Supplementary Material

Table S1. Structure alignment statistics for trichosurin and 26 respresentative structures from the retinol binding-like protein family from SCOP (50815). These alignment statistics along with the dendogram and structure-based sequence alignment were generated using the program MAMMOTH [25].

PDB code	Description	Naln	RMS	P-value	Sequence ID	Z-score
1kt7	Bovine RBP	119	3.43	0.1E-06	11.35	16.82
1aqb	Pig RBP	121	3.39	0.7E-07	13.48	17.16
1rbp	Human RBP	122	3.83	0.6E-07	12.77	17.33
1rlb	Chicken RBP	124	3.51	0.4E-07	11.35	17.67
1iiu	Chicken RBP	125	3.64	0.4E-07	12.06	17.84
1hn2	Bovine OBP	106	2.23	0.8E-06	20.57	14.60
1dzk	Pig OBP	139	2.29	0.4E-08	24.11	20.23
1bj7	Bovine d2	139	2.10	0.4E-08	21.99	20.23
1ew3	Equine c2	141	1.55	0.3E-08	28.37	20.57
1gm6	Boar SLP	140	1.51	0.3E-08	24.82	20.40
1e5p	Hamster aphrodisin	139	2.11	0.4E-08	20.57	20.23
1beb	Bovine BLG	139	2.76	0.4E-08	20.57	20.23
1exs	Pig BLG	134	3.40	0.9E-08	19.15	19.38
1epb	Rat RBP	137	2.71	0.5E-08	15.60	19.89
1jv4	Mouse rMUP	141	1.64	0.3E-08	34.04	20.57
1mup	Mouse MUP	141	1.70	0.3E-08	34.04	20.57
2a2u	Rat MUP	141	1.33	0.3E-08	34.04	20.57
1l6m	Human NGAL	127	2.85	0.3E-07	17.02	18.18
1jzu	Quail q83	123	3.63	0.5E-07	10.64	17.50
1lf7	Human C8γ	124	3.13	0.4E-07	11.35	17.67
1bbp	Butterfly billin BP	107	3.77	0.6E-06	10.64	14.77
1i4u	α-crustacyanin	113	3.35	0.2E-06	11.35	15.80
1qft	Tick histamine BP	94	3.67	0.5E-05	6.38	12.55
1np1	Nitrophorin I	118	3.80	0.1E-06	6.38	16.65
1euo	Nitrophorin II	120	3.86	0.8E-07	4.26	16.99
1d2u	Nitrophorin IV	121	3.83	0.7E-07	7.09	17.16

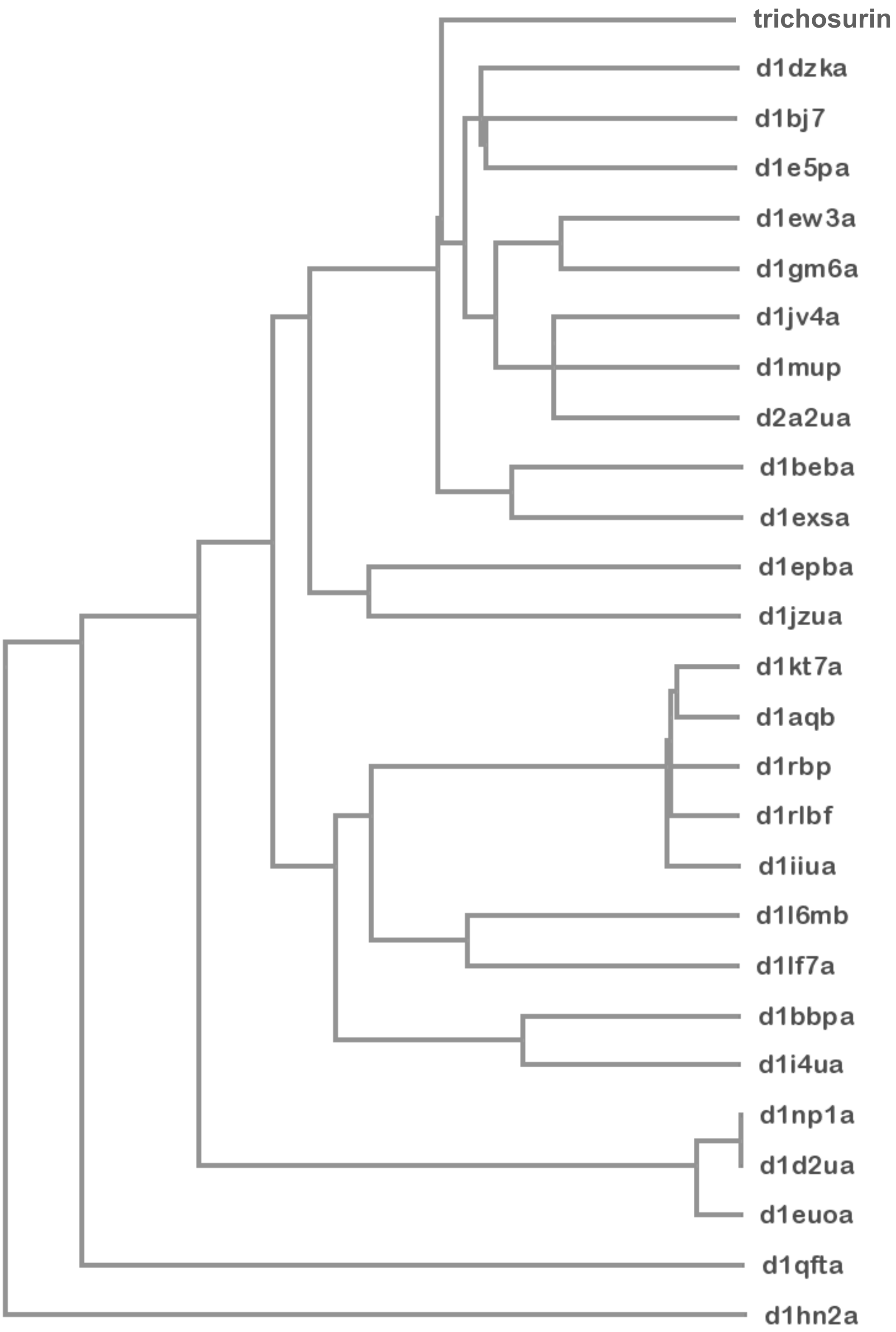

Figure S1. Dendogram showing the relationships between trichorsurin and 26 representative structures from the SCOP retinol binding-like protein family (50815). PDB codes are given at each leaf (these have the prefix ‘d’ and the suffix denoting the chain used in the superposition). The tree was calculated by MAMMOTH incorporating both structure alignment and sequence identity (see table S1) [25].

Figure S2. Sequence alignment derived from the superposition of trichosurin and 26 representative structures from the SCOP retinol binding-like protein family (50815).

```
trichosurin    .......... .......... .......... .L...RHWHT VVLASSDRSL
d1kt7a_.ent    ......ERDC RVSSFRVKE. ..NF..DKAR FA...GTWYA MAKKDPEG..
d1aqb__.ent    ......ERDC RVSSFRVKE. ..NF..DKAR FS...GTWYA MAKKDPEG..
d1rbp__.ent    ......ERDC RVSSFRVKE. ..NF..DKAR FS...GTWYA MAKKDPEG..
d1rlbf_.ent    ......ERDC RVSSFRVKE. ..NF..DKAR FA...GTWYA MAKKDPEG..
d1iiua_.ent    .......MDC RVSSFKVKE. ..NF..DKNR YS...GTWYA MAKKDPEG..
d1hn2a_.ent    .......... ........AQ EEEAEQNLSE LS...GPWRT VYIGSTNPEK
d1dzka_.ent    .......... .......... ........FE LS...GKWIT SYIGSSDLEK
d1bj7__.ent    .......... .......... ...I..DPSK IP...GEWRI IYAAADNKDK
d1ew3a_.ent    .......... ........VA IRNF..DISK IS...GEWYS IFLASDVKEK
d1gm6a_.ent    .......... ........VV TSNF..DASK IA...GEWYS ILLASDAKEN
d1e5pa_.ent    .......... .......... .......FAE LQ...GKWYT IVIAADNLEK
d1beba_.ent    .......... ........QT MKGL..DIQK VA...GTWYS LAMAASDISL
d1exsa_.ent    .......... ....VEVTPI MTEL..DTQK VA...GTWHT VAMAVSDVSL
d1epba_.ent    .......... .......... VKDF..DISK FL...GFWYE IAFASKMGTP
d1jv4a_.ent    .......... ....EEASST GRNF..NVEK IN...GEWHT IILASDKREK
d1mup__.ent    .......... ....EEASST GRNF..NVEK IN...GEWHT IILASDKREK
d2a2ua_.ent    .......... ....EEASST RGNL..DVAK LN...GDWFS IVVASNKREK
d1l6mb_.ent    ..TSDLIPAP PLSKVPLQQ. ..NF..QDNQ FQ...GKWYV VGLAGNAI..
d1jzua_.ent    .......... .......... MTVP..DRSE IA...GKWYV VALASNTEFF
d1lf7a_.ent    ........AS PISTIQPKA. ..NF..DAQQ FA...GTWLL VAVGSAG...
d1bbpa_.ent    ....NVYHDG ACPEVKPVD. ..NF..DWSN YH...GKWWE VAKYPNSV..
d1i4ua_.ent    DKIPDFVVPG KCASVDRNKL WAEQTPNRNS YA...GVWYQ FALTNNPY..
d1qfta_.ent    .........N QPDWADEAAN GAHQDAWKSL KADVENVYYM VKATY...KN
d1np1a_.ent    .........K CTKNALAQT. ..GFNKDKYF NG...DVWYV TDYLDLEPDD
d1euoa_.ent    ........MD CSTNISPKQ. ..GLDKAKYF S....GKWYV THFLDKDPQ.
d1d2ua_.ent    .........A CTKNAIAQT. ..GFNKDKYF NG...DVWYV TDYLDLEPDD

trichosurin    IE.EEGPFRN FIQNIT.VE. .SGNLNGFFL T..RK...NG QCIPLYLTAF
d1kt7a_.ent    ...LFLQDN. IVAEFS.VDE .NGHMSATAK GRVRLLNNWD VCADMVGTFT
d1aqb__.ent    ...LFLQDN. IVAEFS.VDE .NGHMSATAK GRVRLLNNWD VCADMVGTFT
d1rbp__.ent    ...LFLQDN. IVAEFS.VDE .TGQMSATAK GRVRLLNNWD VCADMVGTFT
d1rlbf_.ent    ...LFLQDN. IVAEFS.VDE .NGHMSATAK GRVRLLNNWD VCADMVGTFT
d1iiua_.ent    ...LFLQDN. VVAQFT.VDE .NGQMSATAK GRVRLFNNWD VCADMIGSFT
d1hn2a_.ent    IQ.ENGPFRT YFRELV.FDD EKGTVDFYFS V..KR...DG KWKNVHVKAT
d1dzka_.ent    IG.ENAPFQV FMRSIE.FDD KESKVYLNFF S..KE...NG ICEEFSLIGT
d1bj7__.ent    IV.EGGPLRN YYRRIE.CIN DCESLSITFY L..KD...QG TCLLLTEVAK
d1ew3a_.ent    IE.ENGSMRV FVDVIR.ALD .NSSLYAEYQ T..KV...NG ECTEFPMVFD
d1gm6a_.ent    IE.ENGSMRV FVEHIR.VLD .NSSLAFKFQ R..KV...NG ECTDFYAVCD
d1e5pa_.ent    IE.EGGPLRF YFRHID.CYK NCSEXEITFY V..IT...NN QCSKTTVIGY
d1beba_.ent    LDAQSAPLRV YVEELK.PTP .EGDLEILLQ K..WE...NG ECAQKKIIAE
d1exsa_.ent    LDAKSSPLKA YVEGLK.PTP .EGDLEILLQ K..RE...ND KCAQEVLLAK
d1epba_.ent    GLAHKEEKM. GAMVVE.LKE ..NLLALTTT Y..YS...ED HCVLEKVTAT
d1jv4a_.ent    IE.DNGNFRL FLEQIH.VL. .EKSLVLKFH T..VR...DE ECSELSMVAD
d1mup__.ent    IE.DNGNFRL FLEQIH.VL. .ENSLVLKFH T..VR...DE ECSELSMVAD
d2a2ua_.ent    IE.ENGSMRV FMQHID.VL. .ENSLGFKFR I..KE...NG ECRELYLVAY
d1l6mb_.ent    ...LRPQKM. YATIYE.LKE .DKSYNVTSV LFR.....KK KCDYWIRTFV
d1jzua_.ent    LREKDKMKM. AMARIS.FLG .EDELKVSYA VPKPN...GC RKWETTFKKT
d1lf7a_.ent    ....RRAEA. TTLHVA.PQ. .GTAMAVSTF RKL.....DG ICWQVRQLYG
d1bbpa_.ent    ...EKYGKC. GWAEYT.PE. .GKSVKVSNY HV.I....HG KEYFIEGTAY
d1i4ua_.ent    ...QLIEKC. VRNEYS.FD. .GKQFVIEST GIAY....DG NLLKRNGKLY
d1qfta_.ent    DPVWGNDFTC VGVMANDVNE DEKSIQAEFL FMNNA...DT NMQFATEKVT
d1np1a_.ent    VPKR....YC AALAAG.TAS .GKLKEALYH YDP.....KT QDTFYDVSEL
d1euoa_.ent    VTDQ....YC SSFTPR.ESD .GTVKEALYH YNA.....NK KTSFYNIGEG
d1d2ua_.ent    VPKR....YC AALAAG.TAS .GKLKEALYH YDP.....KT QDTFYDVSEL
```

```
trichosurin    KTEE.....A RQFKL.N... .......... .YYGTNDVYY ESSKPNEYAK
d1kt7a_.ent    DTED.....P AKFKMKYWGV ASF....... LQKGNDDHWI IDTDYETFAV
d1aqb__.ent    DTED.....P AKFKMKYWGV ASF....... LQKGNDDHWI IDTDYDTYAV
d1rbp__.ent    DTED.....P AKFKMKYWGV ASF....... LQKGNDDHWI VDTDYDTYAV
d1rlbf_.ent    DTED.....P AKFKMKYWGV ASF....... LQKGNDDHWI IDTDYETFAV
d1iiua_.ent    DTED.....P AKFKMKYWGV ASF....... LQKGNDDHWV VDTDYDTYAL
d1hn2a_.ent    KQDD...... GTYVAD.... .......... .YEGQNVFKI VSL.SRTHLV
d1dzka_.ent    KQ.E.....G NTYDV.N... .......... .YAGNNKFVV SYA.SETALI
d1bj7__.ent    RQ.E.....G YVYVL.E... .......... .FYGTNTLEV IHV.SENMLV
d1ew3a_.ent    KTEE.....D GVYSL.N... .......... .YDGYNVFRI SEFENDEHII
d1gm6a_.ent    KV.G.....D GVYTV.A... .......... .YYGENKFRL LEVNYSDYVI
d1e5pa_.ent    LK.G.....N GTYET.Q... .......... .FEGNNIFQP LYI.TSDKIF
d1beba_.ent    KTKI.....P AVFKI.D... .......... .ALNENKVLV LDTDYKKYLL
d1exsa_.ent    KTDI.....P AVFKI.N... .......... .ALDENQLFL LDTDYDSHLL
d1epba_.ent    EGDG.....P AKFQVTR... .......... .LSGKKEVVV EATDYLTYAI
d1jv4a_.ent    KTEK.....A GEYSV.T... .......... .YDGFNTFTI PKTDYDNFLM
d1mup__.ent    KTEK.....A GEYSV.T... .......... .YDGFNTFTI PKTDYDNFLM
d2a2ua_.ent    KTPE.....D GEYFV.E... .......... .YDGGNTFTI LKTDYDRYVM
d1l6mb_.ent    PGSQ.....P GEFTLGNIKS Y......... PGLTSYLVRV VSTNYNQHAM
d1jzua_.ent    SDDG.....E VYYSEE.... .......... ...AKKKVEV LDTDYKSYAV
d1lf7a_.ent    DTGV.....L GRFLLQ...A R......... GARGAVHVVV AETDYQSFAV
d1bbpa_.ent    PVGDS...KI GKIYHKLTYG .......... GVTKENVFNV LSTDNKNYII
d1i4ua_.ent    PN.PF...GE PHLSIDYEN. .......... ..SFAAPLVI LETDYSNYAC
d1qfta_.ent    AVKMYGYNRE NAFRYETE.. .......... .DGQVFTDVI AYSD.DNCDV
d1np1a_.ent    QEE.....SP GKYTANFKKV EKNGNVKVDV TSGNYYTFTV MYAD.DSSAL
d1euoa_.ent    KLES....SG LQYTAKYKTV DKKKAVLKEA DEKNSYTLTV LEAD.DSSAL
d1d2ua_.ent    QVE.....SL GKYTANFKKV DKNGNVKVAV TAGNYYTFTV MYAD.DSSAL

trichosurin    FIFYNYH.D. .GKVNVVANL FGRTP.NLSN EIKKRFEEDF .MNRGFRREN
d1kt7a_.ent    QYSCRLLNLD GTCADSYSFV FARDPSGFSP EVQKIVRQRQ .EELCLAR.Q
d1aqb__.ent    QYSCRLQNLD GTCADSYSFV FARDPHGFSP EVQKIVRQRQ .EELCLAR.Q
d1rbp__.ent    QYSCRLLNLD GTCADSYSFV FSRDPNGLPP EAQKIVRQRQ .EELCLAR.Q
d1rlbf_.ent    QYSCRLLNLD GTCADSYSFV FARDPSGFSP QVQKIVRQRQ .EELCLAR.Q
d1iiua_.ent    HYSCRELNED GTCADSYSFV FSRDPKGLPP EAQKIVRQRQ .IDLCLDR.K
d1hn2a_.ent    AHNINVDKH. .GQTTELTEL FVKL..NVED EDLEKFWKLT .EDKGIDKKN
d1dzka_.ent    ISNINVDEE. .GDKTIMTGL LGKGT.DIED QDLEKFKEVT .RENGIPEEN
d1bj7__.ent    TYVENYD.G. .ERITKMTEG LAKGT.SFTP EELEKYQQLN .SERGVPNEN
d1ew3a_.ent    LYLVNFD.K. .DRPFQLFEF YAREP.DVSP EIKEEFVKIV .QKRGIVKEN
d1gm6a_.ent    LHLVDVN.G. .DKTFQLMEF YGRKP.DVEP KLKDKFVEIC .QQYGIIKEN
d1e5pa_.ent    FTNKNXDRA. .GQETNXIVV AGKGN.ALTP EENEILVQFA .HEKKIPVEN
d1beba_.ent    FCMENSAEP. .E.QSLVCQC LVRTP.EVDD EALEKFDKAL .KA..LPMHI
d1exsa_.ent    LCMENSASP. .E.HSLVCQS LARTL.EVDD QIREKFEDAL .KTLSVPMRI
d1epba_.ent    IDITSLV.A. .GAVHRTMKL YSRSL.DDNG EALYNFRKIT .SDHGFSETD
d1jv4a_.ent    AHLINEK.D. .GETFQLMGL YGREP.DLSS DIKERFAQLC .EEHGILREN
d1mup__.ent    AHLINEK.D. .GETFQLMGL YGREP.DLSS DIKERFAQLC .EEHGILREN
d2a2ua_.ent    FHLINFK.N. .GETFQLMVL YGRTK.DLSS DIKEKFAKLC .EAHGITRDN
d1l6mb_.ent    VFFKKVSQ.. .NREYFKITL YGRTK.ELTE .LKENFIRFS .KSLGLPENH
d1jzua_.ent    IYATRVK.D. .GRTLHMMRL YSRSP.EVSP AATAIFRKLA .GERNYTDEM
d1lf7a_.ent    LYLERA.... ...GQLSVKL YARSL.PVSD SVLSGFEQRV .QEAHLTEDQ
d1bbpa_.ent    GYYC.KYDED KKGHQDFVWV LSRSK.VLTG EAKTAVENYL IGSPVVDSQK
d1i4ua_.ent    LYSC.IDYN. FGYHSDFSFI FSRSA.NLAD QYVKKCEAAF .KNINVDTTR
d1qfta_.ent    IYVPGTD... ..GNEEGYEL WTTDYDNIPA NCLNKFNEYA V......GRE
d1np1a_.ent    IHTCLHKGN. .KDLGDLYAV LNRNKDTNAG D...KVKGAV T.AASLKFSD
d1euoa_.ent    VHICLREGS. .KDLGDLYTV LTHQKDAEPS A...KVKSAV T.QAGLQLSQ
d1d2ua_.ent    IHTCLHKGN. .KDLGDLYAV LNRNKDAAAG D...KVKSAV S.AATLEFSK
```

```
trichosurin    ILDISEVDHC .......... .....
d1kt7a_.ent    YRLIPH...N GYCD...... .....
d1aqb__.ent    YRIITH...N GYCD...... .....
d1rbp__.ent    YRLIVH...N GYC....... .....
d1rlbf_.ent    YRLIPH...N GYC....... .....
d1iiua_.ent    YRVIVH...N GFCS...... .....
d1hn2a_.ent    VVNFLENENH PHPE...... .....
d1dzka_.ent    IVNIIERDDC PA........ .....
d1bj7__.ent    IENLIKTDNC PP........ .....
d1ew3a_.ent    IIDLTKIDRC FQLRG..... .....
d1gm6a_.ent    IIDLTKIDRC FQLRG..... .....
d1e5pa_.ent    ILNILATDTC PE........ .....
d1beba_.ent    RLSFNPTQLE EQC....... .....
d1exsa_.ent    LPAQL..EEQ CRV....... .....
d1epba_.ent    LYILKHDLTC VKVLQSAA.. .....
d1jv4a_.ent    IIDLSNANRC .......... .....
d1mup__.ent    IIDLSNANRC .......... .....
d2a2ua_.ent    IIDLTKTDRC L......... .....
d1l6mb_.ent    IVFPVP...I DQCID..... .....
d1jzua_.ent    VAMLPRQEEC TVDEV..... .....
d1lf7a_.ent    IFYFPK...Y GFCEAADQFH VLDEV
d1bbpa_.ent    LVYSDFSEAA CKVN...... .....
d1i4ua_.ent    FVKTVQGSSC PYDTQKTL.. .....
d1qfta_.ent    TRDVFTSACL EIAAA..... .....
d1np1a_.ent    FISTKDNKC. .EYDNVSLKS LLTK.
d1euoa_.ent    FVGTKDLGC. .QYDDQFTSL .....
d1d2ua_.ent    FISTKENNC. .AYDNDSLKS LLTK.
```