Identification and Rational Redesign of Peptide Ligands to CRIP1, A Novel Biomarker for Cancers

Jihua Hao[1], Adrian W.R. Serohijos[4,5], Gail Newton[7], Gina Tassone[7], Zuncai Wang[7], Dennis C. Sgroi[7], Nikolay V. Dokholyan[4,6,*], James P. Basilion[1, 2, 3,*]

[1]Department of Radiology, [2]Department of Biomedical Engineering, and [3]National Foundation for Cancer Research Center for Molecular Imaging, Case Western Reserve University, Cleveland, OH 44106

[4]Department of Biochemistry and Biophysics, [5]Department of Physics and Astronomy, [6]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

[7] Harvard Medical School, Massachusetts General Hospital, Boston, MA 02129

Supplementary Information

Figure S1. The cDNA and amino acid sequences of CRIP1.

After cloning, the insert was confirmed by sequencing and the deduced amino acid sequences for the human CRIP1 shown in Figure S1. Excluding the vector sequence and the poly A region, the cDNA insert is 243 base pairs in length. The start site for transcription is at nucleotide position 73 (not shown in figure) with the start of translation at nucleotide position 162. This open reading frame expresses the amino acids encoding the His-tag (nt: 186-242) and encoding an enterokinase clevage site (nt: 246-260). The sequences encoding the human CRIP1 protein begin at nucleotide 273 and continue through nucleotide 503. Translation of these sequences results in a polypeptide 114 amino acids in length, the majority of which, 77 amino acids, make up CRIP1 protein.

The start (ATG) and stop (TAA) codons are underlined. The sequence of nonadjacent 6 histidines on HAT epitope is in bold. The poly A tail at the end is not shown.

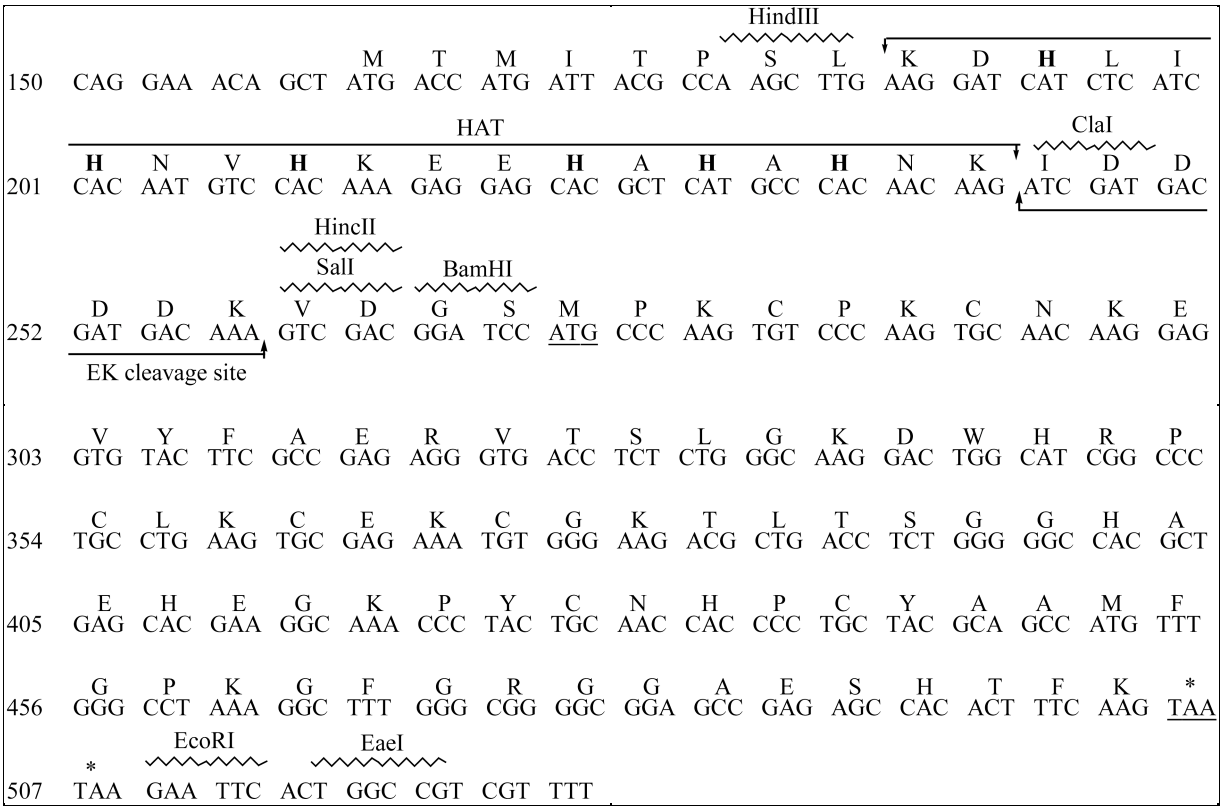
HindIII
HAT
ClaI
HincII
SalI
BamHI
EK cleavage site
EcoRI
EaeI
150 CAG GAA ACA GCT ATG ACC ATG ATT ACG CCA AGC TTG AAG GAT CAT CTC ATC
M T M I T P S L K D H L I
201 CAC AAT GTC CAC AAA GAG GAG CAC GCT CAT GCC CAC AAC AAG ATC GAT GAC
H N V H K E E H A H A H N K I D D
252 GAT GAC AAA GTC GAC GGA TCC ATG CCC AAG TGT CCC AAG TGC AAC AAG GAG
D D K V D G S M P K C P K C N K E
303 GTG TAC TTC GCC GAG AGG GTG ACC TCT CTG GGC AAG GAC TGG CAT CGG CCC
V Y F A E R V T S L G K D W H R P
354 TGC CTG AAG TGC GAG AAA TGT GGG AAG ACG CTG ACC TCT GGG GGC CAC GCT
C L K C E K C G K T L T S G G H A
405 GAG CAC GAA GGC AAA CCC TAC TGC AAC CAC CCC TGC TAC GCA GCC ATG TTT
E H E G K P Y C N H P C Y A A M F
456 GGG CCT AAA GGC TTT GGG CGG GGC GGA GCC GAG AGC CAC ACT TTC AAG TAA
G P K G F G R G G A E S H T F K *
507 TAA GAA TTC ACT GGC CGT CGT TTT
* E F T G R R F

Figure S2. CRIP1 Purity. Comassie Blue Stained-SDS-PAGE analysis of CRIP1 lysate and elutions after purification**.**

Lane 1: standard molecular marker;
Lane 2: lysate before incubation with Resin;
Lane 3: lysate after incubation with Resin;
Lane 4~5: fractions through Clontech TALON CellThru column;

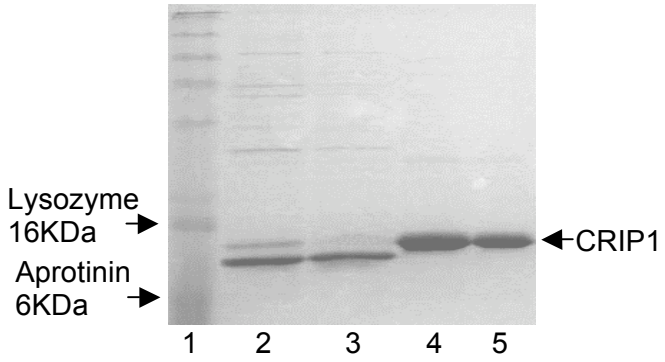

Figure S3. Relative estimates of peptide affinity for CRIP1. Phage binding against immobilized CRIP-1.

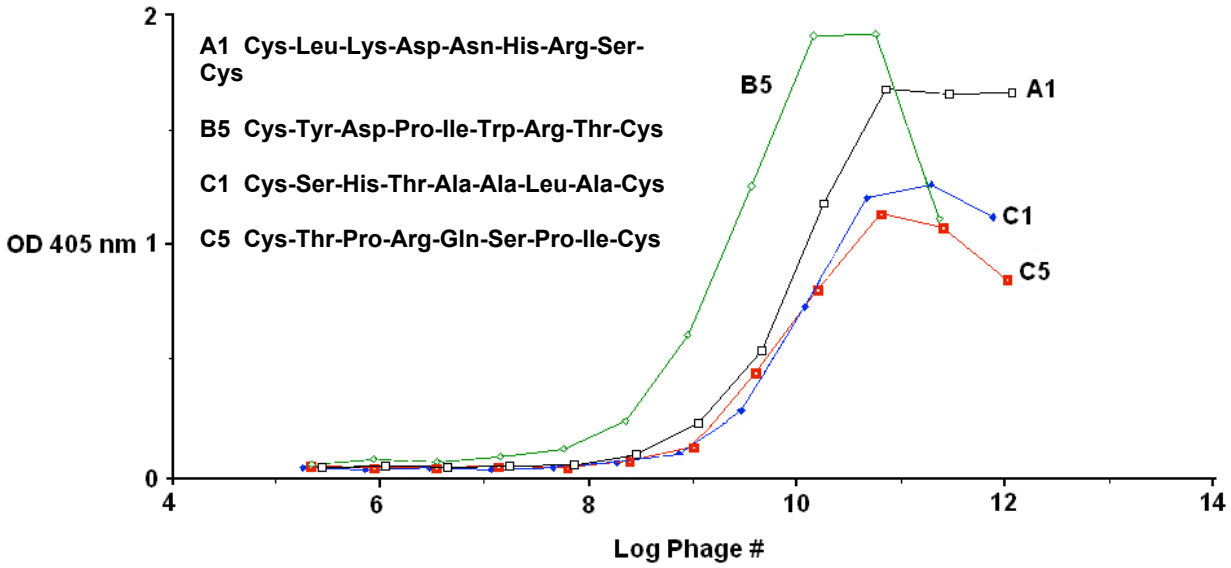

Figure S4. Synthesis of FITC-peptides

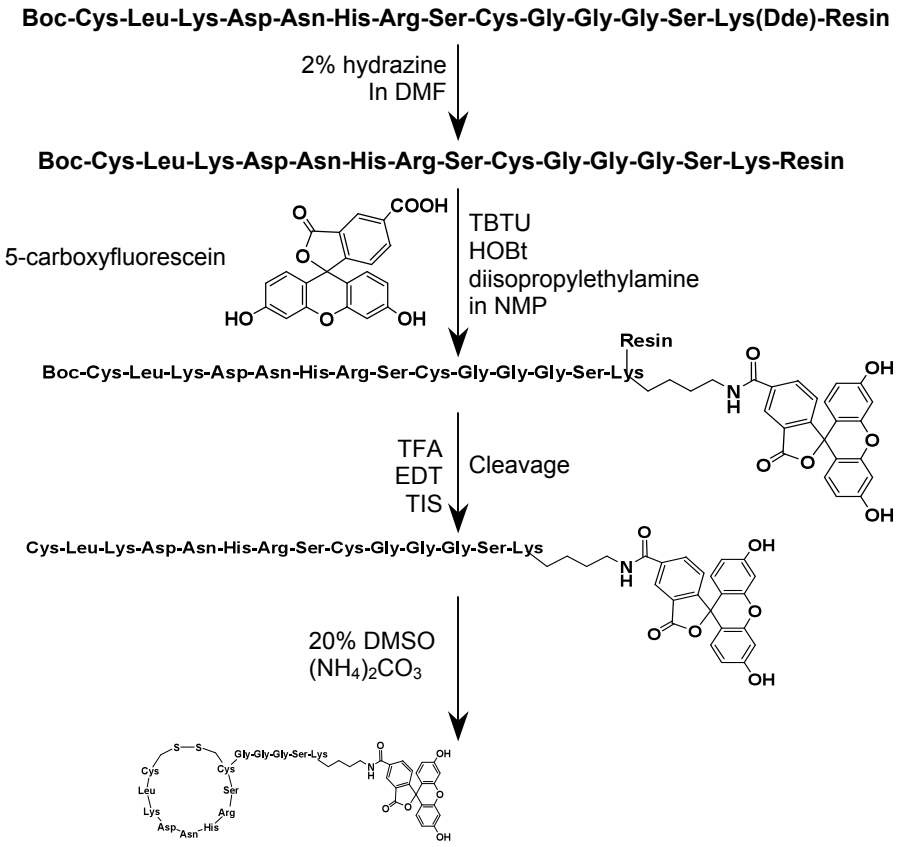

Figure S5. Sequences of redesigned peptides. Shown below are the sequence motifs of the redesigned peptides for the starting peptide structure models 1-ns, 9-ns, and 10-ns. The rank pertains to the order putative the binding site on CRIP1 defined from clustering.

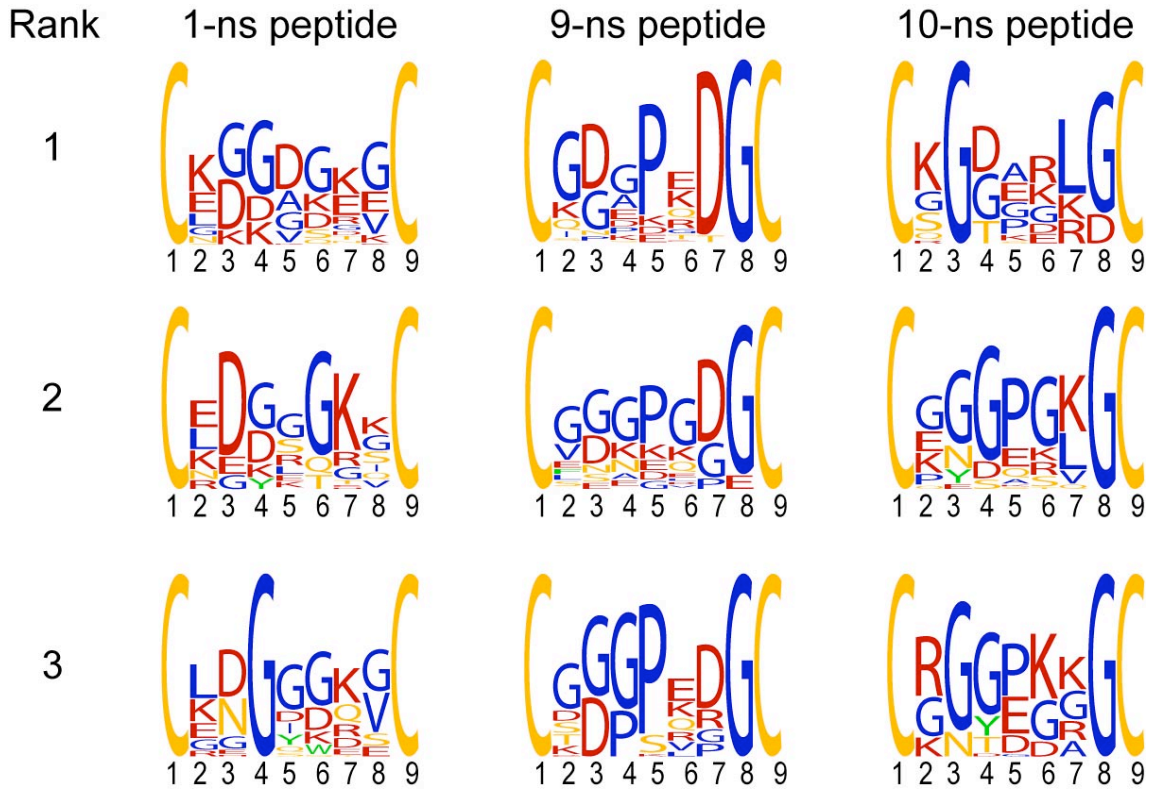

Figure S6. Residue preference without CRIP context. To verify that the observed preference for Gly in some sites in the peptide is not due to a bias in the force field, we employed the protocol to find the optimal peptide sequence when the peptide is not bound to CRIP1. We used 50 independent redesign runs. The preferred sequences are expectedly highly polar which maximize the peptide solvation energy.

Table S1. Analysis of CRIP1

Analysis	Entire Protein
Length	114 aa
Molecular Weight	12743.72
1 microgram	78.470pMoles
Molar Extinction coefficient	10370
1 A[280] corr. To	1.23 mg/ml
A[280] of 1 mg/ml	0.81 AU
Isoelectric point	8.20
Charge at pH 7	2.66

Table S2. Contribution of individual energy terms to the ΔΔG of the redesigned peptide A1M CLDGGGKGC.

Energy Term[a]	**Symbol**	**Value (kcal/mol)**
van der Waals attraction	evdw_a	17.3
van der Waals repulsion	evdw_r	-99.1
Solvation	esolv	-15.6
Hydrogen bond interactions between sidechain and backbone	ehb_sb	1.1
Hydrogen bond interactions between sideschains	ehb_ss	0.0
Internal energy for a rotamer state	efy_chi	1.0
Correction for dependence rotamer energy on backbone conformation and amino acid type	efy_aa	2.7
Reference energy[b]	e_aa_ref	10.0
Total ΔΔG		**-83**

[a]Please see Ding and Dokholyan[39] for detailed calculation of the force field terms.
[b]Reference energy assumes that the polypeptide is completely unfolded.