

文章编号: 1000-5013(2012)01-0044-07

对嗜压蛋白氨基酸压力不对称指数的修订

张光亚

(华侨大学 化工学院, 福建 厦门 361021)

摘要: 选取嗜压菌最多的 4 种希瓦氏菌属(*Shewanella*)微生物共计 473 对同源蛋白,统计了 21 662 个氨基酸突变位点,在此基础上计算氨基酸压力不对称指数(PAI),并将其同氨基酸 70 种理化性质进行相关性分析.结果表明:色氨酸和异亮氨酸的嗜压性能最强,而脯氨酸、甘氨酸和天冬酰胺的嗜压性能最弱,这与之前的报道存在较大差异.此外,70 种氨基酸理化性质中有 13 种同 PAI 存在较强的相关性,其中仅有氨基酸的无规则倾向因子同 PAI 呈负相关,表明高的 PAI 值不易形成卷曲.由于展示的 PAI 值更适合于处于低温环境下的嗜压微生物,因此对以前的 PAI 值进行了修正.

关键词: 嗜压微生物; 压力不对称指数; 氨基酸; 倾向因子; 适压机制

中图分类号: Q 936

文献标志码: A

对多数海洋微生物尤其是深海微生物而言,在高压下生存尤为重要,相比于其他极端微生物,人们对嗜压微生物的关注较少^[1].同温度一样,压力也是严重影响生物体内分子体系的热动力学参数,但压力对生物分子的影响一直没有得到应有的重视^[2].近年来研究发现,一些蛋白质(或者酶)在高压条件下表现出良好的特性,如在最适温度下,生物分子的催化能力明显上升,酶活力提高了约 25%,或者其催化底物的特异性发生改变^[3-4].基因组学和蛋白质组学的迅猛发展,为了解嗜压微生物内生物分子体系提供了全新的视角.通过对一些已完成基因组测序的嗜压微生物基因组或蛋白质组的分析表明,它们在膜蛋白、呼吸链、DNA 损伤修复、核糖体结构及组装等方面都存在明显特点^[5].然而,有关嗜压蛋白质中氨基酸使用的偏好与其理化性质之间联系的研究很少,但在嗜热^[6]、嗜冷^[7]等极端蛋白质中却得到广泛研究. Di Giulio 首次利用蛋白质组信息,系统比较了 *Pyrococcus* 嗜压和非嗜压微生物的 141 个直系同源序列(orthologs),结果发现侧链较小精氨酸(Arg),甘氨酸(Gly),缬氨酸(Val)和天冬氨酸(Asp)在嗜压蛋白质中含量很高,而一些侧链较大的氨基酸如酪氨酸(Tyr)和色氨酸(Trp)则含量很低.在此基础上,他提出了氨基酸压力不对称指数(pressure asymmetry index, PAI),即 PAI 与氨基酸分子量呈反比,而同其极性呈正比^[8].然而,文献[9-10]的研究中却发现,嗜压蛋白中嗜压氨基酸并不一定含量更高,相反非嗜压氨基酸则明显上升.本文从 *Shewanella* 两对嗜压和非嗜压微生物中得到了 473 对同源蛋白,共统计了 185 621 个氨基酸,计算其氨基酸压力不对称指数,发现与 Di Giulio 的结果存在较大差异,并将其同氨基酸的 70 种理化性质进行了相关性分析,找出了与压力显著相关的氨基酸理化特性.

1 材料和方法

1.1 微生物的选取

目前发现的嗜压微生物多为细菌,且同时也为嗜冷菌,*Shewanella* 是目前分布于深海中最为广泛的嗜压微生物^[11].因此,共选取了 *Shewanella* 两对微生物全蛋白质组序列,它们分别为:*S. benthica* KT99 和 *S. frigidimarina* NCMIB400,以及 *S. violacea* DSS12 和 *S. sediminis* HAW-EB3,前者均为嗜压微生物,后者为均非嗜压^[12-13].

收稿日期: 2011-06-11

通信作者: 张光亚(1975-),男,副教授,主要从事酶与生物信息学的研究. E-mail: zhgyghh@hqu.edu.cn.

基金项目: 福建省自然科学基金资助项目(2007J0360); 华侨大学高层次人才科研启动项目(10BS220)

S. benthica KT99 为专性嗜压菌, 分离自海平面下 9 856 m 深处, 其最适生长温度和压力分别为 10 °C 和 50 MPa, 基因组 G+C 的摩尔分数为 46%; *S. violacea* DSS12 分离自海平面下 5 110 m 深处, 最适生长温度和压力分别为 8 °C 和 30 MPa, 基因组 G+C 的摩尔分数为 44.7%; *S. frigidimarina* NC-MIB400 为革兰氏阴性菌, 最适生长温度为 10 °C, 基因组 G+C 的摩尔分数为 41.6%, *S. sediminis* HAW-EB3 为兼性厌氧菌, 最适生长温度为 10 °C, 基因组 G+C 的摩尔分数为 46.1%。这 4 种微生物来源于同一个属, 亲缘关系较近; 其最适生长温度接近, 基因组 G+C 的摩尔分数也很接近, 从而最大程度的降低了其他因素对同源蛋白氨基酸选择的影响, 使生长压力成为影响氨基酸使用偏好的最重要因素。

1.2 同源序列的获取及序列比对

以上 4 种微生物全蛋白质组序列来源于 UniProtKB(<http://www.uniprot.org/>)^[14]。为了减少信息冗余, 剔除了所有长度小于 100 个氨基酸的序列, 因为它们大多为片断(fragment)或者部分(partial)。采用 BlastP 程序^[15]选取直系同源序列(orthologs), 其 *E* 值为 1.0×10^{-10} , 并找出 90% 长度覆盖范围内, 同一性(identity)大于 30% 的同源蛋白序列。样本 1 包含 222 对同源序列, 样本 2 包含 251 对, 共计 473 对, 946 条。该样本量足够用于探测嗜压蛋白在氨基酸使用上的某种趋势。

获得同源序列后, 采用 Clustal X 软件^[16]将 473 对同源蛋白两两进行比对^[8], 以统计同源蛋白在某些位点氨基酸发生突变的情况。

1.3 压力不对称指数的计算

在正常情况下, 两种氨基酸发生互相突变的几率应该是 50 : 50 的比例, 即各占 50% 的几率。若受到某种因素影响, 就会偏离这种几率, 偏离程度越大, 表明该因素的影响越大。压力不对称指数(PAI)就是基于这种理念进行计算的。文中的计算方法同文献[8], 均采用 John H. McDonald 等提出的方法^[17-18]。其计算式为

$$\ln(Q_{AB}/Q_{BA}) = PAI_A - PAI_B.$$

式中: A 为非嗜压蛋白中某位点的氨基酸; B 为嗜压蛋白质对应位点的氨基酸; Q_{AB} 为非嗜压蛋白中氨基酸 A 出现的位点在嗜压蛋白相应位点被 B 氨基酸取代的比例; Q_{BA} 的含义则相反。

需要指出的是, 最近, McDonald 对算法进行了修订, 将指数值的中点数从 1 改为 0, 他认为这能更敏感的反应数据的特性^[6]。因此, 本文计算的 PAI 值不能直接和之前结果比数值大小, 而主要是考察 20 种氨基酸 PAI 值的排序情况。

1.4 氨基酸理化性质及相关性分析

文中使用了 70 种氨基酸的理化性质, 如亲疏水性、极性、分子量、等电点、分子表面可及性、分子体积、热容、吉布斯自由能、形成二级结构的倾向性、电荷分布、侧链相互作用参数、平均柔性指数、长距离相互作用等。这些数据来源于 AA Index 数据库(<http://www.genome.jp/aaindex/>)^[19]。

将计算所得的 20 种氨基酸的压力不对称指数同其 70 种理化性质分别进行相关性分析, 软件为 STATISTICA 9.0(30 天完整功能试用版), 选取相关系数较大, 显著性较高的理化性质进行分析^[8]。

2 结果和分析

2.1 压力对蛋白质氨基酸使用偏性的影响

比较了 473 对嗜压和非嗜压同源蛋白中 20 种氨基酸的替换情况, 结果如表 1 所示。该样本中共统计了 185 621 个氨基酸, 其中有 21 662 个氨基酸位点发生了替换, 样本量足够用于统计氨基酸使用上可能存在的某种趋势。

从表 1 可看出: 在 20 种氨基酸中, 有 6 种氨基酸没有出现差异有统计学意义的氨基酸使用偏性, 占总数量的 30%, 它们分别是 C(半胱氨酸)、D(天冬氨酸)、H(组氨酸)、M(甲硫氨酸)、W(色氨酸)和 Y(酪氨酸); 而出现差异有统计学意义最多的氨基酸是 K(赖氨酸), 其次 I(异亮氨酸)和 N(天门冬酰胺)。

经统计, 在 380 个不同氨基酸对中, 有 12 对存在差异有统计学意义的氨基酸使用偏性, 结果如表 2 所示。表 2 中: Forward 表示非嗜压蛋白突变为嗜压蛋白; Reverse 表示嗜压蛋白突变为非嗜压蛋白。从表 2 可知: 差别最大是 VI(缬氨酸-异亮氨酸), 在嗜压蛋白中有 1 062 个位点是 Ile, 而在非嗜压蛋白对应的位点却为 Val; 同时, 嗜压蛋白中有 763 个位点是 Val, 而对应的位点是非嗜压蛋白却是 Ile, 即嗜压

蛋白有 1 062 个 Val 突变成 Ile, 仅有 763 个从 Ile 突变成 Val, 二者之间差异达极有统计学意义的水平.

表 1 嗜压和非嗜压菌 473 对同源蛋白中氨基酸替换的矩阵

Tab. 1 Amino acid substitution matrix in the 473 orthologous proteins from piezophiles and non-piezophiles

非嗜压蛋白	嗜压蛋白									
	A	C	D	E	F	G	H	I	K	L
A	14 562	21	154	257	13	191	13	44 ^①	104	57
C	23	2 084	2	3	4	5	0	3	0	6
D	133	4	9 762	745	0	87	28	4	60	4
E	295	3	708	10 769	1	63	17	19	196	28
F	17	8	1	0	6 624	2	15	24	2	120
G	223	7	99	55	1	13 906	14	1	52	6
H	12	3	28	24	16	5	3 833	5	30	24
I	25	6	5	16	35	2	5	10 209	21	403
K	85	2	48	131	3	20	35	31	8702	27
L	38	5	8	26	101	11	16	414	33	16 284
M	12	2	0	10	12	1	4	134	28	214
N	46	6	277	81	0	68	78	15	191	9
P	124	1	16	28	0	12	4	8	20	15
Q	69	2	74	290	5	13	85	15	309	87
R	18	9	11	14	0	17	39	11	298	33
S	548	61	102	97	6	128	27	24	115	34
T	250	11	56	94	2	25	12	114	141	36
V	202	10	15	50	21	10	9	1 062	33	205
W	1	0	0	1	5	0	1	2	2	2
Y	6	2	4	2	202	3	53	6	6	19
嗜压蛋白 总数	16 689	2 247	11 370	12 693	7 051	14 569	4 288	12 145	10 343	17 613

嗜压蛋白										非嗜压蛋白 总数
M	N	P	Q	R	S	T	V	W	Y	
18	59	113	69	21	570	284	240	0	2	16 792
0	5	1	3	10	44	13	10	0	5	2 221
0	249	13	72	6	80	51	16	0	4	11 318
8	64	38	326	18	92	87	72	2	1	12 807
12	6	1	3	2	12	7	20	8	214	7 098
2	54	6	17	19	133	28	10	1	2	14 636
2	101	9	98	42	16	12	5	1	48	4 314
112	11	3	14	4	23	65	763	0	7	11 729
24	125	15	271	281	76	118	27	0	2	10 023
248	5	5	80	28	32	38	184	6	18	17 580
4 972	4	1	23	3	5	30	50	1	3	5 509
5	5 821	10	63	23	212	108	7	0	13	7 033
5	6	7165	29	4	86	36	16	0	1	7 576
21	56	30	5 704	73	60	56	22	1	2	6 974
3	30	9	100	8 671	48	26	9	0	3	9 349
8	168	72	45	48	8 492	467	29	2	9	10 482
29	102	27	41	38	456	8 171	111	1	5	9 722
46	11	14	30	8	39	91	11 594	0	4	13 454
0	0	0	0	1	1	0	0	1 818	2	1 836
5	18	0	4	8	6	3	4	1	4 816	5 168
5 520	6 895	7 532	6 992	9 308	10 483	9 691	13 189	1 842	5 161	185 621

① 方框中的数据为显著不对称的氨基酸对

然而, 仅仅从表 2 中这些差异有统计学意义的氨基酸对仍难以得到一个完整的两种类型蛋白中氨基酸使用偏性的模式. 因此, 需要统计其他所有氨基酸向某种氨基酸突变的总趋势. 例如: 嗜压蛋白中除 W(色氨酸) 以外的其他 19 种氨基酸向嗜压蛋白中 W 突变的总数量, 以及嗜压蛋白中除 W(色氨酸) 以外的其他 19 种氨基酸向非嗜压蛋白中 W 突变的总数量. 通过这两组数据的比较, 找出其突变总趋势, 并在此基础上计算 W 的 PAI 值.

表 2 473 对同源蛋白中替换差异有统计学意义的氨基酸对

Tab. 2 Most statistical bias amino acid pairs in 473 homologous proteins

氨基酸对	Forward	Reverse	G 值	P 值	氨基酸对	Forward	Reverse	G 值	P 值
VI	763	1 062	49.194 9	$2.318 \times 10^{-12}^{***}$	IA	44	25	5.262 0	$21.797 \times 10^{-4}^*$
KG	52	20	14.630 3	$1.308 \times 10^{-4}^{***}$	SN	212	168	5.099 5	$23.933 \times 10^{-4}^*$
NK	125	191	13.864 9	$1.965 \times 10^{-4}^{***}$	NF	6	0	—	$31.250 \times 10^{-4}^*$
TI	65	114	13.548 3	$2.326 \times 10^{-4}^{***}$	RQ	73	100	4.219 0	$39.975 \times 10^{-4}^*$
KE	196	131	12.987 1	$3.137 \times 10^{-4}^{***}$	PL	5	15	—	$41.389 \times 10^{-4}^*$
SK	76	115	7.998 7	$46.811 \times 10^{-4}^{**}$	VE	72	50	3.972 7	$46.244 \times 10^{-4}^*$

经统计及计算后, 20 种氨基酸替换总体数据及其 PAI 值如表 3 所示. PAI 值越大, 表明该氨基酸更容易出现在嗜压蛋白中, 其嗜压性能越强. 由表 3 可知: Trp(色氨酸, W) 和 Ile(异亮氨酸, I) 的嗜压性能最强, 其次分别为 Lys(赖氨酸, K)、Cys(半胱氨酸, C), Met(甲硫氨酸, M) 和 Leu(亮氨酸, L); 而 Pro(脯氨酸, P)、Gly(甘氨酸, G) 和 Asn(天冬酰胺) 则为最不嗜压氨基酸. 这与之前报道差异明显.

表 3 20 种氨基酸的 PAI 值

Tab. 3 PAI values of the 20 amino acids

氨基酸	Forward	Reverse	PAI	PAI ^[8]	氨基酸	Forward	Reverse	PAI	PAI ^[8]
W(Trp)	24	18	0.273 2	0.666	D(Asp)	1 608	1 556	-0.046 9	1.290
I(Ile)	1 936	1 520	0.258 2	0.926	S(Ser)	1 991	1 990	-0.055 9	1.342
K(Lys)	1 641	1 321	0.181 2	1.105	E(Glu)	1924	2 038	-0.084 9	1.118
C(Cys)	163	137	0.127 2	0.753	H(His)	455	481	-0.089 9	0.985
M(Met)	548	537	0.120 2	0.974	Y(Tyr)	345	352	-0.089 9	0.531
L(Leu)	1 329	1 296	0.108 2	1.054	A(Ala)	2 127	2 230	-0.094 9	1.076
Q(Gln)	1 288	1 270	0.016 2	0.729	F(Phe)	427	474	-0.103 9	0.869
R(Arg)	637	678	0.000 2	1.361	N(Asn)	1 074	1 212	-0.142 9	1.056
V(Val)	1 595	1 860	-0.021 9	1.131	G(Gly)	663	730	-0.153 9	1.346
T(Thr)	1 520	1 551	-0.041 9	0.871	P(Pro)	367	411	-0.157 9	0.820

例如: 文中 Trp 为嗜压性能最强的氨基酸, 而在之前研究中它确是属于最不嗜压的氨基酸之一; 而文中最不嗜压的 Gly 在之前研究中却是位于第 2 嗜压的氨基酸. Trp 侧链具有芳香结构, 其分子中的 π 电子能够与带正电的离子如钠离子和钾离子相互作用, 被称为“阳离子- π ”相互作用, 这对维系蛋白分子结构稳定性非常重要^[20]. Ile 的侧链为分支的异丙基, 能更有效提高蛋白分子表面的柔性, 增加其稳定性^[21], 故其嗜压性能也较强. Pro 的刚性较强, 过多的 Pro 容易导致蛋白分子刚性太强, 不利于其稳定性^[22]. Gly 虽侧链很小, 易占据蛋白分子内外空间, 方便其折叠, 但它却是非极性氨基酸, 而极性氨基酸之间及其与水分子间较容易形成极性相互作用, 这对维持蛋白质在高压条件下的稳定性非常重要^[5]. Chou 等^[23] 的研究表明, Asn, Gly 和 Pro 很容易出现在转角和无规则转曲中, 是强烈的转角和转曲成子; 过多的转角或无规则转曲会导致蛋白质结构不紧凑, 从而降低其稳定性^[24]. 因此, 这 3 种氨基酸的嗜压性能最低.

2.2 氨基酸压力不对称指数与其理化性质相关性分析

将计算所得 PAI 值与氨基酸的 70 种理化性质分别进行相关性分析, 得到了相关系数较大 ($|r| > 0.49, p < 0.05$) 的 13 个理化性质, 如表 4 所示. 表 4 中的 13 个理化性质, 包括氨基酸固有的理化性质参数(分子体积、折射率等)、热力学参数(热容、水解自由能)、结构参数(螺旋作用面积、螺旋-无规则卷曲平衡常数、无规则转曲形成趋势)、亲疏水性参数(疏水性转移指数、疏水性环绕指数)和可及表面积(accessible surface area, ASA)等. 在这 13 个理化性质中, 只有氨基酸的无规则卷曲倾向性(coil tenden-

cies, CT)与PAI值呈负相关,其他的均为正相关. 其中,氨基酸分子的大小与其PAI值呈正相关($r=0.531, p=0.013$). 这说明氨基酸分子体积越大,其嗜压特性越强,这与之前的报道相反^[8].

表4 与PAI相关系数最大的13个理化性质

Tab. 4 Significant physicochemical properties correlated with PAI

理化性质	r	p	理化性质	r	p	理化性质	r	p
可及表面积(展开)	0.600	0.005	无规则卷曲倾向性	-0.530	0.016	膨松性	0.516	0.020
可及表面积(变性)	0.524	0.018	折射率	0.541	0.014	比容值	0.502	0.024
螺旋作用面积	0.548	0.012	疏水性转移指数	0.496	0.026	热容量	0.537	0.015
螺旋线圈平衡常数	0.513	0.021	水解自由能	0.520	0.019	分子体积	0.531	0.013
疏水性环绕指数	0.494	0.027						

虽然体积小的氨基酸残基更容易占据蛋白分子内外空间,有助于蛋白质折叠^[8],但分子体积更大的氨基酸更容易让蛋白分子中的空穴的数量和体积减少,而这对蛋白质的稳定性具有积极作用^[25]. 氨基酸水合自由能是指其在溶解状态下与其周围一层水分子发生相互作用的强弱,它与氨基酸自身的疏水性及与水分子的接触面积正相关^[26]. 研究表明:高压条件下蛋白质结构的变化主要是由于水合作用的变化所引起的,而可及表面积较大的氨基酸(尤其是位于分子表面)更容易和水分子接触,增加这种水合作用,从而保证其在高压条件下的稳定性^[27]. 这同时也解释了氨基酸分子可及性与PAI之间这种明显的正相关性(表4和图1).

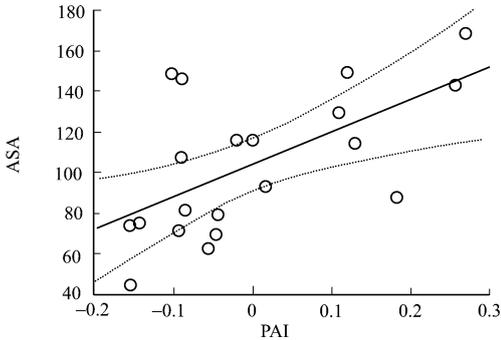


图1 PAI与氨基酸可及表面积的相关性

Fig. 1 Correlation of amino acids accessible surface area with PAI

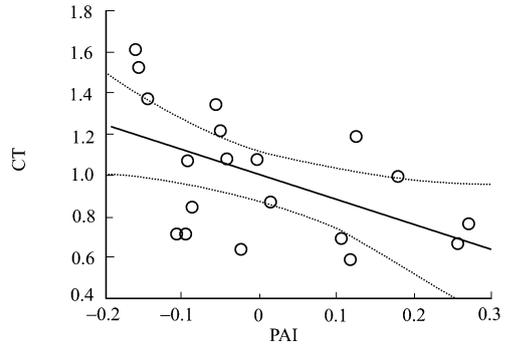


图2 PAI与氨基酸无规则卷曲倾向的相关性

Fig. 2 Correlation of amino acids coil tendencies with PAI

此外,在高压条件蛋白变性过程中,当水分子进入到疏水内核区后会加剧蛋白质变性的过程. 因此,在其分子中需要增加疏水性氨基酸的数目,以抵制水分子的进入,同时也减少分子中空穴的数量和体积,避免水分子快速进入^[28]. 这说明了PAI值和氨基酸疏水性及氨基酸体积大小有较强的正相关性.

在其他一些极端蛋白(如嗜热和嗜盐)中,二级结构的组成含量及不同二级结构中氨基酸组成存在明显差异^[29]. 文献[9]的研究也表明,嗜压和非嗜压蛋白在不同二级结构中氨基酸组成差异明显,这种现象可通过PAI值与氨基酸不同二级结构倾向因子解释. 经检查,PAI值和 α -螺旋倾向因子的相关系数为0.34,和 β -折叠倾向因子的相关系数为0.44,和转角倾向因子的相关系数为-0.43,而和无规则卷曲倾向因子的相关系数为-0.53,4种相关性均达有统计学意义的水平($p<0.05$). 可见,PAI值与 α -螺旋倾向因子相关性最小,与无规则卷曲倾向因子相关性最强(表4和图2). PAI值与 β -折叠倾向因子呈较明显的正相关,即表明越容易形成 β -折叠的氨基酸其嗜压性能越强,如Trp和Ile是强烈的 β -折叠形成子,其嗜压性能最强,因此,在嗜压蛋白的 β -折叠出现的频率也最高^[9]. 另一方面,PAI值和转角及无规则卷曲倾向性呈负相关,尤其是和无规则卷曲. 众所周知,Pro和Gly是强烈的转角及无规则卷曲形成子,其非嗜压性能最强,因此,在非嗜压蛋白的转角或无规则卷曲中含量也最高^[9],而Ile是强烈的转角及无规则卷曲破坏子,故其嗜压性能很强,在嗜压蛋白的转角或无规则卷曲中含量也最高^[9].

3 结论

由于高压对生物分子(尤其是蛋白质或酶)的独特影响使其在不同领域受到了共同的关注^[2]. 但在

真正广泛应用之前,有一个基本的问题摆在面前,那就是高压条件下蛋白质结构与其功能有何关系,是否能在认知其构效关系的基础上设计出耐高压的酶.由此也可见,了解蛋白质在高压条件下稳定性的机理在理论和应用上均具有重要价值.

通过严格选取目前发现嗜压菌最多的 *Shewanella* 的 2 对微生物,获取其全蛋白质组序列,并从中选取 473 对直系同源序列,尽最大可能地减少了其他因素对最终结果的影响,计算出氨基酸压力不对称指数.该结果与之前报道存在较大差异.究其主要原因,可能是由于之前研究选取的是 *Pyrococcus* 属的 2 个微生物,这两种微生物属古菌,同时也是生长在高温条件下的是嗜热菌.本实验选取的是细菌,且生长在低温条件下,也许是其中的嗜压蛋白在不同温度及不同系统发育阶段会有不同适应压力的机制,已有研究者注意到这种现象^[30].但无论如何,研究结果至少可作为对之前所得氨基酸压力不对称指数的补充及修正,这对了解占嗜压菌大多数的低温菌中嗜压蛋白稳定性机制应该更具有参考价值,相关研究仍需进一步深入.

参考文献:

- [1] EI-HAJJ Z W, ALLCOCK D, TRYFONA T, et al. Insights into piezophily from genetic studies on the deep-sea bacterium, *Photobacterium profundum* SS9[J]. Ann NY Acad Sci, 2010, 1189: 143-148.
- [2] ABRAM A, FILIP M, MARC E G H, et al. Biotechnology under high pressure: Applications and implications[J]. Trends Biotechnol, 2009, 27(7): 434-441.
- [3] BUCKOW R, WEISS U, HEINZ V, et al. Stability and catalytic activity of α -amylase from barley malt at different pressure-temperature conditions[J]. Biotechnol Bioeng, 2007, 97(1): 1-11.
- [4] MOMBELLI E, SHEHI E, FUSI P, et al. Exploring hyperthermophilic proteins under pressure: Theoretical aspects and experimental findings[J]. Biochim Biophys Acta, 2002, 1595(1/2): 392-396.
- [5] FRANCESCA S, STEFANO C, FEDERICO M L, et al. Piezophilic adaptation: A genomic point of view[J]. J Biotechnol, 2006, 126(1): 11-25.
- [6] MCDONALD J H. Temperature adaptation at homologous sites in proteins from nine thermophile-mesophile species pairs[J]. Genome Biol Evol, 2010, 2: 267-276.
- [7] GOODARZI H, TORABI N, NAJAFABADI H S, et al. Amino acid and codon usage profiles: Adaptive changes in the frequency of amino acids and codons[J]. Gene, 2008, 407: 30-41.
- [8] DI GIULIO M. A comparison of proteins from *Pyrococcus furiosus* and *Pyrococcus abyssi*: Barophily in the physico-chemical properties of amino acids and in the genetic code[J]. Gene, 2005, 346: 1-6.
- [9] 张光亚, 李红春, 高嘉强, 等. 氨基酸和二肽组成对嗜压微生物蛋白质稳定性的影响[J]. 微生物学报, 2009, 49(2): 198-203.
- [10] 张光亚, 李红春, 高嘉强, 等. 嗜压蛋白氨基酸残基溶剂可及性对其稳定性的影响[J]. 微生物学报, 2010, 50(5): 621-627.
- [11] FANG J, ZHANG L, BAZYLINSKI D A. Deep-sea piezosphere and piezophiles: Geomicrobiology and biogeochemistry[J]. Trends Microbiol, 2010, 18(9): 413-422.
- [12] DELONG E F, FRANKS D G, YAYANOS A A. Evolutionary relationships of cultivated psychrophilic and barophilic deep-sea bacteria[J]. Appl Environ Microbiol, 1997, 63(5): 2105-2108.
- [13] AONO E, BABA T, ARA T, et al. Complete genome sequence and comparative analysis of *Shewanella violacea*, a psychrophilic and piezophilic bacterium from deep sea floor sediments[J]. Mol Biosyst, 2010, 6(7): 1216-1226.
- [14] JAIN E, BAIROCH A, DUVAUD S, et al. Infrastructure for the life sciences: Design and implementation of the UniProt website[J]. BMC Bioinformatics, 2009, 10: 136.
- [15] ALTSCHUL S F, MADDEN T L, SCHAFFER A A, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs[J]. Nucleic Acids Res, 1997, 25(17): 3389-3402.
- [16] LARKIN M A, BLACKSHIELDS G, BROWN N P, et al. Clustal W and Clustal X version 2.0[J]. Bioinformatics, 2007, 23(21): 2947-2948.
- [17] MCDONALD J H, GRASSO A M, REJTO L K. Patterns of temperature adaptation in proteins from *Methanococcus* and *Bacillus*[J]. Mol Biol Evol, 1999, 16(12): 1785-1790.
- [18] MCDONALD J H. Patterns of temperature adaptation in proteins from the bacteria *Deinococcus radiodurans* and

- Thermus thermophilus*[J]. Mol Biol Evol, 2001, 18(5): 741-749.
- [19] KAWASHIMA S, POKAROWSKI P, POKAROWSKA M, et al. AAindex; Amino acid index database, progress report 2008[J]. Nucleic Acids Res, 2008, 36(S1): 202-205.
- [20] MA J C, DOUGHERTY D A. The cation- π interaction[J]. Chemical Reviews, 1997, 97(5): 1303-1324.
- [21] PACK S P, YOO Y J. Protein thermo-stability; Structure-based difference of amino acid between thermophilic and mesophilic proteins[J]. J Biotechnol, 2004, 111(3): 269-277.
- [22] ARNÓRSDÓTTIR J, KRISTJÁNSSON M M, FICNER R. Crystal structure of a subtilisin-like serine proteinase from a psychrotrophic *Vibrio* species reveals structural aspects of cold adaptation[J]. FEBS J, 2005, 272(3): 832-845.
- [23] CHOU P Y, FASMAN G D. Empirical predictions of protein conformation[J]. Ann Rev Biochem, 1978, 47: 251-276.
- [24] SONAN G K, RECEVEUR-BRECHOT V, DUEZET C, et al. The linker region plays a key role in the adaptation to cold of the cellulase from an *Antarctic bacterium*[J]. Biochem J, 2007, 407(2): 293-302.
- [25] ANDRÁS S, PÉTER Z. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: Results of a comprehensive survey[J]. Structure, 2000, 8(5): 493-504.
- [26] OOI T, OOBATAKE M, NMMETHY G, et al. Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides[J]. Proc Natl Acad Sci USA, 1987, 84(10): 3086-3090.
- [27] SILVA J L, FOGUEL D. Hydration, cavities and volume in protein folding, aggregation and amyloid assembly[J]. Phys Biol, 2009, 6(1): 15002.
- [28] NOLWENNIG R, JEAN R, GÉRARD D. Development of high hydrostatic pressure in biosciences; Pressure effect on biological structures and potential applications in biotechnologies[J]. Biotechnol Adv, 2010, 28(6): 659-672.
- [29] SANDIP P, SUMIT K B, SABYASACHI D, et al. Molecular signature of hypersaline adaptation; Insights from genome and proteome composition of halophilic prokaryotes[J]. Genome Biol, 2008, 9R(4): 70.
- [30] RIE S, CHIAKI K, AKIHIKO N. Amino acid substitutions in malate dehydrogenases of piezophilic bacteria isolated from intestinal contents of deep-sea fishes retrieved from the abyssal zone[J]. J Gen Appl Microbiol, 2006, 52(1): 9-19.

Revision of the Amino Acid Pressure Asymmetry Index of Piezophilic Proteins

ZHANG Guang-ya

(College of Chemical Engineering, Huaqiao University, Xiamen 361021, China)

Abstract: A comparison was made between 473 orthologous proteins from *Shewanella*. The pattern of asymmetries in the 21 662 amino acid mutation sites reveals some common trends. The pressure asymmetry index (PAI) identifies the amino acids tryptophan (Trp) and isoleucine (Ile) as those having the most piezophilic behavior, and proline (Pro), glycine (Gly) and asparagines (Asn) as the least piezophilic, which is totally different from the former result. The PAI makes it possible to visualize the amino acid properties that best explain piezophily. Among the 70 properties of amino acids, there are 13 significantly correlated with PAI value. Only the coil tendencies are negatively correlated with PAI value, that is to say, on average, higher PAI values are associated with fewer tendencies to form coil. The PAI we established here may be more suitable to the piezophiles living in the cold environment, and a revision of the previous PAI values was made.

Keywords: piezophilic microbes; pressure asymmetry index; amino acid; coil tendencies; appropriate pressure mechanism

(责任编辑: 黄晓楠 英文审校: 刘源岗)