Biophysical Journal, Volume 97

Supporting Material

Dynameomics: A consensus view of the protein unfolding/folding transition state ensemble across a diverse set of protein folds

Amanda Jonsson, Kathryn Scott, and Valerie Daggett

Supporting Material

Table S1: Properties of the aggregate TS ensemble of 183 proteins relative to the aggregate native state ensemble

Property	Fold Class			
	All	All-α	All-β	Mixed α/β
Polar SASA	1.17 (0.06)	1.14 (0.06)	1.18 (0.09)	1.18 (0.06)
Nonpolar SASA	1.37 (0.11)	1.34 (0.11)	1.37 (0.14)	1.40 (0.10)
Mc* Polar SASA	1.39 (0.12)	1.36 (0.13)	1.40 (0.17)	1.40 (0.11)
Mc Nonpolar SASA	1.28 (0.13)	1.22 (0.12)	1.32 (0.16)	1.29 (0.12)
Sc* Polar SASA	1.07 (0.06)	1.06 (0.06)	1.07 (0.08)	1.08 (0.06)
Sc Nonpolar SASA	1.39 (0.11)	1.35 (0.12)	1.38 (0.14)	1.41 (0.11)
Native Mc-Mc Contacts	0.81 (0.08)	0.88 (0.07)	0.74 (0.09)	0.82 (0.06)
Native Mc-Sc Contacts	0.70 (0.08)	0.75 (0.09)	0.65 (0.10)	0.71 (0.07)
Native Sc-Sc Contacts	0.63 (0.08)	0.63 (0.11)	0.63 (0.11)	0.65 (0.08)
Non-native Mc-Mc Contacts	1.19 (0.49)	1.40 (0.71)	1.22 (0.51)	1.07 (0.30)
Nonnative Mc-Sc Contacts	1.19 (0.44)	1.09 (0.46)	1.34 (0.54)	1.15 (0.33)
Nonnative Sc-Sc Contacts	1.29 (0.54)	1.20 (0.57)	1.43 (0.70)	1.25 (0.40)

For each property, for each protein the ratio of the mean value over the TS ensemble to the mean over the 1 – 21 ns in the 298 K simulation was first calculated – designated ($\langle x_{TS} \rangle / \langle x_N \rangle$). The values given in this table are the mean of $\langle x_{TS} \rangle / \langle x_N \rangle$ over all proteins in the set; the standard deviation is given in parentheses. *Mc indicates main chain, Sc indicates side chain.

Table S2: Per residue properties of the TS

Property	Fold Class			
	All	All-α	All-β	Mixed α/β
Polar SASA (Å^2)	33.59 (5.37)	34.77 (5.24)	34.32 (4.93)	32.18 (4.90)
Nonpolar SASA (Å^2)	49.10 (5.81)	52.04 (5.41)	48.63 (5.42)	47.26 (4.92)
Mc Polar SASA (Å^2)	12.43 (2.39)	11.37 (2.19)	13.78 (2.19)	11.96 (1.73)
Mc Nonpolar SASA (Å^2)	4.87 (0.87)	4.76 (0.68)	4.99 (0.97)	4.74 (0.74)
Sc Polar SASA (Å^2)	21.15 (4.26)	23.40 (4.67)	20.55 (3.64)	20.23 (3.83)
Sc Nonpolar SASA (Å^2)	43.98 (5.43)	47.06 (5.29)	43.30 (4.92)	42.38 (4.72)
# Native Mc-Mc Contacts[2]	1.72 (0.28)	1.97 (0.24)	1.49 (0.27)	1.58 (0.20)
# Native Mc-Sc Contacts[2]	1.53 (0.29)	1.71 (0.29)	1.32 (0.27)	1.58 (0.20)
# Native Sc-Sc Contacts[2]	0.86 (0.19)	0.79 (0.17)	0.85 (0.21)	0.94 (0.15)
# Nonnative Mc-Mc Contacts[2]	0.17 (0.06)	0.15 (0.06)	0.20 (0.06)	0.16 (0.06)
# Nonnative Mc-Sc Contacts[2]	0.30 (0.09)	0.25 (0.09)	0.34 (0.10)	0.29 (0.08)
# Nonnative Sc-Sc Contacts[2]	0.27 (0.07)	0.26 (0.08)	0.27 (0.08)	0.26 (0.07)

For each property, for each protein the mean value over the TS ensemble was calculated and divided by the number of residues in the protein. The values given in this table are the mean over all proteins in the set; the standard deviation is given in parentheses. *Mc indicates main chain, Sc indicates side chain. [1]Contacts counted on a residue-residue basis.[2]Contacts counted on an atom-atom basis.

Figure S1: Pairwise contact difference maps between the TS and the native state for all-α proteins. The number of times each pairwise contact is present in the TS ensemble for one protein is calculated and then divided by the number of structures used in the calculation. The pairwise contact maps for each protein are combined and divided by the number of proteins. The same calculation is then carried out over 1 – 21 ns of all 298 K simulations. The native value is subtracted from that for the TS and the appropriate bin colored according to the magnitude of the difference. Shades of blue represent a reduction in the number of a contact type in the TS, shades of pink an increase. Neighbouring ($i \rightarrow i$ and $i \rightarrow i + 1$) contacts are excluded in all plots. A) All contacts. B) Short range contacts. Only residue pairs separated by no more than 5 amino acids were considered. C) Medium range contacts. Only residue pairs separated by 6 to 15 amino acids were considered. D) Long range contacts. Only residue pairs separated by more than 15 amino acids were considered.

Figure S2: Pairwise contact difference maps between the TS and the native state for all-β proteins. The number of times each pairwise contact is present in the TS ensemble for one protein is calculated and then divided by the number of structures used in the calculation. The pairwise contact maps for each protein are combined and divided by the number of proteins. The same calculation is then carried out over 1 – 21 ns of all 298 K simulations. The native value is subtracted from that for the TS and the appropriate bin colored according to the magnitude of the difference. Shades of blue represent a reduction in the number of a contact type in the TS, shades of pink an increase. Neighbouring ($i \rightarrow i$ and $i \rightarrow i + 1$) contacts are excluded in all plots. A) All contacts. B) Short range contacts. Only residue pairs separated by no more than 5 amino acids

were considered. C) Medium range contacts. Only residue pairs separated by 6 to 15 amino acids were considered. D) Long range contacts. Only residue pairs separated by more than 15 amino acids were considered.

Figure S3: Pairwise contact difference maps between the TS and the native state for mixed α/β proteins. The number of times each pairwise contact is present in the TS ensemble for one protein is calculated and then divided by the number of structures used in the calculation. The pairwise contact maps for each protein are combined and divided by the number of proteins. The same calculation is then carried out over 1 – 21 ns of all 298 K simulations. The native value is subtracted from that for the TS and the appropriate bin colored according to the magnitude of the difference. Shades of blue represent a reduction in the number of a contact type in the TS, shades of pink an increase. Neighbouring ($i \rightarrow i$ and $i \rightarrow i + 1$) contacts are excluded in all plots. A) All contacts. B) Short range contacts. Only residue pairs separated by no more than 5 amino acids were considered. C) Medium range contacts. Only residue pairs separated by 6 to 15 amino acids were considered. D) Long range contacts. Only residue pairs separated by more than 15 amino acids were considered.

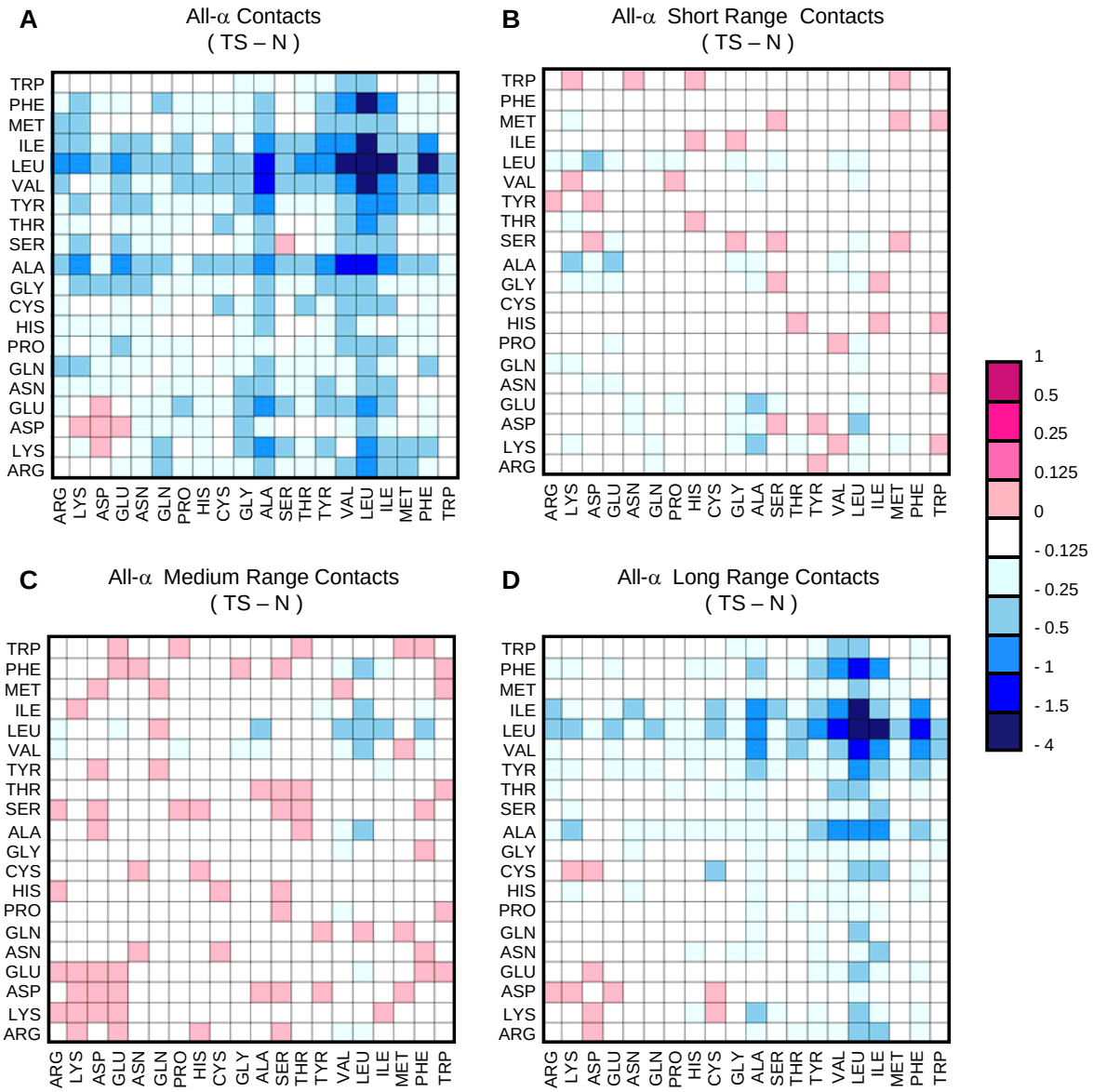
A
All-α Contacts
(TS – N)
B
All-α Short Range Contacts
(TS – N)
C
All-α Medium Range Contacts
(TS – N)
D
All-α Long Range Contacts
(TS – N)
TRP PHE MET ILE LEU VAL TYR THR SER ALA GLY CYS HIS PRO GLN ASN GLU ASP LYS ARG
ARG LYS ASP GLU ASN GLN PRO HIS CYS GLY ALA SER THR TYR VAL LEU ILE MET PHE TRP
1
0.5
0.25
0.125
0
- 0.125
- 0.25
- 0.5
- 1
- 1.5
- 4

Figure S1

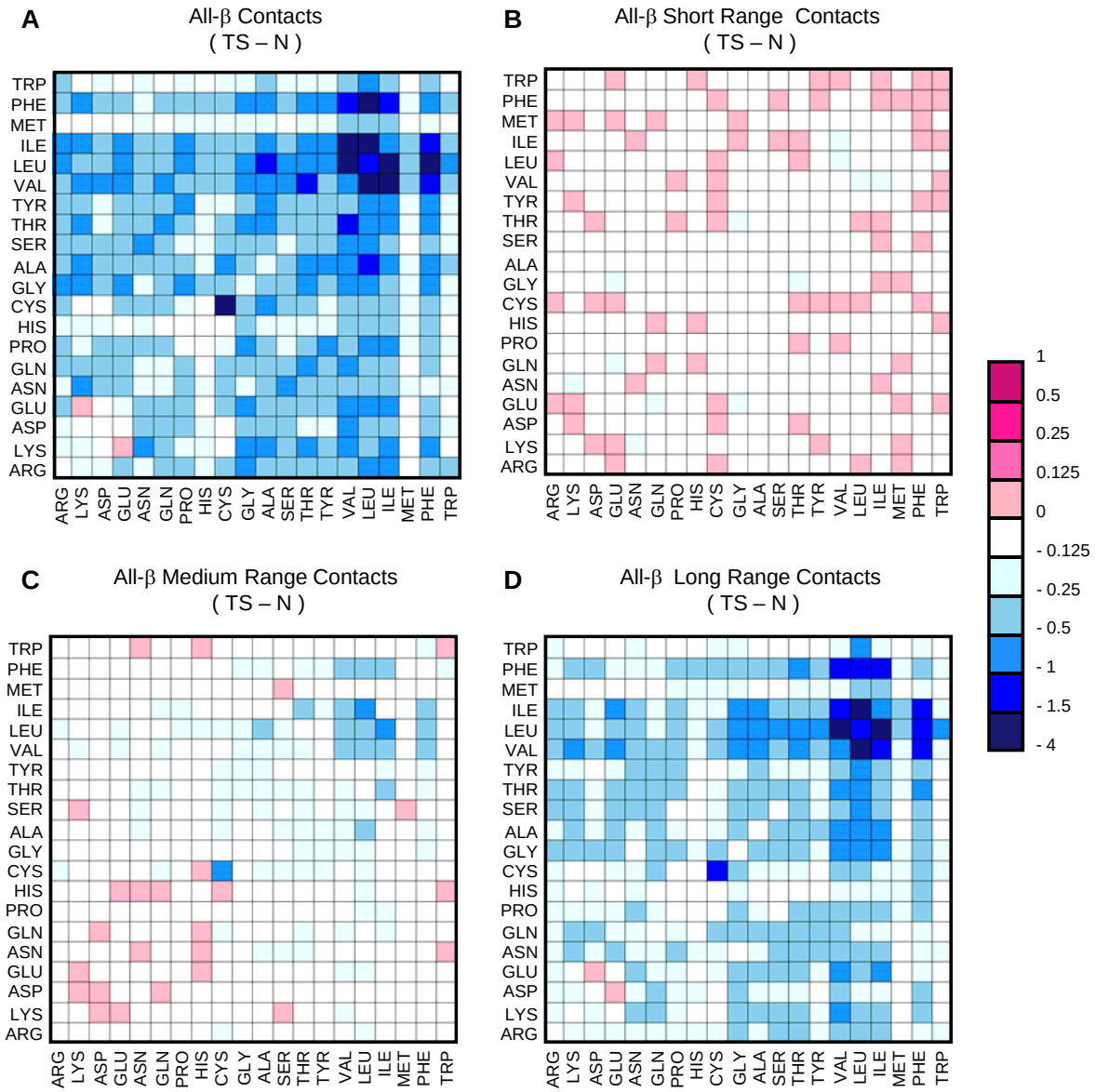
A
All-β Contacts
(TS – N)
B
All-β Short Range Contacts
(TS – N)
C
All-β Medium Range Contacts
(TS – N)
D
All-β Long Range Contacts
(TS – N)
TRP
PHE
MET
ILE
LEU
VAL
TYR
THR
SER
ALA
GLY
CYS
HIS
PRO
GLN
ASN
GLU
ASP
LYS
ARG
1
0.5
0.25
0.125
0
- 0.125
- 0.25
- 0.5
- 1
- 1.5
- 4

Figure S2

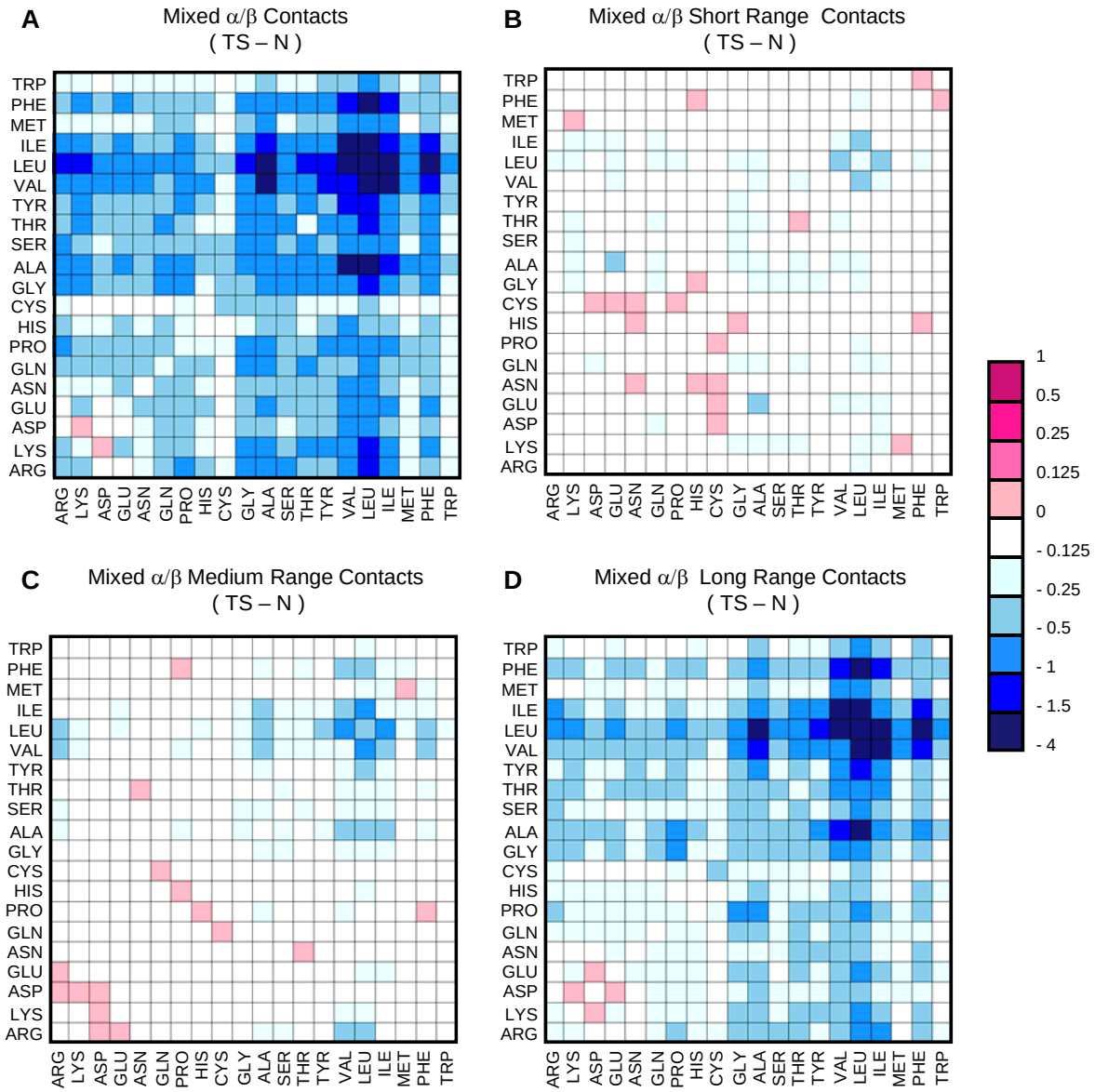
A
Mixed α/β Contacts
(TS – N)
B
Mixed α/β Short Range Contacts
(TS – N)
C
Mixed α/β Medium Range Contacts
(TS – N)
D
Mixed α/β Long Range Contacts
(TS – N)
TRP PHE MET ILE LEU VAL TYR THR SER ALA GLY CYS HIS PRO GLN ASN GLU ASP LYS ARG
ARG LYS ASP GLU ASN GLN PRO HIS CYS GLY ALA SER THR TYR VAL LEU ILE MET PHE TRP
1
0.5
0.25
0.125
0
- 0.125
- 0.25
- 0.5
- 1
- 1.5
- 4

Figure S3