DOI: 10.18621/eurj.1011372

Medical Biology

The in silico interaction analysis of CARMIL1 proteincontaining leucine-rich repeat (LRR) regions with interleukin-1 receptor-associated kinase 1 (IRAK1) protein and LLR peptide

Nail Beşli¹^o, Güven Yenmiş²^o

¹Department of Medical Biology, University of Health Sciences Turkey, Institute of Health Sciences, Istanbul, Turkey; ²Department of Medical Biology, Biruni University, Faculty of Medicine, Istanbul, Turkey

ABSTRACT

Objectives: Capping protein Arp2/3 and myosin-I linker protein 1 (CARMIL1) encoded by the CARMIL, is a major, multidomain, membrane-linked protein regulating actin assembly; however, its function in inflammatory signaling is not fully elucidated. The leucine-rich repeat (LRR) region of CARMIL1 has been associated with interleukin (IL)-1 receptor-associated kinase (IRAK) in fibroblasts by many methods including tandem mass tag mass spectrometry, immunoprecipitation, and CRISPR-Cas9. This study, therefore, set out to assess the interaction of CARMIL1 with each IRAK1 protein and a novel LRR peptide.

Methods: The molecular docking techniques were employed to compare the binding modes and affinities of the 3D structure of CARMIL1 each of LRR peptides and IRAK1 protein. 3D structure model of CARMIL1 protein and LRR peptide was predicted through Robetta tool considering the structures and function of these proteins.

Results: As an overall conclusion of docking, the LRR peptide was observed to contact the residues in the LRR 1-2 of the human CARMIL1, whereas the IRAK1 protein was to interact with the residues in the LRR 1, 2, and 10 regions of the human CARMIL1.

Conclusions: Our computational results suggest that LRRs in CARMIL1 are involved in the formation of protein-peptide binding interfaces with its structural conformation.

Keywords: CARMIL1, IRAK1, LRR peptide, molecular docking simulation

Interleukin-1 (IL-1), a critical intermediate for inflammation, assists the degradation and remodeling of extracellular matrices (ECMs) by enhancing the expression of matrix metalloproteinases (MMPs), thus, plays a central role in various inflammatory diseases including rheumatoid arthritis, chronic periodontitis, and severe lung injury [1]. The binding of IL-1 to its signal receptor [IL-1 receptor type 1 (IL-1R1)], allows the IL1R assistant proteins to be recruited. Activation of the IL-1R complex is followed by recruitment of MyD88 (myeloid differentiation factor 88). This newly formed adapter protein complex then binds to IRAK1 and IRAK2 (Interleukin-1 receptor-associated kinase 1 and 2) [2]. IRAK is quickly phosphorylated, subsequently released from MyD88, and initiates signal transduction pathways, like the mitogen-activated

Received: October 18, 2021; Accepted: June 21, 2022; Published Online: August 3, 2022



How to cite this article: Beşli N, Yenmiş G. The in silico interaction analysis of CARMIL1 protein-containing leucine-rich repeat (LRR) regions with interleukin-1 receptor-associated kinase 1 (IRAK1) protein and LLR peptide. Eur Res J 2022;8(6):810-820. DOI: 10.18621/eurj.1011372 Address for correspondence: Nail Beşli, PhD. Candidate, University of Health Sciences Turkey, Institute of Health Sciences, Department of Medical

Biology, Tibbiye Cad., No:38, 34668 Üsküdar, Istanbul, Turkey. E-mail: beslinail@gmail.com, Phone: +90 216 391 11 91

[®]Copyright © 2022 by Prusa Medical Publishing Available at http://dergipark.org.tr/eurj protein kinase (MAPK) family members ERK [3].

Fibroblasts are a major group of cells in the connective tissues, directly included in the matrix degradation in widespread inflammatory diseases such as periodontitis and rheumatoid arthritis. Concerning IL-1 signal transmission in fibroblasts, IL-1R1 needs to be recruited to focal adhesions (Fas) [4]. Following this recruitment, FA kinase is activated, Ca2+ is released from the endoplasmic reticulum [5], and ERK is activated eventually [3]. Such mechanisms in the central parts of the signaling system increase the level of MMPs expression [6]. Despite the comprehensive examinations, the control of IL-1 signaling in anchorage-dependent stromal cells is not fully elucidated.

Human CARMIL1 (Capping Protein Arp2/3 myosin I linker) from CARMIL-family is a large multidomain protein, profoundly conserved. CARMIL proteins have several membrane-associated functions associated with actin assembly and signaling owing to their structural characteristics. CARMIL1 is expressed in plenty by fibroblasts, which is intensely linked to abnormal inflammatory processes [7].

CARMIL1 is approximately 1370-aa length pro-

tein with 16 leucine-rich repeats (LRRs), which has a non-canonical pleckstrin homology (PH) domain, a long helical domain (HD), and a C-terminus responsible for the interaction with F-actin-capping protein subunit alpha-2 (CAPZA2) [8, 9].

In a recent study, Wang et al. [10] have investigated the function of CARMIL1, its LRR region, and CPI-CSI motifs (CP-binding domain) in controlling the production of IL-1 signaling in circulating fibroblasts. They have concluded that the LRR of CARMIL1 would mediate IL-1-induced collagen degradation in the fibroblasts and might be a goal for the anti-inflammatory drug improvement. In the current study, inspired by the experimental findings from this recent investigation, we computed the interaction of LLR regions in a model of human CARMIL1 protein with IRAK1 protein and the designed LRR peptide by drawing on the tested/trusted bioinformatics tools and databases to find the polar contacts. The cellular and biological role of a protein is broadly direct linked to its 3D structure [11]. In this context, since Carmil1 contains the leucine-rich repeat as a protein recognition motif whose structural characterization of this



Fig. 1. A shortened work-schema of molecular docking of CARMIL1 with each of IRAK1 and LRR peptide.

motif has been implicated as a critical for the development of targeted drugs. Hence, the three-dimensional structure of the CARMIL1 protein and its conformation and binding interfaces at the atomic level with the novel LRR peptide was investigated. Through these findings, we examined what molecular features of LRR regions are responsible for interaction with the IRAK1 (Fig. 1).

METHODS

The Obtaining of the 3D Structure of Protein and Peptide

First, the protein sequence encoded by human CARMIL1 (UniProtKB/Swiss-Prot: Q5VZK9.1) in FASTA format was obtained from NCBI. LRR peptide sequence (GRKKRRQRRRPQTLVHLDLSGNVL-RGDDLSHMYNFLAQPNK) was fetched from the paper of Wang *et al.* [10]. Then, each of the amino acid sequences were then subjected to the Robetta [12, 13], a common tool for homology-based approaches to possess a three-dimensional structure of CARMIL1 and LRR peptide as a model (see Fig. 2). The crystal structure of human IRAK1 (PDB ID: 6BFN) from PDB (Protein Data Bank) at http://www.rcsb.org/ was downloaded in PDB format. In Rosetta server job submission, it was used comparative modeling and all settings left as default and generated five 3D-structure

models. It was selected the most accurate one from model proteins according to confidence scores, which indicates the accuracy of model protein in terms of predicted GDT (1.0 good, 0.0 bad).

Pre-Preparation for Molecular Docking

Energy Minimization and Quality Control of Model Protein Structures

The energy minimization of 3D model protein structures was subjected to the minimization method in chimera 1.14 [13]. The default steepest descent was set to 100 with 0.02 step sizes, without fixing any atoms, and was followed by 10 steps of conjugate gradient steps with 0.02 step size (Å) minimization. To control the quality of the model proteins, we evaluated both the structural quality using Qualitative Model Energy Analysis (QMEANDisCo) [15], as well as Ramachandran plots that were drawn to assign key secondary structures to specific regions in the plot.

Sequence Analysis

The protein sequence encoded by Human CARMIL1 (Reference Sequence: NP_060110.4) was retrieved from NCBI in FASTA format and run via PSI-BLAST (Position-Specific Iterated BLAST) [16]. As search setting, it was selected Protein Data Bank (PDB) proteins, as all general algorithm parameters including MATRIX: BLOSUM62, Existence:11 Extension:1, the threshold value (0.05), world size: 6, and



Fig. 2. The 3D structure of predicted CARMIL1 and LRR peptide. Robetta server predicted the 3D model structure of CARMIL1 protein (1-1000aa) and LRR peptide (1-41aa) by the homology-based model. The model was subjected to a mode of protein-protein and peptide docking. The cartoon model representation and image were produced with Chimera 1.14. Structures are symbolized as interactive colored ribbons to view the strand and helix forms.

Compositional adjustments: conditional compositional score matrix were left as the default settings.

Visualization of Molecular Docking Simulations

In this study, it was employed the method that one of the computer-aided drug design (CADD) approaches is structure-based drug design (SBDD) [17]. SBDD methods analyze macromolecular targeting 3D structural details, commonly of proteins or RNAs, to determine fundamental parts and interactions that are significant for their regarding biological procedures. In the docking process, there is no restrictions were placed between surface-exposed residues of proteinpeptide and protein-protein but conversely performed blind docking. In InterEvDock2 docking server, as demonstration Mode for docking procedure, standard usage (easy level) was opted. Two protein sequences or structures and their respective multiple sequence alignments are employed to estimated binding modes via a free docking process at the web server.

The primary structure of CARMIL1 was colored by drawing on Jalview 2.11.1.3 [18]. PyMOL software [19] was used to illustrate the tertiary structure proteins-peptides and analyze the molecular modeling re-

DEV/TUOM COTA LICO	and the second se	260T	270L			290D	300F	310L	320K	3305
	SNRLEELVLE									
	SNRLEELVLE SNRLEELVLE									
5VZK9/1-1371 ANP	340V	350F			TOL ONTEOD	380D	390C	400S	410K	420
	LTASTLVHLD									
	LTASTLVHLD									
	430N	440P	450C	46	10 A	70G	480E	4905	500T	510S
5VZK9/1-1371 SSL	ALMHINLSGT		1	NLKGVSLDL			1	DISDNGLES	DISTLIV	
	ALMHINLSGT									
nterPro/1-1371 SSL	ALMHINLSGT	KLSPEPLKA	LLLGLACNH	NLKGVSLDL	SNCELRSG	GAQVLEGCI	AEIHNITSL	DISDNGLES	BDLSTLIVW	SKNRSIQ
53	OP 54	400	550S	560V	570N	580G	590M	600	K 61	ON
5VZK9/1-1371 LTP			SLADSKLKT	EVTILINA						
	VLDNL VQMIQ									
nterPro/1-1371 LTP	VLDNL VQM I QI	DEESPLQSL	SLADSKLKT	EVTIIINAL	GSNTSLTK	VD I SGNGMG	DMGAKMLA	KALQINTKLE	RTVIWDKNN	ITAQGFQD
	630R	640S	650T	660Y	670Y	6800	69	OM 7	000	710G
5VZK9/1-1371 NYTI	1 1	1 1	1	LENYLLENH				IDRICVKVO	HUNSIBNO	GOALOED
	REMPIPMYD									
nterPro/1-1371 NYTI	REMPIPMYD	ASQALKTNP	EKTEDALQK	IENYLLRNH	IETRKYLQE	AYRLOOGI	VTSTTQQM	DRICVKVQD	HLNSLRNC	GGDAIQED
	790C	800R	810K	820N	830	1 8	40T	850V	860C	870R
25VZK9/1-1371 DAA	ENLCPNVMKK	AHIRODLIH	HASTEKISIP	RTEVKNVL	LEQSGIDIL	NKISEVKLT	VASFLSDR	IVDEILDAL	SHCHHKLAD	HFSRRGKT
	ENLCPNVMKK									
nterPro/1-1371 DAA	ENLCPNVMKK									
	870R	8805	890P	900E	910L	9201			940F	950P
25VZK9/1-1371 DHF										
	SRRGKTLPQQ SRRGKTLPQQ									
		70F	980K	990P	10000	10105	10200			MACK
tam/1-1371 FPS nterPro/1-1371 FPS	LRQEKRSSGF	ISELPSEE	GKKLEHFTKL	RPKRNKKO						
							QNGLMGRV	DEGVDEFFT	KKVTKMDSK	-
			E: LRR	: Pf	am domains	: Interl	QNGLMGRV Pro annotatio	DEGVDEFFT	KKVTKMDSK	KWSTRGSE
SV7K9/1 1971	MTEESOD	10R	201	<mark>і</mark> : Рf 30К	am domains	: Interl	ONGLMGRV Pro annotatio	DEGVDEFFTI ons	70Y	KWSTRGSE A 80K
25VZK9/1-1371 PDB14k1714K171A/9-668	MTEESSDV		201	30K	am domains 40K EVKGOKVEN	SOS	QNGLMGRV Pro annotatio 6 FAFLVTAR	DEGVDEFFTI ons 01 PTKLELTFS	KKVTKMDSK 70Y	KWSTRGSE A 80K SAGNII
	9	10R PREL I ESTI	201 KDVIGRKIK	30K SVKRKVKL SVKRKVKL	am domains 40K EVKGOK VEN EVKGDR VEN	SOS KVLVLTSC	QNGLMGRV Pro annotatio G CAFLVTAR CAFLLSAR	DEGVDEFFTI ons 01 PTKLELTFS PSKLELTFS	70Y 7ULEIHGVVC	KWSTRGSE A 80K SHSACMI HHPACMV
PDB 4k17 4K17 A/9-668	ааок	10R PREL I EST PRELMEST 100	201 KDVIGRKIKI KDVIGRKIKI	30K SVRRKVKL I SVRRKVKL I SVRRKVKL	am domains 40K EVKGDKVEN EVKGDRVEN 120G	E Interi 50S KVLVLTSCI KVLVLTSCI 130V	ONGLMGRV Pro annotatio AFLVTARI AFLLSARI 140S	DEGVDEFFTI ons 01 PTKLELTFS PSKLELTFS 150V	KKVTKMDSK 70Y YLEIHGVVD YLEIHGVID 160S	KWSTRGSE A 80K SKSROMI HKPAGMV 17
2DB 4k17 4K17 A/9-668 2 6VZK9/1-1371	90K	10R VPREL I EST PREL MEST 100 1 SMRMASPE	201 KDVIGRKIKO KDVIGRKIKO IP 111 DVSEVLAHTO	30K SVRRVKL ISVRRVKL OI STCLRKIPP	am domains 40K EVKGDKVEN EVKGDRVEN 120G GLSEVRIMH	Sos Sos NVLVLTSCH NVLVLTSCH 130V KVSMEPSE	ONGLMGRV Pro annotatio GAFLVT ARI CAFLLS ARI 140S RLABLDALW	DEGVDEFFTI ons IT ALELTFS PSALELTFS 150V	YUEIHGYVD YUEIHGYVD YUEIHGYID 1605 1605	KWSTRGSE 80K SKSROMI HKPROMV 17/ ACVCDWLG
DB 4k17 4K17 A/9-668 26VZK9/1-1371	90K	10R PRECIEST PRELMEST 100 SMKMASPE SMKMVSPE	E LRR 201 KDVIGRKIK KDVIGRKIK P 111 DVSEVLAHIO DVSEVLAHIO	30K SVKKKVKL SVKKKVKL OI STCLRKIPP GTCLRRIPP	am domains 40K EVKGBK VEN EVKGBRVEN 120G GLSEVRIM GLSEVRIM	interi 50S KVLVLTSCH KVLVLTSCH 130V KVSMEPSEH KVSMEPSEH	CONGLMGRV Pro annotatio CAFE VT KR CAFE VT KR CAFE L SAR 140S RLASLOALW RLASLOALW	DEGVDEFFTI ons TILLELTFS FSLLELTFS 150V VOSQTVAEQG VOSQTLAEPG	70Y YCETHGVVC YCETHGVVC YCETHGVI 160S PCGGFSQMY PCGGFSQMY	KWSTRGSE 80K SHSROMI HEPROMV 17/ ACVCDWLC
PDB 4k17 4K17 A/9-668 2 6VZK9/1-1371 PDB 4k17 4K17 A/9-668	9 90K VETERCSI 8 VETERCNN	10R PREL I EST 100 I SMHTA SPE MSMK N V SPE 180W	201 KDV1 GRK I K 1 KDV1 GRK I K 1 KDV1 GRK I K 1 IP 111 DVSEVLAH 10 DVSEVLAH 10 190D	30K SVRRVKL SVRRVKL SVRRVKL SVRRVKL SVRRVKL SVRRVF SVRFF SV	am domains 40K EVKGEKVEN EVKGERVEN 120G GLSEVRIMA GLSEVRIMA GLSELRIMA 2101	Interi SOS KVLVLTSCA NVLVLTSCA 130V KVSMEPSE KVSMEPSE 220F	QNGLMGRV Pro annotatio GAFE VT RR RAFELS AR 140S RLASLOALW RLASLOALW 230	DEGVDEFFTI ons TILLELTFS TSLELTFS SLELTFS 150V CSCTVAEQS CSCTLAEPS DL 2	KKVTKMDSK 70Y YLEIHGVVD YLEIHGVID 160S 160S FCGGF SQMT PCGGF SQMT 40R	KWSTRGSE A 80K SESRON I HPP CONV 17/ AC VC DWLC 250E
2008 4k17 4K17 AV9-668 2 6VZK9/1-1371 2008 4k17 4K17 AV9-668 2 5VZK9/1-1371	90K VETEKČSI B VETEKČNI FSVREVO	10R PREL I EST 100 I SMHTA SPE MSMK N V SPE 180W	201 RDV GRK IN RDV GRK IN P 111 DVSEVLAH IC 190D TODTRELNCT	30K I SVARKVAL I SVARKVAL SVARKVAL OI STOLAR ETOLAR 2005 SDF SHUDAR	am domains 40K EVRGSKVEN EVRGRVEN 120G GLSBVRIME GLSBLRIME 2101 DLIPITAAL	SOS SOS K VL VL T SC K VL VL T SC 130V K VSMEP SEI K V SMEP SEI 220F E YNOWF TH	QNGLMGRV Pro annotatio GRAFE VT FRI RAFE VT FRI RAFE LS BRI 1405 RLASL GALW 230 LSSKOLKLS	DEGVDEFFTI ons FTKLELTFS PSKLELTFS 150V VESOTVEGO CSOTLEPS 0L 2 VEOVCEOTER	KKVTKMDSK 70Y YLEIHGVVD YLEIHGVID 160S 160S FCGGF SQMT PCGGF SQMT 40R	KWSTRGSE A 80K SESRON I HPP CONV 17/ AC VC DWLC 250E
PDB 4k17 4K17 AV9-668 2 5VZK9/1-1371 PDB 4k17 4K17 AV9-668 2 5VZK9/1-1371	9 VETEROSI VETEROSI 8 VETERONI 8 SYREEVC 8 SYREEVC	10R PREL I EST PREL MEST 100 I BANK A SPE 180W 200 VDT I YL 200 VDT I YL	201 RDVIGREIR RDVIGREIR P 111 DVSEVLAHIO DVSEVLAHIO 190D TOOTRELNUG	30K 30K ISVARKVAL SVARKVAL SVARKVAL STOLAK F STOLAK F 2005 SDOLAR SDOLAL F SDOLAL SDOLAL F SDOLAL SDOLAL SDOLAL SDOLAL F SDOLAL SDOLAL SDOLAL SDOLA	am domains 40K EVKGDK VEN EVKGDRVEN 120G GLSPLVB1MB GLSPLS1 AL 2101 DL (PLTAAL	EYNOWFTK	QNGLMGRV Pro annotatio RAFEVT R RAFELS R 1405 RLASLOALW 230 LSSKDLKLS	DEGVDEFFTI ons TKLELTFS TSKLELTFS 150V CSOTVECTER 150VCECTER 150VCECTER	70Y TETHON TETHON 160S 160	KWSTRGSE A 80K SISSON I H PACINY ACVCDVI ACVCDVI 250E ELVLENAC
PDB 4k17 4K17 AV9-666 2 5VZK9/1-1371 DDB 4k17 4K17 AV9-660 2 5VZK9/1-1371 DDB 4k17 4K17 AV9-660	90K VETEKCSI VETEKCNI STREEVC 55YREEVC 55YREEVC 260T	10R PRELIEST 100 1 DEMMA SPE 100W 1 DEMMA SPE 100W 200 VOT 1 YL 200 VOT 1 YL 270L	E LRR 201 KDVIGREIK KDVIGREIK P 111 DVSEVLAH 190D TODTRELNU 190D TODTRELNU 280T	CONTRACTOR CONTRA	Am domains 40K EVKGD K VEN EVKGD R VEN 120G GL SP V S I MH GL SP L S I MH 2101 DL I PI TAAL D 34	E Interi 605 KVLVLTSCI 130V KVSMEPSEI 220F E YNCMFTK E YNCMFTK D0F	QNGLMGRV Pro annotatio REFULSER 1405 RLASLOALV RLASLOALV 230 LSSKDLKLS 310L	DEGVDEFFTI ons TKLEKTFS SKLEKTFS SKLEKTFS 160V CSOTVEC 160V 1	70Y TETHOUV TETHOUV TETHOUV TETHOUV TO TO TO TO TO TO TO TO TO TO TO TO TO	KWSTRGSE A 80K SKSRON I H P KONV 17 AC V COVIC AC V COVIC 250E EL V LENAC 340V
PDB 4k17 4K17 AV9-668 Q5VZK9/1-1371 PDB 4k17 4K17 AV9-668 Q5VZK9/1-1371 PDB 4K17 4K17 AV9-668 X5VZK9/1-1371	90K VETEKOSI 9 VETEKON 9 VETEKON 9 VETEKON 9 SYREEVO 5 SYREEVO	10R VERELIEST 100 I BANKIASPE 180W 2MSVADTIYL 2MOVATIYL 200 VATIYL 270L	LRR 201 KDV I GRK I HX KDV I GRK I HX P 111 OVSEV LAHI 190D TOD TRELNET COTRELNET 2807 PNSGLHTINU	E Pf 30K SXXAK VAL SVAK VAL SVAK VAL SVAK VAL 00 STCL R FP 2005	am domains 40K EVKGD K VEN EVKGD R VEN 120G GLSE V R TMR 2101 OL (P) T RAL 2101 OL (P) T RAL D 3 GVSSLSI DF	E : Interi 50S K VL VL TSCI K VL VL TSCI 130V K VSMEPSEI 220F 220F E Y NGWETK E Y NGWETK E Y NGWETK 6 Y NGWETK 7 Y	Pro annotatio	DEGVDEFFTI Ins OI FTRLELTFS SKLELTFS SKLELTFS 150V ISSOL ABP ISSOL AB	70Y TEETHGAY TEETHGAY 1605 FCGGFSGMY FCGGFSGMY 40R TVY SRSNRU 40R TVY SRSNRU 3305 OSLSANPLT	KWSTRGSE A 80K SHSRON I H P KONV 17 AC VCOV 250E EL VLENA 340V ASTLVHLC
26V2K9/1-1371 PDB 4k17]4K17]4/0-660 DEV2K9/1-1371 DDB 4k17]4K17]4/9-660 DEV2K9/1-1371 PDB 4k17]4/K17]4/9-660 X8V2K9/1-1371 PDB 4k17]4/K17]4/9-660	9	10R PRELIEST 100 1 EMPENASPE 180W 2AOVDTIVL 2AOVDTIVL 2AOVDTIVL 2AOVDTIVL 2AOVDTIVL 2AOVDTIVL 2AOVDTIVL 2AOVDTIVL 2AOVDTIVL	LRR 201 KDV I GRK I HX KDV I GRK I HX P 111 OVSEVU AHI I 190D TDD TRELNU 200 TDD TRELNU 200 FINGLET INU PNS GL HT I NU	30K 30K SVARKVRL SVARKVRL SVARKVRL STCLRK FP 2005 20F5HL0HR 2005 AGNPLEDR AGNS EDR	am domains 40K EVKGDK VEN EVKGDR VEN 120G GLSP VR MH 2101 DL 1P1 TAAL DL 1P1 TAAL DL 3 GVSSLS17 GVSSLS17 GVSSLS17	E Interi 505 RVEVETSCI 130V RVEMEPSEI 130V RVEMEPSEI 220F EVNOMETK EVNOMETK EVNOMETK EVNOMETK RVEVELKER KELKER	Pro annotatio	DEGVDEFFTI ons TILELTFS SKLELTFS SKLELTFS 150V DBDTVRC DDL 2 170VCEGTLR 320K SPKGVNSLS SPKGVNSLS	TUE THEY V TUE THEY I 160S FCGGF SOM FCGGF SOM 40R VV SR SNRLE VV SR SNRLE SI SAMPLT	KWSTRGSE 80K SIS SECON I H P RODUC 17 AC VCCW 250E EL VLENAC 340V ASTLVHLC ASTLVHLC
PDB 4k17 4K17 A/0-668 DDB 4k17 4K17 A/0-668 DDB 4k17 4K17 A/0-688 DDB 4k17 4K17 A/0-688 DDB 4k17 4K17 A/0-688 DDB 4k17 A/0-688	9 9 9 9 9 9 9 9 9 9 9 9 9 9	10R PRELIEST 100 IDMKMASPE MEMANYSPE 180W 240.VDTIYL 24	ERR 201 RDVI GRETIK ROVI GRETIK ROVI GRETIK P 101 P 101 P 101 P 100 T DVIEVLANI 100 T DVIEVLANI 200 ROVI RO	: Pf 30K SVARH VAL SVARH VA	am domains 40K EVKCG K VEN EVKCG K VEN EVKCG R VEN 120G CL SP VE MA 2101 DL TP I TAAL D 31 CV SELSI DT 380D	EInterl 505 RVLVLTSCI 130V RVSMEPSEI 130V RVSMEPSEI 220F EYNGAFTK 220F EYNGAFTK 390C	QNGLMGRV Pro annotatio GAREUVT RRI TABLAS RLASLOALW RLASLOALW RLASLOALW SCHOLKLS SCH	DEGVDEFFTI ons T KLELTFS SKLELTFS SKLELTFS 150V CSOTVEC SCOLER	70Y VLEIHGV 160S 160S 160S 160S 160S 160S 160S 160S	KWSTRGSE A SOK SI SIGON I HI P LON V AS VOLVIL 250E EL VILENAA 340V ASTEL VILE ASTEL VILE ASTEL VILE
DB 4k17 4K17 A/0-668 J5VZK9/1-1371 DB 4k17 A/K17 A/0-688 J5VZK9/1-1371 PDB 4k17 4K17 A/0-668 J5VZK9/1-1371 DB 4k17 4K17 A/0-668 J5VZK9/1-1371	9 9 9 9 9 9 9 9 9 9 9 9 9 9	10R PRELIEST 100 152MBAST 100 152MBAST 100 152MBAST 100 200 200 200 200 200 200 200 200 200	201 KDVIGRINK KDVIGRINK KDVIGRINK KDVIGRINK KDVIGRINK KDVIGRINK KDVIGRINK KDVIGRINK KDVIGRINK KDVIGRINK KDVIGRINK KDVIGRINK	30K 30K SVARA VAL SVARA VAL SVARA VAL SVARA VAL STOLRK 15 2005 20F 5-HD AR 2005 20F 5-HD AR 2005 20F 5-HD AR 2005 20F 5-HD AR 2005 20S 20F 5-HD AR 2005 20S 20S 20S 20S 20S 20S 20S 20S 20S 20S	am domains 40K EVKODK VEN EVKODK VEN EVKODK VEN EVKODK VEN SLIDE VEN 2101 CLIDE VEN 2100 CLIDE VEN 2100 CLIDE VEN 2100 CLIDE VEN 2100 CLIDE V	E Interi 505 R VEVETSCI 130V 130V 130V 130V 130V 130V 130V 130V 130V 130V 130V 140	Pro annotatio Pro an	DEGVDEFFTI nns of TILLETFS SKLETFS 150V 15	70Y TLEIHOV TLEIHOV TICOGESOM FCGGESOM FCGGESOM AUR TV SRSNEL SSS SSS SSS SSS SSS SSS SSS S	KWSTRGSE 80K SS SROW I HE PLOY AC VCDALE 250E 250E 250E 250E 250E 250E 250E 250
208 44:7]4K:17 4/0-668 205/27(59/1-1371 208 44:17]4K:17 4/0-688 25/27(59/1-1371 2028 44:17]4K:17]4/0-668 25/27(59/1-1371 2028 44:17]4K:17]4/0-668 25/27(59/1-1371	9 VETEKSI	10R VPRELISTI PRELISTI PRELISTI PRELISTI PRELISTI PRELISTI SAUDI ASPE 1000 VDTIVL 270L SAUDTIVL 270L CASALAHNIN CASALAHNIN DR 3 SPDISHMYNI	ERR ZOI KDVIGRINK KOVIGRINK KOVIGRINK P 11 OVEVLANIT OVEVLANIT OVEVLANIT SOOF SGLHINK SOOF SGLHINK SOOF ELACERA	30K SVARA VAL SVARA	am domains 40K EVKCG K VEN EVKCG K VEN EVKCG K VEN 200 2100 CLEPTRAL 0 3800 3800 51 0 FVG G K 51 0 FVG G K	interior 505 505 507 507 507 507 507 507	CONGLIMARY TO ANNOTATIO CAPPY TO ANNOTATION CAPPY TO ANNO CAPPY TO ANNO CAPP	DEGVDEFFTI nns ol TILLETFS SKLETFS 150V KSDIVECTFS 100V 10	KKVTKMDSK 70Y TLEIHGVV 160S 160GF50M 40R VVSRSNRUE 330S 0SLSANPLT 0SLSANPLT 15LSANPLT PPSFR0FF5	KWSTRGSE ROK SIS ROBIN SIS ROBIN SIS ROBIN F PLOBIN COMPANY SIS ROBINE SIS RO
DB 4k17]4K17]4V0-668 16V2X9/1-1371 DDB 4k17]4K17]4V3-688 16V2X9/1-1371 DDB 4k17]4K17]4V9-688 16V2X9/1-1371 DB 4k17]4K17]AV9-688 16V2X9/1-1371 DB 4k17]4K17]AV9-688	9 9 9 9 9 9 9 9 9 9 9 9 9 9	10R 100 100 100 100 100 100 100	ERR 201 CONTROL OF A 11 CONTROL OF A 10 OVSEVIANT 1900 1007 ELINU 2007 CONTROL NU 2007 CONTROL NU 3007 FLAGPNA UM 450C	PI 30K SOK SO	am domains 40K EVKGD K VEN EVKGD K VEN 120G GL S VE MI GL S L EL MI 2101 01 IP I AAL 01 IP I AAL 01 IP I AAL 02 I S S S S S S S S 3800 SI 0 KWC G KL S E KVC S L 467	Interl 605 KUNTSCI KUNTSCI 130V KUSMEPSE 220F EVICATINE 2100F EVICATINE AMEPKELKA AMEPKELKA AMEPKELKA 330C LKCL LOVL ISOL LOVL -	CANGL MGR V Pro annotatic FART V TRAS RASE L S RE 140S 140S 140S 140S 23 140L 23 15 KOL K K 310L 11 NL SKTSL 400S VL NL SRTV VL NL SRTV 76G	DEGVDEFFTI ns of TILELTS SILELTS SILELTS SILELTS OU SILELTS OU SILELTS OU SILELTS SILELTS OU SILELTS SILE SILE SILES SILE SILES SILE	70Y VEETHON V YEETHON V 1605 160	KWSTRGSE POK SIS KORI I HIP POB V 17/ AC VODAL AC VODAL 250E EL VLENAC SIS SIS F SIS SIS SIS SIS SIS SIS SIS S
DB 4k17 4K17 A/0-668 06VZK9/1-1371 DB 4k17 4K17 A/0-688 26VZK9/1-1371 PDB 4k17 4K17 A/0-668 16VZK9/1-1371 DB 4k17 4K17 A/0-668 16VZK9/1-1371 DB 4k17 4K17 A/0-668 16VZK9/1-1371	9 VETERISI	10R 100 100 100 100 100 100 100	ERR 200 KDWI GRK IN KDWI GRK IN P 11 DV3EVLANI K VSEVLANI K VSEVLANI K SOD CODTRELNCT SOLHTINL SOLF FLACENA I VF FLACENA I VF LACAMMA	PH SOK SVARAH VILL SVARAH VIL	am domains 40K EVIKGD K VEN EVIKGD K VEN EVIKGD K VEN 120G GUS VEN 120G 1	E Interi 605 NULVITSCH NULVITSCH 130V NUMPSE 220F EVROWTK 200F EVROWTK 200F EVROWTK 200F EVROWTK 200F EVROWTK 200F EVROWTK 200F EVROUTSCH 200F	CANGL MGR VI Pro annotatio CAPE V T RE 1405 RL ASL DALM 231 25 KOL KL 25 KOL KL 23 SID 1405 23 SID 1405 25 SID 140	DEGVDEFFTI ons of TKLELTFS SKLELTFS SKLELTFS SKLELTFS 150V SSOL 4 EP 0L 2 150V SSOL 4 EP 0L 2 150V SSOL 4 EP 0L 2 150V SSOL 4 EP 0L 2 150V SSOL 4 EP 150V SSOL 4 EP 150V 150V SSOL 4 EP 150V	KKVTKMDSK 70Y LETHOV VLETHOV 1605 FCGGF50M 40R 330S 05LEANFL SANFL 55AFL PSFK0F5 PSFK0F5 PSFK0F5 2496S LESDLSTL	KWSTRGSE A 80K SIS 805 V 177 AC VCOM C 250E 250
PDB 4k17 4K17 A/0-668 26V2K9/1-1371 PDB 4K17 4K17 A/0-688 26V2K9/1-1371 PDB 4k17 4K17 A/0-668 26V2K9/1-1371 20B 4k17 4K17 A/0-668 26V2K9/1-1371	3 SOK VETERISI	10R 10R 100 100 100 100 100 100	ERR 200 KDWI GRK IN KDWI GRK IN FP 11 DV3EVLAH IN VSEVLAH IN VSEVLAH IN VSEVLAH IN SOOF FI AG PN A IV FI AG PN FI AG PN	SOK SOK SVARANVEL SVARANVE	am domains 40K EVIKGB K VEN EVIKGB K VEN EVIKGB K VEN 120G GUS VEN 1210 GUS VEN 120G GUS SUS VEN 120G GUS SUS VEN 120G 12	E Interi 605 NULVITSCH NULVITSCH 130V NUMPSE 220F EVROWTK 200F EVROWTK 200F EVROWTK 200F EVROWTK 200F EVROWTK 200F EVROWTK 200F EVROUTSCH 200F	CANGL MGR V/ Pro annotatio CAPE V T R + 1405 RL A5 L C ALV 1405 RL A5 L C ALV 23 15 SKOL R + 23 30L LINE SKT SL 1005 10	DEGVDEFFTI ns 0 1 1 1 5 5 1 1 1 1 1 1 1 1 1 1 1 1 1	KKVTKMDSK 70Y LETHOV VLETHOV 1605 FCGGF50M 40R 330S 05LEANFL SANFL 55AFL PSFK0F5 PSFK0F5 PSFK0F5 2496S LESDLSTL	KWSTRGSE A 80K SIS 805 V 177 AC VCOM C 250E 250
PDB 4k17 4K17 A/0-668 26VZK9/1-1371 PDB 4K17 4K17 A/0-668 26VZK9/1-1371 20B 4K17 4K17 A/0-668 26VZK9/1-1371 20DB 4K17 4K17 A/0-668 26VZK9/1-1371 20DB 4K17 4K17 A/0-668	9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	10R 10R 100W 10KH ASPE 180W 190W 190	ERR 201 KDV GRH IN TODTRELNU 1900	PI 30K SVANK VICL SVANK VICL SVANK VICL SVANK VICL SVANK VICL SVANK S	am domains 40K EVKCG K VEN EVKCG K VEN EVKCG K VEN 2100 CL S VK MR 2100 CL PITAL 0 3300 CVSS S OF 3800 CVSS S OF 467 NCEL GHC K 546L	Interior 505 10 L VI TSCI 130V 10 VI VI TSCI 10 VI	CINGL MGR V. Pro annotatic Far L s annotatic Far	DEGVDEFFTI ms of TRLETTS SRLEITS SRLEITS SRLEITS SRLEITS OCCUPACIENTS SPRCMUS	TOY TETHONY TETHONY TETHONY TETHONY TOOLEDING	KWSTRGSE A SOK SOK SOK SOK SOK SOK SOK SOK
PDB 4k17 4K17 A/0-668 PDB 4k17 A/5-688 PDB 4k17 A/5-688 26V2K9/1-1371 PDB 4k17 4K17 A/9-668 26V2K9/1-1371 PDB 4k17 4K17 A/9-668 26V2K9/1-1371 PDB 4k17 4K17 A/9-668 26V2K9/1-1371	9 9 9 9 9 9 9 9 9 9 9 9 9 9	10R 10R 100W 10KH ASPE 180W 190W 190	ERR 201 KD VI GRK IN KD VI GRK IN FO 11 DV SVI AK TO TRELNET TO TRELNET TO TRELNET TO TRELNET SOF FI AO DNA I VH A SOF GACHTINI SOF	PH 30K 30K SVARAK VILL SVARAK VILL 0 STOLSAN, EFF STOLSAN,	am domains 40K EVIKGG K PEN EVIKGG K PEN 120G GLISE V EVIKI 2101 DLIFITAL 01 PILASL 01 PILASL 01 PILASL 01 PILASL 02 SIGNA 03800 01 PILASL 04 SIGNA 04 SIGNA 05	Interf 505 KULVITSCH KULVITSCH KULVITSCH 130V KUMEPSEI 200 Z000 KULMEPSEI 200 Z000 KULMEPSEI X000 KULMEPSEI	CANGL MGR V Pro annotatic MAFE V T RH AFE V T RH 1405 VE ASL CALV 1405 VE ASL CALV 231 255 KD KH 55 KD KH	DEGVDEFFT ns 0 1 5 5 5 5 5 5 5 5 5 5 5 5 5	70Y TELETHOLY TELETH	KWSTRGSE A 90K SI S 10 K SI S 10 K S
PDB 4k17 4K17 AV9-668 Q5VZK9/1-1371 PDB 4k17 4K17 AV9-668 Q5VZK9/1-1371 PDB 4K17 4K17 AV9-668 X5VZK9/1-1371	9 9 9 9 9 9 9 9 9 9 9 9 9 9	10R 10R 100W 10KH ASPE 180W 190W 190	ERR 201 KD VI GRK IN KD VI GRK IN FO 11 DV SVI AK TO TRELNET TO TRELNET TO TRELNET TO TRELNET SOF FI AO DNA I VH A SOF GACHTINI SOF	PH 30K 30K SVARAK VILL SVARAK VILL 0 STOLSAN, EFF STOLSAN,	am domains 40K EVIKGG K PEN EVIKGG K PEN 120G GLISE V EVIKI 2101 DLIFITAL 01 PILASL 01 PILASL 01 PILASL 01 PILASL 02 SIGNA 03800 01 PILASL 04 SIGNA 04 SIGNA 05	Interf 505 KULVITSCH KULVITSCH KULVITSCH 130V KUMEPSEI 200 Z20	CANGL MGR V Pro annotatic MAFE V T RH AFE V T RH 1405 VE ASL CALV 1405 VE ASL CALV 231 255 KD KH 55 KD KH	DEGVDEFFTI ms of TRLETTS SRLEITS SRLEITS SRLEITS SRLEITS OCCUPACIENTS SPRCMUS	70Y TELETHOLY TELETH	KWSTRGSE A 90K SI S 10 K SI S 10 K S
PDB 4k17 4K17 A/0-668 PDB 4k17 A/5-688 PDB 4k17 A/5-688 26V2K9/1-1371 PDB 4k17 4K17 A/9-668 26V2K9/1-1371 PDB 4k17 4K17 A/9-668 26V2K9/1-1371 PDB 4k17 4K17 A/9-668 26V2K9/1-1371	9 90K 9 VETERG N 9 VETERG N	10R 10R 1000 1000A ASPE 1000A 1000A ASPE 1000A 100A	ERR 201 KDVI GRE IN KDVI GRE IN KDVI GRE IN KDVI GRE IN COTTELNE SOUTHELNE S	PI 30K 30K SVAHA VICL SVAHA VI	am domains 40K EVKCG K VEN EVKCG K VEN EVKCG K VEN 2100 01 F11 AAL 0 32 01 F11 AAL 0 33 00 S1 F1 00 S1 F1	Interf 605 10 - 11 - 11 - 11 - 11 - 11 - 11 -	CINGL MGR V. CINGL MGR V. Pro annotatic General V. RAFE L. S. 1405 RLAGLIGATI 405 SKOLING 310L 4005 VLINLS RTV VLINLS RTV VLINLS RTV VLINLS RTV VLINLS RTV 566A 111 TNA GSN 6N	DEGVDEFFT ms 0 T RLELTS S RLELTS S RLELTS 150V 0501 L SP 00 00 00 00 00 00 00 00 00 0	70Y TETHONY TETHONY TETHONY TOOLE	KWSTRGSE SOK SOK SOK SIS STONE I COMMENT SIS SIS SIS SIS SIS SIS SIS SIS
208 44:7]4K:17]4V:0-668 26V2K9/1-1371 208 44:17]4K:17]4V:0-688 26V2K9/1-1371 208 44:17]4K:17]AV:0-688 16V2K9/1-1371 208 44:17]4K:17]AV:0-688 16V2K9/1-1371 208 44:17]4K:17]AV:0-688 16V2K9/1-1371 208 44:17]AV:0-688 16V2K9/1-1371	9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	10R 10R 10C 10C 10C 10C 10C 10C 10C 10C	ELRR 201 KD VI GRK IN KD VI GRK IN KD VI GRK IN P 11 DV3EVL AN 1900 TOUTRELINE TOUTRELINE SOOF ELACEN A VH ELACEN A 450C EGLACHN GLACHN GLACHN GLACHN GLACHN SON TOUTRELINE SON CONTENLAN SON SON SON CONTENLAN SON	PH 30K 30K 20K 20	am domains 40K EVIKGG K PEN EVIKGG K PEN 120G GL S V S I MI GL S V S I MI 2101 DL TP I AAL 01 TP I A	Interf 505 KULVITSCH KULVITSCH 130V KULVITSCH 130V KULVITSCH 130V KULVITSCH 200E 200E KULVITSCH 200E KULVISCH 200E KULVISCH 300C KULVISCH SGGAL KUL KULVISCH KU	CINGL MGR V Pro annotatic MART V T REL ALL S LS BAN 140S	DEGVDEFFT ns 0 1 5 5 5 5 5 5 5 5 5 5 5 5 5	TOY TOY TE HIDS V TOY TE HIDS V TOY TOY TO TOS TO TO TO TO TO TO TO TO TO TO TO TO TO	KWSTRGSE SOK SOK SOK SIS STONI I AC VCOVC AC VCOVC
DB 4417]4517]40-668 ISVZIS91-1371 DB 4817]4717]49-684 ISVZIS91-1371 DB 4817]4717]49-686 SVZIS91-1371 DB 4817]471749-668 SVZIS91-1371 DB 4817]471749-668 ISVZIS91-1371 DB 4817]471749-668 ISVZIS91-1371	9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	10R 10R 10C 10C 10C 10C 10C 10C 10C 10C	ERR 201 KDVI GRE IN KDVI GRE IN KDVI GRE IN KDVI GRE IN COTTELNE SOUTHELNE S	PH 30K 30K 20K 20	am domains 40K EVIKGG K PEN EVIKGG K PEN 120G GL S V S I MI GL S V S I MI 2101 DL TP I AAL 01 TP I A	Interf 505 KULVITSCH KULVITSCH 130V KULVITSCH 130V KULVITSCH 130V KULVITSCH 200E 200E KULVITSCH 200E KULVISCH 200E KULVISCH 300C KULVISCH SGGAL KUL KULVISCH KU	CINGL MGR V Pro annotatic MART V T REL ALL S LS BAN 140S	DEGVDEFFT ns 0 1 5 5 5 5 5 5 5 5 5 5 5 5 5	TOY TOY TE HIDS V TOY TE HIDS V TOY TOY TO TOS TO TO TO TO TO TO TO TO TO TO TO TO TO	KWSTRGSE A SOK SOK SIS STONE I C VCOW I C VCOW
DB 4k17]4K17]4/0-668 15V2K9/1-1371 DB 4k17]4K17]4/0-688 15V2K9/1-1371 DB 4k17]4K17]4/0-686 15V2K9/1-1371 DB 4k17]4K17]A/0-686 15V2K9/1-1371 DB 4k17]4/0-686 15V2K9/1-1371 DB 4k17]4/0-686	9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	10R 10R 10C 10C 10C 10C 10C 10C 10C 10C	ERR 201 KDV GRE IN KDV GRE IN KDV GRE IN KDV GRE IN COTRELNE 1900 1000 TRELNE 2800 TOOTRELNE 2800 SGLHTINE SGLHTINE SGLF LACTRA SCHT KDK GLACHN SSCHT KDK SSCHT KDK SSCHT SSSCHT SSCHT SSCHT SSCHT SSCHT SSCHT SSCHT	PI 30K SVANK VICL S	am domains 40K EVIKGG K PEN EVIKGG K PEN 120G GL S V S I MI GL S V S I MI 2101 DL TP I AAL 01 TP I A	Interl 505 FULLY INTERL 130V VESSED 100V	CONGLIMGR V. CONGLIMGR V. Pro annotatic Sector V. RAFE L. S BRI 1405 140	DEGVDEFFT ns 0 1 5 5 5 5 5 5 5 5 5 5 5 5 5	TOY TOY TE HIDS V TOY TE HIDS V TOY TOY TO TOS TO TO TO TO TO TO TO TO TO TO TO TO TO	KWSTRGSE SOK SOK SOK SIS STONI I AC VCOVC AC VCOVC

: Increased conservation color visibility

Fig. 3. (A) The visualization of the Pfam domain (purple), InterPro annotations (navy blue), and LRR regions (brown) in the CARMIL1 aa sequences. LRR_6 (274-297aa, 569-592aa), CARMIL_C (786-1081aa) from Pfam domains, Leucine-rich repeat-containing protein 16A (962-1012aa) from InterPro domains, There are 11 parts of known LRR regions from Uniprot. All feature sequences were retrieved from Uniprot and PDBeKB. (B) Representation of the primary structure of Human CARMIL1 protein with Mouse CARMIL1 in similar sequences after alignment. This figure solely provides the LRR region sequences in the human CARMIL1.

sults at the 3D atomic level. All complexes of proteinprotein interactions were performed by a fully automatized InterEvDock2 docking server [20-22] through utilizing the SOAP_PP [23] and FRODOCK2 [24]. In the docking process, there is restrictions were placed between surface-exposed residues of protein-peptide and protein-protein but conversely performed blind docking. In InterEvDock2 docking server, as demonstration Mode for docking procedure, standard usage (easy level) was opted. Two protein sequences or structures and their respective multiple sequence alignments are employed to estimated binding modes via a free docking process at the web server.

RESULTS

The Alignments and Sequence Analysis

According to alignment results obtained through the PSI-BLAST tool, the CARMIL1 sequence demonstrated a 91.45% per identity and a query cover 48% with Chain A, Leucine-rich Repeat-containing Protein 16a [Mus musculus Accession: 4K17_A] as a top homolog protein. Thus, this protein was selected as a template for the homology modeling by the Robetta server. Following the alignments of the sequence of CARMIL1 and the template protein, we detected sim-

ilar sequences in LRR regions (Fig. 3, section B). This finding boosts the accuracy of the prediction of the three-dimensional structure of the LRR regions of the CARMIL model protein.

Docking Consequences of CARMIL1 with Each of LRR Peptide and IRAK1

The docking scores and the residues of polar contacts between CARMIL1 and LRR peptides are listed in Table 1-2. As can be seen from Table 1, the top 10 consensus complexes include IES1 and FRODOCK2 with higher energy scores and SOAP_PP4 with lower energy scores. According to the online prediction tools FRODOCK2 and InterEvDock2, in all top docking complexes, E254 (the LRR1 in CARMIL1), K3, and R2 (LRR peptide), which are the top 5 residues (on each chain) on the consensus of the top 10 models, were predicted to be involved in contacts (see Fig. 4 sec. 3.1, and 4.1, Fig. 5 sec. 7.1 and 10.1). As an overall conclusion of docking, the LRR peptide contacted the residues in the LRR regions (one, seven) of the human CARMIL1.

FRODOCK1 and IES1, and SOAP_PP1 are the top consensus complexes by higher docking scores within the representative models from the 10 best clusters (see Table 2). According to the online prediction tools FRODOCK2 and InterEvDock2, in all top dock-

		• •	1 1		
Docking complexes	LRR peptide	CARMIL1 (Residue number)	LRR region	Docking score	
SOAP_PP9	R5, S20, S30, Q38	H428, S432, G456, T485, R630	LRR6, 7	-16724.74	
FRODOCK6	R25	D488, D577	LRR7, 9	1909.43	
SOAP_PP10	R2, R6, R10, K41	E254, S313, S345, D460, N463, S490, S550	LRR1, 3, 4, 7	-16539.82	
FRODOCK7	K3, R25	E254, D343	LRR1, 4	1888.22	
IES1	Q7, R8, H16	E475, G476, T500	LRR7	33.71	
IES3	Y33, G26, D27	I478, D497 , S499	LRR7	31.7	
SOAP_PP1	R2, H31, N40	E254, N255, N311, S313, D371, D488	LRR1, 3, 7	-16991.95	
SOAP_PP2	K41, S30	E254, K265	LRR1	-16925.85	
FRODOCK2	R6	E254 , D460, K608	LRR1	2060.55	
SOAP_PP4	K3, R10, Y33, P39, N40	E254, N255, S373, S490, K575	LRR1, 7, 9	-16773.15	

Table 1. The results obtained from the analysis of the docking CARMIL1 and LRR peptide

The table provides the residues of polar contacts between the model CARMIL1 and LRR peptide. The residue numbers of LRR regions in CARMIL1 are shown in bold font. The lower score is better in the docked complexes of SOAP_PP whilst the higher score is better in the docked complexes of FRODOCK and IES

presenteu.				
Docking complexes	IRAK1	CARMIL1	LRR	Docking
		(Residue number)	region	score
1.FRODOCK5	K242, N211	E249, E254, R630	LRR1	2335.05
2.IES1	Q254, T258, Y277, R194	V252, A256, D261, E289	LRR1, 2	820.97
3.IES2	Q254, Y277	E250, R259	LRR1	697.85
4.IES3	T383, T431	E254, D261	LRR1	658.16
5.IES4	R363, Q254, N211	E254, N255, Q655	LRR1,10	635.63
6.SOAP_PP1	S331, R194	K644, E651 , E658	LRR10	-34824.7
7.SOAP_PP2	R194, K253, S333, S335	N255 , N374, Y637, S640	LRR1	-34575.19
8.SOAP_PP3	Y277, S193, R366	A256, G257, Y637	LRR1	-34535.14
9.FRODOCK1	E248, S333	T280, N282, R630	LRR2	2548.03
10.FRODOCK2	R366, K253, Y284	T280, H452, L661, L662, R663	LRR2, 10	2431.33

Table 2. The results obtained from the analysis of the docking CARMIL1 and IRAK1 are presented.

The table provides the residues of polar contacts between the model CARMIL1 and IRAK1 protein. The residue numbers of LRR regions in CARMIL1 are shown in bont font. The lower score is better in the docked complexes of SOAP_PP whilst the higher score is better in the docked complexes of FRODOCK and IES.

ing complexes, E250, V252, E254 (the LRR1 in CARMIL1) and R194, Q254, Y277 (IRAK1 protein), which are the top 5 residues (on each chain) on the consensus of the top 10 models, were predicted to be involved in contacts (see Fig. 6 sec. 2.1, 3.1, 3.2, and 5.1). As a general result of the docking process, IRAK1 protein contacted the residues in the LRR regions (one, two, and ten) of the human CARMIL1.

DISCUSSION

This study explores to address the in silico analysis of human CARMIL1 protein, which recently was associated with IL-1-induced ERK activation and MPPs expression and whose 3D structure is not fully uncovered yet. CARMIL1, which has an N-terminal domain-containing LRRs that provide broad contacts with other regulatory proteins, is an actin regulator with diverse functions [7]. Until recently, there has been no reliable evidence concerning the role of CARMIL1 in inflammation through the degradation of ECMs by elevating the expression of MMPs. Given the strong molecular evidence, this protein assures a therapeutic target associated with inflammatory diseases such as rheumatoid arthritis, acute lung injury, and chronic periodontitis [10]. However, much uncertainty still exists about the relation between CARMIL1 and inflammatory disorders regarding targeted therapeutic drug development.

In this study, we aimed to predict the model interactions of CARMIL1 with the IRAK1 that displays a functional feature through interacting with CARMIL1 for ERK activation and MMPs expression, and with a novel LRR peptide.

Robetta is a protein structure prediction service continuously evaluated with CAMEO (Continuous Automated Model Evaluation), which constantly assesses the accuracy and reliability of the prediction. The other prediction tools including CAMEO, Robetta, and QMEAN are among the first-line by timebased statistical confidence and reliable performances. The InterEvScore used in the study has determined the heteromeric protein interfaces (polar contacts) and the integration of the evolutionary information obtained from the multiple sequence alignments of each protein in the clusters with a residual-based multi-body statistical potential. In this prediction server, docking searching is systematically implemented utilizing the FRODOCK2 and the outcomes are re-calculated by InterEvScore [22] and SOAP PP atom-based statistical potential to enhance the confidence level of the



Fig. 4. The InterEvDock2, SOAP_PP, and FRODOCK2 model results of CARMIL1 protein and LRR peptide. The docking complexes are represented in a protein interface representation by PyMOL, colored by (CARMIL1 in green, its interaction residues font in LRR regions in red, and LRR peptide in cyan, and possible polar contacts (dash lines) in red). The sections in the figures are enumerated in the order of the list in Table 1 (Such 1.1-4.1). Only figures from docking pose related to LRR regions are concerned.

predictions.

The structural details present in the predicted protein model are significant for defining the biological function of the modeled protein and its use in future experiments, model verification has a critical step in protein structure prediction. To assess the quality level of secondary structure of model CARMIL1 protein and LRR peptide, we drew the Ramachandran plots through the webserver at https://swissmodel.expasy.org. The Ramachandran plots are represented as an x-y plot of the Phi (ϕ) and Psi (ψ) of the dihedral angles between N-Ca and Ca-C planar peptide bonds in a protein's backbone. The Phi (ϕ) and Psi (ψ) are regarded as conformational angles that determine the conformation of the whole chain of a protein. All theoretically probable secondary structures are shaded in orange and dark blue as the most favored regions in the plots (Figs. 8 and 9). To interpret these plots, we used the outputs of MolProbity [25] tool, which is a structure-validation web service that provides evaluation of model quality. Ramachandran Favoured score is 96.66% for the CARMIL1, while this score is 100.00 % for LRR peptide. The Ideal score for the Ramachandran Favoured in the MolProbity tool has been reported > 98% as an ideal case.

Protein structures can contain multiple intense foldable parts, namely domains. These domains comprise typical hydrophobic cores, and can be folded free of each other, and are nearly always connected to establish diverse roles [26]. In this manner, our first pre-



Fig. 5. The docking complexes are represented in a protein interface representation by PyMOL, colored by (CARMIL1 in green, its interaction residues font in LRR regions in red, and LRR peptide in cyan, and possible polar contacts (dash lines) in red). The sections in the figures are enumerated in the order of the list in Table 1 (Such 5.1-10.3). Only figures from docking pose related to LRR regions are concerned.



Fig. 6. The InterEvDock2, SOAP_PP, and FRODOCK2 model results of the CARMIL1 and IRAK1 protein. Binding modes of the CARMIL1 and IRAK1 in the protein interface representations are shown, colored by (CARMIL1 in green, its interaction residues font in LRR regions in red, and IRAK1 protein in cyan and magenta). The sections in the figures are enumerated in the order of the list in Table 2. (Such 1.1-5.2) Only figures from docking pose related to LRR regions are included.

liminary approach in this study was to analyze the LRR regions in CARMIL1 through Uniprot and PDBe-knowledge base (see Fig. 3, section A) since Wang *et al.* [6] demonstrated that the LRRs of CARMIL1 associate with IL-1 signaling proteins.

LRRs are nearly 21-28 residues (Figs. 6 and 7).s sequence motifs existing in various proteins with distinct functions. The main role of these motifs seems to implement a multidirectional structural framework for the composition of protein-protein interactions. Recent structural information in the LRR proteins has heightened the need for our opinion concerning the structural factors, our experience to model such proteins with uncharacterized structures, and has illuminated how these proteins attend in protein-protein interactions [27]. For this purpose, the molecular docking process may be a satisfactory and robust computational approach to recognize the function of LRRs in protein-protein interaction at the atomic level.

The cell-permeable, CARMIL1 binding LRR peptide, has been reported to inhibit IL-1-induced collagen degradation by MMPs [6]. The novel LRR peptide consists of 41 amino acids, most of which are basic and hydrophilic (see part 2.1.). The results, as shown in Figs. 4 and 5, indicate that lysine(K) and arginine(R) in the LRR peptide are the basic amino acids, whereas glutamate (E) and aspartate(D) in the CARMIL1 are acidic amino acids in the most parts of whose interaction poses in the complexes. Besides, according to the calculation results, the LRR peptide mostly interacts with the residues in the LRR 1 and LRR 7 regions (see Table 1).

Binding modes of CARMIL1 with the IRAK1 protein by molecular docking simulation are displayed in Figs. 3 and 4. In general, IRAK1 interacted with the protein kinase domain between the residue number of 212-521 (see Table 2). As a computational output, residues Q254 and Y277 in the protein kinase domain are the most common polar interactions with CARMIL1, while CARMIL1 binds to IRAK1 through



Fig. 7. The InterEvDock2, SOAP_PP, and FRODOCK2 model results of the CARMIL1 and IRAK1 protein. Binding modes of the CARMIL1 and IRAK1 in the protein interface representations are shown, colored by (CARMIL1 in green, its interaction residues font in LRR regions in red, and IRAK1 protein in cyan and magenta). The sections in the figures are enumerated in the order of the list in Table 2. (Such 6.1-10.2) Only figures from docking pose related to LRR regions are included.



Fig. 8. Ramachandrant and structure quality estimate plot of CARMIL1 from the webserver at https://swissmodel.expasy.org.

V252, E254 residues in LLR1 (see Figs. 6 and 7). According to the results of both the LRR peptide and IRAK1 protein docking with the model of human CARMIL1 protein, E254, N254, and R630 commonly interacted residues. All docked complexes with the LRR peptide and IRAK1 are related to LRRs regions in CARMIL1. These calculations would seem to suggest that the primary function of the LRRs performs a sophisticated structural frame for the formation of protein-protein interactions.

Limitations

As limitations of the study, to be sure and explain the stability of the molecular docking complexes, classical molecular dynamic (MD) process is advised to conduct. InterEvDock2 docking server is presently not able to dock nucleic acids or small molecules. When nucleic acids or ligands are present in a protein chain,

Fig. 9. Ramachandrant and structure quality estimate plot of LRR peptide from the webserver at https://swissmodel.expasy.org.

they will be kept only as steric objects. In this study, as only one ligand is conducted, comparison assessment is recommended by targeting a similar ligand to the LRR region of the CARMIL1 protein just as the LRR peptide is targeted.

CONCLUSION

Inflammation is normally a response of the body to infection and ailment. However, it may be seldomly misled, hereby the immune system preferably attacks healthy tissues. In the current study, an in-silico analysis was performed to predict the polar contacts to evaluate the binding mode and affinities of the CARMIL1 with the IRAK1 protein and the novel LRR peptide. The docking score and protein-protein interaction with significant amino acid residues were identified as hits through the docking server. Both the LRR peptide and IRAK1 protein contacted with the residues of E254 and N255 in the LRR1 region of human CARMIL1. The outcomes indicate that the LRR1 region is more significant in evaluating the LRR peptide for its effectiveness in inhibiting the interaction of IRAK1 and CARMIL1 protein. Taken together, as three scoring programs in the docking process have confirmed each other, we conclude that the models with the highest binding energy are the complexes that interact with residues in the LRRs regions of the CARMIL1. This study may contribute to future studies, as CARMIL1 being a promising target for anti-inflammatory drug development.

Authors' Contribution

Study Conception: NB, GY; Study Design: NB, GY; Supervision: NB, GY; Funding: NB, GY; Materials: NB, GY; Data Collection and/or Processing: NB, GY; Statistical Analysis and/or Data Interpretation: NB, GY; Literature Review: NB, GY; Manuscript Preparation: NB, GY and Critical Review: NB, GY.

Ethical approval

There are no studies with the human participant or animal performed by any of the authors in this paper.

Conflict of interest

The authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during conduction or writing of this study.

REFERENCES

1. Hönig J, Rordorf-Adam C, Siegmund C, Wiedemann W, Erard F. Increased interleukin-1 beta (IL-1 beta) concentration in gingival tissue from periodontitis patients. J Periodontal Res 1989;24:362-7.

2. Muzio M, Ni J, Feng P, Dixit VM. IRAK (Pelle) family member IRAK-2 and MyD88 as proximal mediators of IL-1 signaling. Science 1997;278:1612-5.

3. MacGillivray MK, Cruz TF, McCulloch CAG. The recruitment of the interleukin-1 (IL-1) receptor-associated kinase (IRAK) into focal adhesion complexes is required for IL-1beta-induced ERK activation. J Biol Chem 2000;275:23509-15.

4. Arora PD, Ma J, Min W, Cruz T, McCulloch CAG. Interleukin-1-induced calcium flux in human fibroblasts is mediated through focal adhesions. J Biol Chem 1995;270:6042-9.

5. Lo YYC, Luo L, McCulloch CAG, Cruz TF. Requirements of focal adhesions and calcium fluxes for interleukin-1-induced ERK kinase activation and c-fos expression in fibroblasts. J Biol Chem 1998;273:7059-65.

6. Wang Q, Rajshankar D, Laschinger C, Talior-Volodarsky I, Wang Y, Downey GP, et al. Importance of protein-tyrosine phosphatase-alfa catalytic domains for interactions with SHP-2 and interleukin-1-induced matrix metalloproteinase-3 expression. J Biol Chem 2010;285:22308-17.

7. Stark BC, Lanier MH, Cooper JA. CARMIL family proteins as multidomain regulators of actin-based motility. Mol Biol Cell 2017;28:1713-23.

8. Yang C, Pring M, Wear MA, Huang M, Cooper JA, Svitkina TM, et al. Mammalian CARMIL inhibits actin filament capping by capping protein. Dev Cell 2005;9:209-21.

9. Zwolak A, Yang C, Feeser EA, Ostap EM, Svitkina T, Dominguez R. CARMIL leading edge localization depends on a non-canonical PH domain and dimerization. Nat Commun 2013;4:1-10.

10. Wang Q, Notay K, Downey GP, McCulloch CA. The leucinerich repeat region of CARMIL1 regulates IL-1-mediated ERK activation, MMP expression, and collagen degradation. Cell Rep 2020;31:107781.

11. Akçeşme FB, Beşli N, Peña-García J, Pérez-Sánchez H. Assessment of interaction of human OCT 1-3 proteins and metformin using silico analyses. Acta Chimica Slovenica 2020;67:1202-15.

12. Song Y, DiMaio F, Wang RY-R, Kim D, Miles C, Brunette TJ, et al. High-resolution comparative modeling with RosettaCM. Structure 2013;21:1735-42.

13. Raman S, Vernon R, Thompson J, Tyka M, Sadreyev R, Pei J, et al. Structure prediction for CASP8 with all-atom refinement using Rosetta. Proteins Struct Funct Bioinforma 2009;77(S9):89-99.

14. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605-12.

15. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 2018;46(W1):W296-303.

16. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389-402.

17. Yu W, MacKerell AD. Computer-aided drug design methods. In Antibiotics Humana Press, NY, 2017: pp. 85-106.

18. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 2009;25:1189-91.

19. Schrödinger, LLC. The {PyMOL} Molecular Graphics System, Version~1.8. Nov 2015.

20. Quignot C, Rey J, Yu J, Tufféry P, Guerois R, Andreani J. In-

terEvDock2: an expanded server for protein docking using evolutionary and biological information from homology models and multimeric inputs. Nucleic Acids Res 2018;46(W1):W408-16.

21. Yu J, Vavrusa M, Andreani J, Rey J, Tufféry P, Guerois R. InterEvDock: a docking server to predict the structure of proteinprotein interactions using evolutionary information. Nucleic Acids Res 2016;44(W1):W542-9.

22. Andreani J, Faure G, Guerois R. InterEvScore: a novel coarse-grained interface scoring function using a multi-body statistical potential coupled to evolution. Bioinformatics 2013;29:1742-9.

23. Dong GQ, Fan H, Schneidman-Duhovny D, Webb B, Sali A. Optimized atomic statistical potentials: assessment of protein in-

terfaces and loops. Bioinformatics 2013;29:3158-66.

24. Ramirez-Aportela E, López-Blanco JR, Chacón P. FRODOCK 2.0: fast protein--protein docking server. Bioinformatics 2016;32:2386-8.

25. Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, et al. MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci 2018;27:293-315.

26. Richardson JS. The anatomy and taxonomy of protein structure. In: Advances in protein chemistry Elsevier, 1981: pp. 167-339.

27 Kobe B, Kajava A V. The leucine-rich repeat as a protein recognition motif. Curr Opin Struct Biol 2001;11:725-32.



This is an open access article distributed under the terms of Creative Common Attribution-NonCommercial-NoDerivatives 4.0 International License.