Supporting Information: Hanske *et al.* "An intra-domain allosteric network modulates the calcium affinity of the C-type lectin receptor Langerin"

**Jonas Hanske, Stevan Aleksić, Martin Ballaschk, Marcel Jurk, Elena Shanina, Monika Beerbaum, Peter Schmieder, Bettina G. Keller*, Christoph Rademacher***

Supporting Figures

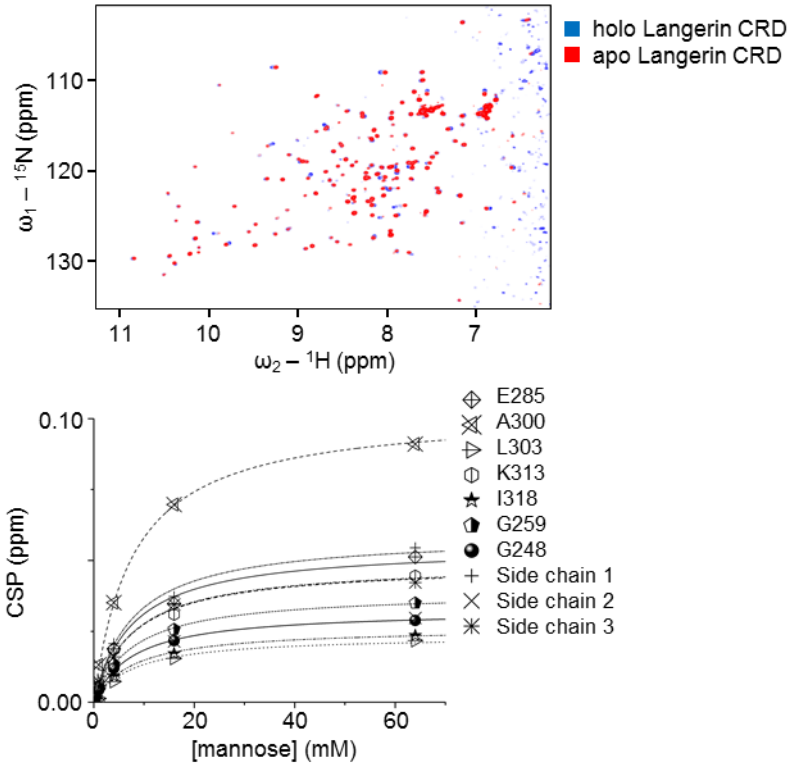



Figure S1: Mannose binding to Langerin CRD at pH 6 measured by $^{1}$H-$^{15}$N HSQC NMR. (Upper panel) Overlay of spectra during addition of mannose starting from the holo form (red) up to 65 mM mannose concentration (blue). Titration points were 1, 4, 16, and 62 mM mannose. (Lower panel) Chemical shift perturbations for selected resonances fitted to a one-site binding model awarding a $K_d$ of 11.4±1.4 mM. All spectra were recorded at 298 K.

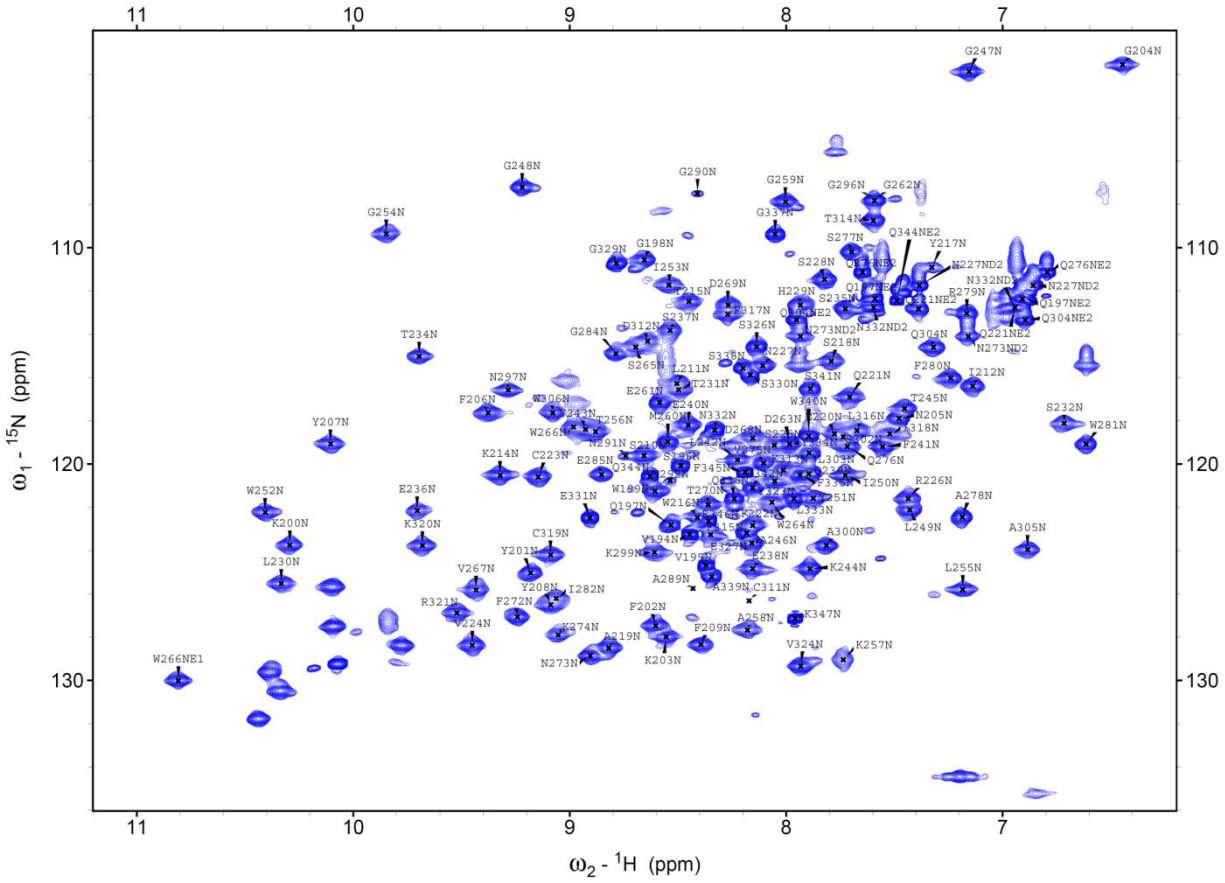



Figure S2: $^{1}H$-$^{15}N$ HSQC NMR spectrum with assigned backbone and side-chain resonances of recombinant Langerin CRD construct with C-terminal Strep tag. 92% of all backbone amide resonances were assigned using standard triple resonance spectra, NOESY spectra, and mutant analysis. In addition, some side chain amides and W266 indol NH resonances were assigned. Backbone resonances are marked with the corresponding residue and N, side chain resonances with the corresponding nitrogen NE1 for tryptophan indole, ND2, or NE2 for asparagine and glutamine side chain amides, respectively. All assignment spectra were recorded at pH 6, 298 K, and in presence of 5 mM $CaCl_2$.

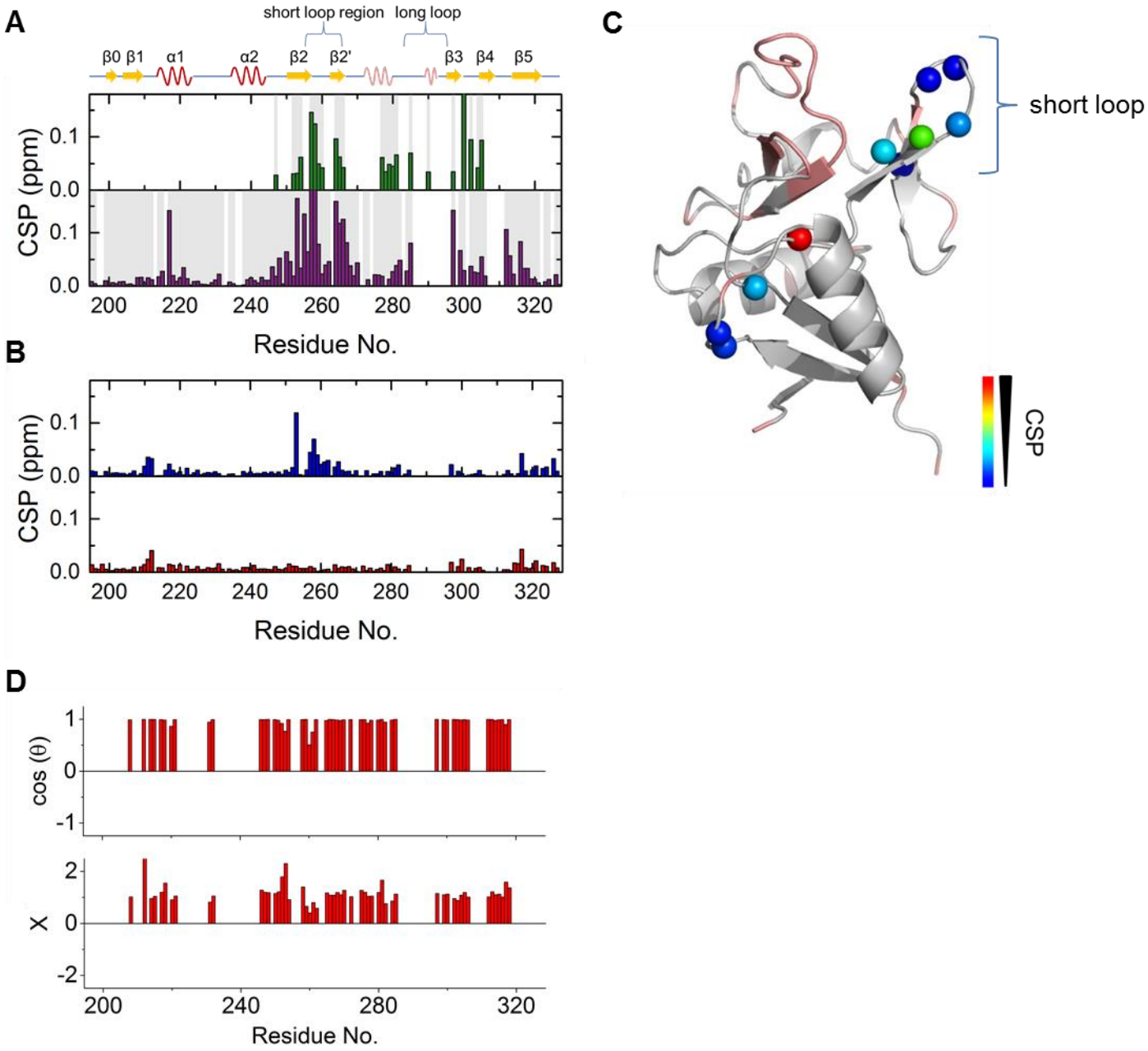



Figure S3: (A) Chemical shift differences observed for the apo cis and trans states (upper panel) and chemical shift perturbation of Langerin CRD after addition of $CaCl_2$ (10 mM) (lower panel) at pH 7. (B) Chemical shift perturbations for apo (lower panel) and holo Langerin CRD (upper panel) upon change of pH. While the apo form remained largely unaffected, in the holo form pronounced CSP were observed in the β2 region and the following short loop. (C) Chemical shift perturbations above 0.02 ppm are heat-map labeled on the protein structure (pdb entry: 3p5f, Feinberg et al., 2011). Unassigned regions are colored salmon. The short loop experiencing most-pronounced changes is labeled. (D) CHESPA analysis of pH-dependent chemical shift changes in Langerin CRD holo form

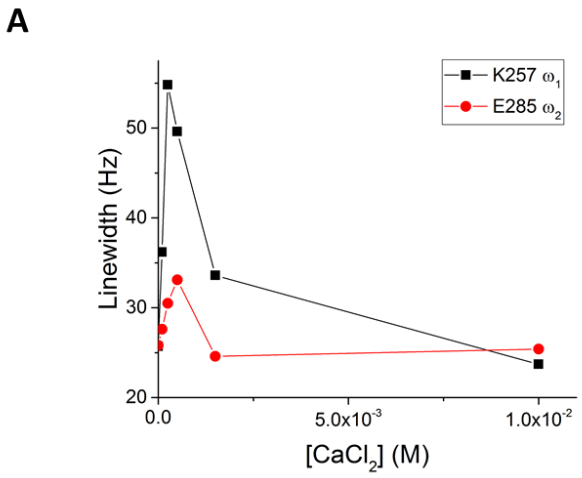
reveal a non-linear relationship. The vector defined by holo and apo Langerin CRD at pH 6 was defined as reference and cos(θ) and relative length of the projection X of the vector defined by the holo form at pH 7 and the apo form at pH 6 calculated. Only residues with a vector length above 0.02 ppm are shown. Although the majority of changes occur along the same axis (cos(θ)≈1), this is not true for some residues in the β2 strand and the sequential short loop. Of note, in the same region the vector projections differ highly in relative length. (E) Chemical shift differences induced by the addition of $Ca^{2+}$ (lower panel) and observed for the trans and cis conformations of the apo form (lower panel). Grey bars indicate assigned resonances in the respective two forms used for CSP calculations.

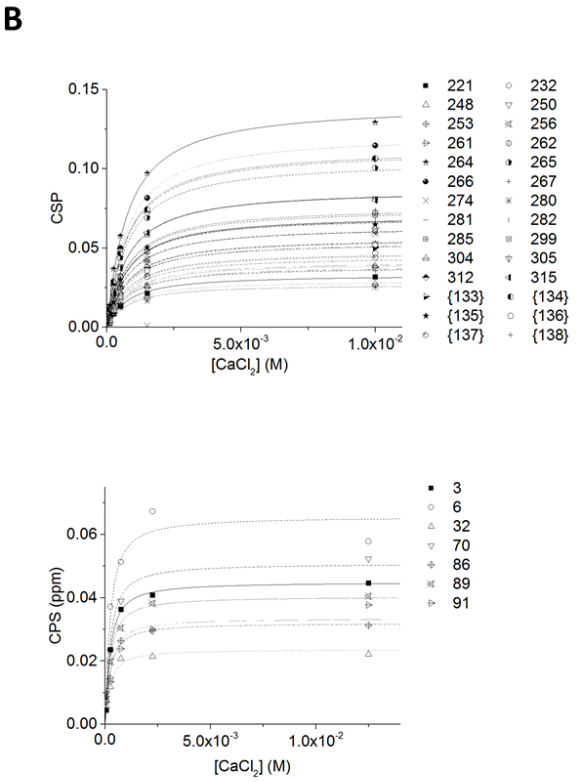
```
SP|Q9UJ71|CLC4K_HUMAN          ------MTVEKEAPDAHFTVDKQNISLWPRE-PPPKSGPSLVPGKTPTVRAALICLTLVL 53
SP|Q8VBX4|CLC4K_MOUSE          --MPEAEMKE-EAPEAHFTVDKQNISLWPREP-PPKQDLSPVLRKPLCICVAFTCLALVL 56
TR|D3ZBX0|D3ZBX0_RAT           --MPEVEMKETEVPDAHFTVDKQNISLWPREP-PPKQDLTPVLRKPHCICAAFICLALVL 57
TR|B3FVQ1|B3FVQ1_PIG           -----MKAVESEVHDAHFTVDKQNISLWPREP-PPKTGPYLVLGRCLTVRAAVVLLTLVL 54
TR|E1BB51|E1BB51_BOVIN         -----MKTADSEVLDAHFTVDKQNVSLWPPDP-PPKTCPSQVLGRLLTVRAVVVFLMLVT 54
TR|F7EFC9|F7EFC9_CALJA         ------MTAEREVPDAHFTVDKQNISLWPRAEPPPKSGSSLVPGKTSTVCAALICLTLVL 54
TR|H0WH23|H0WH23_OTOGA         ---LRMKGTEGEIPDAHFTVDKQNISLWPQAEPPPKSGLCVVLRKAPAIGAVLVCLTLVL 57
TR|G3WUZ1|G3WUZ1_SARHA         VAQARMKENVMDAQVAHFTVDNQDISLWSRDP-PPKKESPIDPRTLPSVYVFIILLALGL 59
TR|H0V494|H0V494_CAVPO         -----MKPEEKEAPDAHFFVDKQNIALWPRDP-PSKQDPPLGPRKSFTIRVVLICLSLIM 54
TR|I3N0C4|I3N0C4_ICTTR         -----MKAAEMEVPDAHFTVDKQNISLWPREP-PPKQDPSLVLRKPLILRVALICLSLVL 54
TR|F6TZE3|F6TZE3_HORSE         -------AAESEVPDAHFTVDKQNISLWPRAPPPPKMGPSLVLRKLLTVRAAVIFLLLVL 53
TR|F6U7K2|F6U7K2_MONDO         VAEARMKENEVDVQVAHFTVDNQDISLWPRDS-PPKKESSISLRTLTSVHVVIILMALVL 59
TR|H2R8S3|H2R8S3_PANTR         ------MTVEREAPDAHFTVDKQNISLWPRE-PPPKSGPSLVPGKTPTVRAALICLTLVL 53
TR|G1MBS3|G1MBS3_AILME         -----MKAAESEVPDVHFTVDKQNVSLWPREP-PPKTGPCLVLRKPLTVQAAVIILTLGL 54
TR|M3WIV5|M3WIV5_FELCA         -----MKAAESEVSDVHFSVDKQNISLWPREP-PPKTGPSLVLRKSLTVRAAVIILTLGL 54
TR|G1RGW2|G1RGW2_NOMLE         ------MTVEREAPDAHFTVDKQNISLWPRE-PPPKSGPSLLPGKTPTVRAALICLTLVL 53
TR|M3Y7D9|M3Y7D9_MUSPF         -----MKAAESEVPDAHFTVDKQNISLWPRDP-PPKTDPCLVLRKPLMVRAAVITLILGF 54
TR|G1PJQ2|G1PJQ2_MYOLU         --------------------------LLPTEP-PPKMGSSPVLRKPCTVRAAFVILTLVL 33
TR|W5PNL3|W5PNL3_SHEEP         -----MKAADSEVPDAHFRVDKQNVSLWPPDP-PPKTCPSQVLGRLLTVHAVVVFLTLVT 54
TR|G3QPX8|G3QPX8_GORGO         ------MTVEREAPDAHFTVDKQNISLWPRE-PPPKSGPSLVPGKTPTVRAALICLTLVL 53
TR|F7HNH2|F7HNH2_MACMU         ------MTAEREVPDAHFTVDKQNISLWPRATPPPKSGPSLFPGKTPTVRAALICLTLVL 54
TR|G3TJF3|G3TJF3_LOXAF         -----RMTAGEELPCAHFEVDKQNISLWPREP-PSKMDRPLALRKPSTLCAVLTFLTLVL 54
TR|A0A0P6J2W2|A0A0P6J2W2_HETGA -----MKAAEKEAHDAHFTVDKQNIALWPREP-PPKQALSLHPRKPLTIRVVLICLSLVV 54
TR|G1TDE9|G1TDE9_RABIT         -----MQAADREALDAHFTVDKQNISLWPREP-SPKPAPSPLLREPVSVRAALICLSLVL 54
TR|A0A0D9RRP0|A0A0D9RRP0_CHLSB ------MTAEREVPDAHFTVDKQNISLWPRAEPPPKSGPSLFSGKTPTVRAALICLTLVL 54
TR|F7A5P2|F7A5P2_MACMU         ------MTAEREVPDAHFTVDKQNISLWPRE-PPPKSGPSLFPGKTPTVRAALICLTLVL 53
TR|F6Y539|F6Y539_CANLF         -----MKAAESEGPDAHFAVDKQSISLRPREP-PPDTGPCPVPRKPLVVQAAAVVLLLGL 54
TR|A0A096N0R1|A0A096N0R1_PAPAN ------------------------------------------------------------

SP|Q9UJ71|CLC4K_HUMAN          VASVLLQAVLYPRFMGTISD--VKTNVQLLKGRVDNIST-LDSEIKKNSDGMEAAGVQIQ 110
SP|Q8VBX4|CLC4K_MOUSE          VTSIVLQAVFYPRLMGKILD--VKSDAQMLKGRVDNIST-LGSDLKTERGRVDDAEVQMQ 113
TR|D3ZBX0|D3ZBX0_RAT           VTSIVLQAVFYPRLMGKILD--VKSDAQMLRGRVDNIST-LGSDLKRERGRVDDAEVQMR 114
TR|B3FVQ1|B3FVQ1_PIG           VASILLQAILYPWFMGTMSD--VKTSAQLLQGRVDNIST-LSSEIKRNSGGVEAAGVQVR 111
TR|E1BB51|E1BB51_BOVIN         ITSVLLQAILYPWFMGTISD--VKTNAQLLKGRVDNISS-LSSEIKRNRGGLVAVGIQVR 111
TR|F7EFC9|F7EFC9_CALJA         VASVLLQAVLYPRFMGTISD--VKTNVQLLKGRVDNIST-LDSEIKKNSDGMEVAGTQIQ 111
TR|H0WH23|H0WH23_OTOGA         VASLLLQGILYPWFMGTISD--VKTNAQLLKGRVDNISA-LGSEIKKNSGSVEAAGLQFH 114
TR|G3WUZ1|G3WUZ1_SARHA         VASLLATFILYPQLLGRIDD--VKTSVHMLKGHAENASVFLSSEIKTLRNKLKDSRTQIQ 117
TR|H0V494|H0V494_CAVPO         VTSIVLQAILYPWFMGTISD--VKANAQWLKGRVDNISI-LGSEMKKDSGGMKAVSVQIE 111
TR|I3N0C4|I3N0C4_ICTTR         VASVVLQAILYPWFMGTISD--IKTDAQLLKGRVNNVSA-LSSKVEKNSRDVEAAGVQIR 111
TR|F6TZE3|F6TZE3_HORSE         VASVTLQAILYSWFMGTISD--VKSNAQLLKGRVDNIST-LSSEIKRNRGGVAATGVQVQ 110
TR|F6U7K2|F6U7K2_MONDO         AAFLLAMAILYPQLLGKIDD--VKTSVHMLKGRAENVSVFLRSEIRKLRNKLEDSHTELQ 117
TR|H2R8S3|H2R8S3_PANTR         VASVLLQAVLYPRFMGTISD--VKTNVQLLKGRVDNIST-LDSEIKKNSDGMEAAGVQIQ 110
TR|G1MBS3|G1MBS3_AILME         TASVLLQAIFYPWFMRTISD--VKTNAQLLKGHVDNIST-LGSEVKRNRGGVEATGIQVR 111
TR|M3WIV5|M3WIV5_FELCA         AASILLQAVLYPWFMRTISN--VKNNAQLLKGRVDNISA-LGSEIKRNRGGVEATGIQVQ 111
TR|G1RGW2|G1RGW2_NOMLE         VASILLQAVLYPWFMGTISD--VKTNVQLLKGRVDNIST-LDSEIKKNSDGMEAAGVQIQ 110
TR|M3Y7D9|M3Y7D9_MUSPF         VASILLQATLYPWFMRTISD--VKTNAQLLKGSVDNIST-LGSEIKRNRGSVEATSLQVR 111
TR|G1PJQ2|G1PJQ2_MYOLU         LASVVLQAILYPWFMGTLSD--VKTNAQLLNGRVDNIST-LSSEVKRNRGGVKAAGIQVQ 90
TR|W5PNL3|W5PNL3_SHEEP         VTSVLLQAILYPWFMGTISD--VKTNAQLLKGRVDNISS-LSSEVKRNRGALVAVGFQVR 111
TR|G3QPX8|G3QPX8_GORGO         VASVLLQAILYPRFMGTISD--VKTNVQLLKGRVDNIST-LDSEIKKNSDGMEAAGVQIQ 110
TR|F7HNH2|F7HNH2_MACMU         VASILLQAVLYPRFMGNISD--VKTNVQLLKGRVDNIST-LDSEIKKNSDGMEAAGVQIQ 111
TR|G3TJF3|G3TJF3_LOXAF         VTSILLQSILYSWLLGTISD--VRINAQLLKGRVDNIST-EGSEIKRNQGGLKTASVQIQ 111
TR|A0A0P6J2W2|A0A0P6J2W2_HETGA VASIVLQAILYPWFMGTISD--VNTNAQMLKSRVDNMST-LGSEIEKNSGSVKVAGVQIQ 111
TR|G1TDE9|G1TDE9_RABIT         AASVVLQAVLYPWFMGRISD--VKTDAQLLKDRVGNISS-LGSEVRKNGVGVQAAGAQIH 111
TR|A0A0D9RRP0|A0A0D9RRP0_CHLSB VASVLLQAVLYPRFMGSISD--VKTNVQLLKGRVDNIST-LDSEIKKNSDGMEAAGVQIQ 111
TR|F7A5P2|F7A5P2_MACMU         VASILLQAVLYPRFMGNISD--VKTNVQLLKGRVDNIST-LDSEIKKNSDGMEAAGVQIQ 110
TR|F6Y539|F6Y539_CANLF         SASVLLQAALCESLGFTVPAGAFSTASQLPQ-HPDPLLP----AVQV----SPTSGFSVG 105
TR|A0A096N0R1|A0A096N0R1_PAPAN --------------MGNISD--VKTNVQLLKGRVDNIST-LDSEIKKNSNGMEAAGVQIQ 43
                                                :   .   :  .            :.           ..

SP|Q9UJ71|CLC4K_HUMAN          MVNESLGYVRSQFLKLKTSVEKANAQIQILTRSWEEVSTLNA--QIPELKSDLEKASALN 168
SP|Q8VBX4|CLC4K_MOUSE          IVNTTLKRVRSQILSLETSMKIANDQLQILTMSWGEVDSLSA--KIPELKRDLDKASALN 171
TR|D3ZBX0|D3ZBX0_RAT           IVNTSLGRVRSKILSLEASMKIVSNQLQVLTMNWGEVDNLNA--KIPELQKDLDKASALN 172
TR|B3FVQ1|B3FVQ1_PIG           MVNASLDHVRSQIQRLEISVKEADARLQKLTSSWEEVDKLNA--QIPELKRDLDKASALN 169
TR|E1BB51|E1BB51_BOVIN         MVNASLDRISSQIRRLETGLKEASAQLQVLTSSWKAVDELNA--QIPELKQDLDKASALN 169
TR|F7EFC9|F7EFC9_CALJA         LVNVSLGYVRSQFLKLKTSVEKANAQIQILTRSWEDVGTLNT--QIPELKRDLEKASALN 169
TR|H0WH23|H0WH23_OTOGA         LVNASLSHVRSQILKLQTSVEKANAQIQTLTRNWEEVDTLNA--QIPELKRDLDKASALN 172
TR|G3WUZ1|G3WUZ1_SARHA         MLNFSMQYSKVHMQMLNAMLQENSNHLQELMSNWEELEHLNA--QIPMLKQDMKKTGELN 175
TR|H0V494|H0V494_CAVPO         TVNASLDSVHSQILILKTNVEKASAQIDNLRKKWEEADHLNT--QIPELKRDLDKASALN 169
TR|I3N0C4|I3N0C4_ICTTR         MVDASLDSVNAQILTLTASIEKATAEIQSLTRNWEEVGHLNA--QIPELKRDLDKASALN 169
TR|F6TZE3|F6TZE3_HORSE         MVNASLDRVRSQMRRLETDVKEANARIQTLTKSWEEVDNLNA--QIPELKRDLDKASTLN 168
TR|F6U7K2|F6U7K2_MONDO         MLNGTLRVSNAQIQLLNSKLKNANNQLQTLMRGWQELDDLNS--QIPMLKQDLAKTGELK 175
TR|H2R8S3|H2R8S3_PANTR         MVNVSLGYVRSQFLKLKTSVEKANAQIQILTRSWEEVSTLNA--QIPELKSDLEKASALN 168
```

```
TR|G1MBS3|G1MBS3_AILME         MVNASLDHVRSQIGKLEAGVKKADAQIQILTRRWEEVIDLNA--QIPALKRDLDKASSLN 169
TR|M3WIV5|M3WIV5_FELCA         KVNASLDHVRSQIWKLETGVKEANAQIQMLTRSWEAVNDLNA--QIPELKRNLDKASSLN 169
TR|G1RGW2|G1RGW2_NOMLE         MVNVSLGYVLSQFLELKTSVEKANAQIQILTRSWEEVSTLNA--QIPELKSDLEKASALN 168
TR|M3Y7D9|M3Y7D9_MUSPF         MVNASLAHVHSQIGKLEAGVRDAAAQIQMLTRRWEEVNDLNA--QIPALKRDLGKASSLN 169
TR|G1PJQ2|G1PJQ2_MYOLU         MVNASLDHVHSQIRRLETGMKAVNAQIQMLTRSWEEVYSLNA--QLPELKRDLDKASALN 148
TR|W5PNL3|W5PNL3_SHEEP         MMNASLGRVSSQIRKLETGLKEASAQLHVLTSSWEAVDELNA--QIPGLKQDLDKASALN 169
TR|G3QPX8|G3QPX8_GORGO         MVNVSLGYVRSQFLKLKTSVEKANAQIQILTRSWEEVSTLNA--QIPELKSDLEKASALN 168
TR|F7HNH2|F7HNH2_MACMU         TVNVSLDYVRSQFLKLKTSVEKANAQIQILTRSWEEVGTLNA--QIPELKSDLEKASALN 169
TR|G3TJF3|G3TJF3_LOXAF         RVNVSLNHVHSQLLTLETGLEKANAQIQMLTKSWEEVDALNA--QIPELKRDLDKASVLN 169
TR|A0A0P6J2W2|A0A0P6J2W2_HETGA MVNASLDNVHSQVLILKTSVEKANAQIQSLRKTWEEVDHLNA--KIPELKRDLDKASALN 169
TR|G1TDE9|G1TDE9_RABIT         TVNASLDRVYSQILTLKTSVERASAQIQILTESWEEVHSLNA--RIPELKSDLDKASTLH 169
TR|A0A0D9RRP0|A0A0D9RRP0_CHLSB TVNVSLGYVRSQFLKLKTSVEKANAQIQILTRSWEEVGTLNA--QIPELKSDLEKASALN 169
TR|F7A5P2|F7A5P2_MACMU         TVNVSLDYVRSQFLKLKTSVEKANAQIQILTRSWEEVGTLNA--QIPELKSDLEKASALN 168
TR|F6Y539|F6Y539_CANLF         L---------------MSPLSRCSHRVPCLQKQREGTSVLSLQGFGMYLVSDLEPLSRVS 150
TR|A0A096N0R1|A0A096N0R1_PAPAN MVNVSLGYVRSQFLKLKTSVEKANAQIQILTRSWEEVGTLNA--QIPELKSDLEKASALN 101
                                                :     .:  *         *.       *  ::   . :

SP|Q9UJ71|CLC4K_HUMAN          TKIRALQGSLENMSKLLKRQNDILQVVSQGWKYFKGNFYYFSLIPKTWYSAEQFCVSRNS 228
SP|Q8VBX4|CLC4K_MOUSE          TKVQGLQNSLENVNKLLKQQSDILEMVARGWKYFSGNFYYFSRTPKTWYSAEQFCISRKA 231
TR|D3ZBX0|D3ZBX0_RAT           TKVQGLQNSLENINKLLKEQSDILEMMSRGWKYFMGNFYYFSRTPKTWYSAEQYCISRKA 232
TR|B3FVQ1|B3FVQ1_PIG           AKVWDLQSGLGNISKMLKQQNDILQMMSQGWKYFRGNVYYFSVMKKTWYSAQQFCVSRES 229
TR|E1BB51|E1BB51_BOVIN         AKVRELQSGLESISKLLQQQKDILQVVSQGWKYFGGHFYYFSKISKTWYSAQQICISRDS 229
TR|F7EFC9|F7EFC9_CALJA         TKIRELQSSLENMSKLLKRQNDILQVVSQGWNFFRGNFYYFSRIPKTWYSAQQFCISRNS 229
TR|H0WH23|H0WH23_OTOGA         AKIRNLQSSLENTSKLLKQQSDILQMVSRGWRYFNGNFYYFSPVLKTWYSAEQFCVSKNS 232
TR|G3WUZ1|G3WUZ1_SARHA         AKIEQLKKDLQNLGASLSQQRYILEMASQNWKFFEGNFYYFSTMKKSWYSAEQFCTTQDS 235
TR|H0V494|H0V494_CAVPO         GMVRTLQSRLDNINKSLRQQRDILQMVSQGWEFFKDNFYYFSRISKTWYSAEQFCVSRDS 229
TR|I3N0C4|I3N0C4_ICTTR         AKVRGLQSSLETISKSLKQQNDILQMVSQDWKYFKGNFYYFSHTPKTWYSAQQFCVSRNS 229
TR|F6TZE3|F6TZE3_HORSE         AKVRALQSSLENVSKLLKQQNDILQVVSQGWKYFKGNFYYFSRVPKTWYSAQQFCISRNS 228
TR|F6U7K2|F6U7K2_MONDO         TKVEQLRKDLQTLGTSLSRQRYILEMASQDWKYFQGNFYYFSTKRKSWYSAEQFCTTRDS 235
TR|H2R8S3|H2R8S3_PANTR         TKIRALQGSLENMSKLLKRQNDILQVVSQGWKYFKGNFYYFSLISKTWYSAEQFCVSRNS 228
TR|G1MBS3|G1MBS3_AILME         AKVRGLQSNLENISKLFKRQNDILQMVSQGWKYFKENFYYFSQIPKTWYSAQQFCMSRDS 229
TR|M3WIV5|M3WIV5_FELCA         AKVQGLQTSLENISKVLKQQNDILQMVSQGWKYFRGNFYYFSQVPKTWYSAQQFCVSRDS 229
TR|G1RGW2|G1RGW2_NOMLE         TKIRALQGSLENMSKLLKRQNDILQVVSQGWKYFKGNFYYFSLIPKTWYSAEQFCVSRNS 228
TR|M3Y7D9|M3Y7D9_MUSPF         AKIQGLQSHLEDISKLLRRQNDILQMVSQGWKYFKENFYYFSQVPKTWYSAQQFCMSRDS 229
TR|G1PJQ2|G1PJQ2_MYOLU         VKVRGLQSSLENISKLLKQQNDTLQMVSQGWKYFKENFYYFSHVQKTWYSAQQFCVSRNS 208
TR|W5PNL3|W5PNL3_SHEEP         AKVRELQSGLESISKLLQQQNDILQVVSQGWKYFRGHFYYFSQISKTWYSAQQMCISRDS 229
TR|G3QPX8|G3QPX8_GORGO         TKIRALQGSLENMSKLLKRQNDILQVVSQGWKYFKGNFYYFSLIPKTWYSAEQFCVSRNS 228
TR|F7HNH2|F7HNH2_MACMU         TKIRALQGSLDNMSKLLKRQNDILQVVSQGWKYFKGNFYYFSLITKTWYSAQQFCVSRNS 229
TR|G3TJF3|G3TJF3_LOXAF         AKVRGLQSSLENMGKSLRQQNDILQMVSQGWKYFKGNFYYFSRVQKTWYSAQQFCKSRNS 229
TR|A0A0P6J2W2|A0A0P6J2W2_HETGA AMVRALQSRLDSINKLLRRQSDILQMVSQGWKFFRENFYYFSHISKTWYSAEQFCMSKDS 229
TR|G1TDE9|G1TDE9_RABIT         AKVRGLQSSLENISEVLRQQNDILQMVYQGWKYFKGNFYYFSRIAKTWYSAQQFCMSRDS 229
TR|A0A0D9RRP0|A0A0D9RRP0_CHLSB TKIRALQGSLDNMSKLLKRQNDILQVVSQGWKYFKGNFYYFSLVTKTWYSAQQFCVSRNS 229
TR|F7A5P2|F7A5P2_MACMU         TKIRALQGSLDNMSKLLKRQNDILQVVSQGWKYFKGNFYYFSLITKTWYSAQQFCVSRNS 228
TR|F6Y539|F6Y539_CANLF         ------SWVVVGLSYLAPKPDDILQMVSQGWKYFRGNFYYFSQVSKTWYSAQQFCMSRDS 204
TR|A0A096N0R1|A0A096N0R1_PAPAN TKIRALQGSLDNMSKLLKRQNDILQAVSQGWKYFKGNFYYFSLVTKTWYSAQQFCVSRNS 161
                                        :        .    *:   : *.:*  ..****   *:****:* * ::.:

SP|Q9UJ71|CLC4K_HUMAN          HLTSVTSESEQEFLYKTAGGLIYWIGLTKAGMEGDWSWVDDTPFNKV--QSVRFWIPGEP 286
SP|Q8VBX4|CLC4K_MOUSE          HLTSVSSESEQKFLYKAADGIPHWIGLTKAGSEGDWYWVDQTSFNKE--QSRRFWIPGEP 289
TR|D3ZBX0|D3ZBX0_RAT           HLTSVSSESEHEFLYKVADGIPHWIGLTKAGSEGDWYWVDQTSFNKE--QSRRFWIPGEP 290
TR|B3FVQ1|B3FVQ1_PIG           HLTSVTSESEQEFLYKMTDGLIYWIGLTKAGSEWDWHWVDDTPFDKL--QSARFWIPGEP 287
TR|E1BB51|E1BB51_BOVIN         HLTSVTSEREQEFLYRTAGGLPYWIGLTKAGSEGDWHWVDGTPYNKV--QSEKFWIPGEP 287
TR|F7EFC9|F7EFC9_CALJA         HLTSVTSESEQEFLYKTAGGRIYWIGLTKAGMEGEWSWVDDTPFNKV--QSERFWIPGEP 287
TR|H0WH23|H0WH23_OTOGA         HLTSVTSDSEQAFLSKSAAEVIRWIGLTKAGTEGGWHWIDDTPFNKA--QTERYWIPGEP 290
TR|G3WUZ1|G3WUZ1_SARHA         HLTSVTSTKEQEFLYKVANGVPYWIGLTKIGSSGTWHWADGTIFIER--ENARFWIKGEP 293
TR|H0V494|H0V494_CAVPO         HLTSVTSQNEQEFLYKTAGGLPHWIGLTKAGSEGDWYWVDDTPFNKD--SVSRFWLPGEP 287
TR|I3N0C4|I3N0C4_ICTTR         HLTSVTSKSEQEFLYKTASDFPFWIGLTKTGSEGNWYWMDNTPFNDPSLRFGRFWIPGEP 289
TR|F6TZE3|F6TZE3_HORSE         HLTSVTSESEQEFLYKAAGGLLHWIGLTKAGSEGDWHWVDDTPFDKV--QSARFWISGEP 286
TR|F6U7K2|F6U7K2_MONDO         HLTSVTSAKEQEFLYKTAAGISYWIGLTKTGSSGTWHWIDGTLYIER--ENARFWAKGEP 293
TR|H2R8S3|H2R8S3_PANTR         HLTSVTSESEQEFLYKTAGGLIYWIGLTKAGMEGDWSWVDDTPFNKV--QSARFWIPGEP 286
TR|G1MBS3|G1MBS3_AILME         HLTSVASENEQEFLYKMAGGLFYWIGLTKAGSEGDWYWVDNTPFDKD--LSARFWIPGEP 287
TR|M3WIV5|M3WIV5_FELCA         HLTSVASESEQEFLYKMAGGLFYWIGLTKAGSEGDWYWVDDTPFNKV--QSARFWIPGEP 287
TR|G1RGW2|G1RGW2_NOMLE         HLTSVTSESEQEFLYKTAGGFIYWIGLTKAGMEGDWSWVDDTPFNKV--QSARFWIPGEP 286
TR|M3Y7D9|M3Y7D9_MUSPF         QLTSVASESEQEFLYKMAGGLFYWIGLTKAGTEGDWYWVDNTPFDKV--QSARFWIPGEP 287
TR|G1PJQ2|G1PJQ2_MYOLU         QLTSVTSDSEQEFLYKTAGGLFYWIGLTKAGTDGDWYWVDETPFNKV--QSDRFWIPGEP 266
TR|W5PNL3|W5PNL3_SHEEP         HLTSVTSEREQEFLYRTAGGLPYWIGLTKAGSEGDWHWVDGTPFNKV--QSEKFWIPGEP 287
TR|G3QPX8|G3QPX8_GORGO         HLTSVTSESEQEFLYKTAGGLIYWIGLTKAGMEGDWSWVDDTPFNKV--QSARFWIPGEP 286
TR|F7HNH2|F7HNH2_MACMU         HLTSVTSESEQEFLYKTAGGLTYWIGLTKAGTEGDWFWVDDTPFDKV--QSAKFWIPGEP 287
TR|G3TJF3|G3TJF3_LOXAF         QLTSVTSESEQEFLYKTAGGISYWIGLTKAGSEGDWSWVDDTPFNKV--QSAKFWIPGEP 287
TR|A0A0P6J2W2|A0A0P6J2W2_HETGA HLTSVTSRDEQEFLYKGAGGLPNWIGLTKAGSEGKWYWVDDTPFNKE--SV-RFWLPGEP 286
TR|G1TDE9|G1TDE9_RABIT         HLTSVTSESEQEFLYKTAGGLVYWIGLSKAGSEGHWYWADGTSFNEA--QSVRFWIPGEP 287
TR|A0A0D9RRP0|A0A0D9RRP0_CHLSB HLTSVTSESEQEFLYKTAGGLTYWIGLTKAGTEGDWFWVDDTPFNKV--QSAKFWIPGEP 287
TR|F7A5P2|F7A5P2_MACMU         HLTSVTSESEQEFLYKTAGGLTYWIGLTKAGTEGDWFWVDDTPFDKV--QSAKFWIPGEP 286
```

```
TR|F6Y539|F6Y539_CANLF         QLTSVTSESEQEFLYRTAGGLSYWIGLTKAGSEGDWSWVDDTPFDKVQ-SAHRFWIPGEP 263
TR|A0A096N0R1|A0A096N0R1_PAPAN HLTSVTSESEQEFLYKTAGGLTYWIGLTKAGTEGDWFWVDDTPFNKV--QSAKFWIPGEP 219
                               :****:*  *: ** : :     ****:* * .  * * * * : .      ::*  ***

SP|Q9UJ71|CLC4K_HUMAN          NNAGNNEHCGNIKAPSLQAWNDAPCDKTFLFICKRPYVPSEP- 328
SP|Q8VBX4|CLC4K_MOUSE          NNAGNNEHCANIRVSALKCWNDGPCDNTFLFICKRPYVQTTE- 331
TR|D3ZBX0|D3ZBX0_RAT           NNVRNNEHCANIRVSALKCWNDSPCDNVYSFICKMPYIRMIT- 332
TR|B3FVQ1|B3FVQ1_PIG           NNSENNEHCANIKRSSLRSWNDAPCDIELLFICKRPYVPSEP- 329
TR|E1BB51|E1BB51_BOVIN         NNVGNNEHCVTLKTSLLRSWNDASCDNTFLFICKRSYKPSEP- 329
TR|F7EFC9|F7EFC9_CALJA         NNYGNNEHCANLKASSLQSWNDAPCDQTFLFICKRPYIPSEP- 329
TR|H0WH23|H0WH23_OTOGA         NNSGFNEHCVSIRVLSLQSWNDSPCDIKYSFICKRPYIPSEP- 332
TR|G3WUZ1|G3WUZ1_SARHA         NDSGQNEHCVTLEQTSLMSWNDVTCDRGLQFICKKGPKSPVID 336
TR|H0V494|H0V494_CAVPO         NNVGNSEHCANIKVSSLRSWNDDSCDIKLFFICKRPYIQSEP- 329
TR|I3N0C4|I3N0C4_ICTTR         NNAGNNEHCADMRFPLLMSWNDASCDKTLPFVCKQPYVPSEP- 331
TR|F6TZE3|F6TZE3_HORSE         NNYGSNEHCANIKLFSLQSWNDASCDITLLFICKRPYTPSEA- 328
TR|F6U7K2|F6U7K2_MONDO         NNGYGNEHCVNLEKSSLMSWNDVNCDGELQFICKKNPKPLEMD 336
TR|H2R8S3|H2R8S3_PANTR         NNAGNNEHCGNIKAPSLQAWNDAPCDITFLFICKRPYVPSEP- 328
TR|G1MBS3|G1MBS3_AILME         NNFGSNEHCANIKASSLQSWNDASCDNKLFFICKRPYIPSEP- 329
TR|M3WIV5|M3WIV5_FELCA         NNFGNNEHCANIKMSSLQSWNDASCDNKLLFICKRPYIPSEP- 329
TR|G1RGW2|G1RGW2_NOMLE         NNAGNNEHCGNIKALSLQAWNDAPCDKTFLFICKRPYVPSEP- 328
TR|M3Y7D9|M3Y7D9_MUSPF         NNYGNNEHCANIKMSSLQSWNDASCDNKLLFICKRPYITSEP- 329
TR|G1PJQ2|G1PJQ2_MYOLU         NNLGNNEHCANLKMSSLQSWNDAPCDHPFLFICKRRYIPSEP- 308
TR|W5PNL3|W5PNL3_SHEEP         NNYGNNEHCVNLKTSSLRSWNDASCDNTFPFICKRSYKPSEP- 329
TR|G3QPX8|G3QPX8_GORGO         NNAGNNEHCGNIKAPSLQAWNDAPCDKTFLFICKRPYVPSEP- 328
TR|F7HNH2|F7HNH2_MACMU         NNAGNNEHCGNIRVSSLQAWNDAQCDKTFLFICKRPYIPSEP- 329
TR|G3TJF3|G3TJF3_LOXAF         NNSGYSEHCAHIRMASLQSWNDASCDNTLPFICKQLYIPSEP- 329
TR|A0A0P6J2W2|A0A0P6J2W2_HETGA NNVGNSEHCASIKVSSLQSWNDDSCDNKLLFICKQPYIQSEP- 328
TR|G1TDE9|G1TDE9_RABIT         NNMGNGEHCVSIKVSLLQSWNDASCDDKLLFVCKRPYTAAGP- 329
TR|A0A0D9RRP0|A0A0D9RRP0_CHLSB NNAGNNEHCGNIRASSLQAWNDAQCDKTFLFICKRPYVPSEP- 329
TR|F7A5P2|F7A5P2_MACMU         NNAGNNEHCGNIRVSSLQAWNDAQCDKTFLFICKRPYIPSEP- 328
TR|F6Y539|F6Y539_CANLF         N----NEHCADIKVSSLQSWNDVSCDSTLFFICKRPYSPSEP- 301
TR|A0A096N0R1|A0A096N0R1_PAPAN NNVGNSEHCGNIRASSLQAWNDAQCDTTFLFICKRPYVPSEP- 261
                               *     ***  :.   * .***  **    *:**
```

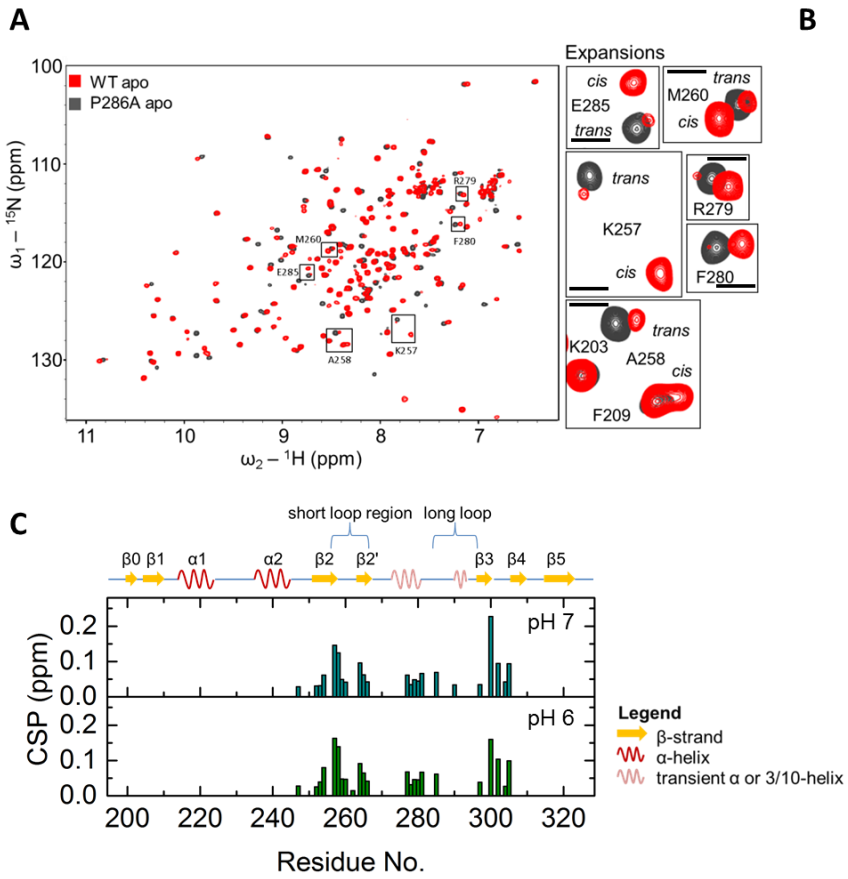
Figure S4: Multiple sequence alignment of 28 mammalian Langerin sequences. Alignment was performed using CLUSTAL O(1.2.1).[1]




| | K257 | | E285 | |
|---|---|---|---|---|
| $[CaCl_2]$ (M) | Line width $\omega_2$ (Hz) | Line width $\omega_1$ (Hz) | Line width $\omega_2$ (Hz) | Line width $\omega_1$ (Hz) |
| 0.0E-00 | 24.9 | 25.7 | 25.8 | 21.0 |
| 1.0E-04 | 21.3 | 36.2 | 27.6 | 21.6 |
| 2.5E-04 | 34.4 | 54.8 | 30.5 | 21.6 |
| 5.0E-04 | 22.4 | 49.6 | 33.1 | 20.5 |
| 1.5E-03 | 23.0 | 33.6 | 24.6 | 20.2 |
| 1.0E-02 | 29.1 | 23.7 | 25.4 | 20.9 |

Figure S5: (A) Binding of $Ca^{2+}$ results in intermediate exchange broadening in several shifting peaks. Linewidths of K257 $^{15}N$ peak dimension and E285 $^{1}H$ peak dimension during $Ca^{2+}$ titration. Both peaks have a maximum between 250 µM and 500 µM ligand concentration roughly corresponding to 1/3 $K_d$. The table summarizes the linewidths of both investigated peaks at different ligand concentrations. (B) $^{1}H$-$^{15}N$ HSQC NMR titrations of Ca2+ interacting with Langerin CRD at pH 6 (lower) and pH 7 (lower panel). Data was fitted to eq. (S3) to obtain Kd values of 620±35 and 160±25 µM, respectively.




**B**

| Residue | Fraction in cis |
|---|---|
| G247 | 76% |
| W252 | 71% |
| I253 | 75% |
| G254 | 72% |
| K257 | 70% |
| A258 | 63% |
| G259 | 71% |
| M260 | 78% |
| G262 | N.A. |
| W264 | 69% |
| S265 | 68% |
| W266 | 68% |
| S277 | 79% |
| A278 | 76% |
| R279 | 80% |
| F280 | 77% |
| W281 | 82% |
| E285 | 69% |
| N297 | 63% |
| A300 | 75% |
| S302 | 79% |
| Q304 | 75% |
| A305 | 78% |
| **Average** | **73%** |
| **STDEV*** | **10%** |

Figure S6: (A) Apo Langerin CRD experiences cis/trans isomerization of P286 prolyl bond. Overlay of $^{1}$H-$^{15}$N HSQC NMR spectra of Langerin CRD WT apo and P286A mutant at pH 6. Expansions of the backbone resonances of representative residues undergoing cis/trans isomerization. In total, 23 residues were identified. Black bar represents 0.1 ppm and 0.75 ppm on the $^{1}$H and $^{15}$N chemical shift scale, respectively. (B) Residues in Langerin CRD apo form at pH 6 that undergo cis trans isomerization as indicated by a second peak. Fraction population in cis conformation was calculated from the peak integrals of cis and trans population. STDEV was determined from three independent measurements. (C) Chemical shift differences between trans and cis states at pH 6 (lower panel) and pH 7 (upper panel). The differences are largely unaffected by pH change.

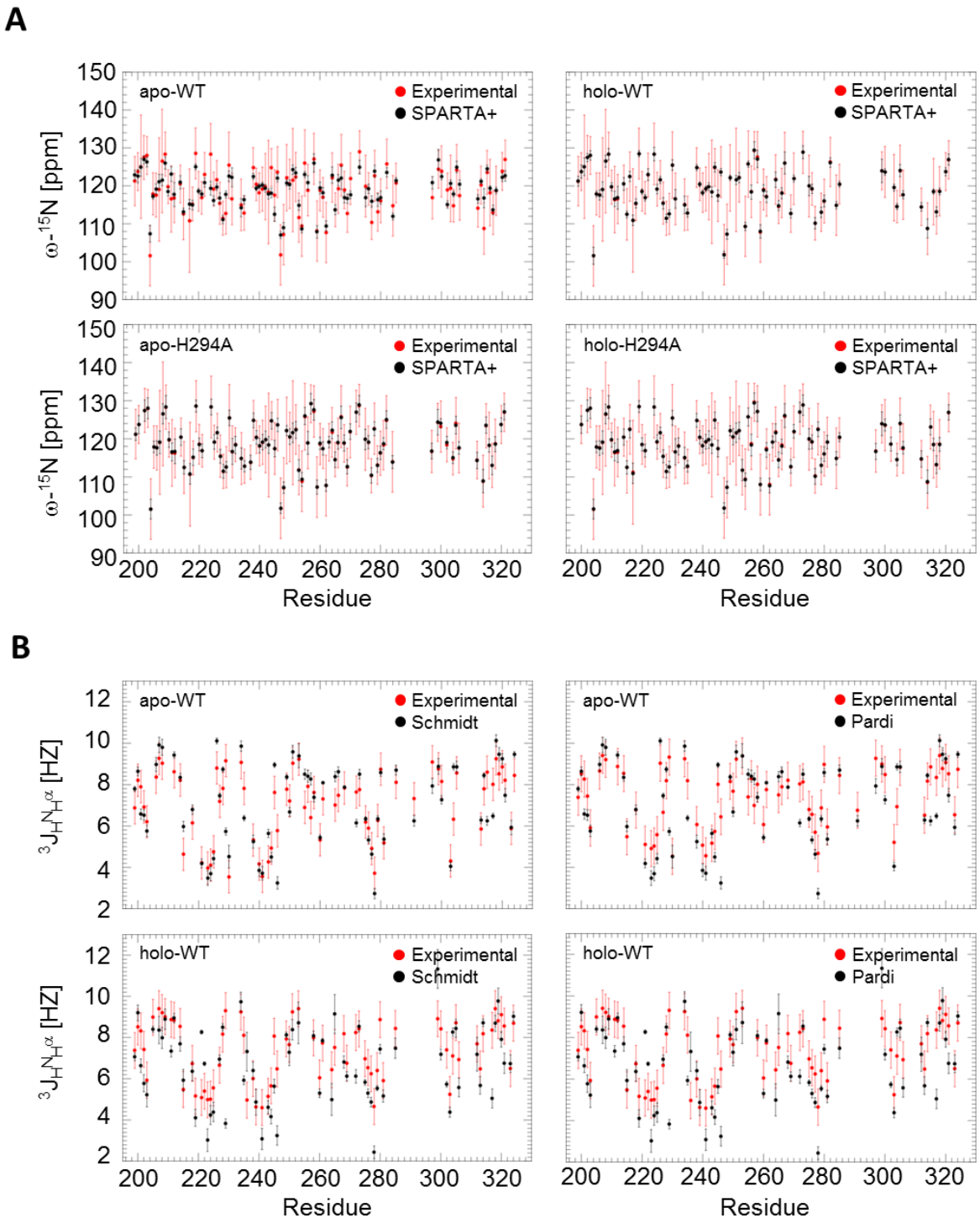



Figure S7: (A) The chemical shifts of WT (upper panel), and H294A mutant forms (lower panel) are calculated from the MD simulations with program SPARTA+ (colored in black), and mapped to the experimentally determined chemical shifts (colored in red). The root mean square deviations (RMSD) values were calculated between the experimental, and computed chemical shifts. In holo WT Langerin CRD, and both mutant forms detected RMSD

values were below 0.2 suggesting almost the perfect match between experimental, and simulation data. However estimate of the chemical shifts in apo WT was significantly worse with RMSD value of 2.9. (B) $^{3}J_{HNHC\alpha}$ coupling constants were estimated from the MD simulations of the apo (upper panel) and holo (lower panel) Langerin CRD by using Karplus equation[2] and its parameters previously described in literature.[3] The RMSD values between experimentally derived (in black), and computed $^{3}J_{HNHC\alpha}$ coupling constants were below 1.5 for both systems, and both combination of the parameters of Karplus equation.

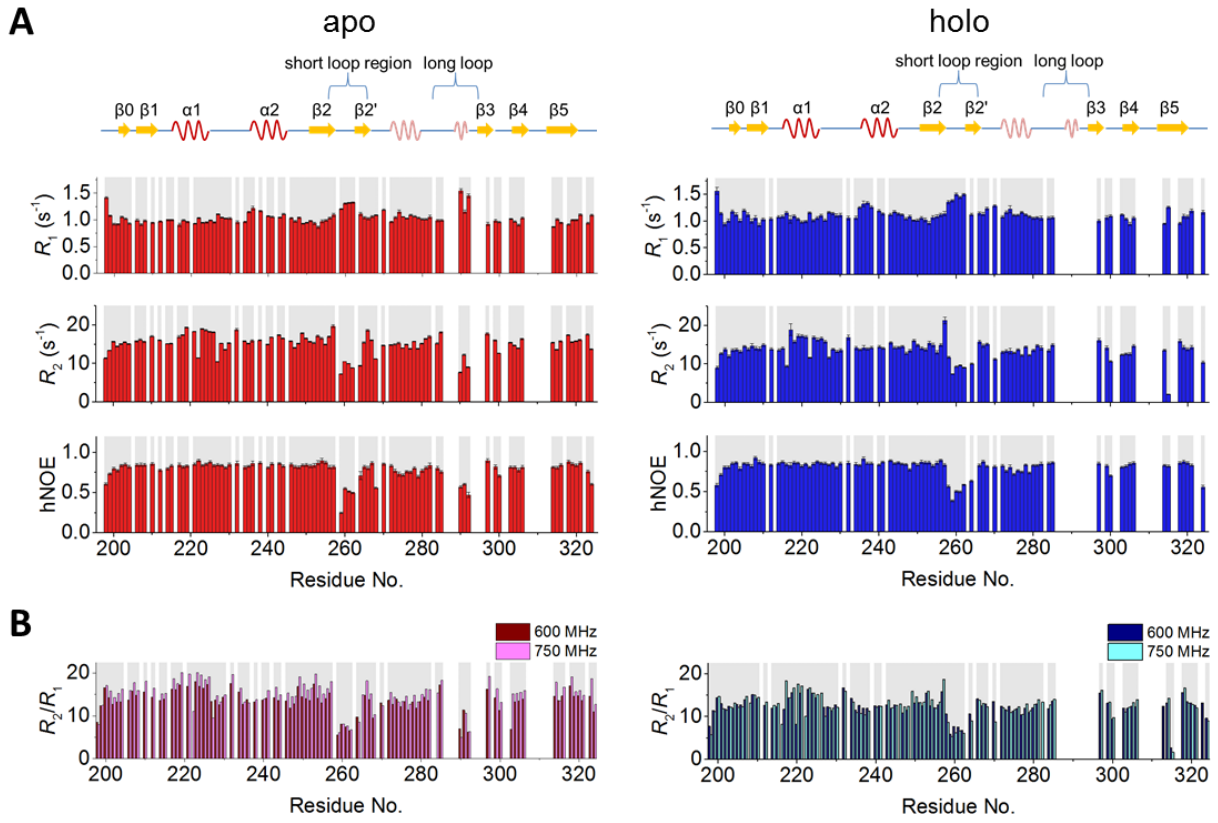
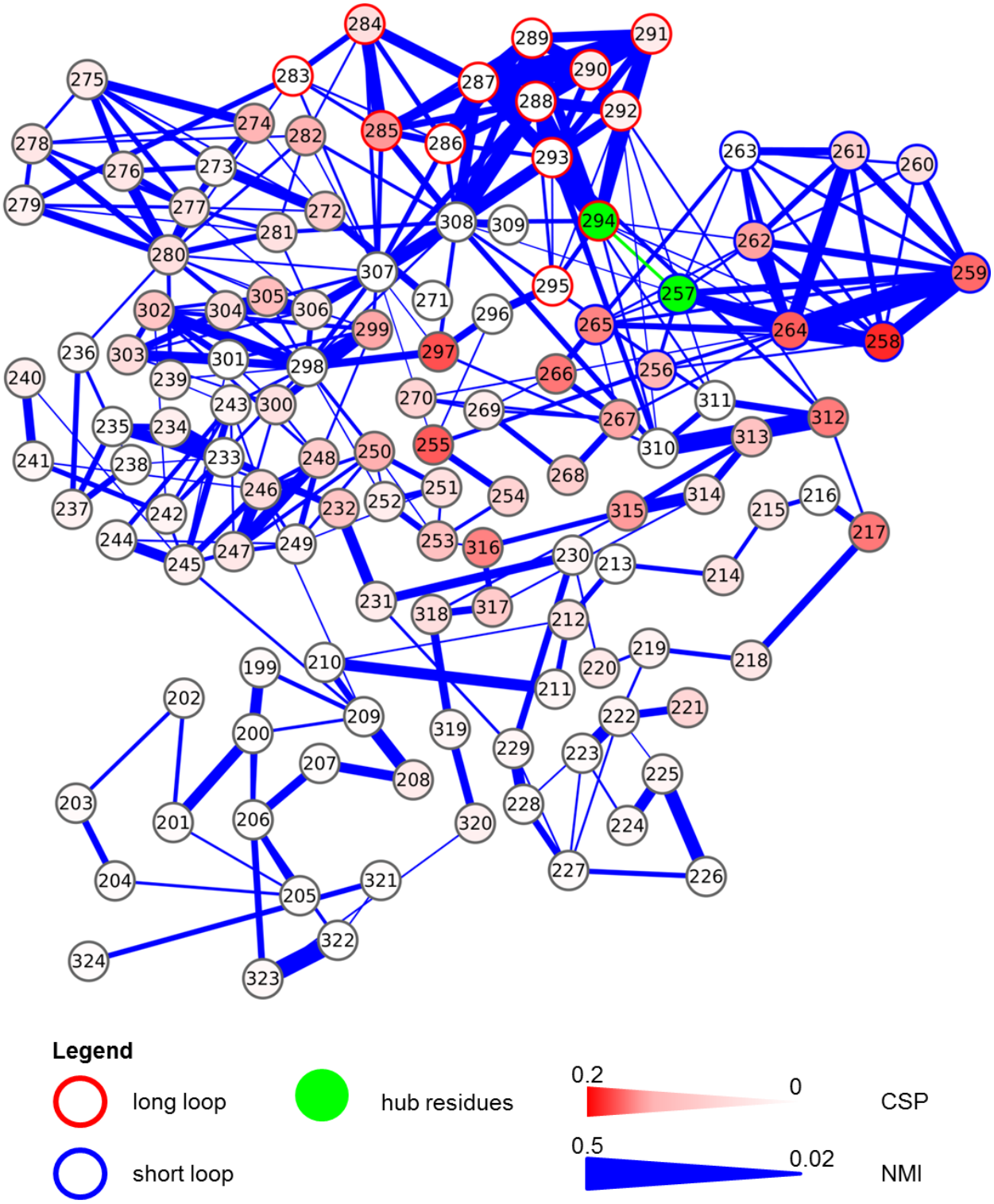
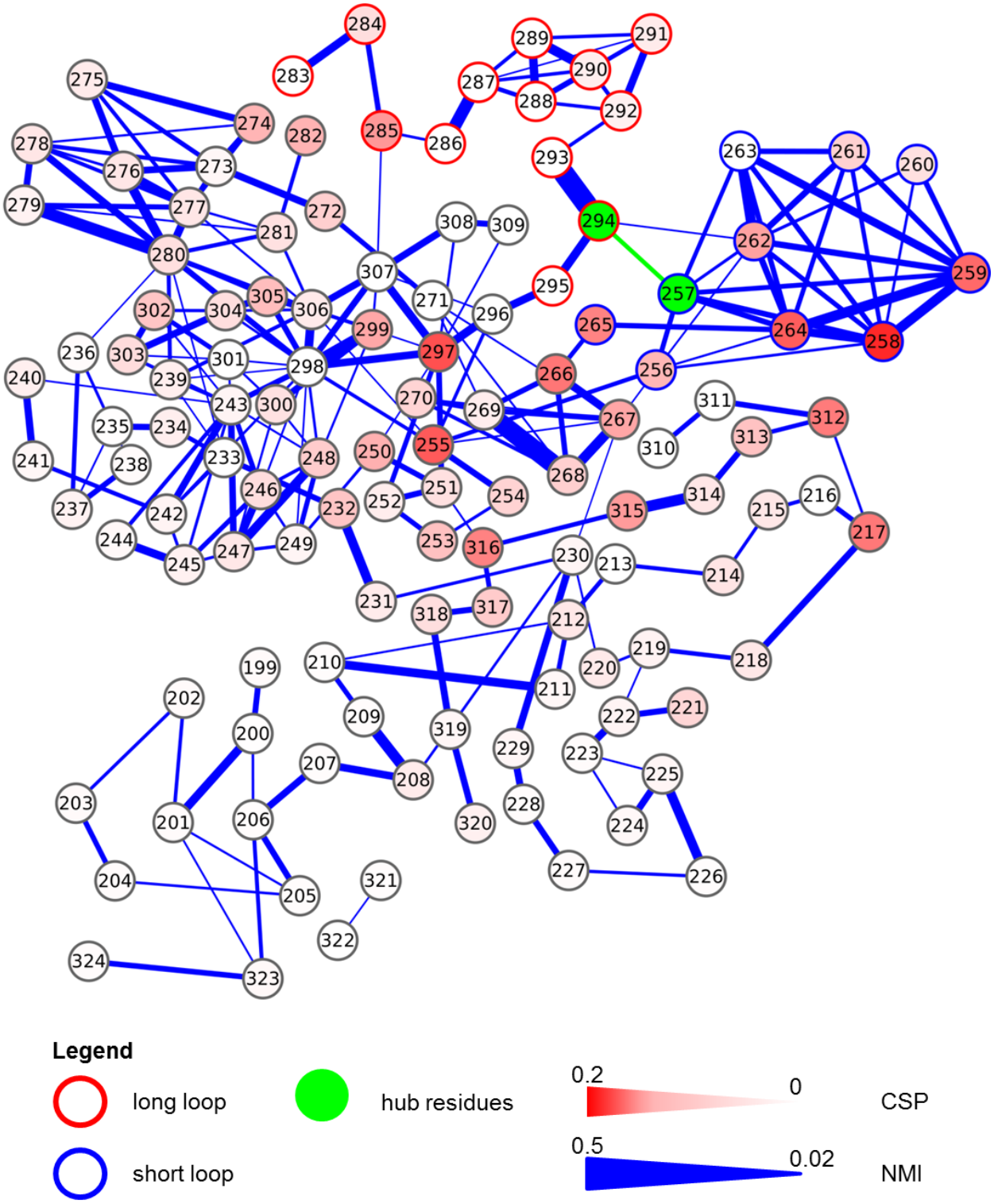



Figure S8: (A) $^{15}$N backbone relaxation rate constants $R_1$ and $R_2$ and hetero NOE of apo (left panel) and holo (right panel) Langerin CRD obtained at 750 MHz. The protein exhibits rather uniform relaxation rate constants with major exceptions in the small loop and long loop region. (B) $R_2/R_1$ ratio per residue to estimate local correlation times $\tau_m$ at 600 MHz and 750 MHz field strengths apo and holo (left and right panel, respectively). The data confirmed fast internal motions on the ps to ns timescale in the short loop. In the apo form, an increased ratio at 750 MHz is observed for the α1-helix and β2-strand indicating motions on the µs to ms timescale, grey bars indicate resolved residues used for analysis. All spectra were recorded at 299.2 K, pH 6. Holo forms were generated by addition of 10 mM $CaCl_2$.

## Trans apo



## Cis apo



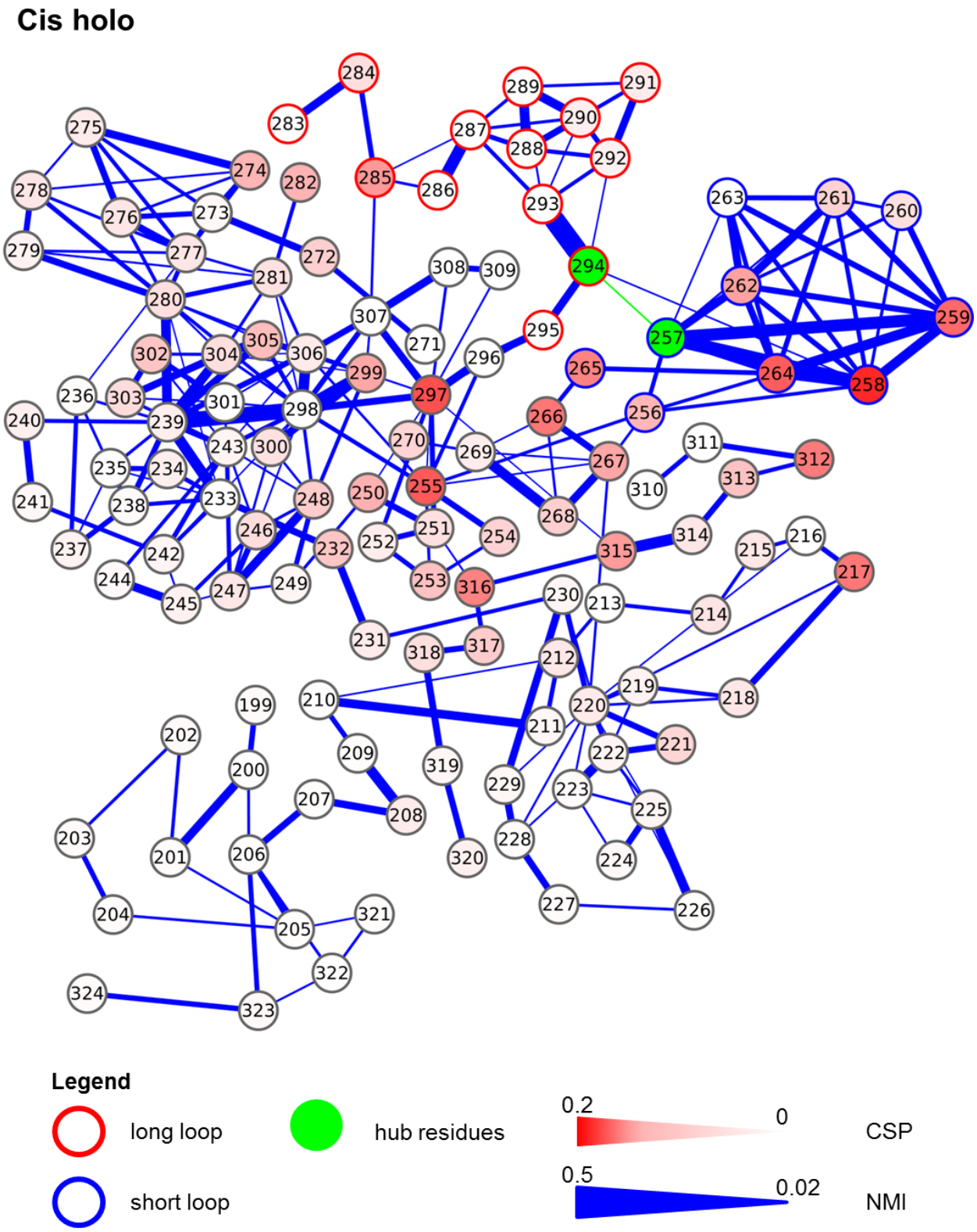



Figure S9: Cytoscape network representation of the mutual dependence graphs computed for the trans apo WT (upper panel), cis apo WT (middle panel), and cis holo WT (lower panel). Residues of the long loop, and the short

loop are marked with red, and blue borders respectively. Node coloring is according the CSP detected upon $Ca^{2+}$ binding to WT Langerin CRD. The edge thickness corresponds to the aggregate NMI values (threshold of 0.02). K257-H294 loops coupling axis is highlighted in green.

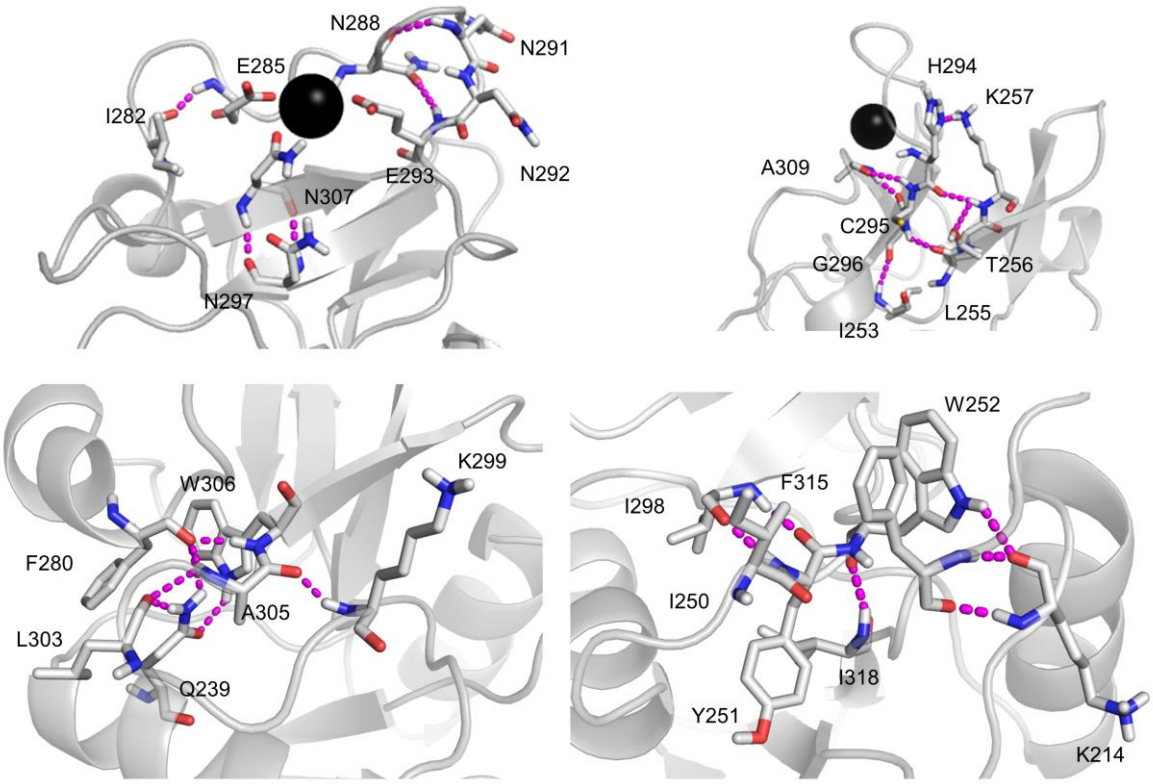



Figure S10: Conserved Hydrogen Bond Network. The allosteric information is propagated from the $Ca^{2+}$ binding site (upper left panel) through the conserved hydrogen bond network downstream to the short loop (upper right panel), then to the core of the protein (lower right panel), α1-helix (lower right panel), and α2-helix (lower left panel).

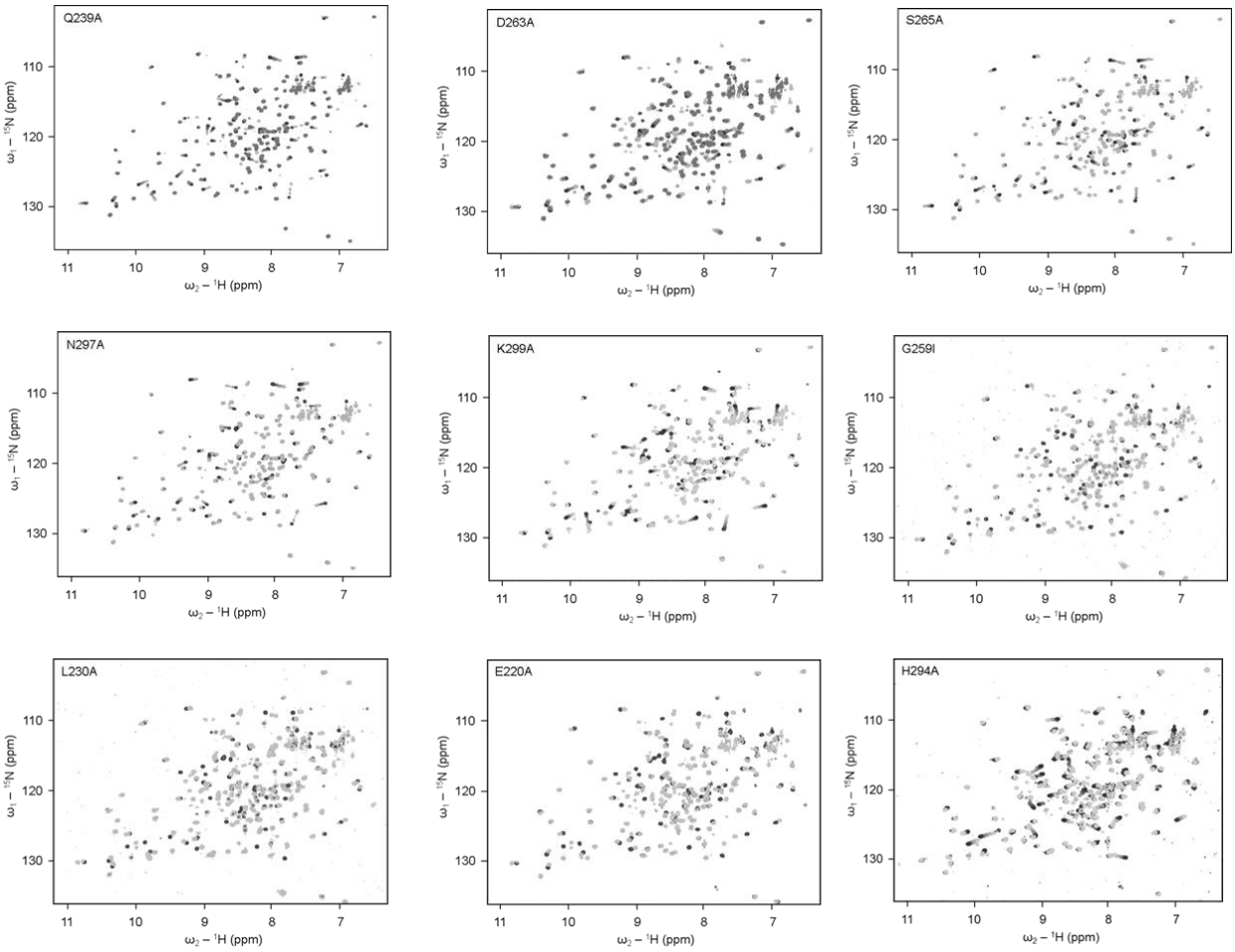



Figure S11: $^{1}$H-$^{15}$N HSQC NMR spectra overlays of Langerin CRD mutants in apo (light grey) and holo forms (dark grey). The name of the mutant is indicated in the upper left corner. In case of titration experiments, titration point experiments are shown in grey shades. All spectra were recorded at pH 6 at 296 K.

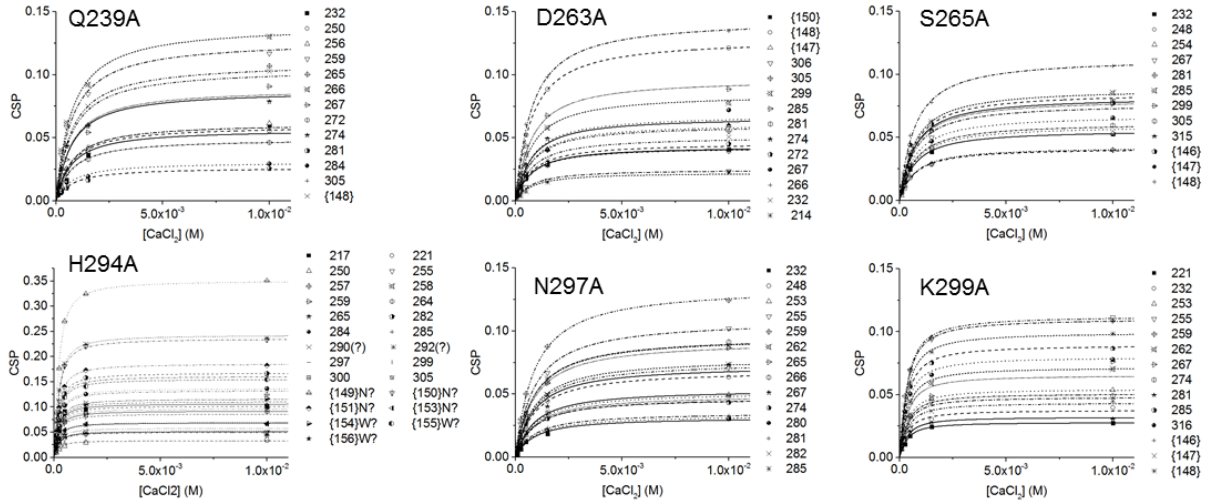



Figure 12: CSP from $^{1}H$-$^{15}N$ HSQC NMR titrations of $Ca^{2+}$ interacting with Langerin CRD mutants at pH 6 plotted against $Ca^{2+}$ concentration. Data was fitted to eq. (S3) to obtain $K_d$ values listed in Table 1.

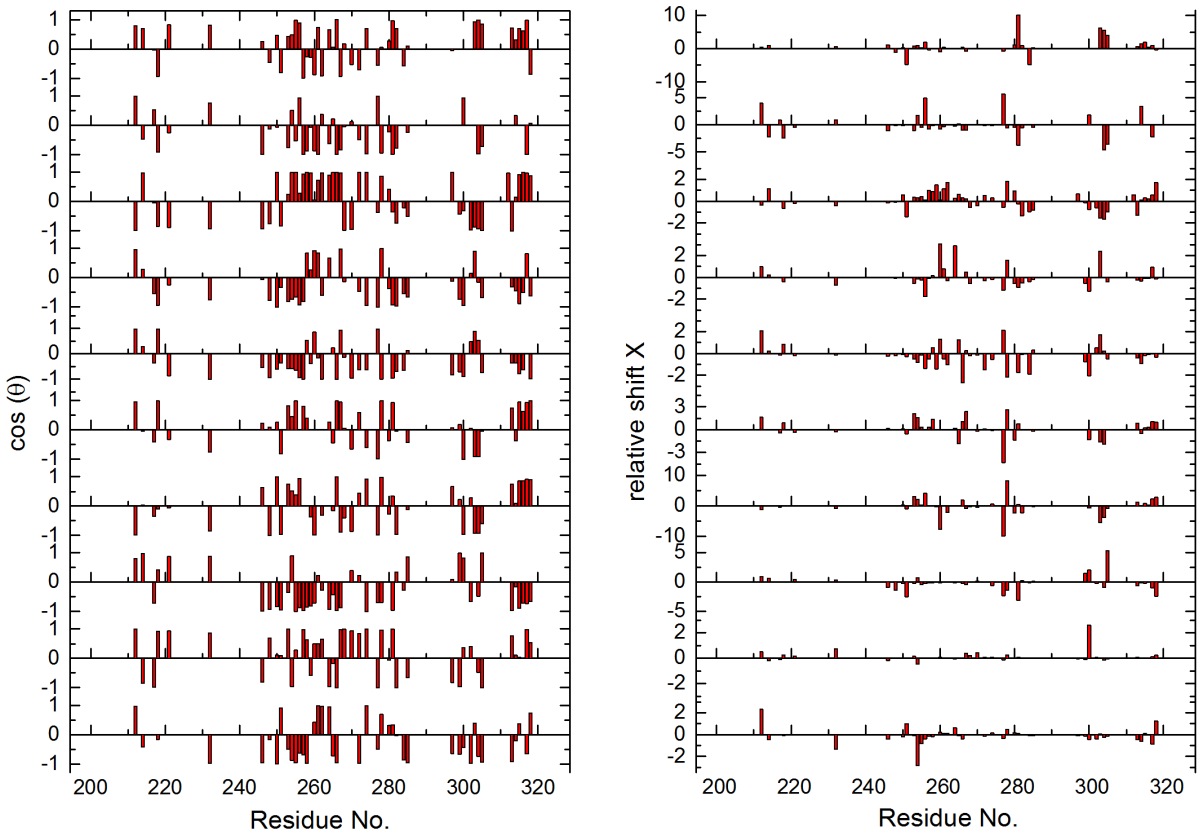



Figure S13: CHESPA analysis of ten Langerin CRD mutants derived from $^{1}H$-$^{15}N$ HSQC NMR measurements at pH 6. With the exception of H294A, no clear pattern of activating and deactivating shifts was discerned.

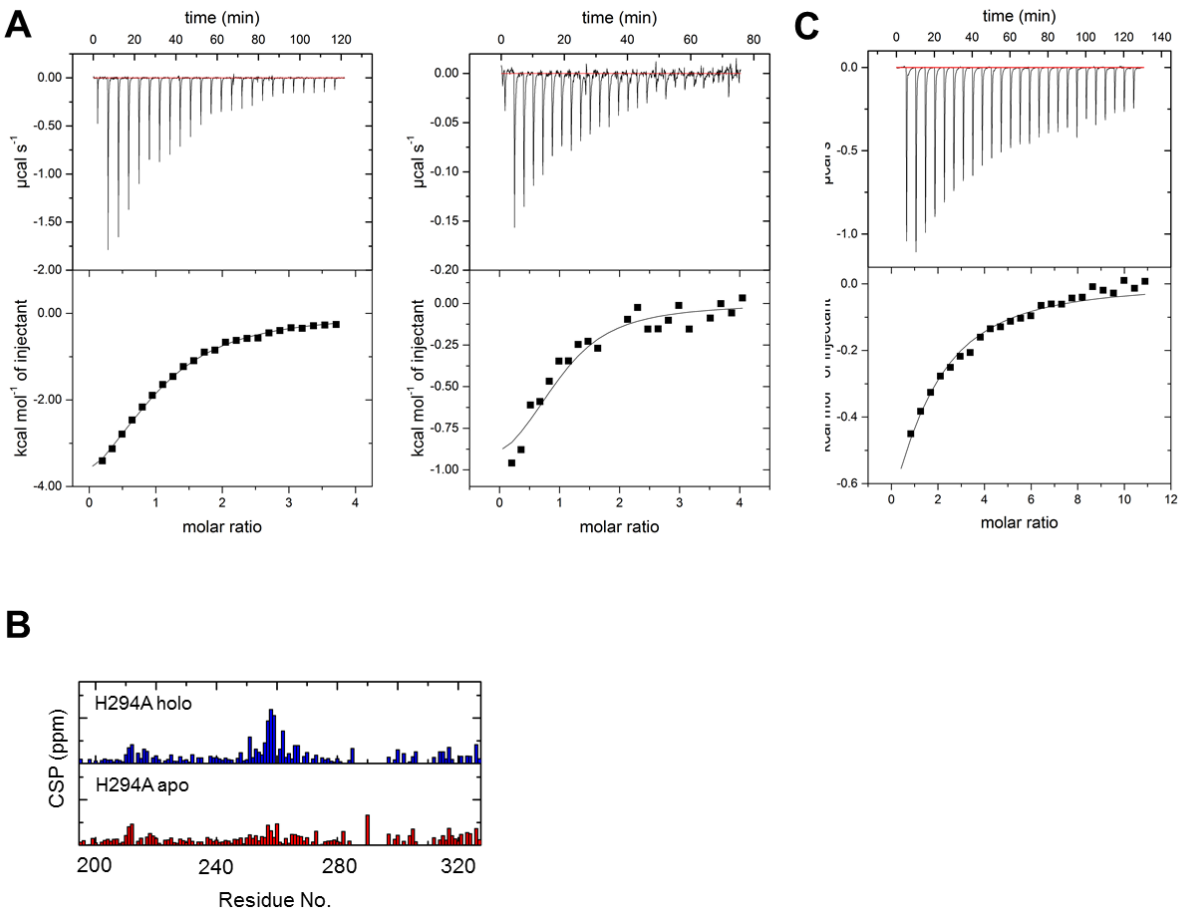
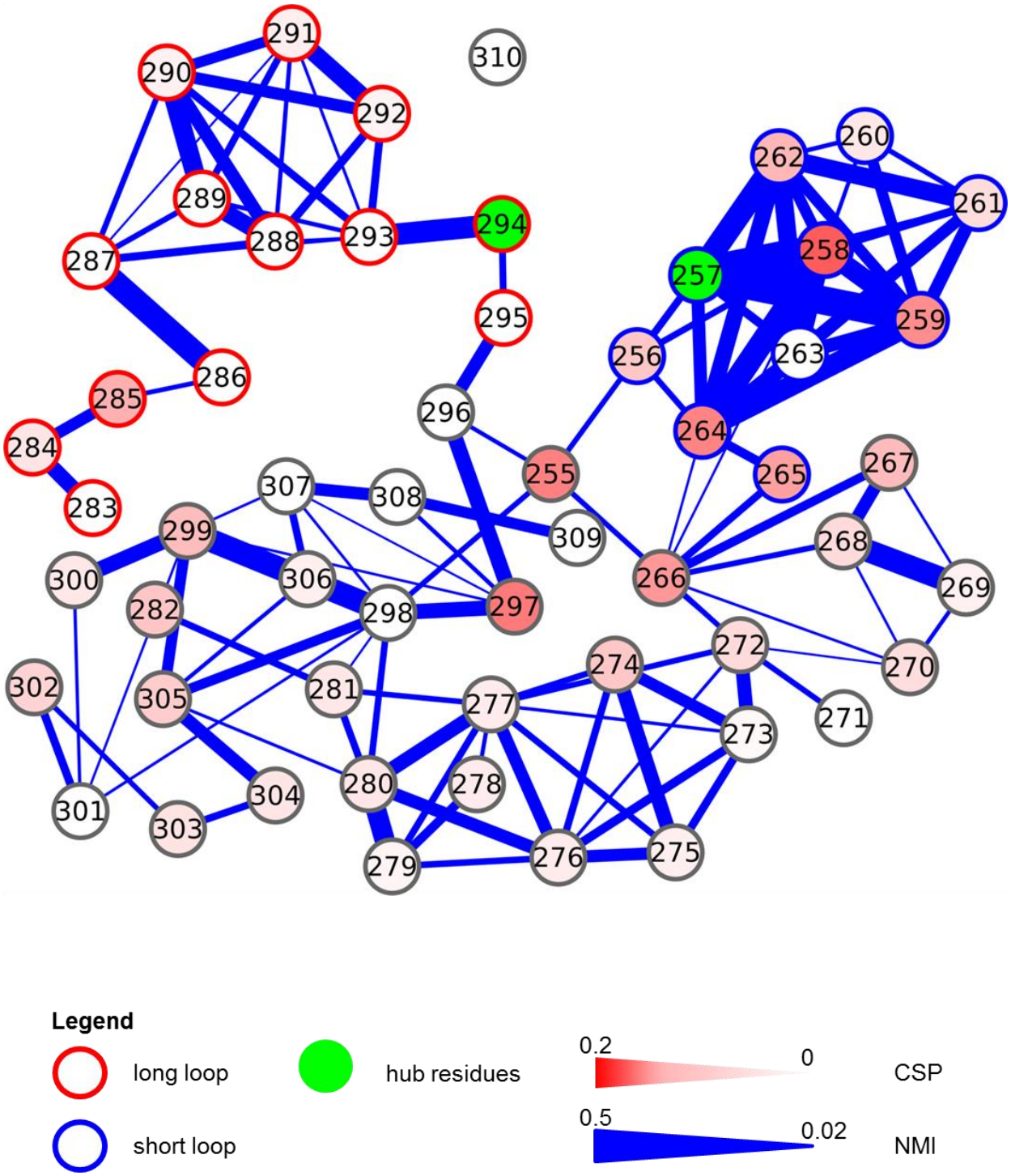
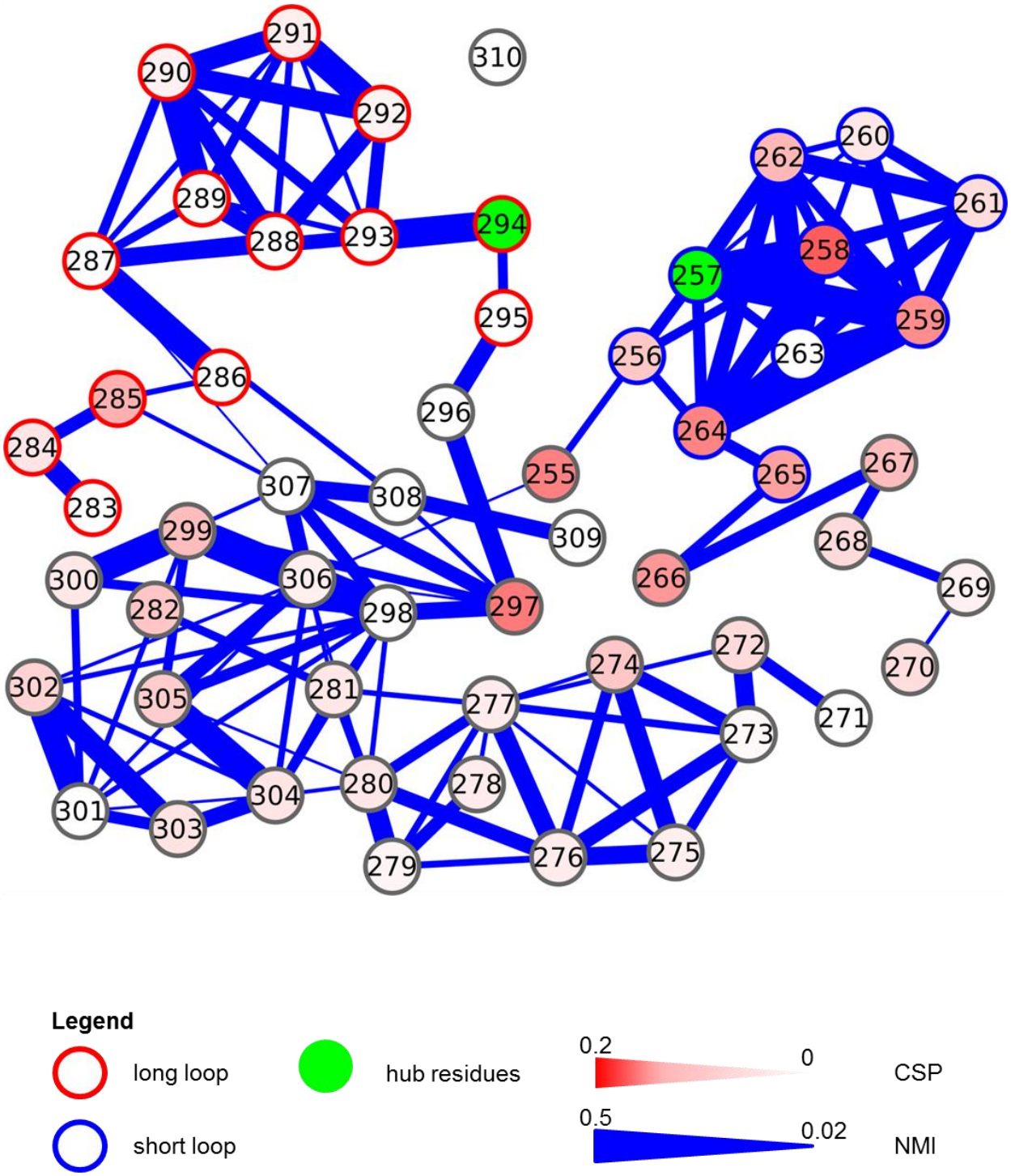



Figure S14: (A) H294A binds Ca2+ with pH-dependent affinity. Isothermal titration calorimetry of Langerin CRD H294A mutant awarded a $K_d$ of 126±5 µM at pH 6 (left panel) and 36±15 µM at pH 7 (right panel) using a one-set-of-sites binding model with a 1:1 stoichiometry. Measurements were performed at 298 K. (B) Chemical shift differences of apo and holo H294A mutants at pH 6 and 7 show that the small loop in the holo form is still affected by pH even in absence of H294 sidechain indicating H294 independent pH mechanism. (C) ITC of Langerin CRD K257A mutant awarded a $K_d$ of 200±20 µM at pH 6 using a one-set-of-sites binding model with a 1:1 stoichiometry. Measurements were performed at 298 K.

## Apo H294A


291
290
292
310
260
262
261
289
288
293
294
287
258
257
259
295
256
263
286
285
296
264
284
255
265
267
307
308
283
299
309
268
300
306
266
269
282
298
297
272
274
270
302
305
281
277
271
273
304
280
278
301
303
276
275
279
Legend
long loop
hub residues
0.2
0
CSP
short loop
0.5
0.02
NMI

## Holo H294A


291
290
292
310
260
262
261
289
294
288
293
287
258
257
295
259
256
263
286
285
296
264
284
255
265
267
307
308
283
309
299
268
300
306
266
269
282
298
297
272
274
270
302
305
281
277
271
273
304
280
278
301
303
276
275
279
Legend
long loop
hub residues
0.2
0
CSP
short loop
0.5
0.02
NMI

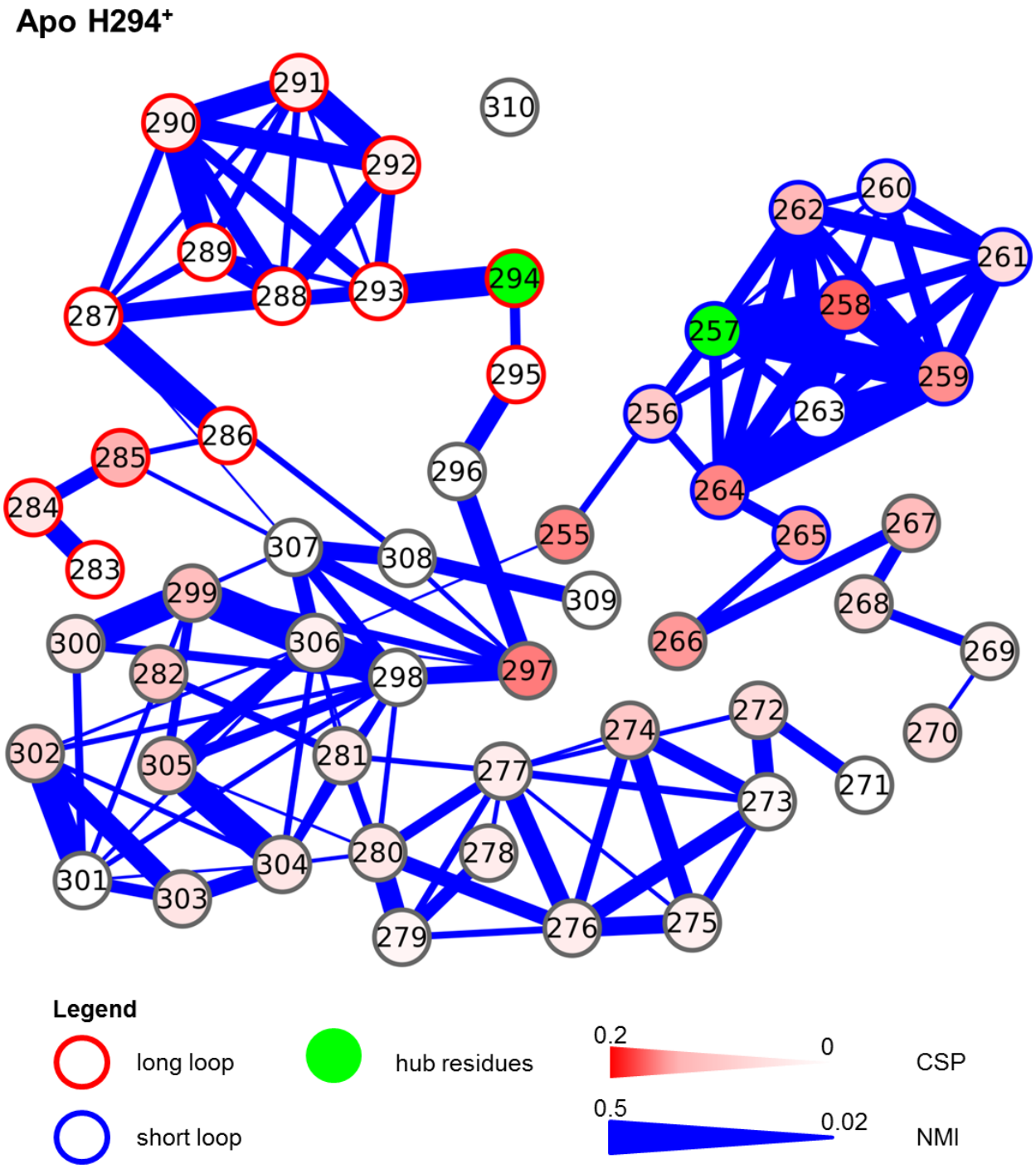



Figure S15: The aggregate NMI was calculated for the residue 255-310 in apo cis H294A mutant, in holo cis H294A mutant, and in cis apo WT with protonated H294 side chain. Residues of the long loop, and the short loop are marked with red, and blue borders respectively. Node coloring is according the CSP detected upon $Ca^{2+}$ binding to WT Langerin CRD. The edge thickness corresponds to the aggregate NMI values (threshold of 0.02). Figure was prepared with Cytoscape.

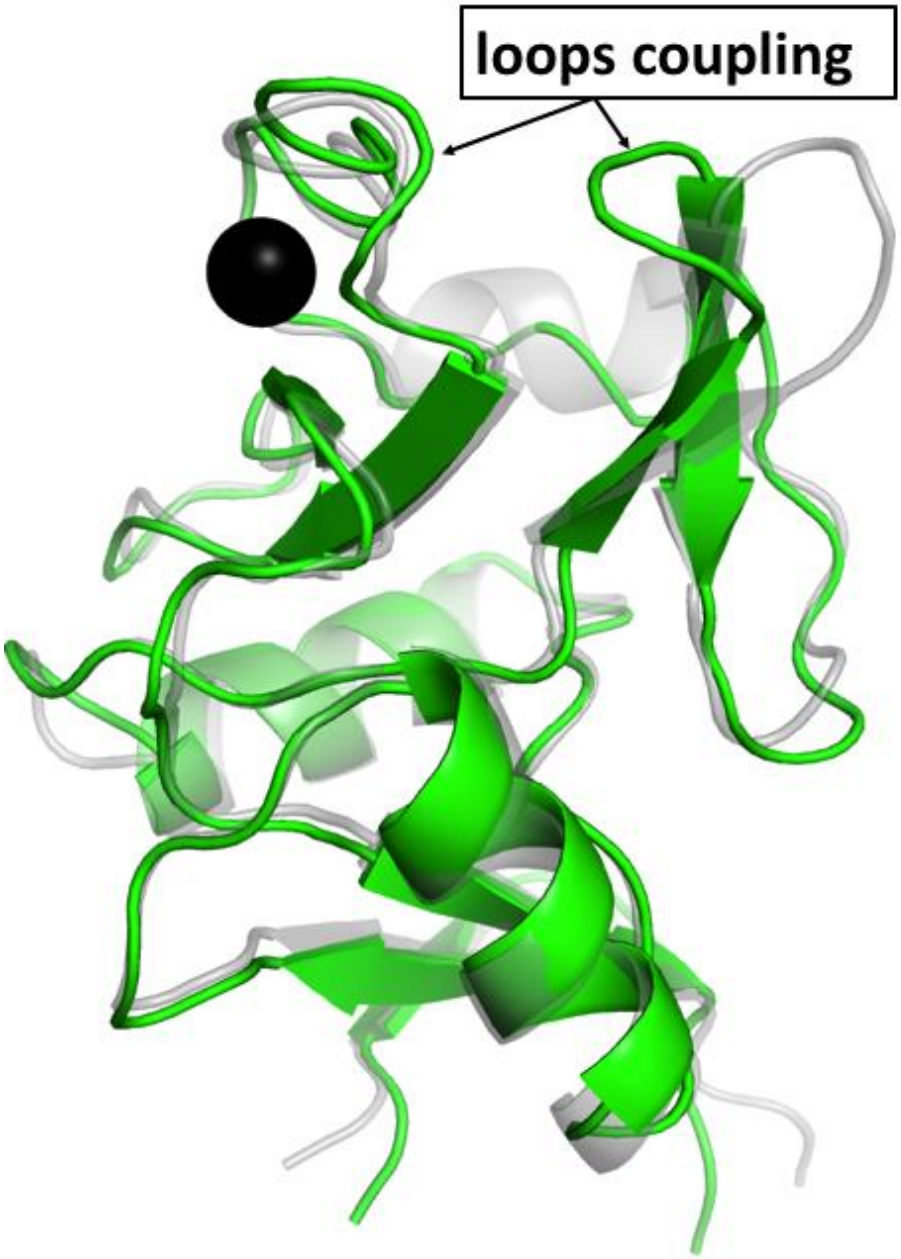



Figure S16: Local structural rearrangements involved in the loops coupling. Cartoon representation of the apo Langerin CRD (in green) superimposed to Langerin CRD crystal structure (3P5H[4]) depicted in gray. The coupling between the long loop, and the short loop caused Langerin CRD to adopt the closed conformation.

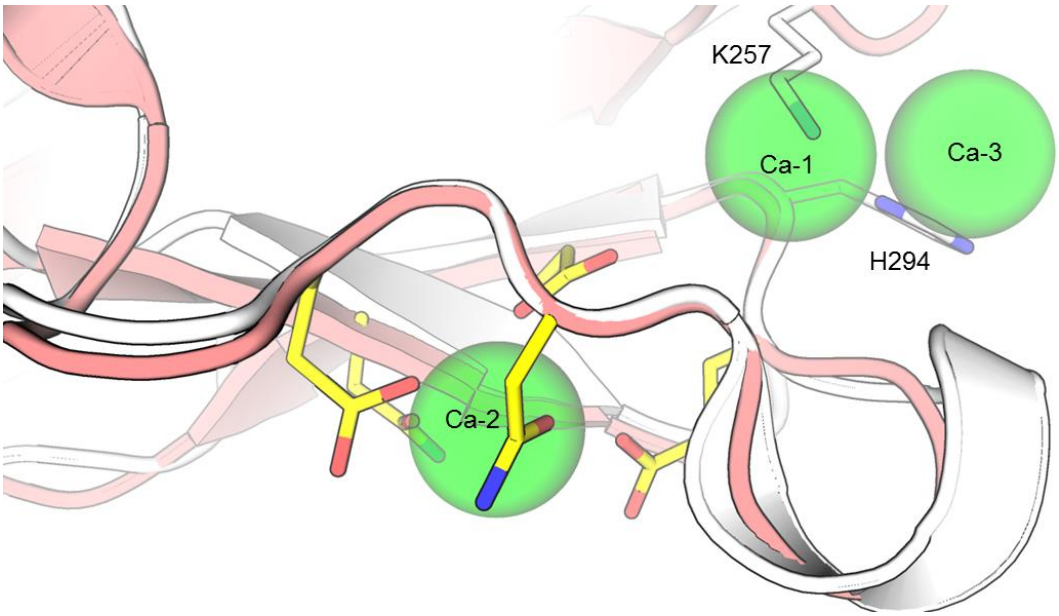



Figure S17: Cartoon representation of Langerin CRD (white) with residues H294 and K257 as white sticks and residues coordinating the $Ca^{2+}$ in the canonic binding site in yellow. Cartoon representation (salmon) of aligned CRD structure of DC-SIGN (salmon, pdb entry: 1SL4[5]) and DC-SIGNR (light magenta. pdb entry 1K9J[4]). The Ca-1 and Ca-3 positions in DC-SIGN are occupied by H294 and K257 side chains in Langerin.

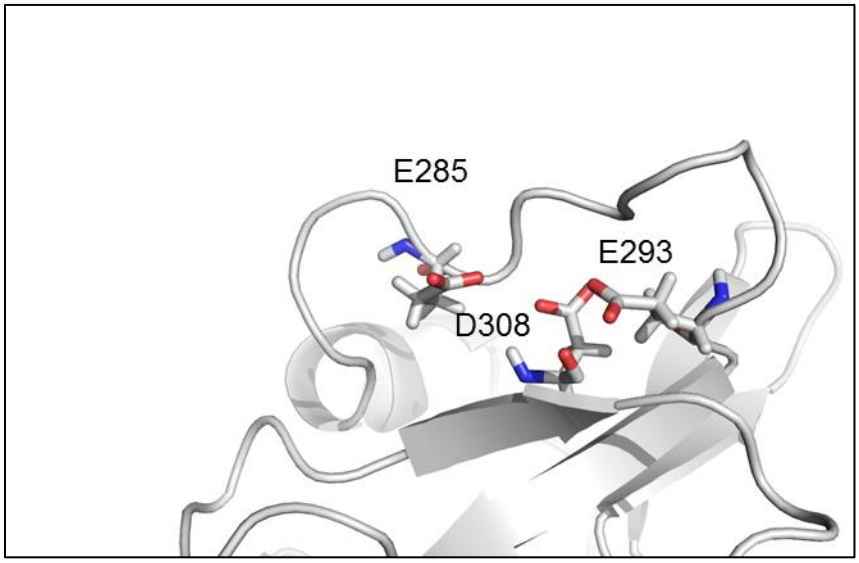



Figure S18: Residues that potentially act as pH sensors in the proximity of the $Ca^{2+}$ cage. Calculated $pK_a$ values of marked residues are listed in Table S5.

Supporting Tables

Table S1: Observed $^{1}H$ and $^{15}N$ chemical shift differences of the apo and cis states in apo Langerin CRD.

| **Residue No.** | **Δδ ($^{1}H$) (Hz)** | **Δδ ($^{15}N$) (Hz)** |
|---|---|---|
| 247 | 23 | 2 |
| 252 | 20 | 6 |
| 253 | 34 | 1 |
| 254 | 47 | 25 |
| 257 | 90 | 76 |
| 258 | 57 | 75 |
| 259 | 38 | 7 |
| 260 | 33 | 16 |
| 262 | 12 | 1 |
| 264 | 73 | 18 |
| 265 | 22 | 36 |
| 266 | 24 | 19 |
| 277 | 50 | 20 |
| 278 | 14 | 16 |
| 279 | 37 | 10 |
| 280 | 38 | 3 |

| | | |
|---|---|---|
| 281 | 23 | 38 |
| 285 | 17 | 36 |
| 297 | 2 | 24 |
| 300 | 113 | 54 |
| 302 | 65 | 43 |
| 304 | 20 | 9 |
| 305 | 71 | 32 |

Table S2: The list of the stable hydrogen bonds connecting the various structural elements of the Langerin CRD in all three WT forms.

| | | **Apo Langerin** | | **Holo Langerin** |
|---|---|---|---|---|
| **Donor** | **Acceptor** | **Cis [%]** | **Trans [%]** | **Cis [%]** |
| A305-Main | F280-Main | 22 | 16 | 20 |
| W306-Main | F280-Main | 97 | | 96 |
| I282-Main | W306-Main | 91 | 84 | 90 |
| E285-Main | I282-Main | 72 | | 72 |
| N291-Main | N288-Main | 35 | | 33 |
| E293-Main | N288-Side | 55 | | 49 |
| N292-Main | A289-Main | 46 | | 49 |
| A289-Main | E293-Side | 60 | | 63 |
| K257-Side | H294-Side | 33 | 35 | 32 |
| K257-Main | H294 -Main | 51 | 56 | 54 |
| A309-Main | C295-Main | 91 | 81 | 89 |
| G296-Main | L255-Main | 59 | 64 | 61 |
| W252-Side | G296-Main | 25 | 24 | 29 |
| I253-Main | G296-Main | 59 | 56 | 65 |
| N297-Main | N307-Main | 90 | 92 | 90 |

| | | | | |
|---|---|---|---|---|
| N307-Side | N297-Side | 88 | 22 | 85 |
| I250-Main | I298-Main | 23 | 25 | 23 |
| I298-Main | Y251-Main | 90 | 89 | 90 |
| I298-Main | A305-Main | 12 | 12 | 16 |
| Y251-Main | I298-Main | 80 | 77 | 84 |
| K299-Main | A305-Main | 89 | 84 | 85 |
| Q239-Side | L303-Main | 63 | 50 | 44 |
| I318-Main | W252-Main | 32 | 29 | 33 |
| W252-Main | L316-Main | 91 | 94 | 90 |
| F315-Main | K214-Main | 87 | 90 | 88 |

Table S3: Average $^{15}N$ relaxation rate constants obtained from NMR measurements at 600 MHz and 750 MHz field strength. Average values with standard deviations are given. The rotational correlation time was estimated from $R_2$ and $R_1$.[6]

| | **apo** | | **holo** | |
|---|---|---|---|---|
| | **600 Mhz** | **750 MHz** | **600 Mhz** | **750 MHz** |
| $\boldsymbol{R_1}$ **(s$^{-1}$)** | 1.2±0.1 | 1.0±0.1 | 1.2±0.1 | 1.1±0.1 |
| $\boldsymbol{R_2}$ **(s$^{-1}$)** | 16.0±2.6 | 15.2±2.5 | 14.8±2.3 | 13.7±2.5 |
| **hetero NOE** | 0.75±0.14 | 0.79±0.11 | 0.74±0.2 | 0.81±0.1 |
| $\boldsymbol{\tau_c}$ **(10$^{-9}$ s)** | 10.8 | 11.6 | 10.0 | 9.6 |

Table S4: P286 prolyl cis/trans ratio observed in single-residue mutants at pH 6.

| Langerin CRD construct | Fraction in cis | STDEV |
|---|---|---|
| WT[+] | 73% | 10% |
| E220A | 78% | 4% |
| L230A | 76% | 3% |
| Q239A | 65% | 9% |
| K257A | 75% | 10% |
| G259I | 77% | 5% |
| D263A | 78% | 5% |
| S265A | 77% | 5% |
| H294A | 80% | 4% |
| N297A | 83% | 5% |
| K299A | 67% | 5% |

[+]cis/trans ratio was determined as average from three independent measurements. The error was propagated from the standard deviations of each measurement. Otherwise the error was calculated from the standard deviation of the cis/trans ratios of the single residues.

Table S5: pKa calculations for H294, E285 and E293 side chains of apo and holo Langerin CRD performed with different snapshots from states observed in the MD simulations using two different computational methods.

| **Structure** | **Method** | **pKa (H294)** | **$pK_a$ (E285)** | **$pK_a$ (E293)** |
|---|---|---|---|---|
| apo open | PROPKA3[7] | 5.87 | 6.35* | 5.29 |
| | Rosetta CL[8] | 3.6 | 6.3 | 5.9 |
| apo intermediate | PROPKA3 | 5.8 | 5.42 | 6.00* |
| | Rosetta CL | 4.8 | 5.9 | 6.2 |
| apo closed | PROPKA3 | 6.56 | 5.47 | 5.17 |
| | Rosetta CL | 6.5 | 6.5 | 4.7 |
| holo open | PROPKA3 | 5.99 | 5.08* | 4.74 |
| | Rosetta CL | -* | -* | -* |
| holo intermediate | PROPKA3 | 5.71 | 5.46* | 4.35 |
| | Rosetta CL | 5.7 | 3.1 | 2.4 |
| holo closed | PROPKA3 | 6.05 | 5.00* | 5.59* |
| | Rosetta CL | 6.3 | 2.8 | 1.9 |

*Calculations did not finish before server timeout.

Table S6: List of $^{1}$H-detected NMR experiments, pulse sequences, and parameters for backbone resonance assignment of Langerin CRD. Number of transients (nt), and complex points in each channel are given corresponding to $^{1}$H ($TD_3$), $^{13}$C ($TD_2$), and $^{15}$N ($TD_1$), respectively.

| **Sample** | **Experiment** | **Pulse sequence** | ***nt*** | ***$TD_3$*** | ***$TD_2$*** | ***$TD_1$*** |
|---|---|---|---|---|---|---|
| Langerin CRD WT (U:$^{15}$N;$^{13}$C, 350 µM) | HNCA | MFhnca_best | 16 | 1024 | 96 | 128 |
| | CT-HNCA | MFhnca_best | 16 | 1024 | 128 | 270 |
| | HNCO | MFhnco_best | 4 | 1024 | 116 | 128 |
| | HN(CA)CO | MFhncaco_best | 24 | 1024 | 96 | 128 |
| | HNCACB | MFhncacb_best | 32 | 1024 | 116 | 110 |
| | HN(CO)CACB | MFhncocacb_best | 32 | 1024 | 116 | 110 |
| | CC(CO)NH | MFccconnhwg | 8 | 1024 | 128 | 128 |
| | NOESY-HSQC | MFnoehsqcwtgf3 | 8 | 1024 | 128 | 128 |
| | HNHA | MFhnhawg | 8 | 1024 | 128 | 128 |
| | NCO | MFnco | 128 | - | 1024 | 128 |

Table S7: List of $^{1}H,^{15}N$-HSQC NMR experiments, pulse sequences and parameters. Number of transients (nt), and complex points in each channel are given corresponding to $^{1}H$ ($TD_2$) and $^{15}N$ ($TD_1$), respectively.

| **Sample** | **Experiment** | **Pulse sequence** | ***nt*** | ***$TD_2$*** | ***$TD_1$*** |
|---|---|---|---|---|---|
| Langerin CRD WT and mutants (U:$^{15}N$; 100-250 µM) | $^{1}H$-$^{15}N$ HSQC | HSQC (WET)[a] | 8-12 | 1404 | 96-128 |
| | | gNfhsqc (WATERGATE)[a] | 4-8 | 1404 | 128 |
| | | MFhsqcwtgf3[b] | 4 | 1024 | 128 |

[a]Performed on Agilent 600 MHz spectrometer with RT probe,

[b]Performed on Bruker 600 MHz spectrometer with cryogenic probe.

Table S8: List of $^{15}N$ relaxation NMR experiments, pulse sequences, and parameters. Number of transients (nt), and complex points in each channel are given corresponding to $^{1}H$ ($TD_2$), $^{15}N$ ($TD_1$), and number of planes respectively.

| **Sample** | **Experiment** | **Pulse sequence** | ***nt*** | $TD_2$ | $TD_1$ | $n_{planes}$ |
|---|---|---|---|---|---|---|
| Langerin CRD WT (U:$^{15}N$; 300 µM) | T1 | mb15nt1wtg_3d_11 | 24 | 1024 | 160 | 10 |
| | T2 | mb15nt2wtg_3d_6 | 56 | 1024 | 160 | 10 |
| | hetero NOE | pF15nnoewtg_3d_6 | 96 | 1024 | 160 | 2 |

Table S9: Volume of the simulation box, number of water molecules per simulation box, number of counter ions per simulation box, and total simulation time for each system.

| **Ref #** | **System** | **Box size** | **# $H_2O$** | **# $Ca^{2+}$** | **# $Cl^-$** | **Simulation Time** |
|---|---|---|---|---|---|---|
| 1 | trans apo Langerin | 238.14 $nm^3$ | 7247 | 0 | 0 | 2 µs |
| 2 | cis apo Langerin | 238.14 $nm^3$ | 7247 | 0 | 0 | 2 µs |
| 3 | holo Langerin | 238.14 $nm^3$ | 7245 | 1 | 2 | 2 µs |
| 4 | apo H294A mutant | 238.14 $nm^3$ | 7247 | 0 | 0 | 2 µs |
| 5 | holo H294A mutant | 238.14 $nm^3$ | 7245 | 1 | 2 | 2 µs |
| 6 | apo H294+ | 238.14 $nm^3$ | 7245 | 0 | 2 | 2 µs |

Supplemental Experimental Procedures

All standard chemicals and buffers used within these work were purchased from Sigma Aldrich (St. Louis, MO, USA) or Carl Roth (Karlsruhe, Germany) if not indicated otherwise. All data analysis, plotting and curve fitting was performed with OriginPro 2015 (OriginLab, Northampton, MA) if not indicated otherwise.

*Protein production*

Human Langerin CRD and ECD were cloned from a codon-optimized Langerin gene for bacterial expression (GenScript, Piscataway, NJ, USA) into a pET30a vector (Merck Millipore, Darmstadt, Germany). The expression constructs comprised a C-terminal Strep-tag 2[9] linked via a TEV cleavage site. Insoluble expression was performed in *E. coli* BL21* (ThermoFisher Scienific, Waltham, MA, USA) at 37 °C in LB medium (ECD) or isotope-labeled M9 medium (CRD). Protein production was induced by addition of 0.5 mM IPTG (Carl Roth, Karlsruhe, Germany). Cells were harvested 4 h after induction. Cell pellets were lysed by incubation with lysozyme (Sigma Aldrich, St. Louis, MO, USA) and DNase I (Applichem, Darmstadt, Germany) in a detergent containing buffer (50 mM Tris, 150 mM NaCl, 10 mM $MgCl_2$, 0.1% Triton-X, pH 8) for at least 3 h at RT. Inclusion bodies were washed once with lysis buffer and three time with 20 mM Tris, pH 8 with centrifugation steps of 10 min at 10,000*g*. IB pellets were solubilized in 6 M guanidinium hydrochloride in 100 mM Tris (pH 8) and 1 mM DTT at 37 °C for at least 2 h by adding 40 mL $L^{-1}$ culture.

Solubilized IBs were centrifuged at 15,000*g* for 90 min to remove insoluble cell debris. Langerin ECD was refolded by adding it drop wise into 0.4 M L-arginine in 50 mM Tris, 20 mM NaCl,, 0.8 mM KCl, pH 7.5 in a ratio of 1:10 while rapidly stirring at 4 °C. Reduced and oxidized glutathione was used in concentrations of 1 mM and 0.2 mM, respectively, to allow for cysteine redox shuffling. Refolded Langerin ECD was dialyzed against the mannan chromatography buffer (50 mM Tris, 150 mM NaCl, 20 mM $CaCl_2$, pH 7.5). Precipitates were removed by centrifugation at 15,000*g*, for 90 min. Protein purification was achieved by affinity chromatography using mannan coupled sepharose beads (Sigma Aldrich) as resin. After application, the resin was washed with at least 5 column volumes chromatography buffer and protein was subsequent eluted with 50 mM Tris, 150 mM Nacl, 5 mM EDTA (pH 7.5). Elution fractions were tested for protein by Bradford (Applichem) and protein-containing fractions were pooled. Pooled protein solution was concentrated to about 2 mL using centrifugal filters (10,000 MWCO, Corning, Corning, NY, USA). Buffer was exchanged against TBS or HBS with centrifugal desalting columns (ThermoFisher Scientific). Langerin CRD was refolded as described for Langerin ECD using 0.8 M L-arginine as the only different

parameter, dialyzed against 50 mM Tris, 150 mM NaCl, 1 mM EDTA, pH 8, centrifuged, and purified using a StrepTactin column (IBA, Göttingen, Germany). Loaded resin was washed with at least 5 column volumes chromatography buffer and protein was eluted with 2.5 mM d-desthiobiotin in chromatography buffer. Protein-containing fractions were pooled and dialyzed against 25 mM MES, 40 mM NaCl, pH 6 to remove soluble protein aggregates by precipitation. Insoluble aggregates were removed by centrifugation at 15,000*g* for 90 min. The protein solution was concentrated by centrifugal filters (10,000 MWCO, Corning).

*Plate-based assays*

NUNC maxisorb© 96-well plates (ThermoFisher Scientific) were coated overnight with 10 µg $mL^{-1}$ mannan in 50 mM carbonate buffer, pH 9.6, and blocked with 2% BSA in TBS-T for 1 h. Wells were washed three times with TBS-T + 5 mM $CaCl_2$, incubated with Langerin ECD in buffers with pH ranging from 4 to 9 (in steps of 0.5 pH units) for at least 90 min at room temperature. Wells were washed and incubated with StrepTactin-HRP conjugate (IBA, Göttingen, Germany) in 2% BSA in TBS-T + 5 mM $CaCl_2$ for 1 h. Plates were developed by adding TMB solution (tebu-bio, Versailles, France). The reaction was quenched by addition of 0.18 M $H_2SO_4$. Absorbance was measured at 450 nm on a SpectraMax M5 plate reader (Molecular Devices, Sunnyvale, CA). Data was fitted to a dose-response model according to the equation:

$$A_{450\text{nm}} = \frac{A_{450\text{nm, max}}}{1+\left(\frac{\text{IC}_{50}}{[\text{I}]}\right)^{p}} \qquad (3\text{-}01)$$

with $A_{450\text{nm}}$ as the absorption at 450 nm corresponding to the readout of the assay, [I] as the inhibitor concentration and $p$ the Hill coefficient. The same equation was utilized to obtain the effective concentration of binding by substitution of the inhibitor concentration with the protein concentration.

*ITC measurements*

Isothermal titration calorimetry experiments were performed using a MicroCal iTC200 or a VP-ITC (Malvern Instruments, Malvern, UK) using either chelex-filtered HBS (25 mM HEPES, 150 mM NaCl, pH 7) or low salt MES buffer (25 mM MES, 40 mM NaCl, pH 6). The titrant was dissolved in the same buffer as was used for dialysis of the protein sample. Using the iTC200, the titrant, either $CaCl_2$ (5 mM to 25 mM depending on the expected affinity) or mannose (50 mM), was added in defined steps of 1-2.5 µL to the protein solution (150 µM to 250 µM in 270 µL total volume) at 298 K while stirring at 750 rpm. For the low affinity systems at pH 6, the VP-ITC was used titrating

$CaCl_2$ (10 mM or 25 mM) or mannose (50 mM) in 20 µL steps to the protein solution (80 µM to 150 µM in 1.4 mL total volume). The differential heat of each injection was measured and plotted against the molar ratio. The data was fitted to a one-set of sites binding model assuming a Hill coefficient of 1.[10] Due to the low c-values of the measurements ($c < 5$), the enthalpy could not be determined reliably.

*NMR measurements and data processing and analysis*

All NMR assignment and relaxation measurements were performed on a 600 MHz Bruker Avance III spectrometer (Bruker, Billerica, MA, USA) equipped with a triple resonance cryogenic probe. Relaxation data was additionally collected on a 750 MHz Bruker Avance II spectrometer equipped with a triple resonance cryogenic probe. $^{1}H$-$^{15}N$-HSQC measurements were either performed on the Bruker 600 MHz (see above) or an Agilent 600 MHz vnmrs machine equipped with a room-temperature double resonance onenmr probe.

For Langerin CRD backbone resonance assignment, a series of standard triple resonance experiments was collected (Table S6) using 350 µM (U:$^{15}N$;$^{13}C$) protein in 25 mM MES, 40 mM NaCl, 5 mM $CaCl_2$, pH 6 containing 100 µM DSS, 0.05% $NaN_3$, and 10% $D_2O$.

$^{1}H$-$^{15}N$ HSQC NMR titrations experiments (Table S7) were performed with 120-250 µM of (U:$^{15}N$) Langerin CRD in 25 mM MES, 40 mM NaCl, pH 6 containing 100 µM DSS, and 10% $D_2O$. For titration of $Ca^{2+}$, $CaCl_2$ was added stepwise to the sample (600 µL initial volume) from a buffer-matched stock (0.1 M or 1 M depending on target concentration) until 10 mM final concentration as endpoint was reached. Step size was varied according to expected $K_d$. For WT and mutant Langerin CRD at pH 6, a five-step protocol of 0, 100, 500, 1500, and 10000 µM $CaCl_2$ was established as standard. Depending on protein concentration and implemented machine, 128 increments and 4-16 transients were recorded at 298 K. For measurements at pH 7, the buffer system was changed to 25 mM HEPES, 150 mM NaCl, pH 7. $Ca^{2+}$ titrations were performed accordingly. Mannose titration was performed analogously using the holo Langerin CRD in 25 mM MES, 40 mM NaCl, pH 6 containing 100 µM DSS, 10% $D_2O$, and 10 mM $CaCl_2$. D-Mannose was added stepwise from a 1 M stock to a final concentration of 100 mM. Measurements were preformed according to Table S6.

$^{15}N$-relaxation experiments were performed with (U:$^{15}N$)-labeled apo Langerin CRD (200 µM) in 25 mM MES, 40 mM NaCl, pH 6 containing 100 µM DSS and 10% $D_2O$ for apo Langerin CRD using the parameters shown in Table S8. The pulse sequences used for $R_2$ determination contained compensation pulses and per-FID interleaving for temperature control. Likewise, for the recording of the hetero NOE, an advanced saturation scheme was used.[11]

Measurement of the holo form was performed on the same sample after addition of 10 mM $CaCl_2$. The temperature was calibrated by measuring a 1% $CD_3OH$ in $D_2O$ prior to the relaxation measurements. The chemical shift difference between methanol and HDO signal is highly temperature sensitive and can be used to calibrate the sample's absolute temperature according to following empirical equation:[12]

$$T = -23.832\Delta\delta^2 - 29.46\Delta\delta + 403 \quad \text{(S1)}$$

The measurements were conducted on both field strengths at 299.2 K.

All assignment spectra were processed in Bruker TopSpin 3.2 and assignment was performed in CCPN Analysis[13] according to standard assignment strategies.

NMR titration data was processed with NMRpipe[14] applying Lorentz-to-Gauss transformation in f2 with 20 Hz line-broadening factor and sine bell functions in f1 dimension. 4x zero-filling and polynomial baseline corrections were applied in both dimensions. The spectra were visualized, referenced and analyzed in CCPN. All spectra were referenced using the internal spectrometer reference and assignments were transferred from the reference spectrum (root spectrum from the resonance backbone assignment) to the nearest neighbor. In case of ambiguities caused by strongly overlapping or disappearing peaks, the resonance assignments were not transferred. For titration experiments, the assignment transfer was facilitated by following shifting peaks in respect to the reference spectrum. Peak lists for each mutant and titration point were exported for CSP calculation according to:

$$\Delta\delta = \sqrt{\frac{1}{2}\left[\Delta\delta_{\mathrm{H}}^2 + \left(\alpha\Delta\delta_{\mathrm{N}}\right)^2\right]} \quad \text{(S2)}$$

with $\Delta\delta_i$ as the difference in chemical shift (in ppm) and $\alpha$ an empirical weighing factor. A weighing factor of 0.2 for glycine and 0.14 for all other amino acid backbone resonances is used.[15]

Dissociation constants were obtained by global fit of peaks that appeared to be near fast exchange limit using equation (S3):

$$\Delta\delta_{obs} = \Delta\delta_{max}\frac{\left([P]_t + [L]_t + K_d\right) - \sqrt{\left([P]_t + [L]_t + K_d\right)^2 - 4[P]_t[L]_t}}{2[P]_t} \quad \text{(S3)}$$

with $\Delta\delta_{obs}$ as the observed chemical shift difference, $\Delta\delta_{max}$ the maximal chemical shift difference, and $[P]_t$ and $[L]_t$ as the total protein and ligand concentrations. $\Delta\delta_{max}$ and $K_d$ can be fitted as free parameters when $\Delta\delta_{obs}$ is measured at different ligand concentrations.

Three independent measurements with different protein batches measured on three spectrometers with three different HSQC pulse sequences awarded a standard deviation of 0.02 ppm for the average chemical shifts over all residues.

CHESPA analysis was performed as described in the literature[16] defining the vector between apo and holo WT as reference and calculating the vector product to determine cos(θ) and projection of the vector between apo mutant and apo WT to determine the fractional vector length X. The same cut-off as defined for the CSP measurements was applied after conversion to vector length (0.025 ppm).

The relaxation data was processed in NMRPipe[14] and peaks were picked and assignment transferred in Sparky.[17] To obtain $R_1$ and $R_2$ relaxation rate constants, peak heights were fitted against the relaxation period using a single exponential decay model with a Newton minimization algorithm implemented in the relax 4.0 analysis package [18]. Errors were estimated by using a Monte Carlo simulation with 500 iterations that tests back-calculated data from the model against the experimental values. The hNOE was determined by the difference in peak intensities of the saturated and reference spectra. The error was estimated based on the experimental noise given by the RMSD of spectral regions without signals. Model-free analysis was performed with the in-built algorithm of relax 4.0 but resulted in overfitting of the parameters.

*Sequence alignment and conservation score*

All mammalian Langerin sequences deposited in UniProtKB were chosen for multiple sequence alignment using CLUSTAL W[19] and subsequent scoring of residue conservation.[20]

*MD Simulations*

All-atom molecular dynamics simulations were performed for the following systems: apo and holo Langerin, apo and holo H294A mutant, and protonated H294 apo Langerin, all of which have P286 in the cis-conformation. Additionally, trans apo Langerin was simulated (see Table S9). All simulations were carried out in explicit water (TIP3P water model[21]) using the GROMACS 4.5.5 simulation package[22] and the AMBER ff99SB-ILDN[23] force field. From the 3P5H[24] crystal structure, the chain A was extracted, while crystal water, and laminaritriose were removed. The initial structure of transP286 apo-Langerin was produced in VMD[25] using the Fix Cis Peptides Bonds tool. The initial structure of the H294A mutant was obtained using the Mutagenesis tool of Pymol (Schrödinger, LLC). In simulation #6, the protonation state of H294 of H229 were changed using the pdb2gmx command of GROMACS. For the apo monomers, the $Ca^{2+}$ ion was removed. The systems were energy minimized in vacuum with

the steepest decent algorithm[27], and then solvated in octahedron boxes with a minimum distance between solute and box walls of 1 nm, followed by another minimization run of the solvated systems. Two Cl- ions were added to all holo systems (#3 und #5) and the system #6 as counter ions to maintain the net charge of the system at zero. The solvated monomers were then equilibrated in the NVT ensemble at 300 K (V-rescale thermostat,[28] time constant = 0.1 ps), and NPT ensemble (Parrinello-Rahman barostat,[29] reference pressure = 1 bar, time constant = 2 ps) for 100 ps respectively. Initial structures for five MD runs were obtained from a short simulation at 350 K (1 ns). Prior to the start of MD runs, the temperature of the system was set back to 300 K. The production runs were simulated in the NPT ensemble (temperature 300 K, pressure 1 bar). Each of the five replicas was simulated for 400 ns, yielding 2 µs of simulation data per system. Position constrains were applied to the $Ca^{2+}$ atom of the holo monomer during the equilibration procedure, but removed in production runs. The LINCS algorithm[30] was used to constrain covalent bonds to hydrogen atoms (lincs iter = 1, lincs order = 4), allowing for an integration time step of 2 fs. Simulations were performed with the leap-frog integrator, the cut-off for Lennard-Jones interactions was set to 1 nm, while electrostatic interactions were treated with the Particle-Mesh Ewald (PME) algorithm[31] with a real space cutoff of 1 nm, a grid spacing of 0.16 nm, and an interpolation order of 4. Periodic boundary conditions were applied in all three dimensions. The solute coordinates were written to the trajectory file every 1 ps.

*Backbone flexibility*

In order to assess the backbone flexibility of the apo- and holo- Langerin monomer (system #2 and #3), the backbone dihedrals were extracted with the GROMACS command g_rama from the simulated trajectories. The in-house developed MATLAB (2011a), The MathWorks, Natick, MA, USA) script, based on the discretization of the Ramachandran plane into 360 x 360 = 129600 bins (the bin width of 1°), was used to project the extracted {ϕ–ψ}-time series onto the grid. The Ramachandran plots were produced for all amino acids for both systems. Additionally, the changes in the {ϕ–ψ}–equilibrium distributions between the apo- and holo- Langerin monomer were observed in the pair-wise difference Ramachandran plots.

*Side chain flexibility*

The side chain dihedral angles were extracted with the GROMACS command g_chi from the apo- and holo-Langerin trajectories (system #2 and #3), and projected onto a one-dimensional histogram with 360 bins (the bin width of 1°),

by the in-house MATLAB script. Side chain equilibrium distributions were plotted for all $\chi_l$ ($l \in \{1,2,3,4\}$) angles of all amino acids, but alanines and glycines.

*Distance Plot*

Taken together NMR data, and MD simulations, we noticed that short loop is the most flexible region in the Langerin structure, and thus we concluded, that a good reaction coordinate to describe the conformational dynamics of short loop, is the distance between Cα atom of G290 of α3 helix and C-α atom of M260 of short loop (both atoms are at the tips of the respective secondary structure elements). This distance was extracted with the GROMACS command g_mindist, and the distance distributions for apo- and holo-Langerin (system #2 and #3) were plotted on the top of each other.

*Normalized Mutual Information*

The possibility of the inter-residue correlation was investigated by employing the Mutual Information (MI) theory. In contrast to the Pearson correlation coefficient, MI also accounts for nonlinear correlations. The MI of two random variables $v_i$, and $v_j$ is given by:

$$\boldsymbol{MI}\left(\boldsymbol{v_i}\,,\ \boldsymbol{v_j}\right) = \int_{v_i}\int_{v_j} \boldsymbol{p_{ij}}\left(\boldsymbol{v_i}, \boldsymbol{v_j}\right) \mathbf{log}\left(\frac{\boldsymbol{p_{ij}}\left(\boldsymbol{v_i}, \boldsymbol{v_j}\right)}{\boldsymbol{p_i}\left(\boldsymbol{v_i}\right)\boldsymbol{p_j}(\boldsymbol{v_j})}\right) \boldsymbol{dv_i dv_j} \qquad (1)$$

where $p_{ij}(v_i,v_j)$ is the joint distribution, while $p_i(v_i)$, and $p_j(v_j)$ are the marginal distributions. Since the values of MI are not restricted to a certain interval, the normalized mutual information (NMI), which has a property to be bound to [0,1] interval is commonly used:

$$\boldsymbol{NMI}\left(\boldsymbol{v_i}, \boldsymbol{v_j}\right) = \frac{\boldsymbol{MI}\left(\boldsymbol{v_i}, \boldsymbol{v_j}\right)}{\mathbf{min}\left(\boldsymbol{H_i}, \boldsymbol{H_j}\right)} \qquad (2)$$

$H_i$, and $H_j$ denote the informational entropies of variables $v_i$, and $v_j$, which are defined as:

$$\boldsymbol{H_i} = -\int_{v_i} \boldsymbol{p_i}(\boldsymbol{v_i})\, \mathbf{log}\, \boldsymbol{p_i}(\boldsymbol{v_i}) \boldsymbol{dv_i} \qquad (3)$$

Two variables are uncorrelated, when their NMI equals to zero. On the other hand, one random variable is fully determined by another, if that their NMI value is 1.[13]

The $\{\phi-\psi\}$ - dihedral time series of the each amino acid of the Langerin sequence was projected onto a regular grid of 36 x 36 = 1296 bins (36 bins per torsion, bin 10° x 10°). The probability distributions are calculated as normalized

histograms in this state space. Side chain dihedral $\chi_i$ (i $\in$ {1,2,3,4}) time series were discretized as a one-dimensional histogram with 36 bins (bin width 10°).

We obtained the normalized mutual information for the following discrete random variables: i.) NMI of $\{\phi-\psi\}_i$ versus $\{\phi-\psi\}_j$, where *i,j* are the amino acid indices (i,j $\in$ {199,200,…,324}); ii.) NMI of $\chi_{ik}$ versus $\chi_{jl}$, where *i,j* are the amino acid indices (i,j $\in$ {199,200,…,324}), and *k,l* are the side chain dihedral indices (k,l $\in$ {1,2,3,4}); iii.) NMI of $\{\phi-\psi\}_j$ versus $\chi_{ik}$, where *i,j* are the amino acid indices (i,j $\in$ {199,200,…,324}), and *k* is the side chain angle index (k $\in$ {1,2,3,4}).

Due to the finite sampling of the {ϕ−ψ}- and χ-dihedral angles and the resulting statistical noise in the histograms, even two fully uncorrelated degrees of freedom exhibit a residual normalized mutual information. This residual NMI, i.e. the significance level of the NMI test, depends on the number of data points used to construct the histograms. In previous studies[14], the significance level has been set arbitrary to 0.1 (where degrees of freedom with a NMI of smaller than 0.1 were regarded as fully uncorrelated). Here we determined the significance level of the NMI test for one MD data set by calculating the NMI of a residue *i* in the apo-Langerin with a residue *j* in the holo-Langerin. Because the simulations are completely independent the NMI should be zero, yet we observed residual NMIs of up to 0.015. We hence set the significance level to 0.02. Residues pairs with a NMI of less than 0.02 were considered uncorrelated.

*Computation of NMR Observables from MD Simulations*

To match NMR data with MD simulations we computed chemical shifts, and $^3$J coupling constants ($H^N H^\alpha$). We used program SPARTA+[15] to obtain chemical shifts from simulations #2, #3, #4, #5. $^3$J coupling constants can be calculated from MD simulations by implementing Karplus equation:[2]

$$^3J(\phi) = A\cos^2(\phi+\theta) + B\cos(\phi+\theta) + C \quad (4)$$

where ϕ is the backbone dihedral angle, θ correction angle (-60°), and A, B, and C are empirically calibrated parameters. We tested various combinations of A, B, and C parameter previously reported in literature[3].

**Supporting References**


(1) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Soding, J.; Thompson, J. D.; Higgins, D. G. *Mol. Syst. Biol.* **2011**, *7*, 539.

(2) Karplus, M. *J. Chem. Phys.* **1959**, *30*, 11.

(3) Beauchamp, K. A.; Lin, Y.-S.; Das, R.; Pande, V. S. *J. Chem. Theory Comput.* **2012**, *8*, 1409.

(4) Feinberg, H.; Mitchell, D. A.; Drickamer, K.; Weis, W. I. *Science* **2001**, *294*, 2163.

(5) Guo, Y.; Feinberg, H.; Conroy, E.; Mitchell, D. A.; Alvarez, R.; Blixt, O.; Taylor, M. E.; Weis, W. I.; Drickamer, K. *Nat. Struct. Mol. Biol.* **2004**, *11*, 591.

(6) Clore, G. M.; Driscoll, P. C.; Wingfield, P. T.; Gronenborn, A. M. *Biochemistry* **1990**, *29*, 7387.

(7) Olsson, M. H. M. *Proteins* **2011**, *79*, 3333.

(8) Kilambi, Krishna P.; Gray, Jeffrey J. *Biophys. J.* **2012**, *103*, 587.

(9) Schlapschy, M.; Grimm, S.; Skerra, A. *Protein Eng. Des. Sel.* **2006**, *19*, 385.

(10) Freyer, M. W.; Lewis, E. A. *Method Cell Biol* **2008**, *84*, 79.

(11) Ferrage, F.; Reichel, A.; Battacharya, S.; Cowburn, D.; Ghose, R. *Journal of Magnetic Resonance (1969)* **2010**, *207*, 294.

(12) Findeisen, M.; Brand, T.; Berger, S. *Magn. Reson. Chem.* **2007**, *45*, 175.

(13) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. *Proteins* **2005**, *59*, 687.

(14) Delaglio, F.; Grzesiek, S.; Vuister, G.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.

(15) Williamson, M. P. *Prog. Nucl. Mag. Res. Sp.* **2013**, *73*, 1.

(16) Selvaratnam, R.; VanSchouwen, B.; Fogolari, F.; Mazhab-Jafari, M. T.; Das, R.; Melacini, G. *Biophys. J.* **2012**, *102*, 630.

(17) Goddard, T. D.; Kneller, D. G.; University of California, San Francisco.

(18) d'Auvergne, E. J.; Gooley, P. R. *J. Biomol. NMR* **2008**, *40*, 107.

(19) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. *Bioinformatics* **2007**, *23*, 2947.

(20) Valdar, W. S. *Proteins* **2002**, *48*, 227.

(21) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(22) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. *Journal of Computional Chemistry* **2005**, *26*, 1701.

(23) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins* **2010**, *78*, 1950.

(24) Feinberg, H.; Taylor, M. E.; Razi, N.; McBride, R.; Knirel, Y. A.; Graham, S. A.; Drickamer, K.; Weis, W. I. *J. Mol. Biol.* **2011**, *405*, 1027.

(25) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics Model.* **1996**, *14*, 33.

(26) Schrodinger, L. 2015.

(27) Dumont, R. S. *J. Comp. Chem.* **1991**, *12*, 391.

(28) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.

(29) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182.

(30) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comp. Chem.* **1997**, *18*, 1463.

(31) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.