROSLCLRITLODFRTKKKKEG FASDNKGVLVPPRROSLCLRITLODFRTKKKKEG FSSDNKGVLVPPRROSLCLRITLODFRTKKKKEG FSSDNKGVLVPPRROSLCLRITLODFRTKKKKEG FASDNKGVLVPPRROSLCLRITLODFRTKKKKEG FASDNKGVLVPPRROSLCLRITLODFRTKKKKEG	DFEKFIYSYASSEARKLRTIHNN DFEKFIYSYASSEARKLRTIHNN DFEKFIYSYASSEAKKLRTIHNN DFEKFIYSYASSEAKKLRTIHNN DFEKFIYSYASSEAKKLRTIHNN DFEKFIYSYASSEAKKLRTIHNN
Isolate097 1 TIN GENKCPTEE Isolate023 1 SNNL GPNKCPVEK Isolate082 1 SNNL GPNKCPVEK Isolate028 1 TIN LGENKCPTEE Isolate020 1 HLEGNHYKCPDKNF	CKDFGPLPQCRKDVHE KNNWLSSNVK CKDFGPLPQCRKDVHE RNNWLSSNVK CKDFSNLPQCRKNVHE RNNWLGSSVK CNGLQNVPNCPLKDFTGTKGDWASSNVR	VF AS DNK GVL V P P R R QS L CL R I T L QDF R T K K K K C VF L T V NK GVL V P P R R K QMC F R I NI NN F P K L K K T E G	DFEKFIYSYASSEARKLRTIHNN KFENFIYS <mark>S</mark> AGSEAKQLIKLYGN
		FASDNR GVL VPPRROSLELRITLOD RITKKKREG HELTYNKGVL VPPRRKONGFRINI TNNFPKLKKREG INDKPKDAKKWITENRHHVWEAMVGGYOS AGKDNG INDKPKDAKKWITENRHHVWEAMVGGYORKKDNE INE-AKDAKKWITENRYHVWEAMVGGYORKKDNE INE-AKDAKKWITENRYHVWEAMVGGYORKKDNE INE-AKDAKKWITENRYHVWEAMVGGYORKKDNE INE-AKDAKKWITENRYHVWEAMVGGYORKKDNE INDFR KDAKKWITENR	

Figure S1. Details of the DBLMSP and DBLMSP2 DBL domains used for functional analysis

A. DNA and protein accession numbers of the different DBL domains used in this study. The aminoacid boundaries of each DBL domain are indicated in the right-hand column. **B.** Sequence alignment of the DBLMSP DBL domains. A common feature of the two DBL domains that do not bind IgM (7G8 and isolate 028) is the presence of a Glu amino-acid residue at position 310, where all other domains contain either a His or an Ala residue. To test whether we could narrow the binding site to IgM even further, we used site-directed mutagenesis and substituted the Glu and Leu present at positions 310-311 in the non-binding domain of isolate 028 for Ala and Gln residues present in the binding domain of isolate 384 (red box). These substitutions were however not sufficient to restore binding of the DBL domain of isolate 028 to human IgM, suggesting that these two residues are not sufficient to confer the IgM-binding property. **C.** Sequence alignment of the DBLMSP2 DBL domains. Sequence alignments of the different DBL variants were performed with ClustalW2 (<u>http://www.ebi.ac.uk/Tools/msa/clustalw2/</u>) and displayed using Boxshade (<u>http://embnet.vitalit.ch/software/BOX_form.html</u>).

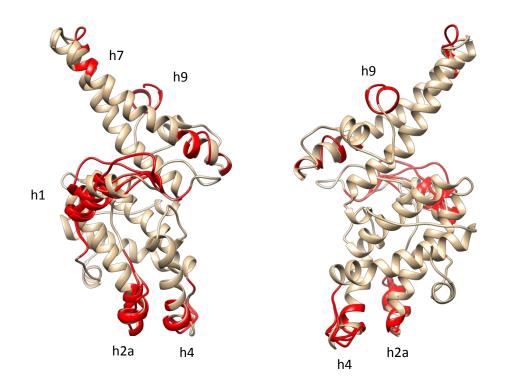


Figure S2. Modelled structural comparisons between IgM-binding and –non-binding DBL domains

The structures of the IgM-binding DBL domains from DBLMSP and DBLMSP2 were modelled using the Phyre2 software and the known crystal structure of the DBL domain from DBLMSP2. The IgM-binding DBL domain models are shown in beige, and the regions that differ in the non-binding DBL domains (isolate 028 and strain 7G8 in DBLMSP), are superimposed and highlighted in red (left panel). The right panel shows a 180° rotation of the same structure. Most differences were observed in helices h1, h2a, h4 and h7. Polymorphisms in h7 are, however, also frequently observed around the same positions amongst DBL domains that have retained their IgM-binding ability. The IgM-binding site is therefore most likely to be in helices h1, h2a or h4.