Empirical Parameters for Estimating Protein-Protein Binding Energies: Number of Short- and Long-Distance Atom-Atom Contacts

Yong Chao Li and Zong-Hao Zeng*

Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

Abstract: The number of atom-atom contacts in long distance can fit to the experimental binding energies in a dataset containing 151 experimental data with the correlation coefficient about 0.68. Based on this factor, a set of distance-dependent empirical potentials for various types of short-distance (2.4 Å-5 Å) contacts was obtained by guided fitting, i.e. a set of two parameters fitting. Incorporation of these short-distance potentials improved the correlation coefficients to 0.881.

INTRODUCTION

How to estimating the free energy change in protein-protein binding is an important problem to understand protein's functions. Chothia and Janin [1] found that the binding energies are directly related to the buried surface areas in complex forming. The importance of hydrophobic interactions was emphasized [2-3], while the importance of polarity and charge must also be considered[4]. In recent investigations, the binding energies are always decomposed into various terms, contributed from static electronic energies, desolvation energy, confomational entropy [5-6], or hydrophobic, hydrophilic surfaces and atomic pairs [7-8]. Based on our recently collected dataset of protein-protein binding energies, here we present an empirical method for estimating the protein-protein binding energies. First we found that the numbers of atom-atom contacts, *TAC*, in a longer distance (2.4 Å -14 Å) are a good zeroth order estimation for binding energies, then we get a set of distance dependent potentials for short-distance atom-atom pairs by guided fitting: a series of two parameter linear fittings.

MATERIALS

All of the protein-protein (peptide) complexes and the experimental affinities are collected by web searching and checked by original literatures. For complexes, which are recorded more than once in the asymmetric unit of a crystal or in different crystals, average values of the multiple interfaces are used. Altogether, the dataset contains 151 binding energies (see Supplemental Materials for details).

METHODS AND RESULTS

The total number of atom-atom contacts, *TAC*, is the number of atom-atom pairs cross the interface with atom-atom distances in range 2.4 Å ~ r_e, or in a "shell" from r_e - 0.6 Å ~ r_e + 0.4 Å, for a specific value of r_e. The results of linear fitting of *TAC* to binding energies are listed in Table **1**. It is worth noting that when r_e > 9 Å, the correlation coefficients of *TAC* with experimental binding energies, ΔG, reach higher values, but when r_e > 14 Å the correlation begin to drop down. If we regard the atom-atom contacts with distances in the range 2.4 Å ~ 14 Å as long-distance contacts, then we can say that *TAC* in long-distance of protein-protein complexes play a major role in determining their binding energies.

If we regard the fitting result of *TAC* as a good zeroth order approximation to ΔG, then we can try to get higher order approximations by guided fitting in the following way.

Firstly, the protein atoms are classified into 15 types (Table **2**), as in CHARMM force field [9], therefore there are 120 types of atom-atom contacts. Denote $AC_n(r)$ as the number of atom-atom contacts of type *n* within the distance from *r* to *r* + 0.1 Å, then the short-distance-dependent potential $SDP_n(r)$ is obtained as the parameter *b*, in the linear two parameters fitting of $a \cdot TAC + b \cdot AC_n(r)$ with the experimental binding energies ΔG (In all the fittings, the range of atom-atom distance is fixed to 2.4 ~ 14 Å for calculations of TAC). Then, for a specific complex, the contributions to the binding energy from long-distance contacts are supposed to be proportional to *TAC*, while contributions from short-distance atom-atom contacts (SDAC) are supposed to be proportional to

$$SDAC = \sum_{n=1}^{120} \sum_{r=2.4}^{r_{end}} SDP_n(r) \cdot AC_n(r)\ .$$

Finally, by methods of two parameters linear fitting, the predicted binding energy can be expressed as

$$\Delta G_{predict} = \alpha \cdot (TAC + 160 \cdot SDAC) + 4.70,$$

where α = 0.000173 kcal/mol. This expression is obtained with the highest correlation coefficient of 0.881, when the *SDAC*'s are calculated with r_{end} = 5.0 Å (Fig. **1**).

*Address correspondence to this author at the Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China; Tel: 86-10-64855592; Fax: 86-10-64871293; E-mail: zzh@ibp.ac.cn

0929-8665/08 $55.00+.00

The Protein Complex Database Containing 151 Protein-Protein Complexes

Protein Complex	PDB code	Interface(s)	Lengths	Type#	k_a (M)	ΔG & kcal/M	Ref $
Chym-Omtky3	1cho	E/I	238-53	P-Pi	1.47×10^{-11}	14.8	1
Chymotrypsin-Eglin	1acb	E/I	241-63	P-Pi	1.49×10^{-12}	16.1	2
Metalloproteinase-inhibitor	1jiw	P/I	470-105	P-Pi	4×10^{-12}	15.5	3
Protease-CA_p2	1bai	AB/C	248-6	P-Pi	2×10^{-8}	10.5	4
Protease-tripeptide	1a30	AB/C	201-3	P-Pi	5×10^{-5}	5.86	5
Proteinase-eglin	1cse, 2sec	E/I	274-63	P-Pi	1×10^{-10}	13.6	6
Proteinase-inhibitor	2sic	E/I	275-107	P-Pi	1×10^{-10}	12.3	7
Proteinase-inhibitor	4sgb	E/I	185-51	P-Pi	4×10^{-9}	11.3	8
Proteinase-peptide	2er6	E/I	330-7	P-Pi	5.95×10^{-8}	9.85	9
SGPB-OMTKY	3sgb	E/I	185-50	P-Pi	1.78×10^{-11}	14.7	10
Thrombin-BPTI	1bth	LH/P, JK/Q	285-58	P-Pi	2.4×10^{-8}	10.4	11
Thrombin-hirudin	1awf*	LH/I	278-9	P-Pi	2.5×10^{-7}	9.00	12
Thrombin-Rhodniin	1tbq*	JK/S, LH/R	308-103	P-Pi	2×10^{-13}	17.3	13
Trypsin-BPTI	2ptc, 3btk	E/I	223-58	P-Pi	1.6×10^{-14}	18.8	14
Trypsinogen-BPTI	3tgk	E/I	217-56	P-Pi	1.2×10^{-5}	6.71	15
Trypsinogen-BPTI	1fy8	E/I	215-56	P-Pi	9×10^{-6}	6.88	16
Thermitase-Eglin_C	1tec	E/I	279-63	P-Pi	5×10^{-11}	14.0	17
Kallikrein_A-BPTI	2kai	AB/I	232-57	P-Pi	1.5×10^{-9}	12.4	18
Trypsinogen-PTI	3tpi	Z/I	223-58	P-Pi	1.8×10^{-6}	7.8	19
Anhydrotrypsin-BPTI	1tpa	E/I	223-58	P-Pi	1.1×10^{-13}	17.8	20
THROMBIN-HIRULOG3	1abi	LH/I	285-19	P-Pi	1×10^{-9}	11.6	21
trypsin-INHIBITOR	1avw	A/B	223-171	P-Pi	1×10^{-9}	12.3	22
HIV-1 protease-MVT-101	4hvp	AB/I	194-6	P-Pi	7.8×10^{-7}	8.33	23
Thermolysin-VM	3tmn	E/I	316-2	P-Pi	1.26×10^{-6}	8.04	24
Amylase-tendamistat	1bvn	P/T	496-71	E-Ei	9×10^{-12}	15.1	25
Angiogenin-inhibitor	1a4y	A/B, D/E	460-123	E-Ei	7.1×10^{-16}	20.7	26
Anhydrase-CAB_CA05	1g6v	A/K	256-126	E-Ei	7.2×10^{-8}	9.73	27
Barnase-barstar	1brs*	A/D, B/E, C/F	110-86	E-Ei	2×10^{-13}	17.3	28
FYN-KINASE	1a0n	A/B	58-14	E-Ei	3×10^{-6}	7.53	29
FYN-peptide	1azg	A/B	58-14	E-Ei	1.6×10^{-5}	6.54	29
Kinase-peptide	1sps	A/D, B/E, C/F	103-7	E-Ei	2×10^{-7}	9.13	30
Kinase-peptide	1bbz	A/B, C/D, E/F, G/H	58-10	E-Ei	1.5×10^{-6}	7.94	31
Kinase-peptide	1lcj	A/B	104-11	E-Ei	6.3×10^{-9}	11.2	32
Kinase-PKI	1fmo, 2erz	E/I	336-20	E-Ei	2.3×10^{-9}	11.8	33
Lactamase-BLIP	1jtg	A/B, C/D	262-165	E-Ei	1.1×10^{-10}	13.6	34

Lactamase-inhibitor	1jtd	A/B	273-262	E-Ei	2.72×10^{-11}	14.4	35
Nef-Fyn	1efn	A/B, C/D	104-57	E-Ei	3.8×10^{-7}	8.75	36
Peroxidase-Cytochrome_C	2pcc	A/B, C/D	294-108	E-Ei	5×10^{-8}	9.95	37
Ribonuclease-inhibitor	1dfj	E/I	456-124	E-Ei	5.9×10^{-14}	18.0	38
Papain-Stefin_B	1stf	E/I	212-98	E-Ei	1.2×10^{-10}	13.5	39
kinase-III_GLC	1gla	G/F	489-161	E-Ei	1.0×10^{-5}	6.7	40
Dehydrogenase-Amicyanin	1mda	LH/A JM/B	489-103	E-Ei	4.5×10^{-6}	7.3	41
kinase-peptide	1lck	A/B	164-9	E-Ei	2.92×10^{-6}	7.0	42
Nef_SH3-Fyn	1avz	B/C	103-57	E-Ei	2.0×10^{-5}	6.4	36
AChE-fasciculin_II	1fss	A/B	532-61	E-Ei	1.1×10^{-11}	14.9	43
Phospholipase_sh2-peptide	2pld	A/B	105-12	E-Ei	3.0×10^{-7}	9.0	42
Deoxyribonuclease_I-actin	1atn	A/D	373-258	E-Ei	2.0×10^{-9}	11.8	44
Penicillopepsin-IvaVVLySta	1apt	E/I	323-4	E-Ei	3.98×10^{-10}	12.8	24
Penicillopepsin-IvaVVSta	1apu	E/I	323-4	E-Ei	2.0×10^{-8}	10.5	24
CPA-GY	3cpa	A/S	307-2	E-Ei	9.5×10^{-5}	5.48	45
Thrombin-Hirudin	4htc	HL/I	291-61	E-Ei	8.0×10^{-11}	15.4	46
5G9-TF	1ahw	AB/C, DE/F	428-200	Ab-Ag	3.4×10^{-9}	11.5	47
Antibody-Lysozyme	1g7j	AB/C	348-129	Ab-Ag	8.7×10^{-7}	8.29	48
Antibody-Lysozyme	1g7m	AB/C	352-129	Ab-Ag	4×10^{-6}	7.36	48
Antibody-Lysozyme	1g7l	AB/C	346-129	Ab-Ag	3.4×10^{-6}	7.45	48
Antibody-Lysozyme	1g7h	AB/C	223-129	Ab-Ag	3×10^{-6}	7.52	48
Antibody-Lysozyme	1g7i	AB/C	223-129	Ab-Ag	1.75×10^{-7}	9.21	48
AP50-peptide	1h6e	A/P	219-10	Ab-Ag	7×10^{-7}	8.39	49
CAB_RN05-Rnase	1bzq	A/L, B/N, C/M, D/K	124-124	Ab-Ag	3.5×10^{-8}	10.2	50
D11.15-Lysozyme	1jhl	LH/A	224-129	Ab-Ag	1.5×10^{-8}	10.7	51
D1.3-Hel	1a2y	AB/C	223-129	Ab-Ag	2.2×10^{-8}	10.4	52
D1.3-Lysozyme	1fdl	LH/Y	432-129	Ab-Ag	3.7×10^{-9}	11.5	53
D1.3-Lysozyme	1vfb	AB/C	223-129	Ab-Ag	5×10^{-9}	11.3	54
D3H44-TF	1jps	LH/T	426-200	Ab-Ag	1.13×10^{-10}	13.6	55
D44.1-F10.6.6	1p2c	AB/C, DE/F	422-129	Ab-Ag	8.3×10^{-9}	13.7	56
FCRIII-IG_GAMMA_1	1iix	AB/C	423-167	Ab-Ag	9×10^{-7}	8.24	57
Fv(L-S91A)-HEL	1j1p	LH/Y	221-129	Ab-Ag	2.75×10^{-9}	11.7	58
Fv(L-S93A)-HEL	1j1x	LH/Y	221-129	Ab-Ag	1.35×10^{-9}	12.1	58
Fv(L-Y50F)-HEL	1j1o	LH/Y	221-129	Ab-Ag	8.9×10^{-9}	11.0	58
Fv-Lysozyme	1ic4	LH/Y	221-129	Ab-Ag	1.0×10^{-8}	10.9	59
Hyhel_63-Lysozyme	1dqj	AB/C	424-129	Ab-Ag	2.77×10^{-9}	11.7	60
HyHEL-Lysozyme	3hfm	LH/Y	429-129	Ab-Ag	6.67×10^{-10}	12.5	61

IgG-ProteinG	1igc	LH/A	435-58	Ab-Ag	5×10^{-11}	14.0	62
Lysozyme-CAB_LYS3	1xfp	A/L	131-129	Ab-Ag	1×10^{-7}	9.54	63
NC10-Neuraminidase	1nmb	LH/N	388-231	Ab-Ag	5×10^{-8}	9.95	64
tAb2-e_pep	1e4x	HL/P, IM/Q	430-7	Ab-Ag	2.5×10^{-8}	10.4	65
tAb2-hc_pep2	1e4w	HL/P	427-7	Ab-Ag	2.5×10^{-8}	10.4	65
Jel42-HPR	2jel	LH/P	435-85	Ag-Ab	2.8×10^{-9}	11.5	66
D44.1-Lysozyme	1mlc	AB/E CD/F	432-129	Ag-Ab	7.69×10^{-8}	9.7	67
E8-Cyt_c	1wej	LH/F	437-104	Ag-Ab	1.0×10^{-7}	9.5	68
Cab_lys3-Lysozyme	1mel	A/L B/M	132-127	Ag-Ab	2.0×10^{-8}	10.5	69
HyHel_5-lysozyme	1yqv	LH/Y	426-129	Ag-Ab	2.5×10^{-11}	14.5	70
N10-Nuclease	1nsn	LH/S	427-138	Ag-Ab	1.0×10^{-10}	13.3	71
Adaptor-peptide	1tce	A/B	107-13	Misc	5×10^{-5}	5.86	72
BCLxl-BAK	1bxl	A/B	181-16	Misc	3.4×10^{-7}	8.82	73
Bungarotoxin-receptor	1kl8, 1kc4	A/B	74-19	Misc	3×10^{-5}	6.17	74
Bungarotoxin-receptor	1l4w, 1ljz	A/B	74-25	Misc	3×10^{-9}	11.6	75
Calmodulin-CaMKK	1ckk	A/B	148-26	Misc	1×10^{-9}	12.3	76
Calmodulin-peptide	1mxe	A/E, B/F	144-25	Misc	1×10^{-12}	16.4	77
CaM-peptide	1iwq	A/B	139-18	Misc	8.8×10^{-9}	11.0	78
CATENIN-TCF4	1jpw	A/D, B/E, C/F	502-24	Misc	3.33×10^{-9}	11.6	79
C_CRK-peptide	1cka	A/B	57-9	Misc	1.9×10^{-6}	7.80	80
C_CRK-peptide	1ckb	A/B	57-8	Misc	5.2×10^{-6}	7.20	80
CHEA-CHEY	1a0o, 1ffg	A/B, C/D, E/F G/H	128-70	Misc	2.0×10^{-6}	7.77	81
Csk-peptide	1jeg	A/B	60-18	Misc	8.0×10^{-7}	8.31	82
Cue-Ubiquitin	1otr	A/B	76-49	Misc	1.55×10^{-4}	5.19	83
DNAK-peptide	1dkz, 1dky	A/B	215-7	Misc	2.0×10^{-7}	9.13	84
EGFR-peptide	1ff1	A/B	95-6	Misc	5.6×10^{-4}	4.43	85
ERBIN-peptide	1mfg	A/B	95-9	Misc	5×10^{-5}	5.86	86
EVH1-peptide	1evh	A/B	111-5	Misc	6.02×10^{-4}	4.39	87
FHA-hNIFK	2aff	A/B	98-37	Misc	7.7×10^{-8}	9.70	88
GRB2L-SLP_76	1h3h	A/B	60-9	Misc	2.2×10^{-7}	9.08	89
GROEL-peptide	1dkd	A/E, B/F, C/G, D/H	146-11	Misc	2.0×10^{-6}	7.77	90
Hormone-Receptor	1a22	A/B	192-180	Misc	3.4×10^{-10}	12.9	91
IGE-E131	1rpq	A/W, C/Y, B/X, D/Z	165-21	Misc	3.5×10^{-8}	10.2	92
KAP95P-NUP1P	2bpt	A/B	860-29	Misc	7.9×10^{-9}	11.0	93
MBP-ANKYRIN	1svx	A/B	369-157	Misc	4.4×10^{-9}	11.4	94
NIDOGEN-PERLECAN	1gl4	A/B	273-89	Misc	3.6×10^{-8}	10.1	95
OppA-KAK	1jet	A/B	517-3	Misc	5.6×10^{-8}	9.89	96

OppA-KCK	1b05	A/B	517-3	Misc	7.5×10^{-8}	9.71	96
OppA-KDK	1b4z	A/B	517-3	Misc	5.9×10^{-6}	7.13	96
OppA-KEK	1jeu	A/B	517-3	Misc	1.5×10^{-7}	9.30	96
OppA-KFK	1b40	A/B	517-3	Misc	5.3×10^{-8}	9.92	96
OppA-KGK	1b3l	A/B, C/D	517-3	Misc	1.3×10^{-6}	8.02	96
OppA-KHK	1b3f	A/B	517-3	Misc	1.3×10^{-7}	9.39	96
OppA-KIK	1b3g	A/B	517-3	Misc	2.0×10^{-7}	9.13	96
OppA-KKK	2olb	A/B	517-3	Misc	2.9×10^{-6}	7.55	96
OppA-KLK	1b9j	A/B	517-3	Misc	1.1×10^{-6}	8.12	96
OppA-KMK	1b32	A/B	517-3	Misc	7.9×10^{-8}	9.68	96
OppA-KNK	1b5i	A/B	517-3	Misc	9.0×10^{-8}	9.60	96
OppA-KPK	1b46	A/B	517-3	Misc	5.2×10^{-6}	7.20	96
OppA-KQK	1b5j	A/B	517-3	Misc	3.7×10^{-8}	10.1	96
OppA-KRK	1qka	A/B	517-3	Misc	1.2×10^{-6}	8.07	96
OppA-KSK	1b51	A/B	517-3	Misc	4.3×10^{-8}	10.0	96
OppA-KTK	1b52	A/B	515-3	Misc	7.6×10^{-8}	9.70	96
OppA-KVK	1qkb	A/B	517-3	Misc	4.5×10^{-8}	10.0	96
OppA-KWK	1jev	A/B	517-3	Misc	1.3×10^{-7}	9.39	96
OppA-KYK	1b58	A/B	517-3	Misc	2.6×10^{-7}	8.98	96
OppA-VKPG	1ola	A/B	517-4	Misc	9.94×10^{-8}	9.55	97
P53-53BP2	1ycs	A/B	193-191	Misc	3.0×10^{-8}	10.3	98
P60_SRC-APP12	1qwe	A/B	56-12	Misc	1.2×10^{-6}	8.07	99
P60_SRC-VSL12	1qwf	A/B	56-12	Misc	4.5×10^{-7}	8.65	99
PI3Kr-Ras.GMPPNP	1he8	A/B	749-166	Misc	1.3×10^{-6}	8.02	100
PLCr1-SLP76	1ywo	A/P	55-10	Misc	1.29×10^{-5}	6.67	101
PTB-peptide	1aqc	A/C, B/D	121-10	Misc	3.2×10^{-7}	8.85	102
PTB-peptide	2nmb	A/B	147-6	Misc	5.3×10^{-7}	8.56	103
RAP1A-RAF1	1gua	A/B	167-76	Misc	4.0×10^{-8}	10.1	104
RGS4-GI_ALPHA_1	1agr	A/E, D/H	350-128	Misc	1.0×10^{-7}	9.54	105
S100B-TRTK_12	1mwn	A/X, B/Y	91-12	Misc	2.5×10^{-7}	9.13	106
Streptavidin-Fshpqnt	1vwa	B/M, D/P	123-6	Misc	1.25×10^{-4}	5.32	107
Streptavidin-peptide	1rsu	B/P	123-8	Misc	7.24×10^{-5}	5.64	108
Streptavidin-peptide	1rst	B/P	123-9	Misc	3.68×10^{-5}	6.04	108
Trypsinogen_PTI-Ile_Val	2tpi	ZI/S	220-2	Misc	1.49×10^{-4}	5.22	109
His binding protein-H	1hsl	AC, B/D	238-1	Misc	6.4×10^{-8}	9.81	110
Cyclophilin_A-Capsid	1ak4	A/D, B/C	165-145	Misc	1.7×10^{-5}	6.5	111
SSI	3ssi	A/B	108-108	Homodimer	1.0×10^{-13}	17.4	112

Hyhel-Lysozyme	1bql	LH/Y	426-129	Ab-Ag		14.5	113
Hemoglobin-Hemoglobin	1hbs	ABCD/EFGH	574-574	Misc		4.8	114
Hormone-Receptor	1hwg	A/C	191-184	Misc		13.0	115
Subtilisin-inhibitor	2sni	E/I	275-64	P-Pi		15.8	116
Trypsinogen-Inhibitor	4tpi	Z/I	223-58	P-Pi		17.7	116

* The following complexes have more than two crystals,
a. 1awf, 1ny2, 1qhr, 1qj1, 1qj6, 1qj7, 1uma and 1way,
b. 1brs, 1b27, 1bgs, 1x1u, 1x1w, and 1x1x
c. 1fmo, 1atp, 2cpk and 2erz
d. 1tbq, 1tbr, and 1toc
P-Pi, protease-protease inhibitor;
E-Ei, enzyme-enzyme inhibitor;
Ab-Ag, antibody-antigen;
Misc, miscellaneous complexes.
& $\Delta G = -0.592\ln(k_a)$, the stabilization energy for binding.

$ REFERENCES

[1] Lu, W., Qasim, M. A., Laskowski, M. Jr and Kent, S. B. (**1997**) *Biochemistry*, *36,* 673-679.
[2] Qasim, M. A., Ganz, P. J., Saunders, C. W., Bateman, K. S., James, M. N., Laskowski, M. Jr. (**1997**) *Biochemistry, 36,* 1598-1607.
[3] Hege, T., Feltzer, R. E., Gray, R. D. and Baumann, U. (**2001**) *J. Biol. Chem., 276,* 35087-35092.
[4] Wu, J., Adomat, J. M., Ridky, T. W., Louis, J. M., Leis, J., Harrison, R. W. and Weber, I. T. (**1998**) *Biochemistry, 37,* 4518-4526.
[5] Louis, J. M., Dyda, F., Nashed, N. T., Kimmel, A. R. and Davies, D. R. (**1998**) *Biochemistry, 37,* 2105-2110.
[6] Seemuller,U. and Fritz,H. (**1981**). *Methods Enzymol., 80,* 804-816.
[7] Uehara, Y., Tonomura, B. and Hiromi, K. (**1978**). *J. Biochem. (Tokyo), 84,* 1195-1202.
[8] Hass, G. M. and Ryan, C. A. (**1981**) *Methods Enzymol., 80,* 778–791.
[9] Foundling, S. I., Cooper, J., Watson, F. E., Cleasby, A., Pearl, L. H., Sibanda, B. L., Hemmings, A., Wood, S. P., Blundell, T. L., Valler, M. J., *et al.* (**1987**) *Nature, 327,* 349-352.
[10] Lu, W., Apostol, I., Qasim, M. A., Warne, N., Wynn, R., Zhang, W. L., Anderson, S., Chiang, Y. W., Ogin, E., Rothberg, I. *et. al.* (**1997**) *J. Mol. Biol. 266,* 441-461.
[11] van de Locht, A., Bode, W., Huber, R., Le Bonniec, B. F., Stone, S. R., Esmon, C. T. and Stubbs, M. T. (**1997**) *EMBO J. 16,* 2977-2984.
[12] Malikayil, J. A., Burkhart, J. P., Schreuder, H. A., Broersma, R. J. Jr, Tardif, C., Kutcher, L. W. 3rd, Mehdi, S., Schatzman, G. L., Neises, B. and Peet, N. P. (**1997**) *Biochemistry. 36,* 1034-1040.
[13] Friedrich, T., Kroger, B., Bialojan, S., Lemaire, H. G., Hoffken, H. W., Reuschenbach, P., Otte, M. and Dodt, J.(**1993**) *J. Biol. Chem., 268,* 16216-16222.
[14] Vincent, J. P. and Lazdunski, M. (**1972**) *Biochemistry***.,** *11,* 2967-2977.
[15] Pasternak, A., White, A., Jeffery, C. J., Medina, N., Cahoon, M., Ringe, D. and Hedstrom, L. (**2001**) *Protein Sci.,* **10**,1331-1342.
[16] Pasternak, A., Liu, X., Lin, T. Y. and Hedstrom, L. (**1998**) *Biochemistry, 37,* 16201-16210.
[17] Seemuller,U. and Fritz,H. (**1981**) *Methods Enzymol., 80,* 804-816.
[18] Dietl, T., Huber, C., Geiger, R., Iwanaga, S. and Fritz, H. (**1979**) *Hoppe-Seyler. Z. Physiol. Chem., 360,* 67-71.
[19] Vincent, J. P., Lazdunski, M. (**1976**) *FEBS Lett., 63,* 240-244.
[20] Vincent, J. P., Peron-Renner, M., Pudles, J. and Lazdunski, M. (**1974**) *Biochemistry, 13,* 4205-4211.
[21] Qiu, X., Padmanabhan, K. P., Carperos, V. E., Tulinsky, A., Kline, T., Maraganore, J. M., Fenton, J. W. (**1992**) *Biochemistry, 31,* 11689-11697.
[22] Baugh, R. J. and Trowbridge, C. G. (**1972**) *J. Biol. Chem., 247,* 7498-7501.
[23] Miller, M., Schneider, J., Sathyanarayana, B. K., Toth, M. V., Marshall, G. R., Clawson, L., Selk, L., Kent, S. B. and Wlodawer, A. (**1989**) *Science,* **246**(4934), 1149 - 1152
[24] Head, R. D., Smythe, M. L., Oprea, T. I., Waller, C. L., Stuart M.Green, S. M. and Marshall, G. R. (**1996**) *J. Am. Chem. Soc., 118,* 3959-3969.
[25] Vertesy, L., Oeding, V., Bender, R., Zepf, K. and Nesemann, G. (**1984**) *Eur. J. Biochem., 141,* 505-512.
[26] Lee, F. S., Shapiro, R. and Vallee, B. L. (**1989**) *Biochemistry, 28,* 225-230.
[27] Desmyter, A., Decanniere, K., Muyldermans, S. and Wynsm L. (**2001**) *J. Biol. Chem., 276,* 26285-26290.
[28] Hartley, R. W. (**1993**) *Biochemistry., 32,* 5978-5984.
[29] Renzoni, D. A., Pugh, D. J., Siligardi, G., Das, P., Morton, C. J., Rossi, C., Waterfield, M. D., Campbell, I. D. and Ladbury, J. E. (**1996**) *Biochemistry, 35,* 15646-15653.
[30] Bradshaw, J. M., Grucza, R. A., Ladbury, J. E. and Waksman, G. (**1998**) *Biochemistry, 37,* 9083-9090.
[31] Pisabarro, M. T., Serrano, L. and Wilmanns, M. (**1998**) *J. Mol. Biol., 281,* 513-521.
[32] Payne, G., Shoelson, S. E., Gish, G. D., Pawson, T. and Walsh, C. T. (**1993**) *Proc. Natl. Acad. Sci. USA, 90,* 4902-4906.
[33] Glass, D. B., Cheng, H. C., Mende-Mueller, L. Reed, J. and Walsh, D. A. (**1989**) *J. Biol. Chem., 264,* 8802-8810.
[34] Petrosino, J., Rudqers, G., Gilbert, H. and Palzkill, T. (**1999**) *J. Bilo. Chem., 274,* 2394-2400.
[35] Lim, D., Park, H. U., De Castro, L., Kang, S. G., Lee, H. S., Jensen, S., Lee, K. J. and Strynadka, N. C. (**2001**) *Nat. Struct. Biol., 8,* 838-852
[36] Lee, C. H., Leung, B., Lemmon, M.A., Zheng, J., Cowburn, D., Kuriyan, J. and Saksela, K. (**1995**) *EMBO J. 14,* 5006-5015
[37] Corin, A. F., McLendon, G., Zhang, Q., Hake, R. A., Falvo, J., Lu, K. S., Ciccarelli, R. B. and Holzschu, D. (**1991**) *Biochemistry, 30,* 11585-11595.
[38] Vicentini, A. M., Kieffer, B., Matthies, R., Meyhack, B., Hemmings, B. A., Stone, S. R. and Hofsteenge, J. (**1990**) *Biochemistry, 29,* 8827-8834.
[39] Turk, B., Krizaj, I., Kralj, B., Dolenc, I., Popovic, T., Bieth, J. and Turk, V. (**1993**) *J. Biol. Chem., 268,* 7323-7329.
[40] Novotny, M. J., Frederickson, W. L., Waygood, E. B. and Saier, M. H. Jr. (**1985**) *J. Bacteriol., 162,* 810-816.
[41] Davidson, V., Graichen, M., Jones, L. (**1993**) *Biochim. Biophys. Acta, 1144,* 39-45.
[42] Zhou, Y., Abagyan, A. R. (**1998**) *Fold. Des., 3,* 513-522.
[43] Eastman, J., Wilson, E., Cerveansky, C., Rosenberry, T. (**1995**). *J. Biol. Chem.,* **270**, 19694-19701.
[44] Mannherz, H., Goody, R., Konrad, M., Nowak, E. (**1980**) *Eur. J. Biochem., 104,* 367-379.
[45] Bunting, J. W. and Myers, C. D. (**1975**) *Can. J. Chem. Rev. Can. Chim., 53,* 1984-1992.
[46] Markwardt F. (**1970**). *Methods Enzymol.,* **19**, 924-932.
[47] Huang, M., Syed, R., Stura, E. A., Stone, M. J., Stefanko, R. S., Ruf, W., Edgington, T. S. and Wilson, I. A. (**1998**). *J. Mol. Biol., 275,* 873-894.

[48] Sundberg, E. J., Urrutia, M., Braden, B. C., Isern, J., Tsuchiya, D., Fields, B. A., Malchiodi, E. L., Tormo, J., Schwarz, F. P. and Mariuzza, R. A. (**2000**) *Biochemistry, 39*, 15375-15387.
[49] Follows, E. R., McPheat, J. C., Minshull, C., Moore, N. C., Pauptit, R. A., Rowsell, S., Stacey, C. L., Stanway, J. J., Taylor, I. W. and Abbott, W. M. (**2001**) *Biochem. J., 359*, 427-434.
[50] Decanniere, K., Desmyter, A., Lauwereys, M., Ghahroudi, M. A., Muyldermans, S. and Wyns, L. (**1999**) *Structure, 7*, 361-370.
[51] Chitarra, V., Alzari, P. M., Bentley, G. A., Bhat, T. N., Eisele, J. L., Houdusse, A., Lescar, J., Souchon, H. and Poljak, R. J. (**1993**) *Proc. Natl. Acad. Sci. USA, 90*, 7711-7715.
[52] Dall'Acqua, W., Goldman, E. R., Lin, W., Teng, C., Tsuchiya, D., Li, H., Ysern, X., Braden, B. C., Li, Y., Smith-Gill, S. J. and Mariuzza, R. A. (**1998**) *Biochemistry, 37*, 7981-7991.
[53] Tello, D., Goldbaum, F. A., Mariuzza, R. A., Ysern, X., Schwarz, F. P. and Poljak, R. J. (**1993**) *Biochem. Soc. Trans., 21*, 943–946.
[54] Verhoeyen, M., Milstein, C. and Winter, G. (**1988**) *Science, 239*, 1534-1536
[55] Faelber, K., Kirchhofer, D., Presta, L., Kelley, R. F. and Muller, Y. A. (**2001**) *J. Mol. Biol., 313*, 83-97.
[56] Cauerhff, A., Goldbaum, F. A. and Braden, B. C. (**2004**) *Proc. Natl. Acad. Sci. USA, 101*, 3539-3544.
[57] Galon, J., Robertson, M. W., Galinha, A., Mazieres, N., Spagnoli, R., Fridman, W. H., Sautes, C. (**1997**) *Eur. J. Immunol., 27*, 1928-1932.
[58] Yokota, A., Tsumoto, K., Shiroishi, M., Kondo, H. and Kumagai, I. (**2003**) *J. Biol. Chem., 278*, 5410-5418.
[59] Shiroishi, M., Yokota, A., Tsumoto, K., Kondo, H., Nishimiya, Y., Horii, K., Matsushima, M., Ogasahara, K., Yutani, K. and Kumagai, I. (**2001**) *J. Biol. Chem., 276*, 23042-23050.
[60] Li, Y., Li, H., Smith-Gill, S. J. and Mariuzza, R. A. (**2000**) *Biochemistry, 39*, 6296-6309.
[61] Padlan, E. A., Silverton, E. W., Sheriff, S., Cohen, G. H., Smith-Gill, S. J. and Davies, D. R. (**1989**) *Proc. Natl. Acad. Sci. USA, 86*, 5938-5942.
[62] Sjobring, U., Bjorck, L. and Kastern, W. (**1991**) *J. Biol. Chem., 266*, 399-405.
[63] De Genst, E., Handelberg, F., Van Meirhaeghe, A., Vynck, S., Loris, R., Wyns, L. and Muyldermans, S. (**2004**) *J. Biol. Chem., 279*, 53593-53601.
[64] Gruen, L. C., McInerney, T. L., Webster, R. G. and Jackson, D, C. (**1993**) *J. Prot. Chem., 12*, 255-259.
[65] Hahn, M., Winkler, D., Welfle, K., Misselwitz, R., Welfle, H., Wessner, H., Zahn, G., Scholz, C., Seifert, M., Harkins, R., J Schneider-Mergener, J. and Hohne, W. (**2001**) *J. Mol. Biol., 314*, 293-309.
[66] Smallshaw, J. E., Brokx, S., Lee, J. S., Waygood, E. B. (**1998**) *J. Mol. Biol., 280*, 765-774.
[67] Schwarz, F. P., Tello, D., Goldbaum, F. A., Mariuzza, R. A. and Poljak, R. J. (**1995**) *Eur. J. Biochem., 228*, 388-394.
[68] Murphy, K. P., Freire, E. and Paterson, Y. (**1995**) *Proteins, 21*, 83-90.
[69] Desmyter, A., Transue, T. R., Ghahroudi, M. A., Thi, M. H., Poortmans, F., Hamers, R., Muyldermans, S. and Wyns, L. (**1996**) *Nat. Struct. Biol., 3*, 803-811.
[70] Hibbits, K. A., Gill, D.S. and Willson, R.C. (**1994**) *Biochemistry, 33*, 3584-3590.
[71] Smith, A. M. and Benjamin, D. C. (**1991**) *J. Immunol., 146*, 1259-1264.
[72] Zhou, M., Meadows, R. P., Logan, T. M., Yoon, H. S., Wade, W. S., Ravichandran, K. S., Burakoff, S. J. and Fesik, R. W. (**1995**) *Proc. Natl. Acad. Sci. USA, 92*, 7784-7788.
[73] Sattler, M., Liang, H., Nettesheim, D., Meadows, R. P., Harlan, J. E., Eberstadt, M., Yoon, H. S., Shuker, S. B., Chang, B. S., Minn, A. J., Thompson, C. B. and Fesik, S. W. (**1997**) *Science, 275*, 983-986.
[74] Moise, L., Piserchio, A., Basus, V. J. and Hawrot, E. (**2002**) *J. Biol. Chem., 277*, 12406-12417.
[75] Tzartos, S. J. and Changeux, J. P. (**1984**) *J. Biol. Chem., 259*, 11512-11519.
[76] Osawa, M., Tokumitsu, H., Swindells, M. B., Kurihara, H., Orita, M., Shibanuma,,T., Furuya, T. and Ikura, M. (**1999**) *Nat. Struct. Biol., 6*, 819-824.
[77] Clapperton, J. A., Martin, S. R., Smerdon, S. J., Gamblin, S. J. and Bayley, P. M. (**2002**) *Biochemistry*, **41**,14669-14679.
[78] Yamauchi, E., Nakatsu, T., Matsubara, M., Kato, H. and Taniguchi, H. (**2003**) *Nat. Struct. Biol., 10*, 226-231.
[79] Knapp, S., Zamai, M., Volpi, D., Nardese, V., Avanzi, N., Breton, J., Plyte, S., Flocco, M., Marconi, M., Isacchi, A. and Caiolfa, V. R. (**2001**) *J. Mol. Biol., 306*, 1179-1189.
[80] Wu, X., Knudsen, B., Feller, S. M., Zheng, J., Sali, A., Cowburn, D., Hanafusa, H. and Kuriyan, J. (**1995**) *Structure, 3*, 215-226.
[81] Li, J., Swanson, R. V., Simon, M. I. and Weis, R. M. (**1995**) *Biochemistry, 34*, 14626-14636.
[82] Ghose, R., Shekhtman, A., Goger, M. J., Ji, H. and Cowburn, D. (**2001**) *Nat. Struct. Biol., 8*, 998-1004.
[83] Kang, R. S., Daniels, C. M., Francis, S. A., Shih, S. C., Salerno, W. J., Hicke, L. and Radhakrishnan, I. (**2003**) *Cell, 113*, 621-630.
[84] Kasper, P., Christen, P. and Gehring, H. (**2000**) *Proteins, 40*, 185-192.
[85] de Beer, T., Carter, R. E., Lobel-Rice, K. E., Sorkin, A. and Overduin, M. (**1998**) *Science, 281*, 1357-1360.
[86] Birrane, G., Chung, J. and Ladias, J. A. (**2003**) *J. Biol. Chem., 278*, 1399-1402.
[87] Prehoda, K. E., Lee, D. J. and Lim, W. A. (**1999**) *Cell, 97*, 471-480.
[88] Byeon, I. J., Li, H., Song, H., Gronenborn, A. M., Tsai, M. D. (**2005**) *Nat. Struct. Mol. Biol., 12*, 987-993.
[89] Liu, Q., Berry, D., Nash, P., Pawson, T., McGlade, C. J. and Li, S. S. (**2003**) *Mol. Cell. Biol., 11*, 471-481.
[90] Chen, L. and Sigler, P. B. (**1999**) *Cell. 99*, 757-768.
[91] Clackson, T., Ultsch, M. H., Wells, J. A. and de Vos, A. M. (**1998**) *J. Mol. Biol., 277*, 1111-1128.
[92] Nakamura, G. R., Reynolds, M. E., Chen, Y. M., Starovasnik, M. A., Lowman, H. B. (**2002**) *Proc. Natl. Acad. Sci. USA, 99*, 1303-1308.
[93] Pyhtila, B. and Rexach, M. (**2003**) *J. Biol. Chem. 278*, 42699-42709.
[94] Binz, H. K., Amstutz, P., Kohl, A., Stumpp, M. T., Briand, C., Forrer, P., Grutter, M. G. and Pluckthun, A. (**2004**) *Nat. Biotechnol., 22*, 575-582.
[95] Hopf, M., Gohring, W., Mann, K. and Timpl, R. (**2001**) *J. Mol. Biol., 311*, 529-541.
[96] Sleigh, S. H., Seavers, P. R., Wilkinson, A. J., Ladbury, J. E. and Tame, J. R. (**1999**) *J. Mol. Biol., 291*, 393-415.
[97] Puvanendrampillai, D., Mitchell, J. B. (**2003**) *Bioinformatics, 19*, 1856-1857.
[98] Gorina, S. and Pavletich, N. P. (**1996**) *Science, 274*, 1001-1005.
[99] Feng, S., Kasahara, C., Rickles., R. J. and Schreiber, S. L. (**1995**) *Proc. Natl. Acad. Sci. USA, 92*, 12408-12415.
[100] Pacold, M. E., Suire, S., Perisic, O., Lara-Gonzalez, S., Davis, C. T., Walker, E. H., Hawkins, P. T., Stephens, L., Eccleston, J. F. and Williams, R. L. (**2000**) *Cell, 103*, 931-943.
[101] Deng, L., Velikovsky, C. A., Swaminathan, C. P., Cho, S. and Mariuzza, R. A. (**2005**) *J. Mol. Biol., 352*, 1-10.
[102] Zhang, Z., Lee, C. –H., Mandiyan, V., Borg, J. –P., Margolis, B., Schlessinger, J. and Kuriyan, J. (**1997**) *EMBO J., 16*, 6141-6150.
[103] Li, S. C., Zwahlen, C., Vincent, S. J., McGlade, C. J., Kay, L. E., Pawson, T. and Forman-Kay, J. D. (**1998**) *Nat. Struct. Biol., 5*, 1075-1083.
[104] Nassar, N., Horn, G., Herrmann, C., Block, C., Janknecht, R. and Wittinghofer, A. (**1996**) *Nat. Struct. Biol., 3*, 723-729.
[105] Tesmer, J. J., Berman, D. M., Gilman, A.G. and Sprang, S. R. (**1997**) *Cell, 89*, 251-261.
[106] Inman, K. G., Yang, R., Rustandi, R. R., Miller, K. E., Baldisseri, D. M. and Weber, D. J. (**2002**) *J. Mol. Biol., 324*, 1003-1014.
[107] Giebel, L. B., Cass, R. T., Milligan, D. L., Young, D. C., Arze, R. and Johnson, C. R. (**1995**) *Biochemistry, 34*, 15430-15435.
[108] Schmidt, T. G., Koepke, J., Frank, R. and Skerra, A. (**1996**) *J. Mol. Biol., 255*, 753-766.
[109] Bode, W. (**1979**) *J. Mol. Biol. 127*, 357-374.
[110] Yao, N., Trakhanov, S. and Quiocho, F. A. (**1994**) *Biochemistry, 33*, 4769-4779.
[111] Gamble, T. R., Vajdos, F. F., Yoo, S., Worthylake, D. K., Houseweart, M., Sundquist, W. I. and Hill, C. P. (**1996**) *Cell, 87*, 1285-1294.
[112] Akasaka, K., Fujii, S., Hayashi, F., Rokushika, S., and Hatano, H. (**1982**) *Biochem. Internat., 5*, 637–642.

[113] Stites, W. (**1997**) *Chem. Rev., 97,* 1233-1250.
[114] Ross, P. D., Hofrichter, J. and Eaton, W. A. (**1977**) *J. Mol. Biol., 115,* 111-134.
[115] Guinto, E. R., Ye, J., Bonniec, B. F. L. and Esmon, C. T. (**1994**) *J. Biol. Chem., 269,* 18395-18400.
[116] Wallqvist, A., Jernigan, R. L. and Covell, D. G. (**1995**) *Prot. Sci., 4,* 1881-1903.

Table 1. Correlation Coefficients of the Total Numbers of Atom-Atom Contacts TAC with Experimental Binding Energies ΔG

r_e (Å)	3	4	5	6	7	8	9
CC1	0.231	0.519	0.608	0.627	0.642	0.650	0.659
CC2	0.317	0.616	0.624	0.636	0.649	0.647	0.666
r_e (Å)	10	11	12	13	14	15	16
CC1	0.665	0.673	0.677	0.679	0.680	0.679	0.675
CC2	0.667	0.678	0.672	0.674	0.667	0.657	0.643

CC1: the correlation coefficients of *TAC* , from 2.4 ~ r_e , with ΔG
CC2: the correlation coefficients of *TAC* , from r_e - 0.6 ~ r_e + 0.4, with ΔG

Table 2. Atom Classification Scheme in the 15 Atom-Types Model

Atom-types	Atoms	Atom-types	Atoms
C	Planar C not connected to H	NH2	Amine group with two H
CH1E	Tetrahedron C with one H	NH3	NZ in LYS
CH2E	Tetrahedron C with two H	O	Carbonyl O
CH3E	Tetrahedron C with three H	OC	Carboxyl O
CR1E	Planar C with one H	OH1	hydroxyl O
N	N in PRO	SH1E	SG in CYS
NC2	NH1 & NH2 in ARG	S	SD in MET
NH1	Amine group with one H	Het	Others

DISCUSSIONS AND CONCLUSION

The total atom-atom contacts, either in a larger range, e.g from 2.4 Å to 14 Å, or in a shell, e.g from r_e - 0.6 Å ~ r_e + 0.4 Å with r_e = 10 Å, 12 Å, or 13 Å, display moderately high correlations to the experimental ΔG. This illustrates that the numbers of contacts in different shells represent a common quantity, which essentially reflects the extent of contacting in a complex. As has long been known that binding energies are related to the buried surface areas [1], we also have checked the latter's behavior in our dataset and found a slightly lower correlation of 0.62 (data not shown). This is in agreement with the above interpretation of *TAC* as reflecting the extent of contacting, but *TAC* contains more features, e.g. the tightness and complement of the interface.

As listed in Table **1**, the number of atom-atom contacts in short-distance alone does not fit to the experimental data, but its importance is exhibited after classification of the atom-atom contacts and trained by method of guided fitting. The guided fitting can be looked upon as a way of getting the expanding coefficients of the binding energies on terms of short-distance contacts. These expanding coefficients give the relative importance of contributions from various atom-atom types of short-distance contacts. As there are many types of atom-atom types, whereas still less than the number of experimental data, we feel that much more experimental data are needed to go deeper into the core of the problem. At present we can confidently say that the incorporation of the short-distance atom-atom contacts in a proper manner will greatly improve the estimation of protein-protein binding energies.

In conclusion, the analysis on 151 experimental protein-protein binding energies revealed that, the single quantity, i.e. the total number of atom-atom contacts in long-distance, either in a larger distance range or in a distance shell of shorter width, alone gives a moderately high correlation coefficient to the experimental binding energies. The contributions from short-distance contacts can be estimated by the

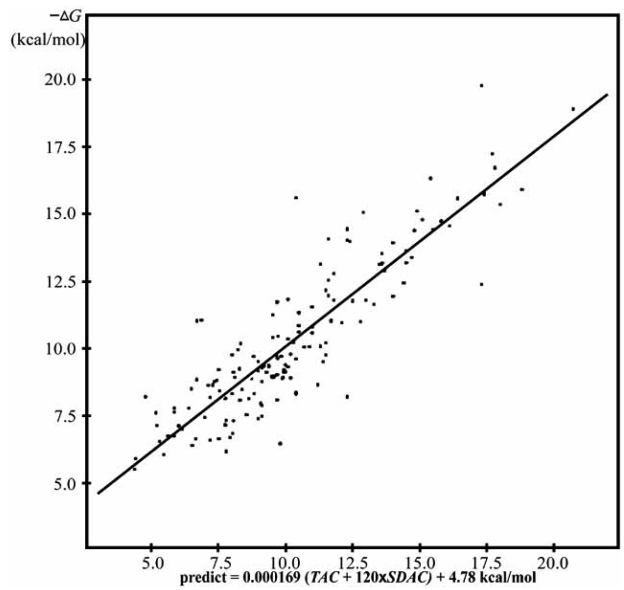

Fig. (1). Fitting of predicted binding energies, Δ*G*, to that of experiment with a correlation coefficient of 0.881.

Received: June 23, 2007 Revised: August 07, 2007 Accepted: August 08, 2007

method of guided fitting in the form of a set of distance-dependent potentials. Incorporation of the contributions from short distance contacts results in a correlation coefficient of 0.881, which greatly improved the fitting of atom-atom contacts to the experimental binding energies.

ACKNOWLEDGEMENTS

Work supported by "973" project 2004CB520801 of China.

REFERENCES

[1] Chothia, C. and Janin, J. (**1975**) *Nature, 256,* 705-708.
[2] Janin, J. and Chotothia, C. (**1978**) *Biochemistry*. *17,* 2943-2948.
[3] Goto, K. (**1995**) *Biochem. Biophys. Res. Commun*., *206,* 497-501.
[4] Brooijmans, N.; Sharp, K.A. and Kuntz, I.D. (**2002**) *Proteins: Structure, Function and Genetics, 48*, 645-653.
[5] Vajda, S.; Weng, Z.P. and DeLisi, C. (**1995**) *Protein Sci*., *8,* 1081-1092.
[6] Zhang, C.; Vasmatzis, G.; Cornette, J.L. and DeLisi, C. (**1997**) *J. Mol. Biol.*, *267*, 707-726.
[7] Xu, D.; Lin, S.L. and Nussinov, R. (**1997**) *J. Mol. Biol., 265,* 68-84.
[8] Ma, X.H.; Wang, C.X.; Li, C.H. and Chen, W.Z. (**2002**) *Prot. Eng.*, *15,* 677-81.
[9] Brooks, B.R.; Bruccoleri, R.E.; Olafson, B.D.; States, D.J.; Swaminathan, S. and Karplus, M. (**1983**) *J. Comp. Chem.*, *4*, 187-217.