

The Crystal Structure of Glycyl-phenylalanyl-glycine*

BY RICHARD E. MARSH AND JENNY PICKWORTH GLUSKER†

Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California, U.S.A.

(Received 11 January 1961)

The crystal structure of the tripeptide glycyl-phenylalanyl-glycine 'mono'-hydrate has been determined by analysis of three-dimensional intensity data from copper radiation. The crystals are orthorhombic with space group $P2_12_12_1$; the unit-cell dimensions, as reported previously by Dégeilh and Pickworth (1956), are:

$$a = 29.72, b = 9.98, c = 4.90 \text{ \AA}.$$

The peptide chain is in an extended configuration and adjacent molecules are hydrogen-bonded together to form a parallel-chain pleated sheet. The benzene ring of the phenylalanine residue protrudes laterally from the pleated sheet and packs in an interlocking fashion with benzene rings of neighboring sheets.

Approximately 50% of a water molecule of crystallization is present for each molecule of tripeptide.

Introduction

As part of a continuing program of research on the structure of amino acids and peptides, Dégeilh and Pickworth (now Glusker) (1956) determined the unit-cell dimensions, space groups and densities of six peptides. Of these, the tripeptide glycyl-phenylalanyl-glycine (GPG) was chosen as being the most advantageous subject for a complete structure determination.

Experimental

Dégeilh and Pickworth (1956) carried out preliminary crystallographic studies on GPG. They obtained crystals in the form of long, thin needles, elongated along c , by slow cooling of an aqueous solution of DL-glycyl-phenylalanyl-glycine, and reported the following unit-cell dimensions, space group, and density:

$$\begin{aligned} a &= 29.72, b = 9.98, c = 4.90 \text{ \AA} \\ &(\text{all } \pm 0.5\%) \\ \rho_{\text{obs.}} &= 1.334 \text{ g. cm.}^{-3} \\ \text{Space group, } &P2_12_12_1 (D_2^4). \end{aligned}$$

They reported the compound to be a monohydrate with four molecules of GPG.H₂O in the unit cell.

The density calculated on the basis of four molecules of the monohydrate per unit cell is $1.36 \pm 0.02 \text{ g. cm.}^{-3}$, slightly larger than the observed value. During the course of the refinement of the structure it became evident that the actual amount of water of crystallization present corresponds to approximately 0.5 molecules per molecule of GPG, or two molecules

per unit cell. On this basis the calculated density is 1.32 g. cm.^{-3} , in somewhat better agreement with the observed value.

Since the space group $P2_12_12_1$ contains no symmetry operation of the second kind, resolution of the original DL-solution must have occurred during crystallization. We have not determined the absolute configurations of the molecules within the particular crystals we used; the parameters we have derived, when referred to a right-handed coordinate system, define the L configuration of the phenylalanine residue.

A preliminary set of multiple-film equi-inclination Weissenberg photographs about the c axis was prepared from a very thin crystal. The equator and first three layer lines were photographed with copper radiation and exposures of approximately 100 hours; even with exposures of this length, reflections could be observed only to $\sin \theta = 0.7$ ($d = 1.1 \text{ \AA}$). Later we were able to grow a somewhat larger crystal from which additional Weissenberg intensity photographs were prepared. The new $hk0$ and $hk1$ photographs were quite satisfactory and reflections were observed out to the edge of the film; the $hk2$ photograph, however, was of poor quality due to partial decomposition of the crystal and no $hk3$ photograph was obtained. Subsequently it became impossible to grow any satisfactory crystals in this laboratory as the compound chose to crystallize in a different modification with the same space group and similar unit-cell dimensions but apparently containing two molecules of water of crystallization per molecule of tripeptide.*

* Contribution No. 2658 from the Gates and Crellin Laboratories of Chemistry. The work described in this article was carried out in part under a contract (Nonr-220(05)) between the Office of Naval Research and the California Institute of Technology and in part by a research grant No. H-2143 from the National Heart Institute, National Institutes of Health, Public Health Service.

† Present address: Institute for Cancer Research, Philadelphia 11, Pennsylvania.

* The unit-cell dimensions and density of this second modification, which was originally crystallized from ethanol-water solutions and later from aqueous solutions, were kindly measured by Dr E. L. Eichhorn. The values are:

$$\begin{aligned} a &= 31.49 \pm 0.04, b = 10.15 \pm 0.02, c = 4.838 \pm 0.007 \text{ \AA}; \\ \rho_c &= 1.354 \pm 0.002 \text{ g.cm.}^{-3}. \end{aligned}$$

The density calculated on the basis of two water molecules per molecule of peptide is $1.353 \pm 0.006 \text{ g.cm.}^{-3}$.

Attempts to cleave the small crystals perpendicular to the needle axis were unsuccessful; nevertheless, some Weissenberg photographs were taken with needle crystals rotating about the *b* axis. Because of absorption effects—and, more important, because in certain orientations not all of the crystal was irradiated—the intensity data from these photographs were unreliable; they served mainly as an aid in correlating data obtained from the *c*-axis photographs.

Intensities were estimated visually and corrected for Lorentz and polarization effects. Altogether, 255 non-equivalent *hk0* reflections were observed out of a total of 377 within the copper sphere, 314 *hkl* reflections out of a possible 395, 190 *hk2* reflections out of a possible 360, and 78 *hk3* reflections out of a possible 314. In spite of these restrictions on the number of data available, it seemed worthwhile to pursue the structure determination of the compound.

Determination of the structure

A sharpened three-dimensional Patterson function, from which a spherical origin peak was subtracted, was calculated. The sharpening factor was modified by a function with a maximum at $\sin \theta = 0.5$ so that contributions from reflections with very high and low scattering angles were considerably reduced.

We first attempted a direct analysis of the Patterson map, concentrating particularly on vectors which might determine the positions of the benzene rings. The presence of a high peak along the line $u=0$, $v=\frac{1}{2}$ suggested that the benzene rings might be packed parallel to the *b* axis along the lines $x=0$ and $x=\frac{1}{2}$, and a large number of structures based on this principle were formulated and tested. Although some of the structures were quite close to the correct one, none could be refined. Attempts to determine the precise orientation of the benzene ring from the Patterson peaks near the origin were unsuccessful. Furthermore, the Harker section at $w=\frac{1}{2}$ could not be interpreted, as there were too many non-Harker peaks on or near this section. Finally, after many months of failure, attempts at a direct interpretation of the Patterson map were abandoned.

The key to the solution of the structure was the short *c* axis whose length of 4.9 Å is close to the value 4.85 Å predicted by Pauling and Corey (1953) for the distance between two polypeptide chains hydrogen-bonded together in a parallel-chain pleated-sheet structure. As supporting evidence for a structure based on this configuration, there is a maximum in the Patterson function along the line $u=v=0$ which could be interpreted as the C–O interactions of the two carbonyl groups.

A model of GPG based on the parallel-chain pleated-sheet configuration was built with the backbone of the tripeptide chain extended in a direction perpendicular to *c* and with the planes of the two amide groups parallel to *c*. In such a configuration the length

of the three amino-acid residues is nearly 10 Å (Pauling & Corey, 1953), which suggested that the molecule is extended in the *b* direction. The benzene ring of the phenylalanyl residue would then extend in the *a* direction.

These principles led to a satisfactory trial structure. Preliminary refinements of the *x* and *y* parameters were carried out on the basis of the *hk0* data; approximate *z* parameters were derived from a consideration of the Patterson projection onto (010). An ambiguity concerning the orientation of the benzene ring was resolved and initial refinement of the *z* parameters was undertaken by a number of electron density projections onto (010). Attention was then turned to the three-dimensional data.

Refinement of the structure

All calculations were carried out on a Burroughs 205 digital computer, using a structure-factor least-squares

Table 1. *Calculated hydrogen atom coordinates*
($B=6 \text{ \AA}^2$)

Atom	Attached to	<i>x</i>	<i>y</i>	<i>z</i>
H ₁	C ₂	0.1063	0.569	1.008
H ₂	C ₂	0.1101	0.611	0.686
H ₃	N ₁	0.1254	0.786	1.118
H ₄	C ₄	0.1391	0.993	1.069
H ₅	N ₂	0.1852	1.058	0.571
H ₆	C ₆	0.2768	1.133	0.804
H ₇	C ₆	0.2407	1.182	0.574
H ₈	N ₃	0.2708	1.364	0.745
H ₉	N ₃	0.2663	1.320	1.030
H ₁₀	N ₃	0.2261	1.359	0.874
H ₁₁	C ₇	0.1038	1.170	0.806
H ₁₂	C ₇	0.0976	1.062	0.557
H ₁₃	C ₉	0.0524	1.188	1.143
H ₁₄	C ₁₀	−0.0209	1.131	1.340
H ₁₅	C ₁₁	−0.0601	0.945	1.178
H ₁₆	C ₁₂	−0.0299	0.814	0.822
H ₁₇	C ₁₃	0.0446	0.868	0.631

Table 2. *Atomic coordinates*

Atom	<i>x</i>	<i>y</i>	<i>z</i>	<i>B</i> (Å ²)
C ₁	0.1690	0.5447	0.850	4.7
C ₂	0.1254	0.6153	0.869	4.3
C ₃	0.1366	0.8492	0.755	4.4
C ₄	0.1389	0.9941	0.866	3.8
C ₅	0.2100	1.1126	0.931	4.1
C ₆	0.2480	1.1810	0.772	4.2
C ₇	0.0989	1.0730	0.760	4.6
C ₈	0.0543	1.0319	0.877	4.3
C ₉	0.0353	1.1084	1.074	4.9
C ₁₀	−0.0072	1.0757	1.191	5.8
C ₁₁	−0.0300	0.9674	1.097	5.5
C ₁₂	−0.0126	0.8940	0.892	5.4
C ₁₃	0.0309	0.9236	0.779	5.2
N ₁	0.1286	0.7560	0.946	4.0
N ₂	0.1794	1.0553	0.752	4.0
N ₃	0.2532	1.3192	0.865	4.4
O ₁	0.2051	0.6071	0.881	5.0
O ₂	0.1678	0.4228	0.787	6.1
O ₃	0.1432	0.8240	0.512	5.4
O ₄	0.2064	1.1181	1.176	6.6
O _w	0.1276	0.3115	0.323	7.4

Table 3. Observed and calculated structure factors

The five columns in each group contain the values, reading from left to right, of h , $10|F_o|$, $10|F_c|$, $10A_0$, and $10B_0$. Reflections indicated with an asterisk were given zero weight in the least-squares calculations

h	$10 F_o $	$10 F_c $	$10A_0$	$10B_0$
0	0	0	0	0
1	53	59	53	53
2	28	36	28	28
3	118	142	118	118
4	77	77	77	77
5	35	35	35	35
6	12	12	12	12
7	15	15	15	15
8	18	18	18	18
9	21	21	21	21
10	24	24	24	24
11	27	27	27	27
12	30	30	30	30
13	33	33	33	33
14	36	36	36	36
15	39	39	39	39
16	42	42	42	42
17	45	45	45	45
18	48	48	48	48
19	51	51	51	51
20	54	54	54	54
21	57	57	57	57
22	60	60	60	60
23	63	63	63	63
24	66	66	66	66
25	69	69	69	69
26	72	72	72	72
27	75	75	75	75
28	78	78	78	78
29	81	81	81	81
30	84	84	84	84
31	87	87	87	87
32	90	90	90	90
33	93	93	93	93
34	96	96	96	96
35	99	99	99	99
36	102	102	102	102
37	105	105	105	105
38	108	108	108	108
39	111	111	111	111
40	114	114	114	114
41	117	117	117	117
42	120	120	120	120
43	123	123	123	123
44	126	126	126	126
45	129	129	129	129
46	132	132	132	132
47	135	135	135	135
48	138	138	138	138
49	141	141	141	141
50	144	144	144	144
51	147	147	147	147
52	150	150	150	150
53	153	153	153	153
54	156	156	156	156
55	159	159	159	159
56	162	162	162	162
57	165	165	165	165
58	168	168	168	168
59	171	171	171	171
60	174	174	174	174
61	177	177	177	177
62	180	180	180	180
63	183	183	183	183
64	186	186	186	186
65	189	189	189	189
66	192	192	192	192
67	195	195	195	195
68	198	198	198	198
69	201	201	201	201
70	204	204	204	204
71	207	207	207	207
72	210	210	210	210
73	213	213	213	213
74	216	216	216	216
75	219	219	219	219
76	222	222	222	222
77	225	225	225	225
78	228	228	228	228
79	231	231	231	231
80	234	234	234	234
81	237	237	237	237
82	240	240	240	240
83	243	243	243	243
84	246	246	246	246
85	249	249	249	249
86	252	252	252	252
87	255	255	255	255
88	258	258	258	258
89	261	261	261	261
90	264	264	264	264
91	267	267	267	267
92	270	270	270	270
93	273	273	273	273
94	276	276	276	276
95	279	279	279	279
96	282	282	282	282
97	285	285	285	285
98	288	288	288	288
99	291	291	291	291
100	294	294	294	294
101	297	297	297	297
102	300	300	300	300
103	303	303	303	303
104	306	306	306	306
105	309	309	309	309
106	312	312	312	312
107	315	315	315	315
108	318	318	318	318
109	321	321	321	321
110	324	324	324	324
111	327	327	327	327
112	330	330	330	330
113	333	333	333	333
114	336	336	336	336
115	339	339	339	339
116	342	342	342	342
117	345	345	345	345
118	348	348	348	348
119	351	351	351	351
120	354	354	354	354
121	357	357	357	357
122	360	360	360	360
123	363	363	363	363
124	366	366	366	366
125	369	369	369	369
126	372	372	372	372
127	375	375	375	375
128	378	378	378	378
129	381	381	381	381
130	384	384	384	384
131	387	387	387	387
132	390	390	390	390
133	393	393	393	393
134	396	396	396	396
135	399	399	399	399
136	402	402	402	402
137	405	405	405	405
138	408	408	408	408
139	411	411	411	411
140	414	414	414	414
141	417	417	417	417
142	420	420	420	420
143	423	423	423	423
144	426	426	426	426
145	429	429	429	429
146	432	432	432	432
147	435	435	435	435
148	438	438	438	438
149	441	441	441	441
150	444	444	444	444
151	447	447	447	447
152	450	450	450	450
153	453	453	453	453
154	456	456	456	456
155	459	459	459	459
156	462	462	462	462
157	465	465	465	465
158	468	468	468	468
159	471	471	471	471
160	474	474	474	474
161	477	477	477	477
162	480	480	480	480
163	483	483	483	483
164	486	486	486	486
165	489	489	489	489
166	492	492	492	492
167	495	495	495	495
168	498	498	498	498
169	501	501	501	501
170	504	504	504	504
171	507	507	507	507
172	510	510	510	510
173	513	513	513	513
174	516	516	516	516
175	519	519	519	519
176	522	522	522	522
177	525	525	525	525
178	528	528	528	528
179	531	531	531	531
180	534	534	534	534
181	537	537	537	537
182	540	540	540	540
183	543	543	543	543
184	546	546	546	546
185	549	549	549	549
186	552	552	552	552
187	555	555	555	555
188	558	558	558	558
189	561	561	561	561
190	564	564	564	564
191	567	567	567	567
192	570	570	570	570
193	573	573	573	573
194	576	576	576	576
195	579	579	579	579
196	582	582	582	582
197	585	585	585	585
198	588	588	588	588
199	591	591	591	591
200	594	594	594	594
201	597	597	597	597
202	600	600	600	600
203	603	603	603	603
204	606	606	606	606
205	609	609	609	609
206	612	612	612	612
207	615	615	615	615
208	618	618	618	618
209	621	621	621	621
210	624	624	624	624
211	627	627	627	627
212	630	630	630	630
213	633	633	633	633
214	636	636	636	636
215	639	639	639	639
216	642	642	642	642
217	645	645	645	645
218	648	648	648	648
219	651	651	651	651
220	654	654	654	654
221	657	657	657	657
222	660	660	660	660
223	663	663	663	663
224	666	666	666	666
225	669	669	669	669
226	672	672	672	672
227	675	675	675	675
228	678	678	678	678
229	681	681	681	681
230	684	684	684	684
231	687	687	687	687
232	690	690	690	690
233	693	693	693	693
234	696	696	696	696
235	699	699	699	699
236	702	702	702	702
237	705	705	705	705
238	708	708	708	708
239	711	711	711	711
240	714	714	714	714
241	717	717	717	717
242	720	720	720	720
243	723	723	723	723
244	726	726	726	726
245	729	729	729	729
246	732	732	732	732
247	735	735	735	735
248	738	738	738	738
249	741	741	741	741
250	744	744	744	744
251	747	747	747	747
252	750	750	750	750
253	753	753	753	753
254	756	756	756	756
255	759	759	759	759
256	762	762	762	762
257	765	765	765	765
258	768	768	768	768
259	771	771	771	771
260	774	774	774	774
261	777	777	777	777
262	780	780	780	780
263	783	783	783	783
264	786	786	786	786
265	789	789	789	789
266	792	792	792	792
267	795	795	795	795
268	798	798	798	798
269	801	801	801	801
270	804	804	804	804
271	807	807	807	807
272	810	810</		

program developed by Lavine and Rollett (1956) and a Fourier routine which calculates electron density at intervals of 1/38 or 1/76 of the unit cell. Four parameters for each atom—three positional coordinates and one isotropic temperature factor—were refined. Hydrogen parameters were calculated assuming tetrahedral or plane trigonal configurations and C–H and N–H distances of 1.0 and 0.9 Å, respectively; the hydrogen atoms of the terminal NH_3^+ group were placed in a tetrahedral configuration such that the three N–H...O hydrogen bonds were, on the average, as linear as possible. The hydrogen atoms of the water molecule were not included. The calculated hydrogen parameters are listed in Table 1. These positions were checked on two- and three-dimensional maps and without exception fell on pronounced maxima.

During the course of refinement of the atomic parameters, approximately thirty cycles of structure-factor least-squares or structure-factor difference-map calculations were carried out, many of the cycles being based on data from only one or two layer lines. The primary reason for this inordinate number of cycles lay in the poor quality of the intensity photographs and especially in the lack of adequate data for reflections with high values of l ; as a result, the rate of convergence of the z parameters was quite slow. An additional difficulty arose in the refinement of the parameters of the carbon and oxygen atoms of the two carbonyl groups; within each group the carbon and oxygen atoms have nearly the same x and y parameters and their z parameters differ by almost exactly $\frac{1}{4}$, and the coupling between the indicated parameter shifts of the two atoms in each pair was extensive. This difficulty, too, can be traced to the lack of satisfactory high-layer-line data.

A series of difference projections on to (001) calculated at an intermediate stage of the refinement indicated the presence of approximately one-half a water molecule per asymmetric unit; no further attempt was made to refine the compositional parameter of the water. The contributions of the hydrogen atoms were included at an early stage, but the hydrogen parameters were not adjusted by least-squares. The weighting function of Hughes (1941) was used for the least-squares adjustment of the positional and individual isotropic temperature factors of the heavy atoms; the $hk3$ data, as well as a few low-order reflections which were observed only on the $h0l$ Weissenberg photograph, were given an additional weight of $\frac{1}{2}$. Missing reflections were included in the least squares only if the calculated structure factor was greater than the minimum observable, in which case ΔF was taken as the difference between the two quantities. Scale factors for the various layer lines were adjusted after every few cycles; since the data from the different layer lines were not well correlated, the individual scale factors undoubtedly compensate for thermal anisotropy in the c direction which, in view

of the geometry of the molecule, might be expected to be large.

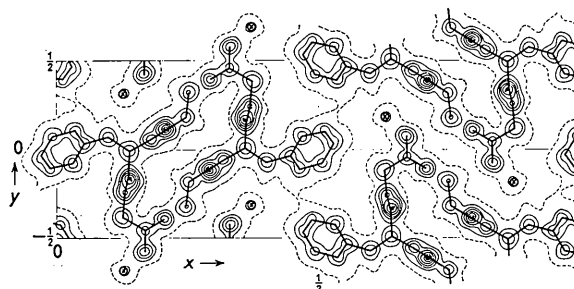


Fig. 1. The final electron-density projection on (001). Contours are at intervals of $2 \text{ e.}\text{\AA}^{-2}$; the dashed line is the $1 \text{ e.}\text{\AA}^{-2}$ contour.

The final positional and temperature-factor parameters for the twenty-one heavy atoms are given in Table 2; the observed and calculated structure factors are listed in Table 3. (The calculated values listed in Table 3 were obtained from the penultimate set of parameters which differed from the final ones by a maximum of 0.007 \AA). An electron density projection on to (001), based on signs for the $hk0$ reflections calculated from the final parameters, is shown in Fig. 1.

Accuracy of the results

The R factors for the individual layer lines are listed in Table 4; the over-all factor for 885 reflections is 0.126. All observed reflections which were given non-zero weight in the least-squares treatment were included in the computation of the R factor; the unobserved reflections for which the calculated structure factor was greater than the threshold value were also included, the observed structure factor being set equal to the threshold value. The principal sources of error between calculated and observed structure factors are undoubtedly two-fold: (1) the poor quality of the data, particularly for the second and third layer lines; (2) the failure to compensate for anisotropic thermal effects.

Table 4. *Final R factors for the individual layer lines*

l	R	ΣF_o	ΣF_c	n
0	0.115	3602	3658	276
1	0.119	3796	3901	331
2	0.123	2242	2241	196
3	0.202	942	917	82
All	0.126	10582	10717	885

The standard deviations in the positional parameters were calculated from the residuals and diagonal elements of the normal equations of the least-squares procedure; they range from 0.006 to 0.010 \AA for the x and y parameters and from 0.013 to 0.022 \AA for the z parameters of the atoms in the tripeptide mole-

cule. The standard deviations in the x , y , and z parameters of the oxygen atom of the water molecule are 0.015, 0.016, and 0.033 Å, respectively. It seems appropriate, then, to quote standard deviations of about 0.02 Å in the bond distances and 1.5° in the bond angles; for reasons discussed earlier, it seems likely that the uncertainties in the distances involving the carbon and oxygen atoms of the carbonyl groups are somewhat larger.

The calculated standard deviations in the temperature-factor parameters B range from 0.14 to 0.24 Å² for the atoms in the main molecule; it is 0.42 Å² for the oxygen atom of the water molecule.

Discussion of the structure

(i) Bond distances and angles

The bond distances and angles calculated from the parameters in Table 2 are shown in Fig. 2 and listed in Table 5. The equations for the best planes of the carboxyl group, of the two peptide groups, and of the benzene ring are given in Table 6, together with the deviations of the individual atoms from these best planes. None of these deviations is significant.

In view of the relatively large uncertainties in the atomic parameters, it is probably not fruitful to undertake a detailed discussion of the individual bond distances and angles. In general, the values are in good agreement with those found in analogous compounds. The average of the five single-bond C-C distances is 1.51 Å, which is smaller than the value 1.54 usually assigned to a C-C single bond but close to the average values found in leucyl-prolyl-glycine (Leung & Marsh, 1958) and in various amino acids (see Donohue & Trueblood, 1952). The average C-C distance within the benzene ring is 1.38 Å, which is

Table 5. Bond distances and angles

Distance		Angle	
C ₁ -C ₂	1.48 Å	O ₁ -C ₁ -O ₂	122°
C ₃ -C ₄	1.55	O ₁ -C ₁ -C ₂	120
C ₅ -C ₆	1.53	O ₂ -C ₁ -C ₂	117
C ₄ -C ₇	1.52	C ₁ -C ₂ -N ₁	115
C ₇ -C ₈	1.50	C ₂ -N ₁ -C ₃	120
		N ₁ -C ₃ -O ₃	124
C ₈ -C ₉	1.35	N ₁ -C ₃ -C ₄	114
C ₉ -C ₁₀	1.42	O ₃ -C ₃ -C ₄	122
C ₁₀ -C ₁₁	1.36	C ₃ -C ₄ -C ₇	109
C ₁₁ -C ₁₂	1.35	C ₃ -C ₄ -N ₂	107
C ₁₂ -C ₁₃	1.44	C ₇ -C ₄ -N ₂	107
C ₁₃ -C ₈	1.37	C ₄ -N ₂ -C ₅	118
		N ₂ -C ₅ -O ₄	126
C ₂ -N ₁	1.46	N ₂ -C ₅ -C ₆	110
C ₄ -N ₂	1.46	O ₄ -C ₅ -C ₆	123
C ₆ -N ₃	1.46	C ₅ -C ₆ -N ₃	110
		C ₄ -C ₇ -C ₈	115
N ₁ -C ₃	1.34	C ₇ -C ₈ -C ₉	119
N ₂ -C ₅	1.39	C ₇ -C ₈ -C ₁₃	122
		C ₈ -C ₉ -C ₁₀	122
C ₁ -O ₁	1.25	C ₉ -C ₁₀ -C ₁₁	119
C ₁ -O ₂	1.26	C ₁₀ -C ₁₁ -C ₁₂	120
		C ₁₁ -C ₁₂ -C ₁₃	121
C ₃ -O ₃	1.23	C ₁₂ -C ₁₃ -C ₈	119
C ₅ -O ₄	1.21	C ₁₃ -C ₈ -C ₉	119

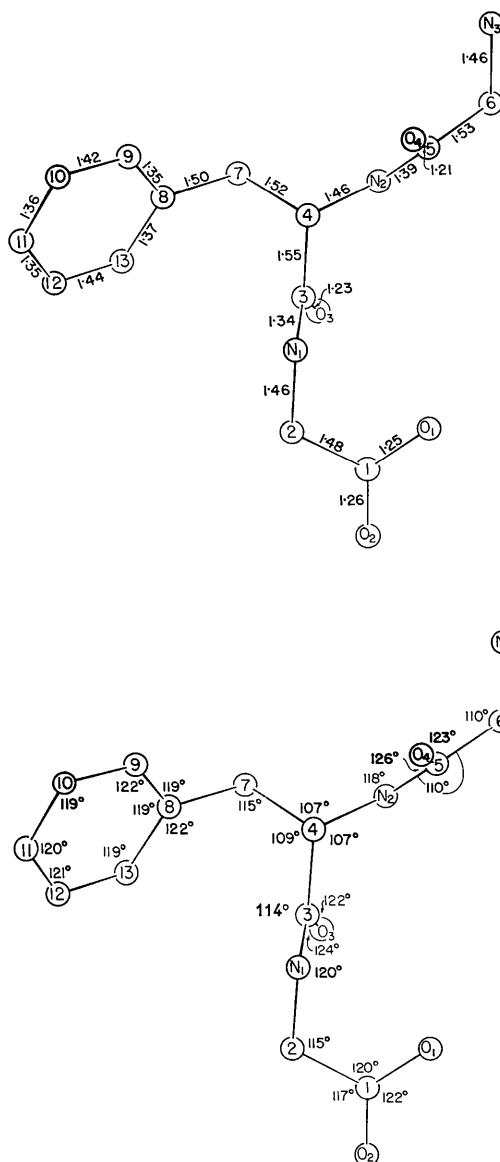


Fig. 2. The bond distances (a) and bond angles (b).

close to the accepted value of 1.39 Å. With the exception of the N₂-C₅ distance, which appears to be about 0.07 Å longer than usual, the distances within the peptide and carboxyl groups are close to the expected values (Pauling & Corey, 1953); the surprising length of the N₂-C₅ bond is, we feel, an artifact associated with the difficulty discussed earlier in positioning the carbon and oxygen atoms of the peptide groups. Additional evidence that the parameters of the atoms in the N-terminal peptide group are especially poorly determined is furnished by the small value of the N₂-C₅-C₆ bond angle and the general lack of coplanarity of the five atoms (Table 6).

(ii) Intermolecular packing

The most important result of this structure inves-

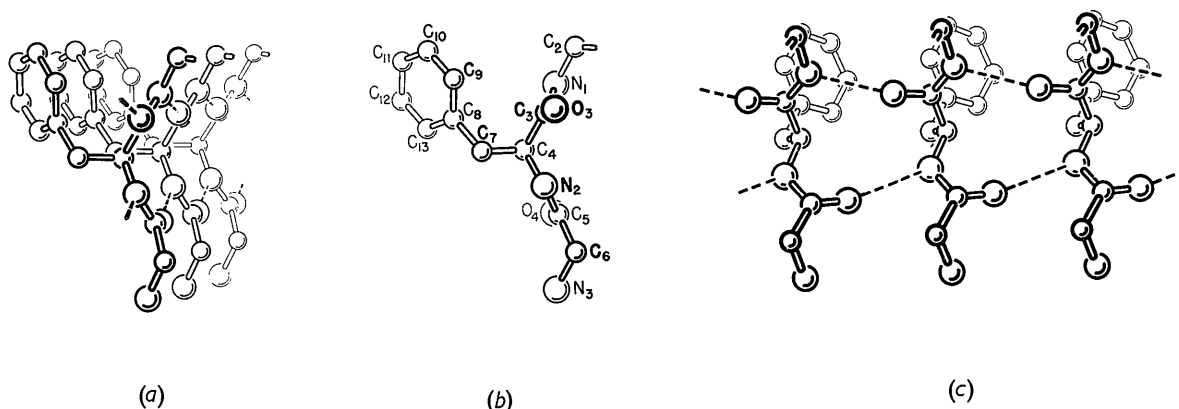


Fig. 3. Drawings showing the pleated-sheet arrangement of GPG molecules. (a) Portions of three molecules viewed at an angle of 15° from the c axis. (b) A single molecule viewed down c . (c) Three molecules viewed along a .

tigation is, we feel, the observation that the parallel-chain pleated sheet configuration of a polypeptide

Table 6. *Best planes through various groups of atoms*

The equations of the planes are expressed in the form $A(a_0x) + B(b_0y) + C(c_0z) = D$, where D is the origin-to-plane distance

Plane	Atom	Deviation	A	B	C	D
Carboxyl group	C ₁	0.020 Å	-0.032	-0.224	0.974	0.121
	C ₂	-0.006				
	O ₁	-0.007				
	O ₂	-0.007				
Phe-gly peptide group	C ₂	-0.004	0.980	-0.102	0.169	3.941
	C ₃	-0.011				
	C ₄	0.000				
	N ₁	0.009				
	O ₃	0.006				
Gly-phe peptide group	C ₄	0.025	-0.503	0.863	-0.051	-2.119
	C ₅	-0.034				
	C ₆	0.027				
	N ₂	-0.025				
	O ₄	0.006				
Benzene ring	C ₇	0.011	0.448	-0.578	0.682	2.324
	C ₈	-0.024				
	C ₉	0.012				
	C ₁₀	-0.002				
	C ₁₁	-0.009				
	C ₁₂	0.021				
	C ₁₃	-0.010				

chain occurs in crystals of a simple peptide. This configuration was first formulated by Pauling & Corey (1951), who later (Pauling & Corey, 1953) modified the atomic parameters and, on the basis of the close agreement between the predicted repeat distance along the polypeptide chain and the observed fiber-axis identity distances in the β -keratin proteins, suggested that it may be the basic structural feature of these proteins. In glycyl-phenylalanyl-glycine we now find direct evidence of the stability of this configuration.

Views of the pleated sheet structure of GPG are shown in Fig. 3. The dimensions of the two-residue fragment occurring in GPG differ somewhat from

those proposed by Pauling and Corey (1953) for the parallel-chain pleated sheet. The distance between C₂ and C₆, which corresponds to the repeat distance in the pleated sheet configuration, is 6.74 Å compared with the proposed value of 6.50 Å; the distance between adjacent chains within the sheet (the c -axis identity distance) is 4.90 Å compared with the proposed value 4.85 Å, and the dihedral angle at C₄ between the chains of the two peptide groups is 126.2° compared with 117.8°. Thus, the peptide chain in GPG is slightly more extended than in the parallel-chain pleated sheet configuration; this extension, which causes some distortion of the N-H...O hydrogen bonds between chains, might easily be accounted for by considering other intermolecular forces, particularly those involving the terminal groups.

Besides the N-H...O hydrogen bonds between peptide groups of adjacent molecules, there is a network of hydrogen bonds between the terminal NH₃⁺ group and the carboxylate ions of adjacent molecules; this network is shown in Fig. 4. The distances and angles involving the hydrogen bonds are listed in Table 7; these values are similar to those found in other peptides and amino acids (see, for example, Fuller, 1959).

The water molecule, when present, is surrounded

Table 7. *Hydrogen-bond distances and angles*

N ₁ ...O ₃	2.89 Å	C ₂ -N ₁ ...O ₃	119°
		C ₃ -N ₁ ...O ₃	119
N ₂ ...O ₄	3.00	C ₄ -N ₂ ...O ₄	132
		C ₅ -N ₂ ...O ₄	110
N ₃ ...O ₁	2.78	C ₆ -N ₃ ...O ₁	124
N ₃ ...O ₁ '	2.91	C ₆ -N ₃ ...O ₁ '	92
N ₃ ...O ₂	2.77	C ₆ -N ₃ ...O ₂	102
		O ₁ ...N ₃ ...O ₁ '	119
		O ₁ ...N ₃ ...O ₂	101
		O ₁ '...N ₃ ...O ₂	115
O _w ...O ₂	2.80	O ₂ ...O _w ...O ₂ '	113
O _w ...O ₂ '	3.09	O ₂ ...O _w ...O ₄	97
O _w ...O ₄	3.12	O ₂ '...O _w ...O ₄	75

