SUPPORTING INFORMATION

Elucidating the molecular interactions of chemokine CCL2 orthologs with flavonoid baicalin

Nidhi Joshi¹, Dinesh Kumar² and Krishna Mohan Poluri^{1*}

¹Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee – 247667, Uttarakhand, India

²Centre of Biomedical Research, SGPGIMS Campus, Lucknow - 226014, Uttar Pradesh, India

Running Title: Interaction of baicalin with CCL2 chemokines

*Corresponding Author: Dr. Krishna Mohan Poluri, Department of Biotechnology, Indian Institute of Technology Roorkee (IIT-Roorkee), Roorkee – 247667, Uttarakhand, India Tel: +91-1332-284779; E-mail: krishfbt@iitr.ac.in / mohanpmk@gmail.com **Table S1:** List of primers used for site directed mutagenesis of mCCL2-P8A and hCCL2-P8A mutant proteins.

Name of Mutant	Primer Sequence						
mCCL2-P8A	Fwd: 5'- GATGCAGTTAACGCCGCACTCACCTGCTGC -3'						
	Rev: 5'- GCAGCAGGTGAGTGCGGCGTTAACTGCATC -3'						
hCCL2-P8A	Fwd: 5'- GCAATCAATGCCGCAGTCACCTGCTG -3'						
	Rev: 5'- CAGCAGGTGACTGCGGCATTGATTGC -3'						

Table S2: Total amount of protein obtained for mCCL2-P8A and hCCL2-P8A after each step of purification.

Purification steps	mCCL2-P8A			
After 1 st Ni-NTA column chromatography	65 ± 5			
After Dialysis	50 ± 5			
After ion-exchange chromatography	25 ± 3			
After reverse Ni NTA chromatography	10± 2			
Total mCCL2-P8A protein obtained after final	purification 8± 2 mg/L			

Purification steps	hCCL2-P8A			
After 1st Ni-NTA column chromatography	60 ± 5			
After Dialysis	45 ± 5			
After ion-exchange chromatography	18 ± 3			
After reverse Ni NTA chromatography	8 ± 2			
Total hCCL2-P8A protein obtained after final p	urification 6 ± 2 mg/L			

Protein	Helix (%)	Sheet (%)	Coil (%)
mCCL2-WT	17.3	14.6	68.1
hCCL2-WT	17.2	14.9	67.9
mCCL2-P8A	17.9	14.6	67.5
hCCL2-P8A	17.8	14.8	67.4

Table S3: Percentage of secondary structural elements in CCL2 wild type and monomeric proteins (for both human and murine orthologs) as estimated using DICHROWEB-K2D software.

Figure S1: Multiple sequence alignment of CCL2 chemokine from rodent and primate families. The sequence alignment was generated from CLUSTAL W software. The secondary structural elements are shown on the top of the sequences and the presence of β c-stand at C terminal, which is specifically observed in murine family is represented in cyan color.

		β _N — —		β1 — ·		β2 -	<u>β3</u>	α-he	elix	- B -
HUMAN	QPDAINA	PVTCCYNFTN	RKISVQRLA	SYRRITSS	K <mark>C</mark> PKEAVI	FKTIVAK	EICADPK	QKWVQDSMD	HLDKQTQ	TPKT
ANGOLAN COLOBUS	QPDAINA	<mark>PVTCC</mark> YNFTN	RKISVQRLA	SYRRITSS	K <mark>C</mark> PKEAVI	FKTIVAK	EICADPK	QKWVQDSMD	HLDKQIQ	TPKP
B.S.N. MONKEY	QPDAINA	PVTCCYNFTN	RKISVQRLA	SYRRITSS	K <mark>C</mark> PKEAVI	FKTIVAK	EICADPK	QKWVQDSMD	HLDKQIQ	TPKP
BO.S. MONKEY	QPDAINA	PVTCCYIFTN	KKISLQRLV	SYRRITSS	K <mark>C</mark> PKEAVI	FKTILAK	EICADPK	QKWVEDSID	HLDKQIQ	TPKP
CHIMPANZEE	QPDAINA	PVTCCYNFTN	RKISVQRLA	SYRRITSS	K <mark>C</mark> PKEAVI	FKTIVAK	EICADPK	QKWVQDSMD	HLDKQTQ	TPKT
COQUERELS SIFAKA	AQ PDAVNA <mark>I</mark>	PVTCCYTFSS	RKISLQRLM	SYRRVANS	K <mark>C</mark> PKEAVI	FKTILAK	EICANPE	2NWVKDYIA	KLDKKTQ	TPKP
C.E. MACAQUE	QPDAINA	PVTCCYNFTN	RKISVQRLA	SYRRITSS	K <mark>C</mark> PKEAVI	FKTIVAK	EICADPK	QKWVQDSMD	HLDKQIQ	TPKP
DRILL	QPDAINA	VTCCYNFTN	RKISVQRLA	SYRRITSS	K <mark>C</mark> PKEAVI	FKTIVAK	EICADPK	QKWVQDSMD	HLDKQIQ	TPKP
GELADA	QPDAINA	VTCCYNFTN	RKISVQRLA	SYRRITSS	KCPKEAVI	FKTIVAK	EICADPK	QKWVQDSMD	HLDKQIQ	TPKP
G.S.N. MONKEY	QPDAINA	VTCCYNFTN	RKISVQRLA	SYRRITSS	KCPKEAVI	FKTIVAK	EICADPK	QKWVQDSMD	HLDKQIQ	TPKP
GORILLA	QPDAINA	PVTCCYNFTN	RKISVQRLA	SYRRITSS	KCPKEAVI	FKTIVAK	EICADPK	QKWVQDSMD	HLDKQTQ	TPKT
GREEN MONKEY	QPDAINA	VTCCYNFTN	RKISVQRLA	SYRRITSS	KCPKEAVI	FKTIVAK	EICADPK	QKWVQDSMD	HLDKQIQ	TPKP
GREY MOUSE LEMUR	RQPDAVNA	VTCCYTFSS	RKISLQRLT	SYRRVSNS	KCPKEAVI	FKTILAK	EICANPE	2NWVKDYIA	KLDKKTQ	TPTP
MAS NIGHT MONKEY	QPDAINA	PVTCCYVFTN	KRISLQRLA	SYRRITSS	KCPKEAVI	FKTILAK	EICADPK	QKWVEDSID	HLDKQIQ	TPKP
N.W.C. GIBBON	QPDAINA	PVTCCYNFTN	RKISVQRLA	SYRRITSS	KCPKEAVI	FKTIVAK	EICADPK	QKWVQDSMD	HLDKQTQ	LRSL
OLIVE BABOON	QPDAINA	VTCCYNFTN	RKISVQRLA	SYRRITSS	KCPKEAVI	FKTIVAK	EICADPK	QKWVQDSMD	HLDKQIQ	TPKP
PHILIPPINE TARIS	QPDAINA	PVTCCYKFTN	KRISVQRLA	SYRRVTSS	KCPKEAVI	FKTILDK	EICADPN	QKWVQDSMA	HLDKKTQ	TPKP
RHESUS MONKEY	QPDAINA	PVTCCYNFTN	RKISVQRLA	SYRRITSS	KCPKEAVI	FKTIVAK	EICADPK	2KWVQDSMD	HLDKQIQ	TPKP
PYGMY CHIMPANZEE	QPDAINA	PVTCCYNFTN	RKISVQRLA	SYRRITSS	KCPKEAVI	FKTVVAK	EICADPK	QKWVQDSMD	HLDKQTQ	TPKT
S.E. GALAGO	QPDAVNA	PVTCCYSFSS	GKIVFKRLV	SYRRVTSS	KCPSEAVI	FKTKLAR	EICADPEI	KEWVKEYIA	RLDKKTQ	TP
SOOTY MONKEY	QPDAINA	PVTCCINFTN	RKISVQRLA	SYRRITSS	KCPKEAVI	FKTIVAK	EICADPK	ZKWVQDSMD	HLDKQIQ	TPKP
SUMATRAN ORANGUI	QPDAINA	PVTCCYNFTN	RKISVQRLA	SYRRITSS	KCPKEAVI	FKTIVAK	EICADPK	2KWVQDSMD	HLDKQTQ	TLKT
U.R. COLOBUS	QPDAINA	PVTCCINFTN	KKISVQRLA:	SIRKITSS	KCPKEAVI VCDVEAVI	FKTIVAK	EICADPK	2KWVQDSMD 2KWVEDSTD	HLDKQIQ	TPKP
W.F. CAPUCHIN	OPDAINA	PVTCCIIFTN	KRISLQRLV:	SIKKITSS	KCPKEAVI	FKTKLAK	EICADPK	2KWVEDSID	HLDKQIQ	TPKP
W.T.E. MANMOSET	QPDAINA	PVICCIIFIN	KKI SLQKLA	51661155	KCPKEAVI	FRIKLAK	CADPR	ZKWVEDSID	HIDKQIQ	IPRP
ALPINE MARMOT	QPDAVNS	<mark>PVTCCYTLTS</mark>	KKIPMKRLM	SYRRVTSS	K <mark>C</mark> PKEAVI	FTTVLNK	EICADPL	QKWVEDYVA	KMDQKTE	RNQN
HOUSE MOUSE	QPDAVNA	LTCCYSFTS	KMIPMSRLE	SYKRITSS	R <mark>C</mark> PKEAVV	FVTKLKR	EVCADPK	KEWVQTYIK	NLDRNQM	RSEP
A.G. SQUIRREL	QPDAVNS	vtccytlts	RKIPMKRLM	SYRRVTSS	K <mark>C</mark> PKEAVI	FTTVLNK	EICADPL	QKWVEDYVA	KMDQKTE	GNRN
ARID THICKET RAT	Q PDAVNA <mark>l</mark>	PLTCCYSFTG	KMIPMNRLE	SYKRITSS	R <mark>C</mark> PKEAVV	FTTKLKR	EICADPKI	KEMAŐTAIK	NLDQNKM	RSET
CHINESE HAMSTER	QPDAVNS	PLTCCYSFTA	KRIPEKRLE	SYKRITSS	K <mark>C</mark> PKEAVI	FITKLKR	EICADPK	QDWVQTYTK	KLDQSQA	KSEA
DAMARA MOLE RAT	QPDGVNA	<mark>P-ACCYKF-N</mark>	KRIPLRKVV	SYVKVTSS	R <mark>C</mark> PKEAVI	FQTIANR	EICADPTI	EKWVKDYIA	KVDQNTQ	LKQNST
DEGU	QPDGVNA	<mark>P – V C C Y T F – N</mark>	KKIPLKRVR	SYERITSS	R <mark>C</mark> PQEAVI	FKTVSNK	EV <mark>C</mark> ANPN	QSWVKEYIA	KLDQKSQ	KKQNSA
GOLDEN HAMSTER	QPDSVNS	PLTCCYSFNA	RKIPEKRLE	SYKRITSS	K <mark>C</mark> PKEAVI	FITKLKR	EICADPK	QDWVQAYTK	KLDQSQA	KSET
GUINEA PIG	QPDGVNT	<mark>P – T C C Y T F – N</mark>	KQIPLKRVK	GYERITSS	R <mark>C</mark> PQEAVI	FRTLKNK	EVCADPT	QKWVQDYIA	KLDQRTQ	QKQNST
L.E. JEROBA	QPDALNI	<mark>PVTCCYTVTN</mark>	KKIPGKRLK	SYS-VASS	K <mark>C</mark> SKKAVI	FLTVLDK	EICADPT	QKWVQEQIA	RLDQKNQ	LNRNPP
L.T. CHINCHILLA	QPDGVNA	P-VCCYTF-N	KKIPLKRVM	SYERITSS	R <mark>C</mark> PQEAVI	FKTVMNR	EICADPK	QSWVEDYIA	RLDQKTQ	QKQNST
MONGOLIAN GERBII	Q PEGVT S <mark>I</mark>	PLTCCYSFTS	RKIPEKRLE	SYKRITSS	K <mark>C</mark> PKEAIV	FVTKLKR	EICANPN	EKWVQSYIQ	NLDQNQM	RSET
NAKED MOLE RAT	QPDGVNA	P-VCCYIF-N	KKIPLKRVV	SYERITSS	R <mark>C</mark> PKEAVI	FQTIANR	EICADPT	QKWVKDYIA	KLDQKTQ	LKQNST
NORWAY RAT	QPDAVNA	PLTCCYSFTG	KMIPMSRLEI	NYKRITSS	R <mark>C</mark> PKEAVV	FVTKLKR	EICADPNI	KEWVQKYIR	KLDQNQV	RSET
PRAIRIE DEER MOU	JQ PDAVNA <mark>I</mark>	PLTCCYMFAI	KKIPEKRLE	SYKRITNS	K <mark>C</mark> PREAVI	FITKLKR	EICADPK	QEWVQTYTR	KLDQNQA	RSET
PRAIRIE VOLE	Q PDAVN S	PLTCCFSFTI	KKIPERRLE	SYKRITSS	K <mark>C</mark> PKEAVI	FITKLKK	EICADPK	QAWVQAYMN	KLDQNQA	RSET
RYUKYU MOUSE	QPDAVNA	PLTCCYSFTS	KMIPMSRLE	SYKRITSS	R <mark>C</mark> PKEAVV	FVTKLKR	EVCADPKI	KEWVQTYIK	NLDRNQM	RSEP
SHREW MOUSE	QPDAVNA	PLTCCYSFTS	KMIPMSRLE	SYKRITSS	RCPKEAVV	FVTKLKR	EVCADPK	KEWVQTYIK	NLDQNQM	RSET
T.L.G. SQUIRREL	Q PDAVN S	PVTCCYTLTS	RKIPMKRLM	SYRRVTSS	K <mark>C</mark> PKEAVI	FTTVLNK	EICADPL	QKWVEDYVA	KMDQKTE	GNQN
U.G.M.B. MOLE RA	Q PDAVN S	VTCCYTFNK	RKIPENRLV	SYRRITSS	K <mark>C</mark> PKEAVI	LTTKQKK	EICADPK	QNWVQDYIA	KLEQKSQ	QNPT
W.F. MOUSE	QPDAVNA	PLTCCYTFTS	KKIPEKRLE	SYKRITNS	KCPREAVI	FITKLKR	EICADPK	QEWVQTYTR	KLDÖNÖA	RSET

Figure S2: (A) Site directed mutagenesis; Lane M - DNA marker; Lane 2- mCCL2-P8A mutant gene; Lane 3- hCCL2-P8A mutant gene. (B) Protein overexpression profile: Lane M- Marker (Bovine serum albumin (BSA)-66 kDa and hen egg lysozyme (HEL)-14 kDa); Lane 2uninduced sample of mCCL2-P8A protein; Lane 3- induced sample of mCCL2-P8A protein. (C) Ni-NTA affinity chromatography profile: Lane M- marker; Lane 2 – supernatant obtained after lysis, Lane 3 – flow through; Lane 4 and 5 - wash I (20 mM imidazole + 20 mM Tris + 500 mM NaCl), wash II (50 mM imidazole + 20 mM Tris + 500 mM NaCl); and Lane 6-9 - Elution fractions (400 mM imidazole + 20 mM Tris + 500 mM NaCl) having purified fusion mCCL2-P8A protein. (D) TEV digestion profile: Lane M – marker; Lane 2 - undigested mCCL2- protein; Lane 3 - digested mCCL2-P8A protein containing the left over undigested protein, thioredoxin tag, and mCCL2-P8A protein. (E) Ion-exchange chromatography profile for tag (TRX) separation: Lane M-marker; Lane 2- supernatant (collected after TEV digestion); Lane 3- flow through; Lane 4-5 - wash I (50 mM NaCl + 20 mM Tris), wash II (100 mM NaCl + 20 mM Tris); Lane 6-8- Elution fractions (500 mM NaCl + 20 mM Tris) having purified mCCL2-P8A protein. (F) Reverse Ni-NTA affinity chromatography profile: Lane M – Marker; Lane 2- pure mCCL2-P8A protein (MW-9 kDa).



Figure S3: (**A**) Protein overexpression profile for hCCL2-P8A mutant: Lane M- Marker (Bovine serum albumin (BSA) – 66 kDa and hen egg lysozyme (HEL)-14 kDa); Lane 2- uninduced sample of hCCL2-P8A protein; Lane 3- induced sample of hCCL2-P8A protein. (**B**) Ni-NTA affinity chromatography profile: Lane M- marker; Lane 2 - supernatant obtained after lysis, Lane 3 - flow through; Lane 4 and 5 - wash I (20 mM imidazole + 20 mM Tris + 500 mM NaCl), wash II (50 mM imidazole + 20 mM Tris + 500 mM NaCl) having purified fusion hCCL2-P8A protein. (**C**) TEV digestion profile: Lane M - marker; Lane 2 - undigested monomeric protein; Lane 3 - digested hCCL2-P8A protein containing the left over undigested protein, thioredoxin tag, and hCCL2-P8A protein. (**D**) Ion-exchange chromatography profile for tag (TRX) separation: Lane M-marker; Lane 2- supernatant (collected after TEV digestion); Lane 3- flow through; Lane 4-5 - wash I (50 mM NaCl + 20 mM Tris), wash II (100 mM NaCl + 20 mM Tris); Lane 6-8 - Elution fractions (500 mM NaCl + 20 mM Tris) having purified hCCL2-P8A protein. (**E**) Reverse Ni-NTA affinity chromatography profile: Lane M - Marker; Lane 2 - pure hCCL2-P8A protein (MW-9 kDa).



Figure S4: Overlay of tryptophan fluorescence of WT (mCCL2 and hCCL2) and monomeric (mCCL2-P8A and hCCL2-P8A) CCL2 orthologs.









