

Supplementary Material

Energy evaluation of β -strand packing in a fibril-forming SH3 domain

Sichun Yang, Krishnakumar M Ravikumar, and Herbert Levine

The sequence of segments of A β , HET-s, and SH3 proteins used for packing energy evaluations are shown below.

Sequence of A β segments used in packing energy evaluation

S_4	DAEFRHD	S'_4	DHRFEAD
S_5	AEFRHDS	S'_5	SDHRFEA
S_6	EFRHDSG	S'_6	GSDHRFE
S_7	FRHDSGY	S'_7	YGSDHRF
S_8	RHDSGYE	S'_8	EYGSDHR
S_9	HDSGYEV	S'_9	VEYGSDH
S_{10}	DSGYEVH	S'_{10}	HVEYGSD
S_{11}	SGYEVHH	S'_{11}	HHVEYGS
S_{12}	GYEVHHQ	S'_{12}	QHHVEYG
S_{13}	YEVHHQK	S'_{13}	KQHHVEY
S_{14}	EVHHQKL	S'_{14}	LKQHHVE
S_{15}	VHHQKLV	S'_{15}	VLKQHHV
S_{16}	HHQKLVF	S'_{16}	FVLKQHH
S_{17}	HQKLVFF	S'_{17}	FFVLKQH
S_{18}	QKLVFFA	S'_{18}	AFFVLKQ
S_{19}	KLVFFAE	S'_{19}	EAFFVLK
S_{20}	LVFFAED	S'_{20}	DEAFFVL
S_{21}	VFFAEDV	S'_{21}	VDEAFFV
S_{22}	FFAEDVG	S'_{22}	GVDEAFF
S_{23}	FAEDVGS	S'_{23}	SGVDEAF
S_{24}	AEDVGSN	S'_{24}	NSGVDEA
S_{25}	EDVGSNK	S'_{25}	KNSGVDE
S_{26}	DVGSNKG	S'_{26}	GKNSGVD
S_{27}	VGSNKGA	S'_{27}	AGKNSGV
S_{28}	GSNKGAI	S'_{28}	IAGKNSG
S_{29}	SNKGAIH	S'_{29}	IIAGKNS
S_{30}	NKGAIHG	S'_{30}	GIIAGKN
S_{31}	KGAIHGL	S'_{31}	LGIAGK
S_{32}	GAIHGLM	S'_{32}	MLGIIAG
S_{33}	AIHGLMV	S'_{33}	VMLGIIA
S_{34}	IIHGLMVG	S'_{34}	GVMLGII
S_{35}	IGLMVGG	S'_{35}	GGVMLGI
S_{36}	GLMVGGV	S'_{36}	VGGVMLG
S_{37}	LMVGGVV	S'_{37}	VVGGVML
S_{38}	MVGGVVI	S'_{38}	IVVGGVM
S_{39}	VGGVVIA	S'_{39}	AIVVGGV

Sequence of HET-s prion segments used in packing energy evaluation

S_{221}	KIDAIVG	S'_{221}	GVIADIK
S_{222}	IDAIVGR	S'_{222}	RGVIADI
S_{223}	DAIVGRN	S'_{223}	NRGVIAD
S_{224}	AIVGRNS	S'_{224}	SNRGVIA
S_{225}	IVGRNSA	S'_{225}	ASNRGVI
S_{226}	VGRNSAK	S'_{226}	KASNRGV
S_{227}	GRNSAKD	S'_{227}	DKASNRG
S_{228}	RNSAKDI	S'_{228}	IDKASNR
S_{229}	NSAKDIR	S'_{229}	RIDKASN
S_{230}	SAKDIRT	S'_{230}	TRIDKAS
S_{231}	AKDIRTE	S'_{231}	ETRIDKA
S_{232}	KDIRTEE	S'_{232}	EETRIDK
S_{233}	DIRTEER	S'_{233}	REETRID
S_{234}	IRTEERA	S'_{234}	AREETRI
S_{235}	RTEERAR	S'_{235}	RAREETR
S_{236}	TEERARV	S'_{236}	VRAREET
S_{237}	EERARVQ	S'_{237}	QVRAREE
S_{238}	ERARVQL	S'_{238}	LQVRARE
S_{239}	RARVQLG	S'_{239}	GLQVRAR
S_{240}	ARVQLGN	S'_{240}	NGLQVRA
S_{241}	RVQLGNV	S'_{241}	VNGLQVR
S_{242}	VQLGNVV	S'_{242}	VVNGLQV
S_{243}	QLGNVVT	S'_{243}	TVVNGLQ
S_{244}	LGNVVTAA	S'_{244}	ATVVNGL
S_{245}	GNVVTAA	S'_{245}	AATVVNG
S_{246}	NVVTAAA	S'_{246}	AAATVVN
S_{247}	VVTAAAL	S'_{247}	LAAATVV
S_{248}	VTAAALH	S'_{248}	HLAAATV
S_{249}	TAAALHG	S'_{249}	GHLAAAT
S_{250}	AAALHGG	S'_{250}	GGHLAAA
S_{251}	AALHGGI	S'_{251}	IGGHLAA
S_{252}	ALHGGIR	S'_{252}	RIGGHIA
S_{253}	LHGGIRI	S'_{253}	IRIGGHL
S_{254}	HGGIRIS	S'_{254}	SIRIGGH
S_{255}	GGIRISD	S'_{255}	DSIRIGG
S_{256}	GIRISDQ	S'_{256}	QDSIRIG
S_{257}	IRISDQT	S'_{257}	TQDSIRI
S_{258}	RISDQTT	S'_{258}	TTQDSIR
S_{259}	ISDQTTN	S'_{259}	NTTQDSI
S_{260}	SDQTTNS	S'_{260}	SNTTQDS
S_{261}	DQTTNSV	S'_{261}	VSNTTQD
S_{262}	QTTNSVE	S'_{262}	EVSNTTQ
S_{263}	TTNSVET	S'_{263}	TEVSNTT
S_{264}	TNSVETV	S'_{264}	VTEVSNT
S_{265}	NSVETVV	S'_{265}	VVTEVSN
S_{266}	SVETVVG	S'_{266}	GVVTEVS
S_{267}	VETVVGK	S'_{267}	KGVVTEV

S_{268}	ETVVGKG	S'_{268}	GKGVVTE
S_{269}	TVVGKGE	S'_{269}	EGKGVVT
S_{270}	VVGKGES	S'_{270}	SEGKGVV
S_{271}	VGKGESR	S'_{271}	RSEGKGV
S_{272}	GKGESRV	S'_{272}	VRSEGKG
S_{273}	KGESRVL	S'_{273}	LVRSEK
S_{274}	GESRVLI	S'_{274}	ILVRSEG
S_{275}	ESRVLI	S'_{275}	GILVRSE
S_{276}	SRVLIG	S'_{276}	NGILVRS
S_{277}	RVLIGNE	S'_{277}	ENGILVR
S_{278}	VLIGNEY	S'_{278}	YENGILV
S_{279}	LIGNEYG	S'_{279}	GYENGIL
S_{280}	IGNEYGG	S'_{280}	GGYENGI
S_{281}	GNEYGGK	S'_{281}	KGGYENG
S_{282}	NEYGGKG	S'_{282}	GKGGYEN
S_{283}	EYGGKGF	S'_{283}	FGKGGYE
S_{284}	YGGKGF	S'_{284}	WFGKGGY
S_{285}	GGKGFWD	S'_{285}	DWFGKGG
S_{286}	GKGFWDN	S'_{286}	NDWFGKG

Sequence of SH3 domain segments used in packing energy evaluation

S_4	GSMSAEG	S'_4	GEASMSG
S_5	SMSAEGY	S'_5	YGEASMS
S_6	MSAEGYQ	S'_6	QYGEASM
S_7	SAEGYQY	S'_7	YQYGEAS
S_8	AEGYQYR	S'_8	RYQYGEA
S_9	EGYQYRA	S'_9	ARYQYGE
S_{10}	GYQYRAL	S'_{10}	LARYQYG
S_{11}	YQYRALY	S'_{11}	YLARYQY
S_{12}	QYRALYD	S'_{12}	DYLARYQ
S_{13}	YRALYDY	S'_{13}	YDYLARY
S_{14}	RALYDYK	S'_{14}	KYDYLAR
S_{15}	ALYDYKK	S'_{15}	KKYDYLA
S_{16}	LYDYKKE	S'_{16}	EKKYDYL
S_{17}	YDYKKER	S'_{17}	REKKYDY
S_{18}	DYKKERE	S'_{18}	EREKKYD
S_{19}	YKKEREE	S'_{19}	EEREKKY
S_{20}	KKEREED	S'_{20}	DEEREKK
S_{21}	KEREEDI	S'_{21}	IDEEREK
S_{22}	EREEDID	S'_{22}	DIDEERE
S_{23}	REEDIDL	S'_{23}	LDIDEER
S_{24}	EEDIDLH	S'_{24}	HLDIDEE
S_{25}	EDIDLHL	S'_{25}	LHLDIDE
S_{26}	DIDLHLG	S'_{26}	GLHLDID
S_{27}	IDLHLGD	S'_{27}	DGLHLDI
S_{28}	DLHLGDI	S'_{28}	IDGLHLD
S_{29}	LHLGDIL	S'_{29}	LIDGLHL
S_{30}	HLGDILT	S'_{30}	TLIDGLH
S_{31}	LGDILTV	S'_{31}	VTIDGL
S_{32}	GDILTVN	S'_{32}	NVTIDG
S_{33}	DILTVNK	S'_{33}	KNVTID
S_{34}	ILTVNKG	S'_{34}	GKNVTI
S_{35}	LTVNKGS	S'_{35}	SGKNVTL
S_{36}	TVNKGSL	S'_{36}	LSGKNVT
S_{37}	VNKGSLV	S'_{37}	VLSGKNV
S_{38}	NKGSLVA	S'_{38}	AVLSGKN
S_{39}	KGSLVAL	S'_{39}	LAVLSGK
S_{40}	GSLVALG	S'_{40}	GLAVLSG
S_{41}	SLVALGF	S'_{41}	FGLAVLS
S_{42}	LVALGFS	S'_{42}	SFGLAVL
S_{43}	VALGFSD	S'_{43}	DSFGLAV
S_{44}	ALGFSDG	S'_{44}	GDSFGLA
S_{45}	LGFS DGQ	S'_{45}	QGDSFGL
S_{46}	GFSDGQE	S'_{46}	EQGDSFG
S_{47}	FSDGQEA	S'_{47}	AEQGDSF
S_{48}	SDGQEAK	S'_{48}	KAEQGDS
S_{49}	DGQEAKP	S'_{49}	PKAEQGD
S_{50}	GQEAKPE	S'_{50}	EPKAEQG

S_{51}	QEA KP EE	S'_{51}	EE PK AE Q
S_{52}	EAK P EE I	S'_{52}	IE EP KA E
S_{53}	AK P EE IG	S'_{53}	GI EP KA
S_{54}	K P EE IG W	S'_{54}	WG I EP K
S_{55}	P EE IG WL	S'_{55}	LW G I EP
S_{56}	EE IG WLN	S'_{56}	NL WG I EE
S_{57}	E IG WLN G	S'_{57}	GN LW G IE
S_{58}	I GWLN GY	S'_{58}	YG NLW GI
S_{59}	G WLN GYN	S'_{59}	NY GN LWG
S_{60}	WLN GYNE	S'_{60}	EN YGN LW
S_{61}	LNG YNET	S'_{61}	TEN YGNL
S_{62}	NGYN ET T	S'_{62}	TTEN YGN
S_{63}	GYN ET TG	S'_{63}	GT TEN YG
S_{64}	YN ET TG E	S'_{64}	EG TTEN Y
S_{65}	NET TGER	S'_{65}	REG TTEN
S_{66}	ET TGER G	S'_{66}	GRE GTTE
S_{67}	TTGER GD	S'_{67}	DG REG TT
S_{68}	TGER GD F	S'_{68}	FD GRE GT
S_{69}	GER GD FP	S'_{69}	PF DG REG
S_{70}	ER GD FP G	S'_{70}	GP FD GRE
S_{71}	RG DF PG T	S'_{71}	TG PF DG R
S_{72}	G DF PG TY	S'_{72}	YT GP FD G
S_{73}	DF PG TY V	S'_{73}	VY TG PF D
S_{74}	FP GT YVE	S'_{74}	EV YT GP F
S_{75}	PG TY VEY	S'_{75}	YE VY TG P
S_{76}	GT YVEY I	S'_{76}	IY EVY TG
S_{77}	TY VEY IG	S'_{77}	GI YEVY T
S_{78}	Y VEY IGR	S'_{78}	RG IYEVY
S_{79}	VEY IGR K	S'_{79}	KR GIYEV
S_{80}	EY IGR KK	S'_{80}	KK RGIYE
S_{81}	Y IGR KKI	S'_{81}	IK KRGIY
S_{82}	IGR KKI S	S'_{82}	SI KKRGI
S_{83}	GR KKI SP	S'_{83}	PS IKKRG

Estimating uncertainty in packing energy

The uncertainty in the packing energy was evaluated by comparing the change when different length of the segments were used. Three different lengths of segment (5, 7, and 9) were used and the energy was evaluated using Eq.1 (shown below) for A β , HET-s and SH3 peptide systems as follows.

$$E_m(S_A, S_B) = \frac{1}{N_m} \sum_{i=1}^m \sum_{j=\max\{1, i-1\}}^{\min\{i+1, m\}} \epsilon^{MJ}(i, j) \quad (1)$$

where $E_m(S_A, S_B)$ is the packing energy between segments S_A , S_B , and $\epsilon^{MJ}(i, j)$ is the Miyazawa-Jernigan (MJ) statistical energy between residues i, j within segments of S_A , S_B respectively. The segment length (m), is chosen to be 5, 7, or 9 depending on the case (see Fig. S1). N_m , the number of terms in the summation (Eq.1), is a normalization factor which depends on the value of m . Specifically, $N_m = 13, 19$, or 25 when $m = 5, 7$, or 9 respectively (Fig. S1).

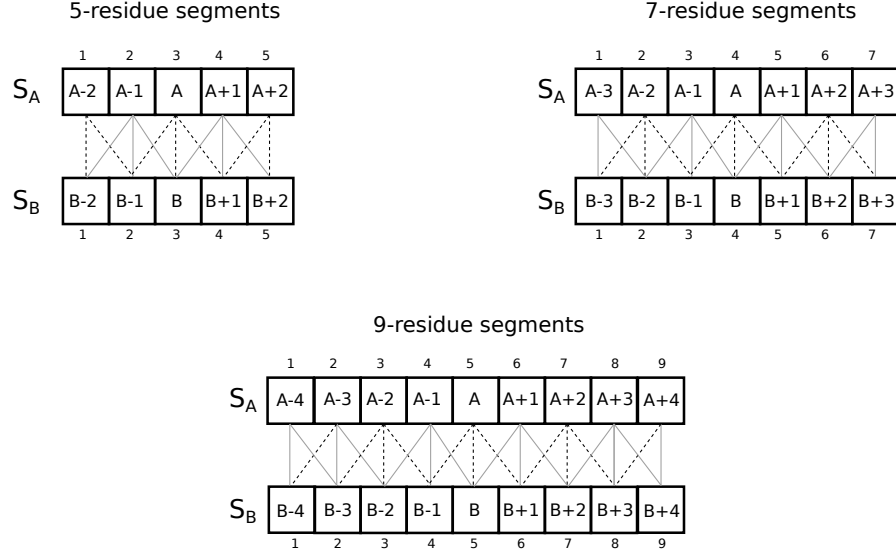


Fig. S1: Shown are 5-, 7-, and 9-residue long segment pairs (S_A, S_B) used for packing energy evaluation. Each line joining the residues between S_A and S_B (dotted and gray) represents a term in the summation in Eq.1. A total of 13, 19, and 25 terms are seen for 5-, 7-, and 9-residue long segments respectively.

The difference between $E_5(S_A, S_B)$ and $E_7(S_A, S_B)$ was calculated as

$$\chi_5(S_A, S_B) = \frac{\sqrt{(E_5(S_A, S_B) - E_7(S_A, S_B))^2}}{E_7(S_A, S_B)}.$$

Similarly, the difference between $E_9(S_A, S_B)$ and $E_7(S_A, S_B)$ was calculated as

$$\chi_9(S_A, S_B) = \frac{\sqrt{(E_9(S_A, S_B) - E_7(S_A, S_B))^2}}{E_7(S_A, S_B)}.$$

We found that the average $\bar{\chi}_5 = 7.8\%$ and $\bar{\chi}_9 = 6.1\%$ for all segment pairs in A β , HET-s and SH3 peptide systems (both one-way and two-way alignments). This suggests that the expected uncertainty in the energy evaluations is $\approx (\bar{\chi}_5 + \bar{\chi}_9)/2 = 6.9\%$.