Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 9297–9301, September 1994
Biophysics

Geometry of interplanar residue contacts in protein structures

LUCIANO BROCCHIERI AND SAMUEL KARLIN

Department of Mathematics, Stanford University, Stanford, CA 94305-2125

Contributed by Samuel Karlin, June 9, 1994

ABSTRACT **The relative spatial disposition of interacting side-chain planar groups (aromatic, guanidinium, amide, carboxyl, imidazole) is analyzed for 186 non-homologous well-resolved protein structures. The dihedral angle of amide or carboxyl planar groups with other planar groups accords with a random distribution of planes. By contrast, the dihedral angle of the planes between close aromatic rings or of the histidine ring interacting with aromatic residues is significantly nonrandom, showing an approximately uniform distribution. Our results indicate that edge-to-edge and edge-to-center spatial dispositions of residue planar sections are prevalent, while complete stacking configurations are uncommon. The hypothesis that electrostatic forces are a major determinant of the geometry of interactions between side-chain planar groups is discussed.**

Protein structures embodying vital contacts between planar groups are increasingly recognized. For example, planar interactions of amide and amino groups with aromatic rings occur for bovine pancreatic trypsin inhibitor in complexes and for hemoglobin with various drugs (e.g., refs. 1 and 2). Cation–aromatic interactions are conspicuous in the context of protein crystal structures (3), as a force of molecular recognition and catalysis (1, 4, 5), as a mechanism of aromatic guidance (6, 7), as a means of ion selectivity in channels and related structures (8), in interpreting conformational changes of guanine nucleotide-binding protein (G-protein)-coupled receptors (9, 10), and generally in stabilizing protein structures (1). Aromatic interactions featuring tryptophan and tyrosine residues are prevalent in immunoglobulin antibody with antigen-binding regions (11) and are found in excess in the groove of the major histocompatibility complex structure (12).

There has been considerable attention in recent years to the geometry of interactions between close aromatic residues, to the preferred orientations of aromatic rings in contact with an arginine guanidinium group, and to the spatial disposition of an amide–aromatic complex (e.g., refs. 2, 3, 13–21, and 24). The dihedral angles between interacting phenylalanine side-chain rings in globular proteins showed an approximate random distribution (15, 16). In contrast, aromatic rings tend to approach the planar guanidinium group of arginine in a parallel stacking fashion, where the line connecting the centers of the planar sections is approximately orthogonal to each plane (19). Amide group–phenyl planar arrangements strive to be parallel, but the nitrogen atom of the amide group is located at or beyond the ring edge (2).

Natural side-chain planar groups occur with the following amino acid groups: the aromatic rings of phenylalanine (F), tyrosine (Y), and tryptophan (W) (in the one-letter code), the imidazole ring of histidine (H), the guanidinium group of arginine (R), the amide groups of asparagine (N) and glutamine (Q), and the carboxyl groups of aspartate (D) and glutamate (E). It is of interest to compare and contrast the interplanar side-chain geometry observed in protein structures among these residue types. Important factors underlying interplanar interactions in protein structures pertain to steric and tertiary constraints, electrostatic orientations, hydrogen bonding, π-cloud electron distributions, functional roles, and hydrophobic effects.

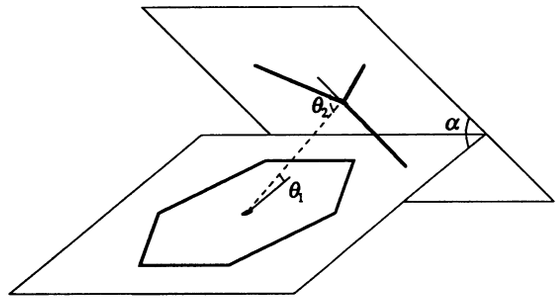

FIG. 1. Geometry of interaction between planar groups, showing angles used to classify the geometry of interaction between planar sections of contacting residues. α is the dihedral angle between the two extended planes, and θ_1 and θ_2 refer to the angles determined by the line joining the centers of the two planar sections and the two planes, respectively.

METHODS AND DATA

The spatial organization for interacting planar groups of two residues in a protein environment is aptly characterized by three angles (16). For a pair of planar sections, let α be the dihedral angle between the two extended planes and let θ_1 and θ_2 refer to the angles between the line joining the centers of the two planar sections and the two respective planes (see Fig. 1). It is convenient to partition interplanar angles into three categories: $0° < \alpha < 30°$ referred to as parallel type, $30° < \alpha < 60°$ referred to as oblique type, and $60° < \alpha < 90°$ referred to as orthogonal type. As noted in Singh and Thornton (15) and Blundell *et al.* (16), a random distribution of this dihedral angle varies having a density equal to $\sin(\alpha)$. Thus, the probability for a random dihedral angle of parallel type is 0.134, of oblique type is 0.366, and of orthogonal type is 0.500. In particular, two random planes tend to be more orthogonal than coplanar.

From a broad data collection of 186 nonhomologous well-resolved protein structures, all nearest neighbor residue pairings were compiled. These pairs are formed by each residue and the residue closest to it on the basis of the minimum distance between their side-chain atoms. For listings and characteristics of the data set and formal definitions of distances, see Karlin *et al.* (21). Our data set cumulates 7002 pairs of interacting planar side chains involving the residues F, Y, W, H, R, N, Q, D, and E (see Table 1). These reflect a plethora of R residues in side-chain proximity to E and D, significant relative abundance of neighboring aromatics, a prominence of Y and W among cation–aromatic interactions, versatility of H in associations with D, E, and itself, and distinctive overrepresentations of amide residues with R and themselves (21).

RESULTS

Distribution of Interplanar Angles. The proportion of closest side-chain distances pairing planar groups of specified residue types are described in Table 1. These, except for Y, are persistently high, whereas Y often involves its hydroxyl group in these interactions, particularly with anionic residues. Y and W (not F) are significantly overrepresented (in

Table 1. Statistics of dihedral angles of interacting residues characterized by side-chain planar groups

α	FF	FW	FY	WW	WY	YY	AROMATIC-AROMATIC
0°–30°	54	29	46	2	31	21	183 (*91*)
30°–60°	51	37	63	9	40	42	242 (*249*)
60°–90°	67	50	63	16	22	36	254 (*340*)
Total	172	116	172	27	93	99	679
Rel. freq. of inter-planar contacts (%)*	*75.2*	*82.7*	*59.3*	*77.8*	*48.4*	*38.1*	*63.5*

α	HF	HW	HY	HISTIDINE-AROMATIC
0°–30°	47	16	40	103 (*38*)
30°–60°	24	15	43	82 (*103*)
60°–90°	29	11	56	96 (*141*)
Total	100	42	139	281
Rel. freq. of inter-planar contacts (%)*	*75.2*	*88.1*	*29.5*	*54.5*

α	HH
0°–30°	11
30°–60°	19
60°–90°	19
Total	49
Rel. freq. of inter-planar contacts (%)*	*79.2*

α	RF	RW	RY	GUANIDINIUM-AROMATIC
0°–30°	46	32	43	121 (*56*)
30°–60°	57	21	59	137 (*154*)
60°–90°	44	32	87	163 (*211*)
Total	147	85	189	421
Rel. freq. of inter-planar contacts (%)*	*46.7*	*54.1*	*26.3*	*39.0*

α	RH
0°–30°	18
30°–60°	35
60°–90°	34
Total	87
Rel. freq. of inter-planar contacts (%)*	*67.8*

α	RR
0°–30°	15
30°–60°	25
60°–90°	22
Total	62
Rel. freq. of inter-planar contacts (%)*	*50.0*

α	NF	NW	NY	QF	QW	QY	AMIDE-AROMATIC
0°–30°	9	8	22	13	14	27	93 (*92*)
30°–60°	51	26	66	45	25	60	273 (*253*)
60°–90°	56	28	89	51	34	66	324 (*345*)
Total	116	62	177	109	73	153	690
Rel. freq. of inter-planar contacts (%)*	*82.8*	*96.8*	*46.3*	*75.1*	*79.4*	*83.7*	*73.3*

α	NH	QH	AMIDE-HISTIDINE
0°–30°	11	12	23 (*22*)
30°–60°	28	26	54 (*60*)
60°–90°	50	36	86 (*82*)
Total	89	74	163
Rel. freq. of inter-planar contacts (%)*	*82.4*	*87.6*	*84.8*

α	NR	QR	AMIDE-GUANIDINIUM
0°–30°	19	37	56 (*47*)
30°–60°	46	76	122 (*130*)
60°–90°	98	78	176 (*177*)
Total	163	191	354
Rel. freq. of inter-planar contacts (%)*	*61.3*	*77.3*	*69.9*

α	NN	NQ	QQ	AMIDE-AMIDE
0°–30°	24	34	12	70 (*59*)
30°–60°	55	75	21	151 (*161*)
60°–90°	67	111	40	218 (*220*)
Total	146	220	73	439
Rel. freq. of inter-planar contacts (%)*	*100.*	*87.7*	*73.2*	*89.4*

α	DF	DW	DY	EF	EW	EY	CARBOXYL-AROMATIC
0°–30°	16	7	28	13	7	19	90 (*120*)
30°–60°	42	35	83	46	33	94	333 (*328*)
60°–90°	66	45	124	59	38	140	472 (*448*)
Total	124	87	235	118	78	253	895
Rel. freq. of inter-planar contacts (%)*	*74.2*	*87.4*	*17.9*	*70.0*	*92.4*	*20.5*	*46.6*

α	DH	EH	CARBOXYL-HISTIDINE
0°–30°	26	23	49 (*53*)
30°–60°	77	65	142 (*145*)
60°–90°	114	90	204 (*198*)
Total	217	178	395
Rel. freq. of inter-planar contacts (%)*	*92.2*	*83.1*	*88.1*

α	DR	ER	CARBOXYL-GUANIDINIUM
0°–30°	84	91	175 (*158*)
30°–60°	241	192	433 (*433*)
60°–90°	284	290	574 (*591*)
Total	609	573	1182
Rel. freq. of inter-planar contacts (%)*	*86.1*	*83.2*	*84.7*

α	DN	DQ	EN	EQ	CARBOXYL-AMIDE
0°–30°	42	22	51	29	144 (*124*)
30°–60°	120	72	82	50	324 (*338*)
60°–90°	149	90	128	89	456 (*462*)
Total	311	184	261	168	924
Rel. freq. of inter-planar contacts (%)*	*98.7*	*78.9*	*76.7*	*76.1*	*84.4*

α	DD	DE	EE	CARBOXYL-CARBOXYL
0°–30°	20	31	16	67 (*51*)
30°–60°	36	59	31	126 (*139*)
60°–90°	61	85	41	187 (*190*)
Total	117	175	88	380
Rel. freq. of inter-planar contacts (%)*	*98.3*	*73.7*	*64.8*	*79.2*

Numbers of nearest neighbor interactions between residue side chains involving planar groups from a data set of 186 well-resolved structures. The interplanar angles α are categorized as of parallel type ($0° < \alpha < 30°$), oblique type ($30° < \alpha < 60°$), and orthogonal type ($60° < \alpha < 90°$). Expectations based on the distribution of random planes are shown in italic for groupings. *Proportion among side-chain nearest neighbor interactions of given residues involving closest atoms in respective planar groups.

side-chain minimum distances) as nearest neighbors to R, but not to K, probably because of its favorable cation–aromatic interaction (21). K associates with aromatic rings often via a hydrophobic interaction involving one or several of its methylene groups.

Table 1 reports statistics on the interplanar angles achieved between the planar groups of minimum side-chain distance in our 186 representative structures. The following observations stand out:

(*i*) The dihedral angles between interacting phenyl rings of aromatic residue pairings ([W,W] excepted) favor $\alpha < 60°$, signifying a nonrandom distribution according to the criteria above (see also ref. 19). The herringbone (orthogonal) pattern observed for benzene crystals (14) is not prominent among protein phenyl ring contacts.

(*ii*) The interplanar angles of the H ring with the rings of Y, W, and H show an approximate uniform distribution—i.e., about equally frequent in parallel, oblique, and orthogonal intersections and therefore significantly nonrandom. The planar orientation of H with F is the most deviant from random expectations among all residue planar pairs.

(*iii*) The interplanar angles of R with the aromatic residues F and W are uniformly (hence nonrandomly) distributed, while the interplanar angles of R with Y follow approximately the distribution of random planes (compare refs. 2 and 19).

(*iv*) The interplanar dispositions for the residue pairings [R,R], [H,R], [Q,R], [Q,Y], [N,W], [Q,F], and [N,F] are predominantly of orthogonal or oblique types ($\alpha > 30°$). All other interacting residue pairings involving planar groups favor a dihedral angle tending to orthogonality as for a random distribution. In particular, the plane of the carboxyl groups of D and E with other planar residue types are rarely parallel. The planar side chain of the carboxyl groups of D and E and the oppositely charged guanidinium residue tend to be positioned in an orthogonal orientation. In summary, the dihedral angles formed by D and E with all planar groups show mostly a random distribution. The same applies to the amide side chains.

Spatial Arrangements Among Residue Planar Sections. Table 2 reports counts among interacting planar groups of the centering parameters measured by θ_1 and θ_2 in relation to the interplanar angle α. Values of each panel of Table 2 summarize the counts for each dihedral category, $0° < \alpha < 30°$, $30° < \alpha < 60°$, and $60° < \alpha < 90°$. Blundell *et al.* (16) and Mitchell *et al.* (19) emphasize their nonrandom nature relative to aromatic–aromatic and cationic–aromatic planar dispositions. We highlight (large boldface numbers) and venture some interpretations on the statistics of Table 2.

(*i*) Aromatic–aromatic geometries adopt parallel phenyl rings significantly more than expected by chance, rarely fully stacked but mostly positioned in an edge-to-center manner as exemplified in Fig. 2 *d* and *l* and especially in *i*. Most aromatic contacts are from oblique to orthogonal, with either an above-ring "herringbone" geometry ($0° < \theta_1 < 30°$ and $60° < \theta_2 < 90°$) as in Fig. 2*h* or displaced in an edge-to-center setting as in Fig. 2 *e* and *f*. Overall, the number of edge-to-center geometries represents about 75% of the cases compared with about 25% of edge-to-edge geometries.

(*ii*) Guanidinium groups interacting with aromatic rings are largely arranged in stacking formations (Fig. 2 *i*, *l*, and *o*). The stacking geometries of Fig. 2 *i*, *l*, and *o* are consistent with the observations of Mitchell *et al.* (19) and Flocco and Mowbray (2). However, we also find many edge-to-edge and edge-to-center geometries (Fig. 2 *e*, *f*, *c*, and *h*).

(*iii*) There are two significant geometries of amide and aromatic planar interactions: one involves stacking and the other subtends a dihedral interplanar angle of about 45°. In both cases the line of centers is approximately orthogonal to one of the planes. In most examples the geometry involves edge-to-edge displacements.

(*iv*) The spatial relationships between contacting H residues and aromatic rings is primarily that of partial stacking. The numbers of Table 2 indicate that the planar groups of H with aromatic residues often project a parallel configuration coupled to an edge-to-center displacement.

(*v*) The interplanar organization of a carboxyl side chain with phenyl rings tends to establish orthogonal planes in a side-by-side position, as in Fig. 2 *b*, *c*, and *f*. Edge-to-edge configurations cover more than 73% of the cases.

Table 2. Distribution of interplanar angles in relation to the centers of the planar sections of interacting residues

Aromatic ring (F,Y,W)-Aromatic ring (F,Y,W)

θ_1 \ θ_2	0°–30°	30°–60°	60°–90°	Total
0°–30°	33 (23)	**48** (16)	0 (0)	**114** (81)
	39 (42)	112 (90)	**44** (26)	234 (223)
	31 (39)	86 (140)	112 (130)	260 (304)
	103 (136)	246 (238)	**156** (98)	608 (679)
30°–60°		**76** (26)	**18** (6)	**218** (71)
		25 (34)	20 (22)	182 (195)
		20 (61)	6 (14)	132 (266)
		121 (105)	44 (84)	532 (497)
60°–90°			8 (4)	34 (29)
			1 (5)	66 (80)
			0 (0)	118 (109)
			9 (17)	218 (182)
Total				**183** (91)
				242 (249)
				254 (340)
				679

Guanidinium group (R)-Aromatic ring (F,Y,W)

θ_1 \ θ_2	0°–30°	30°–60°	60°–90°	Total
0°–30°	10 (12)	7 (5)	0 (0)	17 (21)
	21 (23)	36 (29)	3 (4)	60 (57)
	25 (21)	36 (45)	19 (18)	80 (78)
	56 (48)	79 (75)	22 (35)	157 (211)
30°–60°	2 (2)	**31** (16)	**16** (8)	**49** (21)
	22 (16)	12 (21)	13 (22)	47 (56)
	19 (25)	30 (37)	7 (6)	56 (77)
	43 (47)	73 (73)	36 (33)	152 (154)
60°–90°	0 (0)	**24** (10)	**31** (23)	**55** (15)
	7 (7)	21 (28)	2 (10)	30 (41)
	23 (23)	4 (11)	0 (0)	27 (56)
	30 (34)	49 (54)	**33** (23)	**112** (56)
Total	12 (17)	**62** (27)	**47** (12)	**121** (56)
	50 (47)	69 (74)	18 (33)	137 (154)
	67 (64)	70 (100)	26 (46)	163 (211)
	129 (211)	**201** (154)	**91** (56)	421

Amide group (N,Q)-Aromatic ring (F,Y,W)

θ_1 \ θ_2	0°–30°	30°–60°	60°–90°	Total
0°–30°	23 (37)	5 (5)	0 (0)	28 (42)
	68 (66)	24 (31)	13 (12)	105 (113)
	74 (62)	54 (47)	49 (50)	177 (155)
	165 (141)	83 (91)	62 (78)	310 (345)
30°–60°	4 (7)	17 (19)	13 (12)	34 (36)
	47 (42)	26 (23)	51 (47)	124 (98)
	59 (61)	40 (41)	11 (16)	110 (134)
	110 (120)	83 (78)	75 (70)	268 (253)
60°–90°	0 (0)	7 (5)	24 (25)	**31** (15)
	8 (7)	19 (21)	17 (16)	44 (41)
	27 (28)	10 (10)	0 (0)	37 (56)
	35 (50)	36 (33)	41 (30)	112 (92)
Total	27 (42)	29 (27)	**37** (24)	93 (92)
	123 (113)	69 (74)	**81** (65)	273 (253)
	160 (155)	104 (101)	60 (89)	324 (354)
	310 (345)	202 (253)	**178** (92)	690

Histidine ring (H)-Aromatic ring (F,Y,W)

θ_1 \ θ_2	0°–30°	30°–60°	60°–90°	Total
0°–30°	18 (15)	4 (3)	0 (0)	22 (15)
	22 (25)	20 (15)	2 (2)	44 (41)
	22 (22)	11 (17)	12 (12)	45 (56)
	62 (48)	35 (45)	14 (18)	111 (140)
30°–60°	1 (1)	**37** (13)	**10** (4)	**48** (14)
	14 (11)	5 (18)	8 (12)	27 (38)
	13 (16)	15 (26)	2 (4)	30 (53)
	28 (40)	**57** (44)	20 (21)	105 (103)
60°–90°	0 (0)	**18** (6)	15 (11)	**33** (9)
	3 (4)	7 (15)	1 (5)	11 (24)
	19 (18)	2 (6)	0 (0)	21 (33)
	22 (24)	27 (30)	16 (11)	65 (38)
Total	19 (15	**59** (16)	**25** (7)	**103** (38)
	39 (41)	32 (44)	11 (18)	82 (103)
	54 (56)	28 (60)	14 (25)	96 (140)
	112 (140)	119 (103)	**50** (38)	281

Carboxyl group (D,E)-Aromatic ring (F,Y,W)

θ_1 \ θ_2	0°–30°	30°–60°	60°–90°	Total
0°–30°	58 (70)	3 (9)	0 (0)	61 (60)
	141 (129)	38 (39)	6 (6)	185 (164)
	119 (118)	65 (58)	19 (19)	203 (224)
	318 (297)	106 (115)	25 (37)	449 (448)
30°–60°	7 (17)	7 (17)	4 (5)	18 (44)
	56 (78)	26 (22)	18 (22)	100 (121)
	150 (118)	48 (41)	14 (8)	**212** (165)
	213 (211)	81 (89)	36 (30)	330 (328)
60°–90°	0 (0)	4 (8)	7 (10)	11 (16)
	11 (9)	27 (26)	10 (7)	48 (42)
	40 (42)	17 (14)	0 (0)	57 (58)
	51 (74)	48 (31)	17 (11)	116 (120)
Total	65 (78)	14 (31)	11 (10)	90 (120)
	208 (213)	91 (86)	34 (29)	333 (328)
	309 (291)	130 (118)	33 (39)	472 (448)
	582 (448)	235 (328)	78 (120)	895

Carboxyl group (D,E)-Guanidinium group (R)

θ_1 \ θ_2	0°–30°	30°–60°	60°–90°	Total
0°–30°	137 (115)	12 (10)	0 (0)	**149** (88)
	227 (206)	36 (43)	4 (7)	267 (240)
	144 (187)	70 (65)	26 (23)	240 (328)
	508 (493)	118 (128)	30 (34)	**656** (591)
30°–60°	11 (24)	7 (19)	2 (4)	20 (59)
	111 (126)	22 (25)	14 (14)	147 (163)
	206 (178)	62 (47)	9 (6)	**277** (222)
	328 (334)	91 (87)	25 (23)	444 (433)
60°–90°	0 (0)	3 (3)	3 (3)	6 (11)
	3 (11)	12 (11)	4 (4)	19 (30)
	50 (42)	7 (8)	0 (0)	57 (41)
	53 (62)	22 (16)	7 (4)	82 (158)
Total	**148** (119)	22 (31)	5 (8)	175 (158)
	341 (325)	70 (85)	22 (23)	433 (433)
	400 (444)	139 (116)	35 (31)	574 (591)
	889 (591)	231 (433)	62 (158)	1182

Counts are conditioned on the centering angles θ_1 and θ_2, determined by the line connecting the centroids of the associated planar groups and the respective planes. The centering angles are grouped by values (0°–30°), (30°–60°), and (60°–90°); for each pair of θ_1 and θ_2, in each panel the counts correspond from top to bottom to the interplanar angles $0°<\alpha<30°$, $30°<\alpha<60°$, and $60°<\alpha<90°$. Expectations are shown in parentheses based on a random distribution of the planes conditioned on the observed centering angles.

(*vi*) The oppositely charged planar groups (guanidinium with carboxyl) feature two principal configurations. The first is coplanar, with the centers considerably displaced as in Fig. 2 *a* and *b*, suggesting a double contact in such a manner that two nitrogens of the guanidinium group bond ionically with the two oxygens of the carboxyl group (compare ref. 18). The second has planar contacts subtending an orthogonal dihedral angle as in Fig. 2 *c* and *f*, suggesting that only one nitrogen points to one oxygen, while the other charged atoms are free to interact with other residues or other proteins and/or nucleic acids.

Other interplanar relationships not shown in Table 2 are here briefly described:

(*vii*) The dihedral angle between amide–histidine, amide–amide, amide–carboxyl, and carboxyl–carboxyl pairs conform to the distribution of random planes. In particular, these generate about 50% orthogonal intersections and planar section displacements corresponding to either θ_1 or $\theta_2 < 30°$. These conformations likely reflect the amphoteric character of the amide group.

(*viii*) The interplanar geometry of the H residue relative to the anionic residues D and E has predominantly (almost 80% of cases) the carboxyl group adjacent to the edge of the H ring. Again the planar orientations tend to a random distribution. A similar geometry prevails between the guanidinium group and the amide residues N and Q.

Associations of H with H, H with R, and R with R are too rare to allow a confident statistical summary.

Interplanar nearest neighbor contacts for tryptophan involve either its five-atom or its six-atom ring. The contrasting contacts across the 186 protein structures are as follows:

Residue	W	Y	F	H	R	Q	N	D	E	Total
Five-atom ring	16	61	57	26	37	45	35	79	53	409
Six-atom ring	38	38	59	16	48	28	27	8	25	281.

The preponderance of the W five-atom ring versus the six-atom ring in proximity to an acidic residue planar group reflects putatively a hydrogen bond of the electrophilic imino group in the five-atom ring of W with the negatively charged carboxyl group. The equal numbers of F contacting the two rings of W putatively reflect on hydrophobic interactions. The preference of R for the W six-atom ring suggests a cation–aromatic, π-cloud electrostatic interaction. Overall, planar interactions of W involve mostly the five-atom ring, which is versatile in its capacity of hydrogen bonding (involving its imino group) and in engaging π-cloud electrostatic interactions as well as hydrophobic interactions.

DISCUSSION

The nature and importance in protein structures of the interplanar spatial organization of aromatic residues relative

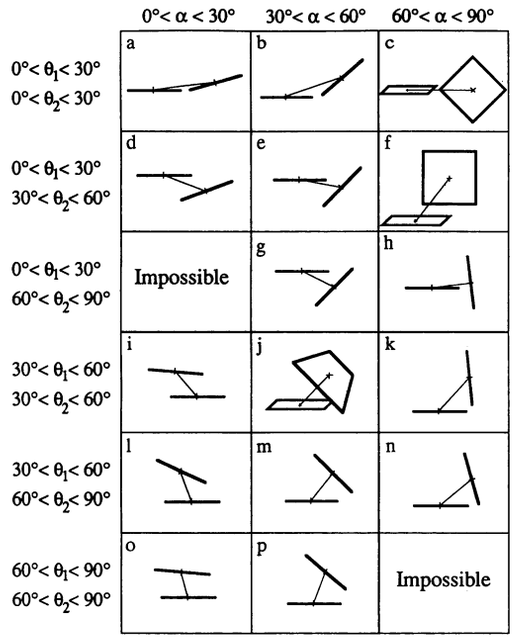

FIG. 2. Examples of possible geometrical arrangements of two neighboring planar sections. α is the dihedral angle of the two planar groups. θ_1 and θ_2 are the angles determined by the line connecting the centroids of the two planar groups and each of the two planes, respectively.

to aromatic, cationic, and amide residues has long-standing and renewed interest (2, 13–16, 19, 20). The geometry of these interplanar contacts can be reasonably parameterized by three angles (α,θ_1, and θ_2), where α is the dihedral angle between the planes and θ_1 and θ_2 specify the angles defined by the line joining the centers of the two planar sections and the extended planes. A full description requires six parameters. Whereas the dihedral angle α assesses the relative orientation of the planes (contrasting parallel, oblique, and orthogonal), the angles θ_1 and θ_2 help in characterizing the relative spatial displacement of the planar sections, distinguishing among configurations such as partial to full stacking, edge to edge, and edge to center.

The results of our analyses support the hypothesis of a major influence of electrostatic forces in determining the geometry of interactions among planar side-chain groups. In this context, the spatial dispositions of the planar sections appear to be more important than the orientation defined by the corresponding dihedral angle.

Electrostatic charges associated with planar groups are well recognized. These include a weak negative charge concentrated about the center of the aromatic ring and a weak positive charge projected at the ring periphery (14), the delocalized negative charge of the carboxyl group and delocalized positive charge of the guanidinium group, the potential of amide residues for hydrogen bonding, and the versatile charge repertoire of the imidazole ring of the H residue. These charge capacities are reflected in most cases (54%) in an edge-to-edge displacement for the carboxyl and guanidinium planes independent of their interplanar angle. Contiguous aromatic rings in proteins arrange the proximal contacts between the positive edge and negative center in 73% of the cases. The H ring in relation with aromatic rings show mostly an edge-to-center disposition, whereas carboxyl groups interact with H in an edge-to-edge arrangement (74% of cases) as might be expected by their opposite charge. Amide side-chain planar sections with respect to all other planar groups have lines of center mostly orthogonal, suggestive of a hydrogen bond.

It is also recognized that the aromatic residues Y or W, or both, can be involved in important internal cross-linking hydrogen bonds and effect conformational changes. Chemical modification experiments on the binding sites in the α subunit of nicotinic acetylcholine that are located near the extracellular face on the membrane spanning helices also point to involvement of aromatic residues in the nicotinic acetylcholine binding site (9). It is proposed that cation–aromatic planar interactions and hydrogen bonds play a role in the conformational change of the receptor, modulating the coupling of many guanine nucleotide-binding proteins. For example, five conserved aromatic residues are thought to be especially important in muscarinic neurotransmitter receptors (10).

Dihedral angles among interacting aromatics ([W,W] excepted) and cation–aromatic contacts ([R,Y] excepted) strongly favor $\alpha < 60°$ contrary to expectations of a random dihedral angle distribution. The variegated nature of R and Y planar section contacts reflects on the two major ways R and Y interact: the first way projecting the guanidinium/π-cloud electrostatics and the second emphasizing ionic interactions of the guanidinium group with the hydroxyl group of Y. In fact, investigation of the proximal atoms in R and Y contacts reveals that 26% of the cases have the guanidinium group closest to the Y phenyl ring and 38% of the cases have the guanidinium group closest to the Y side-chain hydroxyl group. Corresponding statistics of R and W close atoms relate the guanidinium group proximal to the five-atom ring of W in 32% of the cases and proximal to the six-atom ring of W in 26% of cases. With respect to R and F contacts, 47% of cases place the guanidinium group closest to the F phenyl ring, consequently, more than 50% of the cases apparently are involved in a hydrophobic relationship.

The spatial arrangement of the charged residues of D or E with R feature a pronounced edge-to-edge nexus generally involving both oxygens of D and E in an ionic bond with two amino components. The numbers of close D or E with Y planar interactions is significantly in excess of the numbers of D and E contacts with F and W (Table 1, 488 compared with 242 and 165). In these cases, the bulk of the close atoms of Y are represented by the hydroxyl group.

It is of interest to review the statistics on the closest atoms in aromatic–aromatic interactions (detailed data to be presented elsewhere). Thus, for interactions of F with F, in 75% of the cases the closest atoms of both residues involve ring components; with respect to interactions of F with Y, the close atoms consist of ring-with-ring contacts in 60% of the cases and ring-with-OH group contacts in 15% of the cases. The planar orientations are parallel significantly more than expected by chance, and their spatial dispositions are predominantly partial stacking. The close atoms of H and Y consist of ring-with-ring contacts in 30% of the cases and ring-with-OH group contacts in 52% of the cases, featuring a nitrogen atom of the H residue in 39% of all cases. The corresponding data on H and F contacts are of opposite character, emphasizing a ring-with-ring closeness. H and F are the only pair of residues involving an interplanar set of contacts for which parallel planes are prevalent, having $\alpha < 30°$ more frequent than $\alpha > 60°$. H-residues confer stability in α-helices mostly when present at the C-cap, where they compensate the helix dipole. A substantial positive correlation among occurrences of H at the C-cap of α-helices with F and Y (see also refs. 22 and 23) at corresponding positions four residues upstream from H would implicate a spatial affinity of the F and Y rings to H.

The results of our analysis underscore edge-to-edge and edge-to-center geometries as the most relevant to side-chain planar arrangements in protein structures. These spatial arrangements are in the main uncorrelated with the interplanar dihedral angle of the extended planes. While energetic

minima of pairwise interactions might correspond to precise angles between the planar groups, these are generally subject to structural constraints and the influence of ambient residues. The impression from our data is that electrostatic interactions are an important determinant of the interplanar geometry.

We are happy to acknowledge helpful comments on the manuscript from B. E. Blaisdell, F. E. Cohen, M. Gerstein, and J. Griffin. S.K. was supported in part by National Institutes of Health Grants 2R01HG00335-06 and 5R01GM10452-30 and by National Science Foundation Grant DMS91-06974.

1. Perutz, M. F. (1993) *Philos. Trans. R. Soc. London A* **345**, 105–112.
2. Flocco, M. M. & Mowbray, S. L. (1994) *J. Mol. Biol.* **235**, 709–717.
3. Burley, S. K. & Petsko, G. A. (1986) *FEBS Lett.* **203**, 139.
4. Dougherty, D. A. & Stauffer, D. A. (1990) *Science* **250**, 1558–1560.
5. McCurdy, A., Jimenez, L., Stauffer, D. A. & Dougherty, D. A. (1992) *J. Am. Chem. Soc.* **114**, 10314–10321.
6. Sussman, J. L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L. & Silman, I. (1991) *Science* **253**, 872–879.
7. Ripoll, D. R., Faerman, C. H., Axelsen, P. H., Silman, I. & Sussman, J. L. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 5128–5132.
8. Kumpf, R. A. & Dougherty, D. A. (1993) *Science* **261**, 1708–1710.
9. Trumpp-Kallmeyer, S., Hoflack, J., Bruinvels, A. & Hibert, M. (1992) *J. Med. Chem.* **35**, 3448–3462.
10. Changeux, S. P., Devillers-Thiery, A., Galzi, J. L. & Bertrand, D. (1992) *Trends Pharmacol. Sci.* **13**, 299–301.
11. Padlan, E. A. (1990) *Proteins Struct. Funct. Genet.* **7**, 112–124.
12. Prochnicka-Chalufour, A., Casanova, J., Avrameas, S., Claverie, J. & Kourilsky, P. (1991) *Int. Immunol.* **3**, 853–864.
13. Burley, S. K. & Petsko, G. A. (1985) *Science* **229**, 23–28.
14. Burley, S. K. & Petsko, G. A. (1988) *Adv. Protein Chem.* **39**, 125–189.
15. Singh, J. & Thornton, J. M. (1985) *FEBS Lett.* **191**, 1–6.
16. Blundell, T., Singh, J., Thornton, J., Burley, S. K. & Petsko, G. A. (1986) *Science* **234**, 1005 (lett.).
17. Hunter, C. A., Singh, J. & Thornton, J. M. (1991) *J. Mol. Biol.* **218**, 837–846.
18. Mitchell, J. B. O., Thornton, J. M., Singh, J. & Price, S. L. (1992) *J. Mol. Biol.* **226**, 251–262.
19. Mitchell, J. B. O., Nandi, C. L., Ali, S., McDonald, I. K., Thornton, J. M., Price, S. L. & Singh, J. (1993) *Nature (London)* **366**, 413 (lett.).
20. Nandi, C. L., Singh, J. & Thornton, J. M. (1993) *Protein Eng.* **6**, 247–259.
21. Karlin, S., Zuker, M. & Brocchieri, L. (1994) *J. Mol. Biol.* **239**, 227–248.
22. Armstrong, K. M., Fairman, R. & Baldwin, R. L. (1993) *J. Mol. Biol.* **230**, 284–291.
23. Bradley, E. K., Thomason, J. F., Cohen, F. E., Kosen, P. A. & Kuntz, I. D. (1990) *J. Mol. Biol.* **215**, 607–622.
24. Mitchell, J. B. O., Nandi, C. L., McDonald, I. K., Thornton, J. M. & Price, S. L. (1994) *J. Mol. Biol.* **239**, 315–331.